

UNIVERSIDAD DE SALAMANCA
ESCUELA DE DOCTORADO STUDII SALAMANTINI
FORMACIÓN EN LA SOCIEDAD DEL CONOCIMIENTO



**ANÁLISE DE CONFIABILIDADE NOS REPOSITÓRIOS DIGITAIS BRASILEIROS: ESTUDO
DE CASO DA REDE CARINIANA**

MARCOS VINÍCIUS BITTENCOURT DE SOUZA

Salamanca, Espanha
2023

UNIVERSIDAD DE SALAMANCA
ESCUELA DE DOCTORADO STUDII SALAMANTINI
FORMACIÓN EN LA SOCIEDAD DEL CONOCIMIENTO



**ANÁLISE DE CONFIABILIDADE NOS REPOSITÓRIOS DIGITAIS BRASILEIROS: ESTUDO
DE CASO DA REDE CARINIANA**

Tese apresentada à *Escuela de Doctorado Studii Salamantini* da *Universidad de Salamanca*, como requisito parcial para a obtenção do título Doutor em Formação na Sociedade do Conhecimento.

Marcos Vinícius Bittencourt de Souza
Orientadora: Profa. Dra. María Manuela Moro Cabero
Co-Orientador: Prof. Dr. José Luis Bonal Zazo

Salamanca, Espanha
2023

Souza, Marcos Vinícius Bittencourt

ANÁLISE DE CONFIABILIDADE NOS REPOSITÓRIOS DIGITAIS
BRASILEIROS: ESTUDO DE CASO DA REDE CARINIANA / Marcos
Vinícius Bittencourt de Souza.- 2023.

473 p.; 30 cm

Orientadora: María Manuela Moro Cabero

Coorientador: José Luis Bonal Zazo

Tese (doutorado) - Universidad de Salamanca, Escuela de
Doctorado Estudios Salamantini, Formación en la Sociedad del
Conocimiento, 2023.

1. Documentos eletrônicos 2. Normas de auditoria e
certificação 3. Rede Cariniana 4. Repositórios digitais I.
Moro Cabero, María Manuela II. Bonal Zazo, José Luis III.
Título.

Sistema de geração automática de ficha catalográfica da UFSM. Dados fornecidos
pelo autor(a). Sob supervisão da Direção da Divisão de Processos Técnicos da
Biblioteca Central. Bibliotecária responsável Paula Schoenfeldt Patta CRB
10/1728.

Declaro, MARCOS VINÍCIUS BITTENCOURT SOUZA, para os devidos fins e sob as penas da lei, que a pesquisa constante neste trabalho de conclusão de curso (Tese) foi por mim elaborada e que as informações necessárias objeto de consulta em literatura e outras fontes estão devidamente referenciadas. Declaro, ainda, que este trabalho ou parte dele não foi apresentado anteriormente para obtenção de qualquer outro grau acadêmico, estando ciente de que a inveracidade da presente declaração poderá resultar na anulação da titulação pela Universidade, entre outras consequências legais.

**ANÁLISE DE CONFIABILIDADE NOS REPOSITÓRIOS DIGITAIS BRASILEIROS:
ESTUDO DE CASO DA REDE CARINIANA**

Marcos Vinícius Bittencourt de Souza

Tese apresentada à *Escuela de Doctorado Studii Salamantini* da *Universidad de Salamanca*, como requisito parcial para a obtenção do título Doutor em Formação na Sociedade do Conhecimento.

**María Manuela Moro Cabero, Dra. (Universidad de Salamanca)
(Orientadora)**

**José Luis Bonal Zazo Dr. (Universidad de Extremadura)
(Co-Orientador)**

Salamanca, Espanha
2023



Doña María Manuela Moro Cabero, profesora titular de la Universidad de Salamanca y Don José Luis Bonal Zazo, profesor titular de la Universidad de Extremadura, en calidad de directora y codirector del trabajo de tesis que presenta para su defensa don Marcos Vinicius Bittencourt de Souza, titulado *Análise de confiabilidade nos repositórios digitais brasileiros: Estudo de caso da Rede Cariniana*

Autorizan la realización de dicha defensa, al considerar que el trabajo cumple con todos los requisitos de una investigación de tesis doctoral.

En Salamanca, a 10 de noviembre de 2023

Visto Bueno

María Manuela Moro Cabero

José Luis Bonal Zazo

AGRADECIMENTOS

Agradeço a Deus por minha vida, por mais essa oportunidade de evolução. Por nos proteger e guardar durante todo o período da pandemia de COVID-19.

Agradeço à *Universidad de Salamanca*, seus funcionários e corpo docente, pela oportunidade de realizar os estudos de doutorado.

Agradeço à Rede Cariniana, sob a coordenação do Sr. Miguel Angel Márdero Arellano, que propiciou a realização desse estudo com a Rede e não mediu esforços para auxiliar na comunicação com os repositórios.

Agradeço à UFSM por me oportunizar realizar o doutorado com afastamento parcial das minhas atividades. Agradeço à toda a equipe do Centro de Processamento de Dados (CPD) da UFSM pelo apoio nas demandas necessárias para a realização desse trabalho. Agradeço à equipe do Departamento Geral de Arquivo (DAG) da UFSM, em especial à Débora Flores, pelo apoio e incentivo na realização desse trabalho.

Agradeço à Manuela e José Luis por confiarem em mim para a realização desse trabalho, por suas colaborações, contribuições e apoio fundamentais para a realização desse trabalho. Não poderia ter escolhido melhores pessoas para me guiarem nessa jornada.

Agradeço ao professor Daniel Flores pelo incentivo, mentoria e colaboração na escrita dos artigos que contribuíram para esse trabalho.

Agradeço à minha esposa Luciana por seu apoio incondicional durante a realização desse trabalho e por compreender as minhas tantas faltas. Ainda que todos os planos tenham se modificado, ela sempre esteve ao meu lado.

Agradeço à minha filha Maria Valentina por compreender, de maneira tão adulta, as minhas faltas durante a realização desse trabalho. Agradeço ao meu filho Caetano por vir a nós durante esse período.

SUMÁRIO GERAL

LISTA DE FIGURAS	XII
LISTA DE QUADROS E TABELAS.....	XV
LISTA DE SIGLAS E ABREVIATURAS.....	XVIII
0 INTRODUÇÃO À PESQUISA	1
0.1 Contextualização e Justificativa	1
0.2 Definição do Tema, Alcance e Limitações	8
0.3 Objetivos.....	10
0.4 Hipóteses	12
0.5 Estrutura do Texto	14
 PARTE I. MARCO TEÓRICO: A PRESERVAÇÃO DOS OBJETOS DIGITAIS	
1 A PRESERVAÇÃO DOS OBJETOS DIGITAIS: DESAFIOS E FERRAMENTAS	16
1.1 Introdução.....	16
1.2 O Documento Digital como Objeto de Preservação.....	18
1.2.1 A Perspectiva Tradicional	18
1.2.2 A Perspectiva Digital	22
1.2.3 Análise Desde uma Perspectiva Integral.....	36
1.3 Os Desafios da Preservação de Documentos Digitais	42
1.3.1 O Desafio de Preservar as Propriedades dos Documentos	42
1.3.2 O Desafio de Escolha das Estratégias de Preservação	52
1.3.3 O Desafio de Dispor de uma Infraestrutura Institucional de Preservação	56
1.3.4 O Desafio de Adaptar os Processos de Gestão Documental	57
1.3.5 O Desafio de Garantir a Cadeia de Custódia e a Cadeia de Preservação.....	61
1.3.6 O Desafio de Preservar Recursos Especiais.....	63
1.4 As Normas como Ferramentas para Enfrentar os Desafios da Preservação	67
1.4.1 Um Marco Normativo Complexo	67
1.4.2 ISO 15489.....	70
1.4.3 ISO 30300.....	73
1.4.4 e-Arq Brasil	81
1.4.5 ISO 23081.....	85
1.4.6 PREMIS	87
2 O REPOSITÓRIO DIGITAL: NOVO DEPÓSITO PARA PROBLEMAS HETEROGÊNEOS NA TAREFA DE PRESERVAR CONTEÚDOS DIGITAIS.....	95

2.1	Introdução.....	95
2.2	Dos Arquivos aos Repositórios Digitais.....	96
2.3	O Modelo de Referência OAIS	105
2.3.1	A Origem do Modelo	106
2.3.2	Conceitos e Requisitos Básicos do Modelo	106
2.3.3	Arquitetura do Modelo.....	109
2.4	Exemplos de Ferramentas para Repositórios.....	119
2.4.1	FEDORA	119
2.4.2	Archivematica.....	120
2.4.3	RODA	122
2.4.4	DSpace	124
2.5	A Experiência do Brasil: Repositório Arquivístico Digital Confiável – RDC-Arq	126
2.6	A Contribuição dos Repositórios para o Acesso Aberto	129
2.6.1	Os Fundamentos do Acesso Aberto	129
2.6.2	Iniciativas Institucionais	134
3	AVALIAR E CERTIFICAR A CONFIABILIDADE NOS REPOSITÓRIOS DIGITAIS	137
3.1	Introdução.....	137
3.2	Avaliar e Alcançar a Confiança na Preservação Digital	138
3.2.1	O Conceito de Confiança.....	138
3.2.2	Repositórios Digitais de Confiáveis	143
3.2.3	Auditorias e Certificações: Mecanismos de Avaliação para Assegurar a Confiança	147
3.3	A Linha do Tempo dos Métodos e Ferramentas de Auditoras e Certificações	148
3.3.1	O Período de 1996 a 2010.....	149
3.3.2	O Período de 2011 a 2025.....	160
3.4	Identificação e Caracterização de Ferramentas de Avaliação	171
3.4.1	FAIR	171
3.4.2	NDSA Levels of Digital Preservation.....	173
3.4.3	ISO 16363.....	179
3.4.4	ISO 17068.....	183
3.4.5	Nestor Seal.....	185
3.4.6	CoreTrust Seal	189
3.4.7	DRAMBORA	194
3.4.8	Guia de Avaliação de Repositórios Institucionais de Investigação REBIUN	201
 PARTE II. ESTUDO APLICADO: AVALIAÇÃO DE REPOSITÓRIOS DIGITAIS DA REDE CARINIANA		
4	METODOLOGIA.....	206
4.1	Abordagem Metodológica	206
4.2	Fases de Desenvolvimento do Trabalho	210
4.2.1	Fase 1: Seleção da População Objeto de Avaliação Exploratória: Os Repositórios da Rede Cariniana	210
4.2.1.1	Distribuição Geográfica das Instituições por Regiões.....	214
4.2.1.2	Natureza das Instituições	215
4.2.1.3	População Geral dos Repositórios de Instituições de Ensino Superior	215
4.2.2	Fase 2: Seleção de Fontes e Identificação de Critérios de Avaliação para a Certificação de Repositórios Confiáveis	218

4.2.2.1	Seleção de Fontes Normativas.....	219
4.2.2.2	Identificação, Seleção e Classificação de Critérios	221
4.2.3	Fase 3: Elaboração do Formulário de Coleta de Dados	237
4.2.3.1	Área 0: Informações Sobre o Respondente	239
4.2.3.2	Área 1: Contextualização do Repositório	240
4.2.3.3	Área 2: Estratégia e Responsabilidade	242
4.2.3.4	Área 3: Certificações e Normas.....	245
4.2.3.5	Área 4: Infraestrutura Organizacional	247
4.2.3.6	Área 5: Infraestrutura Técnica e Tecnológica.....	252
4.2.3.7	Área 6: Segurança	256
4.2.3.8	Área 7: Integridade dos Dados	258
4.2.3.9	Área 8: Metadados.....	261
4.2.3.10	Área 9: Acesso e Consulta.....	265
4.2.3.11	Área 10: Visibilidade.....	268
4.2.4	Fase 4: Coleta de Dados.....	270
4.2.5	Fase 5: Elaboração da Ferramenta de Pontuação.....	272
4.2.6	Fase 6: Análise de Dados	274
5	RESULTADOS	277
5.1	Análise do Respondente	277
5.2	Análise Global dos Repositórios	279
5.3	Resultados Gerais por Áreas.....	282
5.4	Resultados Específicos por Áreas.....	287
5.4.1	Resultados da Área 1: Contextualização do Repositório	289
5.4.2	Resultados da Área 2: Estratégia e Responsabilidade	296
5.4.3	Resultados da Área 3: Certificações e Normas.....	302
5.4.4	Resultados da Área 4: Infraestrutura Organizacional	307
5.4.5	Resultados da Área 5: Infraestrutura Técnica e Tecnológica	316
5.4.6	Resultados da Área 6: Segurança.....	326
5.4.7	Resultados da Área 7: Integridade dos Dados	332
5.4.8	Resultados da Área 8: Metadados.....	339
5.4.9	Resultados da Área 9: Acesso e Consulta	344
5.4.10	Resultados da Área 10: Visibilidade dos Repositórios	353
5.5	Análises de Fatores Externos.....	360
5.5.1	Resultados por Região Geográfica Brasileira	360
5.5.2	Resultados por Tipo de Parceria	367
5.5.3	Resultados por Tipo de Instituição.....	368
5.5.4	Resultados por Tempo de Criação dos Repositórios.....	370
5.5.5	Resultados dos Repositórios Universitários.....	372
5.5.5.1	Por Orçamento.....	372
5.5.5.2	Por Tamanho das Comunidades	377
5.5.5.3	Por Ranking do MEC	381
5.5.5.4	Por Ranking do SCImago.....	382
5.6	Avaliação Geral dos Resultados	385
6	CONCLUSÕES	391
6.1	Nível de Cumprimento dos Objetivos	392

6.2	Avaliação das Hipóteses	395
6.3	Conclusões Gerais e Recomendações.....	398
6.4	Linhas Futuras de Pesquisa.....	406
REFERÊNCIAS		409
ANEXO A		431
ANEXO B		445

LISTA DE FIGURAS

Figura 1 – Conceitos da ISO 30300 relacionados com documento.	35
Figura 2 – Diagrama de propriedades relacionadas ao Documento Digital.	51
Figura 3 – Estratégias de preservação.	56
Figura 4 – Conceitos relacionados aos instrumentos de gestão de documentos.	72
Figura 5 – Modelo de dados PREMIS	89
Figura 6 – Registro do evento de modificação de nome de arquivo.	93
Figura 7 – Modelo de ambiente de um repositório OAIS.	110
Figura 8 – Estrutura do Pacote de Informação e seus relacionamentos.	112
Figura 9 – Modelo funcional de entidades de um arquivo OAIS.	114
Figura 10 – Linha do tempo da preservação digital de 1996 a 2010.	149
Figura 11 – Linha do tempo da preservação digital (a partir de 2011).	161
Figura 12 – Matriz de Níveis de Preservação NDSA.	178
Figura 13 – Totais de critérios atendidos pelos repositórios REBIUN.	204
Figura 14 – Distribuição das instituições parceiras da Rede Cariniana pelas regiões brasileiras.	214
Figura 15 – Natureza das instituições parceiras da Rede Cariniana.	215
Figura 16 – Público geral das IES parceiras da Rede Cariniana distribuído pelas regiões brasileiras.	218
Figura 17 – Cobertura de critérios	235
Figura 18 – Cargo ou função desempenhada pelo respondente no repositório	278
Figura 19 – Formação acadêmica dos respondentes	278
Figura 20 – <i>Ranking</i> geral dos repositórios pelas médias de pontuações.	281
Figura 21 - <i>Ranking</i> das áreas pela Média Aritmética de pontuação.	283
Figura 22 – Pontuações médias, Desvios Médios e Desvios Padrão das áreas avaliadas.	288
Figura 23 - Exemplo de classificação dos repositórios de acordo com as suas pontuações nas áreas.	289
Figura 24 – Pontuações individuais dos repositórios na Área 1: Contextualização do repositório.	290
Figura 25 – Repositórios com missão, visão e valores declarados.	291
Figura 26 – Utilização de termos para a definição da missão do repositório.	293
Figura 27 – Delimitação da comunidade-alvo dos repositórios.	294
Figura 28 – Natureza dos materiais custodiados pelos repositórios.	296
Figura 29 – Pontuações individuais dos repositórios na Área 2: Estratégia e Responsabilidade.	297
Figura 30 – Existência de Política de Preservação nos repositórios.	298
Figura 31 – Presença de Plano de Continuidade nos repositórios.	300
Figura 32 – Pontuações individuais dos repositórios na Área 3: Certificações e Normas.	302
Figura 33 – Interesse dos repositórios em obter certificações.	303
Figura 34 – Tipos de certificações de interesse dos repositórios.	304

Figura 35 – Adoção de certificações e normas pelos repositórios.	306
Figura 36 – Pontuações individuais dos repositórios na Área 4: Infraestrutura Organizacional.....	308
Figura 37 – Adequação da estrutura organizacional dos repositórios.	309
Figura 38 – Adequação dos orçamentos dos repositórios à preservação de longo prazo.....	310
Figura 39 – Adequação da quantidade de pessoas nos repositórios.	311
Figura 40 – Adequação de capacitações e conhecimentos das equipes dos repositórios	314
Figura 41 – Existência de um plano de capacitação nos repositórios.	314
Figura 42 – Pontuações individuais dos repositórios na Área 5: Infraestrutura Técnica e Tecnológica.	317
Figura 43 – Adequação da estrutura física dos repositórios.....	318
Figura 44 – Adequação da infraestrutura técnica e tecnológica dos repositórios.	319
Figura 45 – Existência de um plano de atualização de <i>hardware</i> e <i>software</i> nos repositórios.....	320
Figura 46 – Plataformas (<i>softwares</i>) utilizados pelos repositórios.....	321
Figura 47 – Pontuações individuais dos repositórios na Área 6: Segurança.....	326
Figura 48 – Análise dos riscos da preservação digital.	328
Figura 49 – Adoção de medidas de segurança contra ataques.	329
Figura 50 – Adoção de medidas na estrutura física contra acessos indevidos.	331
Figura 51 – Pontuações individuais dos repositórios na Área 7: Integridade dos Dados.....	332
Figura 52 – Monitoramento e tratamento dos formatos de arquivos.	334
Figura 53 – Adoção de medidas contra o corrompimento de dados.	335
Figura 54 – Presença de planos de <i>backup</i> nos repositórios.	336
Figura 55 – Estratégias de <i>backup</i> adotadas pelos repositórios.	337
Figura 56 – Pontuações individuais dos repositórios na Área 8: Metadados.....	340
Figura 57 – Adoção de um padrão de metadados na ingestão dos documentos.....	341
Figura 58 – Recolhimento de metadados de contexto dos documentos.....	342
Figura 59 – Pontuações individuais dos repositórios na Área 9: Acesso e Consulta.	345
Figura 60 – Existência de política de acesso nos repositórios.	346
Figura 61 – Existência de um plano de monitoramento das necessidades das Comunidades-Alvo.....	348
Figura 62 – Frequência de formatos de arquivos disponibilizados	350
Figura 63 – Pontuações individuais dos repositórios na Área 10: Visibilidade.	354
Figura 64 – Acesso ao repositório a partir da página institucional.	355
Figura 65 – Existência de iniciativas de promoção dos repositórios.	356
Figura 66 – Iniciativas de promoção dos repositórios.	357
Figura 67 – Formas de contato dos repositórios	358
Figura 68 - Pontuações absolutas e médias entre as regiões geográficas brasileiras.	361
Figura 69 – Regiões geográficas frente às maiores e menores médias.	366
Figura 70 – <i>Ranking</i> dos tipos de parceria com a Rede Cariniana.....	368
Figura 71 – Média de pontuação dos repositórios em relação ao ano de criação.	371
Figura 72 – Orçamento das universidades com repositórios.	374
Figura 73 – Pontuação média dos repositórios por faixa de orçamento.....	375
Figura 74 – Pontuações médias das regiões em relação aos orçamentos médios.	377
Figura 75 – Pontuação média dos repositórios e suas populações totais.	379
Figura 76 – Orçamento per capita dos repositórios e pontuações totais.	380

Figura 77 – Pontuação média dos repositórios pelo <i>ranking</i> MEC das suas instituições.	382
Figura 78 – Pontuação média dos repositórios de acordo com suas colocações no SIR.....	384
Figura 79 – Pontuações totais dos repositórios e suas colocações no SIR.....	385
Figura 80 – Instrumento quantitativo de avaliação dos repositórios com perguntas, respostas e pontuações.	446

LISTA DE QUADROS E TABELAS

Quadro 1 – Definições dos termos relacionados a documentos.....	38
Quadro 2 – Requisitos do Anexo A da norma ISO30301:2019 referentes aos processos de gestão.	76
Quadro 3 – Requisitos do Anexo A da norma ISO30301:2019 referentes aos instrumentos utilizados na gestão documental.....	78
Quadro 4 – Requisitos do Anexo A da norma ISO30301:2019 referentes às aplicações de gestão documental.....	79
Quadro 5 – Trecho do Dicionário de Dados para a unidade semântica “size”.....	91
Quadro 6 – Elementos de avaliação de maturidade do DPC RAM.	166
Quadro 7 – Ferramentas de auditoria, certificação e maturidade de repositórios digitais.....	170
Quadro 8 – Atributos dos princípios FAIR.	172
Quadro 9 – Tabela de pontuação Nestor Seal.	188
Quadro 10 – Perguntas de contextualização do repositório do critério CoreTrust Seal R0.	190
Quadro 11 – Respostas aos critérios das perguntas da certificação CoreTrust Seal.....	192
Quadro 12 – Requisitos CoreTrust Seal.....	192
Quadro 13 – Pontuação DRAMBORA de probabilidade de ocorrência de um risco.	198
Quadro 14 – Pontuação DRAMBORA de impacto dos riscos.....	198
Quadro 15 – Áreas e evolução de critérios do Guia	201
Quadro 16 – Áreas e critérios de avaliação dos repositórios REBIUN.....	202
Quadro 17 – Instituições integrantes da Rede Cariniana.	213
Quadro 18 – Seleção de critérios de avaliação e suas áreas funcionais.	222
Quadro 19 – Análise de critérios e similaridades da área “1) Contexto Geral”.	223
Quadro 20 – Análise de critérios e similaridades da área “2) Estratégia e responsabilidade”.	225
Quadro 21 – Análise de critérios e similaridades da área “3) Infraestrutura organizacional”.....	226
Quadro 22 – Análise de critérios e similaridades da área “4) Infraestrutura física e técnica”.	228
Quadro 23 – Análise de critérios e similaridades da área “5) Segurança, integridade e autenticidade dos dados”.	230
Quadro 24 – Análise de critérios e similaridades da área “6) Metadados”.	233
Quadro 25 – Análise de critérios e similaridades da área “7) Acesso e visibilidade”.	234
Quadro 26 – Apresentação do questionário de avaliação.....	238
Quadro 27 – Resumo das seções do questionário.....	239
Quadro 28 – Perguntas da área 0) Informações sobre o respondente.	239
Quadro 29 – Perguntas da área 1) Contextualização do repositório.	242
Quadro 30 – Perguntas da área 2) Estratégia e responsabilidade.....	244
Quadro 31 – Respostas das perguntas da Área 2) Estratégia e responsabilidade.....	245
Quadro 32 – Perguntas da área 3) Certificações e normas.....	247
Quadro 33 – Perguntas da área 4) Infraestrutura organizacional.	252

Quadro 34 – Perguntas da área 5) Infraestrutura técnica e tecnológica.	256
Quadro 35 – Perguntas da área 6) Segurança.	258
Quadro 36 – Perguntas da área 7) Integridade dos dados.	261
Quadro 37 – Perguntas da área 8) Metadados.....	265
Quadro 38 – Perguntas da área 9) Acesso e consulta.....	268
Quadro 39 – Perguntas da seção 10) Visibilidade.....	270
Quadro 40 - Exemplo de pontuação de perguntas das áreas de avaliação.	273
Quadro 41 – Exemplo de cálculos de MA, DM, DP e CV.	276
Quadro 42 – Repositórios, suas pontuações totais e Médias Aritméticas.	280
Quadro 43 – Classificação dos repositórios pelos seus desempenhos gerais.....	282
Quadro 44 – Pontuações dos repositórios por área.	285
Quadro 45 – Áreas de avaliação e pontuações totais.	286
Quadro 46 – Pontuações e indicadores gerais das áreas avaliadas.	287
Quadro 47 – Classificação dos repositórios pelos seus desempenhos na Área 1: Contextualização do repositório	291
Quadro 48 – Classificação dos repositórios pelos seus desempenhos na Área 2: Estratégia e Responsabilidade.	298
Quadro 49 – Classificação dos repositórios pelos seus desempenhos na Área 3: Certificações e Normas.	303
Quadro 50 – Classificação dos repositórios pelos seus desempenhos na Área 4: Infraestrutura Organizacional.	308
Quadro 51 – Proporção de funcionários em relação ao público total das IES.	313
Quadro 52 – Classificação dos repositórios pelos seus desempenhos Área 5: Infraestrutura Técnica e Tecnológica. ...	317
Quadro 53 – Classificação dos repositórios pelos seus desempenhos Área 6: Segurança.....	327
Quadro 54 – Classificação dos repositórios pelos seus desempenhos Área 7: Integridade dos Dados.....	333
Quadro 55 - Cumprimento de critérios da área funcional “Armazenamento”.	339
Quadro 56 – Classificação dos repositórios pelos seus desempenhos Área 8: Metadados.	340
Quadro 57 – Cumprimento de critérios da área funcional “Metadados”.	344
Quadro 58 – Classificação dos repositórios pelos seus desempenhos Área 8: Metadados.	346
Quadro 59 – Classificação dos repositórios pelos seus desempenhos Área 10: Visibilidade.	354
Quadro 60 – Cumprimento de critérios de visibilidade dos repositórios.	360
Quadro 61 - Desempenho dos repositórios por área e região geográfica brasileira.	362
Quadro 62 - Maiores e menores médias por área e região geográfica.	364
Quadro 63 – Pontuações dos repositórios por tipo de parceria com a Rede Cariniana.	367
Quadro 64 – Repositórios, seus tipos de instituição e pontuações totais.	369
Quadro 65 – Pontuações totais dos repositórios por tipo de instituição.....	369
Quadro 66 – Repositórios e seus anos de criação.	370
Quadro 67 – Repositórios, seus respectivos orçamentos e pontuações totais.	373
Quadro 68 – Orçamentos das instituições por região geográfica brasileira e suas pontuações médias.	376
Quadro 69 - Repositórios, tamanhos das suas comunidades e pontuações totais.	378
Quadro 70 – Repositórios, seus IGC (<i>ranking</i> do MEC) e pontuações totais.....	381
Quadro 71 – Repositórios, ranking SIR, faixa de avaliação e pontuações totais.	383
Quadro 72 – Desempenho dos repositórios em relação ao alcance efetivo ou em teoria dos requisitos das perguntas..	387
Quadro 73 – Cumprimento dos objetivos específicos do Objetivo Geral 1.....	392
Quadro 74 – Cumprimento dos objetivos específicos do Objetivo Geral 2.....	393
Quadro 75 – Cumprimento dos objetivos específicos do Objetivo Geral 3.....	394

Quadro 76 – Hipóteses e suas avaliações.	395
Quadro 77 – Questionário completo de avaliação dos repositórios com perguntas e justificativas.	432
Tabela 1 - Dados das Instituições de Ensino Superior (IES) parceiras da Rede Cariniana.....	217

LISTA DE SIGLAS E ABREVIATURAS

ABNT	Associação Brasileira de Normas Técnicas
ACTDR	<i>Audit and Certification of Trustworthy Digital Repositories</i>
AENOR	<i>Asociación Española de Normalización y Certificación</i>
AIP	<i>Archival Information Package</i>
CCD	Código de Classificação de Documentos
CCDA	Cadeia de Custódia Digital Arquivística
CCSDS	<i>Consultative Committee for Space Data Systems</i>
CIA	Conselho Internacional de Arquivos
CMS	<i>Content Management System</i>
CNEDA	<i>Comisión de Normas Españolas de Descripción Archivística</i>
COAR	<i>Confederation of Open Access Repositories</i>
CONARQ	Conselho Nacional de Arquivos
CRC	<i>Cyclic Redundancy Check</i>
CriB	<i>Conversion and Recommendation of Digital Objects Format</i>
CRL	<i>Center for Research Libraries</i>
CRUE	<i>Conferencia de Rectores de las Universidades Españolas</i>
CSS	<i>Cascading Style Sheets</i>
CTNDA	Câmara Técnica de Descrição Arquivística
CV	Coeficiente de Variação
DANS	<i>Data Archiving and Networked Services</i>
DCC	<i>Digital Curation Centre</i>
DGLAB	Direção Geral do Livro, dos Arquivos e das Bibliotecas
DIN	<i>Deutsches Institut für Normung</i>
DIP	<i>Dissemination Information Package</i>
DM	Desvio Médio
DP	Desvio Padrão
DPC	<i>Digital Preservation Coalition</i>
DPC-RAM	<i>Digital Preservation Coalition Rapid Assessment Model</i>
DPE	<i>DigitalPreservationEurope</i>
DRAMBORA	<i>Digital Repository Audit Method Based on Risk Assessment</i>
DRIVER	<i>Digital Repository Infrastructure Vision for European Research</i>
DSA	<i>Data Seal of Approval</i>
EAD	<i>Encoded Archival Description</i>
EGAD	<i>Experts Group on Archival Description</i>
FAIR	<i>Findable, Accessible, Interoperable, Recoverable</i>
FECYT	<i>Fundación Española para la Ciencia y Tecnología</i>
FEDORA	<i>Flexible Extensible Digital Object Repository Architecture</i>

GPDR	<i>General Data Protection Regulation</i>
HDD	<i>Hard Disk Drive</i>
HTML	<i>Hypertext Markup Language</i>
HTTP	<i>Hypertext Transfer Protocol</i>
IBICT	Instituto Brasileiro de Informação em Ciência e Tecnologia
ICA	<i>International Council on Archives</i>
ICS-WDS	<i>International Council for Science's World Data System</i>
IES	Instituição de Ensino Superior
IGC	Índice Geral de Cursos
INCCA	<i>International Network for the Conservation of Contemporary Art</i>
IP	<i>Information Packages</i>
ISAAR(CPF)	<i>International Standard Archival Authority Record for Corporate Bodies, Persons and Families</i>
ISAD(G)	<i>General International Standard Archival Description</i>
ISDF	<i>International Standard for Describing Functions</i>
ISDIAH	<i>International Standard for Describing Institutions with Archival Holdings</i>
ISO	<i>International Organization for Standardization</i>
JISC	<i>Joint Information Systems Committee</i>
KNAW	<i>Royal Netherlands Academy of Arts and Sciences</i>
LGPD	Lei Geral de Proteção de Dados
MA	Média Aritmética
MEC	Ministério da Educação e Cultura
METS	<i>Metadata Encoding and Transmission Standard</i>
MSR	<i>Management Systems for Records</i>
MSS	<i>Management System Standards</i>
NARA	<i>National Archives and Records Administration</i>
NARS	<i>National Archives and Records Services</i>
NDSA	<i>National Digital Stewardship Alliance</i>
NESTOR	<i>Network of Expertise in Long-term STorage of Digital Resources</i>
NOBRADE	Norma Brasileira de Descrição Arquivística
NTI	Norma Técnica de Interoperabilid
NWO	<i>Dutch Research Council</i>
OAI-PMH	<i>Open Archives Initiative Protocol for Metadata Harvesting</i>
OAIS	<i>Open Archival Information System</i>
OCLC	<i>Online Computer Library Center</i>
OECD	<i>Organization for Economic Co-operation and Development</i>
OpenDOAR	<i>Directory of Open Access Repositories</i>
OSTP	<i>Office of Science and Technology Policy</i>
PAI	Pacote de Arquivamento de Informação
PCD	Plano de Classificação de Documentos
PDI	Pacote de Disseminação de Informação
PDI	<i>Preservation Description Information</i>
PLATTER	<i>Planning Tool for Trusted Electronic Repositories</i>
PREMIS	<i>Preservation Metadata: Implementation Strategies</i>
RAM	Random Access Memory
RDC-Arq	Repositório Arquivístico Digital Confiável

REBIUN	<i>Red de Bibliotecas Universitarias</i>
RECOLECTA	<i>Recolector de Ciencia Abierta</i>
RiC	<i>Records in Contexts</i>
RLG	<i>Research Libraries Group</i>
RODA	Repositório de Objetos Digitais Autênticos
RTC	Repositório de Terceiros Confiável
SAAI	Sistemas Abertos de Arquivamento de Informações
SGD	Sistemas Gerenciadores de Documentos
SGDA	Sistemas de Gestão de Documentos de Arquivo
SIGA	Sistema de Gestão de Documentos de Arquivo
SINAR	Sistema Nacional de Arquivos
SIP	<i>Submission Information Package</i>
SIR	<i>SCImago Institutions Ranking</i>
SLA	<i>Service Level Agreement</i>
SMART	<i>Specific Measurable Attainable Realistic Time-related</i>
SOP	<i>Strategic Objective Plans</i>
SSD	<i>Solid State Drive</i>
TDR	<i>Trusted Digital Repositories, Attributes and Responsibilities</i>
TI	Tecnologia da Informação
TRAC	Trustworthy Repositories Audit & Certification, Criteria and Checklist
TRUST	Transparency, Responsibility, User focus, Sustainability and Technology
TTD	Tabela de Temporalidade de Documentos
TTPR	Trusted Third Party Repository
UC3	University of California Curation Center
URL	Uniform Resource Locator
WDS	World Data System
XML	eXtensible Markup Language

RESUMO

ANÁLISE DE CONFIABILIDADE NOS REPOSITÓRIOS DIGITAIS BRASILEIROS: ESTUDO DE CASO DA REDE CARINIANA

AUTOR: Marcos Vinícius Bittencourt de Souza

DIRETORA: María Manuela Moro Cabero

CO-DIRETOR: José Luis Bonal Zazo

A tarefa de preservar materiais digitais traz consigo uma série de desafios a serem superados para que esses materiais permaneçam autênticos e utilizáveis em um momento futuro, muitas vezes inalcançável à imaginação humana. Dentre esses desafios está a superação da obsolescência tecnológica, própria desse ambiente, a manutenção da interpretabilidade e autenticidade dos materiais ao longo do tempo, necessitando de planejamentos, estruturas organizacionais e tecnológicas adequadas. Assim, se utilizam repositórios digitais cuja gestão engloba desde aspectos organizacionais até a utilização de tecnologias voltadas a esse fim, tendo as suas estruturas baseadas em normas e padrões nacionais e internacionais. A justificativa para a realização do presente estudo recai no fato de se atestar que a preservação digital dos repositórios institucionais da Rede Cariniana possui as características necessárias para a preservação plena e confiável. A partir desse ponto, elaboraram-se hipóteses para a comprovação da presença das estruturas necessárias, existência de diretrizes de regulação da preservação, sobre a existência de avaliações de confiabilidade e o grau de utilidade de uma ferramenta de avaliação de qualidade dos repositórios baseada em padrões e normas internacionais. Dentre os objetivos propostos, está a identificação do estado da preservação digital dos repositórios institucionais, apresentando melhorias a serem aplicadas, o estudo de ferramentas e métodos de avaliação de confiabilidade de repositórios e a elaboração de um instrumento de avaliação de critérios de qualidade para aplicação em contexto nacional. De uma forma geral, a metodologia do trabalho possui natureza exploratória e descritiva, utilizando um questionário elaborado para aplicação junto aos repositórios da Rede Cariniana, tendo sido feita a análise dos resultados de forma qualitativa e quantitativa. Os resultados colhidos permitiram identificar que os repositórios participantes do estudo realizam a preservação digital dos materiais de forma empírica, sem embasamentos em diretrizes apropriadas, focando em conhecimentos específicos das equipes sobre as suas comunidades. Constata-se ainda a inexistência de processos periódicos de avaliação dos repositórios, justificando a criação de uma ferramenta específica, quantitativa, para aplicação nos repositórios. A validação dessa ferramenta se deu através das análises realizadas sobre as informações recolhidas no questionário de avaliação, sendo possível utilizá-la para futuras novas análises e disponibilização aos próprios repositórios como forma de realizarem autoavaliações.

Palavras-chave: Documentos eletrônicos. Normas de auditoria e certificação. Rede Cariniana. Repositórios digitais.

ABSTRACT

RELIABILITY ANALYSIS OF BRAZILIAN DIGITAL REPOSITORIES: *REDE CARINIANA* CASE STUDY

AUTHOR: Marcos Vinícius Bittencourt de Souza

ADVISOR: María Manuela Moro Cabero

CO-ADVISOR: José Luis Bonal Zazo

The task of preserving digital materials brings with it a series of challenges to be overcome so that these materials remain authentic and usable at a future time, often unreachable to the human imagination. Among these challenges is overcoming technological obsolescence, which is typical of this environment, and maintaining the interpretability and authenticity of the materials over time, requiring appropriate planning, organizational and technological structures. Thus, digital repositories are used that encompass everything from organizational aspects to the use of technologies for this purpose, with their structures based on national and international norms and standards. The justification for carrying out this study lies in the fact that the digital preservation of the institutional repositories of the *Rede Cariniana* has the necessary characteristics for full and reliable preservation. From this point, hypotheses were developed to prove the presence of the necessary structures, the existence of preservation regulation guidelines, the existence of reliability assessments and the usefulness of a repository quality assessment tool based on international standards and norms. Among the proposed objectives is the identification of the state of digital preservation of institutional repositories, presenting improvements to be applied, the study of tools and methods for evaluating the reliability of repositories and the development of an instrument for evaluating quality criteria for application in a national context. In general, the methodology of the work is exploratory and descriptive in nature, using a survey designed to be applied to the repositories of the *Rede Cariniana*, and the results were analyzed qualitatively and quantitatively. The results showed that the repositories participating in the study carry out the digital preservation of materials in an empirical way, without any basis in appropriate guidelines, focusing on the specific knowledge of the teams about their communities. There is also no periodic process for evaluating the repositories, justifying the creation of a specific, quantitative tool to be applied to the repositories. This tool was validated through analysis of the information collected in the evaluation survey, and can be used for future new analyses and made available to the repositories themselves as a way of carrying out self-evaluations.

Keywords: Electronic records. Audit and certification standards. *Rede Cariniana*. Digital repositories.

RESUMEN

ANÁLISIS DE LA FIABILIDAD DE LOS REPOSITARIOS DIGITALES BRASILEÑOS: ESTUDIO DE CASO DE LA *REDE CARINIANA*

AUTOR: Marcos Vinícius Bittencourt de Souza

DIRECTORA: María Manuela Moro Cabero

CO-DIRECTOR: José Luis Bonal Zazo

La tarea de preservar los materiales digitales conlleva una serie de retos que hay que superar para que estos materiales sigan siendo auténticos y utilizables en un tiempo futuro, a menudo inalcanzable para la imaginación humana. Entre estos retos está superar la obsolescencia tecnológica, propia de este entorno, y mantener la interpretabilidad y autenticidad de los materiales a lo largo del tiempo, lo que requiere planificación, estructuras organizativas y tecnológicas adecuadas. Así, se recurre a repositorios digitales cuya gestión abarca desde aspectos organizativos hasta el uso de tecnologías para este fin, con sus estructuras basadas en normas y estándares nacionales e internacionales. La justificación para la realización de este estudio radica en comprobar que la preservación digital de los repositorios institucionales de la *Rede Cariniana* posee las características necesarias para una preservación plena y confiable. A partir de ahí, se elaboraron hipótesis para comprobar la presencia de las estructuras necesarias, la existencia de directrices de regulación de la preservación, la existencia de evaluaciones de confiabilidad y la utilidad de una herramienta de evaluación de la calidad de los repositorios basada en estándares y normas internacionales. Entre los objetivos propuestos se encuentra la identificación del estado de la preservación digital de los repositorios institucionales, presentando mejoras a aplicar, el estudio de herramientas y métodos de evaluación de la fiabilidad de los repositorios y el desarrollo de un instrumento de evaluación de criterios de calidad para su aplicación en un contexto nacional. En general, la metodología del trabajo es de carácter exploratorio y descriptivo, utilizando un cuestionario diseñado para ser aplicado a los repositorios de la *Rede Cariniana*, y los resultados fueron analizados cualitativamente y cuantitativamente. Los resultados mostraron que los repositorios participantes en el estudio realizan la preservación digital de materiales de forma empírica, sin base en directrices apropiadas, centrándose en el conocimiento específico de los equipos de sus comunidades. Tampoco existe un proceso periódico de evaluación de los repositorios, lo que justifica la creación de una herramienta cuantitativa específica para su aplicación a los repositorios. Esta herramienta fue validada mediante el análisis de la información recogida en el cuestionario de evaluación y podrá ser utilizada para futuros nuevos análisis y puesta a disposición de los propios repositorios como forma de realizar autoevaluaciones.

Palabras-clave: Documentos electrónicos. Normas de auditoría y certificación. *Rede Cariniana*. Repositorios digitales.

0 INTRODUÇÃO À PESQUISA

0.1 CONTEXTUALIZAÇÃO E JUSTIFICATIVA

Nas últimas décadas, a criação e uso de documentos digitais se tornou realidade presente no cotidiano de pessoas, instituições e empresas, trazendo à tona as preocupações inerentes à preservação desses documentos. Uma vez que a natureza dos documentos digitais é efêmera e frágil, surge a necessidade de se planejar e adotar processos e ferramentas que permitam preservar os documentos digitais pelo tempo que for necessário, garantindo os seus usos em um horizonte de tempo muitas vezes inalcançável pela imaginação humana.

O decaimento e falhas das mídias de armazenamento, além das constantes desatualizações de tecnologias levam as instituições a estarem sempre um passo à frente da obsolescência do *hardware* e, não raramente, alguns efeitos só são detectados tardiamente, necessitando uma resposta ágil e eficaz das organizações que preservam documentos digitais. Há poucas décadas, a principal mídia de armazenamento para preservação era a fita magnética, ao que já havia problemas de compatibilidade de formatos proprietários e dispositivos de leitura e escrita. Uma organização ao escolher um determinado padrão de fita magnética, ligado ao seu fabricante, adquiria os conhecimentos necessários e estruturava seus processos baseando-se nessa solução. Em uma eventual necessidade de mudança de padrão, por exemplo, todos os processos de preservação e acesso necessitavam ser revistos, o que tornava as organizações rígidas quanto às mudanças e, em determinados casos, reféns dos fabricantes de tecnologia.

À medida em que a tecnologia se desenvolveu, as fitas magnéticas foram se tornando obsoletas, tornando raros os dispositivos de leitura e gravação, o que levou as organizações a realizarem processos de migração para mídias mais modernas. Mesmo antes do abandono do uso das fitas, a ausência de fabricantes, assistências técnicas e pessoas qualificadas já apresentava obstáculos às organizações, levando-as a adoção de novas tecnologias e, conforme o caso, a perda de dados armazenados que não podiam mais ser acessados.

Com o passar dos anos, novas mídias de armazenamento foram surgindo, se popularizando e sendo adotadas na preservação de documentos. O surgimento dos HDD (*Hard Disk Drive*) magnéticos representou uma evolução na indústria de computação como um todo, com uma maior densidade de dados, o que possibilitou mais armazenamento em um menor espaço físico,

além de melhorar a velocidade de acesso e escrita dos dispositivos em relação às fitas magnéticas. Os avanços dessa tecnologia se basearam no desenvolvimento de semicondutores mais eficientes e menores, entre outras tecnologias, o que facilitou e popularizou o seu uso. Como melhorias, os próprios dispositivos passaram a contar com mecanismos de detecção e correção automáticas de falhas, aumentando a sua confiabilidade.

A mídia ótica foi outra tecnologia utilizada para o armazenamento de dados, que possuía maiores capacidades que os disquetes e preços menores em relação aos HD. No entanto, fazia-se necessária a implementação de processos bem estabelecidos para os seus armazenamentos, leitura e gravação que, de certa forma, eram semelhantes aos usados nas fitas magnéticas.

Além dos HD, outras tecnologias também utilizam o meio magnético como modo de operação para leitura e escrita de dados, citando-se como exemplo os SSD (*Solid State Drive*). Nesse caso, o dispositivo utiliza transistores para o armazenamento persistente dos dados, de forma magnética, mas não possui componentes móveis como os HD, diminuindo a propensão a falhas de componentes físicos. A utilização de HD ainda hoje é largamente utilizada para o armazenamento de dados, possuindo grandes capacidades de armazenamento e vidas úteis mais longas.

Independente da mídia que se utilize, a preservação digital requer um olhar cauteloso sobre os meios de armazenamento, procurando identificar condições e pontos de falhas que causem corrompimentos ou adulterações nos materiais. Ao nível mais físico, devem-se adotar processos de verificação que permitam identificar o momento antes da falha para a substituição dos equipamentos. Abstraindo-se um nível, deve-se monitorar as tecnologias empregadas nesse armazenamento, observando os sistemas de arquivos utilizados e, da mesma forma, buscando se antecipar às mudanças tecnológicas que podem tanto beneficiar quanto colocar em risco a preservação digital.

Subindo-se para um nível mais abstrato, tem-se os formatos de arquivos utilizados para armazenar textos, imagens, músicas, vídeo, etc. Ao passo em que as mídias físicas vão se deteriorando com o tempo, os formatos de arquivos vão se modificando, seja para adição de funcionalidades, melhoria de desempenho ou para a correção de bugs identificados. Os *softwares* utilizados para a produção desses materiais evoluem com o tempo e, em alguns casos, a compatibilidade com os formatos antigos deixa de existir, levando à necessidade de se realizar a migração para as versões mais atualizadas. Nos casos de descontinuação de suporte de um formato, deve-se ainda realizar a conversão para o formato mais equivalente possível, evitando-se ao máximo perdas de dados.

Nota-se que, mesmo no nível mais abstrato dos formatos de arquivos, existe a obsolescência, que deve ser alvo de estratégias de preservação que se antecipem a possíveis

problemas futuros de acesso, interpretabilidade, uso e reuso dos materiais. A escolha por uso de formatos sustentáveis, por exemplo, é vista como uma dessas estratégias, buscando-se formatos abertos e largamente aceitos na comunidade. No entanto, conforme as necessidades e a base de conhecimento da comunidade, nem sempre essa é uma possibilidade, o que requer procedimentos e esforços minuciosos de acompanhamento e prevenção da obsolescência.

Há ainda os casos em que a preservação utiliza estratégias de emulação, com as quais equipamentos atuais emulam um ambiente já obsoleto e de difícil reprodução, sendo útil para a preservação de programas de computador desenvolvidos ou jogos de videogame, por exemplo. Em outros casos, fazem-se necessárias técnicas de encapsulamento, quando se preserva todo o contexto tecnológico requerido para o acesso e uso dos materiais preservados.

A escolha da estratégia de preservação deve avaliar o seu propósito, optando por preservar todas as características do contexto original de criação e uso dos objetos digitais ou priorizar o uso do material através de migrações e conversões de formatos, abrindo mão do contexto tecnológico original. Thibodeau (2002, p. 15) identifica um filtro de quatro critérios a ser aplicado na escolha de uma estratégia de preservação: ser viável, sustentável, praticável e apropriado. Ser viável significa haver equipamentos e softwares capazes de efetivamente aplicar a estratégia. Ser sustentável significa que o método ainda poderá ser usado no futuro a longo prazo ou oferecer alternativas que o substituam. Ser praticável significa que o método deve estar dentro das restrições de orçamento e não possuir grandes dificuldades na sua aplicação. Ser apropriado significa que o método deve poder ser aplicado aos materiais que o repositório se propõe preservar, considerando os requisitos impostos e regras de negócio.

Comparando-se as realidades, no caso dos documentos físicos, a informação está intimamente ligada ao suporte utilizado, fazendo com que os processos de preservação se voltem principalmente à manutenção desses suportes, uma vez que a utilização e acesso aos conteúdos desses documentos não requer mecanismos adicionais tecnológicos. No caso dos materiais digitais, o conteúdo não se prende ao suporte, facilitando as suas cópias e transferências entre mídias e locais *ad infinitum*, porém a obsolescência tecnológica detém uma grande fatia de preocupação e esforços na preservação digital, requerendo conhecimentos específicos, técnicas especializadas e profissionais de perfis diferentes dos anteriores.

O armazenamento físico dos arquivos requer cuidados especiais para evitar que os dados sejam corrompidos, uma vez que os equipamentos sofrem desgaste ao longo de sua utilização, necessitando rotinas que verificam a integridade dos documentos periodicamente, assim como procedimentos agendados para a substituição dos equipamentos mais utilizados, evitando que se chegue ao ponto de haver perda documental. Fazendo-se um paralelo com os documentos em

suporte papel, estes também necessitam verificações periódicas a fim de se detectar ou amenizar deteriorações ocorridas que também podem gerar perdas irreparáveis.

Uma vez que a obsolescência tecnológica tenha sido tratada através das estratégias e processos adequados, existe a necessidade de se preservar a autenticidade dos materiais digitais. O acúmulo de materiais digitais sem garantias de autenticidade acaba por gerar o armazenamento de *bits* e *bytes* inúteis, uma vez que não seja possível atestar que eles são de fato evidências dos fatos que relatam. Nesse âmbito se inserem tanto os documentos arquivísticos digitais quanto os dados de pesquisa que se originam e fundamentam pesquisa científicas, por exemplo.

A autenticidade vai se construindo desde o momento da gênese dos materiais digitais ao se analisar os sistemas e processos envolvidos em sua produção. A captura de elementos que permitam se conhecer o seu contexto de produção, atores envolvidos com suas responsabilidades e papéis, e suas relações com outros materiais permite se obter uma espécie de fotografia daquele instante, um *snapshot*, que dá embasamento à verificação de autenticidade.

A preocupação com a autenticidade é um tema central no gerenciamento e preservação de documentos. Enquanto a presunção de autenticidade de documentos em suporte papel pode ser realizada através da análise do próprio documento e seus elementos, no mundo digital existe a necessidade de se capturar informações adicionais, chamados metadados, que embasem o processo de verificação de autenticidade. A facilidade de criação e adulteração de documentos digitais é desafiante para a disciplina da diplomática, pois um usuário de um *software* de edição de textos, por exemplo, pode redigir um documento fazendo parecer que o documento foi originalmente escrito por outra pessoa.

Dessa forma, a verificação de autenticidade no universo digital passa pela análise dos metadados dos documentos, que registram detalhes do arquivo criado, o ambiente de produção, permissões do usuário e demais interações sofridas ao longo de sua vida. A análise criteriosa dos metadados relacionados aos documentos torna possível a presunção de autenticidade e a identificação de adulterações intencionais ou não.

Na realidade dos documentos arquivísticos digitais, os sistemas de gestão de documentos desempenham papel central na composição da autenticidade ao controlar os papéis e permissões dos atores envolvidos, permitindo ou negando operações de acordo com as suas atribuições, em geral relacionadas aos seus cargos ocupados nas organizações. À medida em que ocorrem os usos e acessos desses documentos, mais elementos vão se adicionando aos seus blocos de informações, construindo uma estrutura sólida que permite atestar as suas autenticidades. Esses elementos dão detalhes sobre a posição dos documentos dentro da organização, suas finalidades e funções, relacionando-se às atividades refletidas no Planos de Classificação de Documentos dessa

organização. Constrói-se assim a relação entre os documentos que apoiam as autenticidades uns dos outros, permitindo atestar que eles são testemunhas verídicas dos fatos que relatam.

Nesse âmbito, é necessário ressaltar que a Organização Internacional de Normalização (*International Organization for Standardization – ISO*) desenvolveu nas últimas décadas um amplo conjunto de normas vinculadas, de uma forma ou outra, com os documentos arquivísticos digitais. Com relação à gestão documental, pode-se citar a ISO 15489:2016¹ e a família de normas ISO 30300:2020², que estabelecem processos e requisitos concernentes à gestão documental na fase ativa. Ainda a ISO 23081:2017³ que estabelece um arcabouço para a criação, gestão e uso de metadados de documentos arquivísticos, alinhado à ISO 15489, a ISO 13008:2022⁴ que estabelece diretrizes a serem observadas nas conversões de formatos de arquivos e migrações dos documentos e a ISO/TR 18492:2005⁵ com um guia prático para a preservação a longo prazo de documentos autênticos, definindo longo prazo como “[...] o período de retenção excede a vida útil esperada da tecnologia (*hardware* e *software*) usada para criar e manter as informações” (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2005). Ao se focar na preservação digital, têm-se as normas ISO 14721:2012⁶, que estabelece requisitos para a implementação de um Sistema Aberto de Arquivamento de Informações (*Open Archival Information System – OAIS*) e define um modelo de referência com um fluxo de informações realizado através de pacotes, a norma ISO 16363:2012⁷, que é utilizada como certificação e estabelece mecanismos e critérios para se atestar que um Arquivo pode se denominar OAIS e a norma ISO 17068:2012⁸, que trata sobre a contratação de repositórios de terceiros.

Adicionalmente, pode-se citar a família de normas ISO 9000⁹, que trata sobre Sistemas de Gestão e Garantia da Qualidade, tendo um escopo bastante amplo e passível de aplicação nos mais diversos nichos e a ISO 16919¹⁰, que estabelece requisitos para a composição e funcionamento de órgãos de auditoria e certificação de repositórios digitais confiáveis. Tem-se ainda a família de normas ISO 27000, que trata sobre os Sistemas de Gestão de Segurança da Informação (*Information Security Management System – ISMS*) e assume que toda organização, independente do seu tamanho, recolhe, trata, gerencia e armazena informações e, por isso, devem haver regulações e mecanismos que garantam a segurança dessas informações, conforme o definido na norma

¹ (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016)

² (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2020)

³ (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2017)

⁴ (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2022)

⁵ (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2005)

⁶ (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012)

⁷ (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012)

⁸ (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012)

⁹ (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2015)

¹⁰ (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2014)

(INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2018, p. 11, tradução nossa): “O termo segurança da informação geralmente se baseia no fato de a informação ser considerada um ativo que tem um valor que requer proteção adequada, por exemplo, contra a perda de disponibilidade, confidencialidade e integridade”. A partir da necessidade de se manter a segurança das informações das organizações, define-se um ISMS como (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2018, p. 11, tradução nossa): “[...] consiste em políticas, procedimentos, diretrizes e recursos e atividades associados, gerenciados coletivamente por uma organização, com o objetivo de proteger seus ativos de informação”. Logo, identifica-se uma intersecção de interesses entre os objetivos de se estabelecer um ISMS e a garantia de autenticidade dos documentos gerenciados por esses sistemas, característica fundamental para a preservação confiável de documentos digitais.

No Brasil, no âmbito da gestão documental, cita-se ainda o Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos, o e-ARQ, que possui requisitos, processos e metadados a serem implementados por sistemas de gestão documental arquivística (CONSELHO NACIONAL DE ARQUIVOS, 2022).

O recebimento dos documentos digitais e sua guarda é de responsabilidade das instituições conhecidas como Arquivos ou Repositórios Digitais, representando uma empresa, órgão ou organização que possui políticas de preservação e acesso, além de estruturas físicas e organizacionais adequadas para a realização da preservação digital. Essas instituições têm o dever de estabelecer os critérios necessários de curadoria dos documentos recebidos assim como definir os processos internos que garantirão a guarda adequada pelo tempo necessário de preservação, que muitas vezes pode ser para sempre.

Essa definição é correspondente ao conceito de Arquivo OAIS, criado pelo CCSDS em 2002 e depois transformado em norma internacional ISO 14721:2012 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012). A norma descreve um conjunto de entidades do arquivo com responsabilidades delimitadas, um fluxo de dados entre as entidades e os padrões de pacotes de informações que fazem parte de algumas dessas comunicações. A norma especifica também os requisitos e condições a serem satisfeitos para que se determine que um arquivo está em conformidade com a norma ou não. À medida em que um arquivo de uma instituição adere à norma ISO 14721, pode-se dizer que a preservação digital confiável é realizada plenamente.

A confiabilidade sobre o repositório e, por consequência, seus materiais preservados recai, em alto nível, na constatação de que a cadeia de custódia não foi corrompida ou comprometida. Uma vez que se detectem falhas nessa cadeia já no ambiente de gestão, toda a

seqüência de preservação e acesso se torna frágil, não sendo possível atestar que os materiais são de fato autênticos. A extração dos documentos dos seus sistemas de gestão para ingestão no repositório apresenta-se como um ponto crítico a ser analisado, buscando-se que ela seja realizada através de sistemas desenvolvidos e especializados nesse processo e não por pessoas que possam colocar em risco a autenticidade desses documentos.

Uma vez que sejam implantados os arquivos digitais, faz-se necessário reconhecer quais realizam, de forma autêntica, a preservação dos seus materiais. Não basta que os arquivos se autodenominem confiáveis, há que se estabelecer critérios que validem os seus processos e ferramentas, criando-se assim, um ambiente de confiança. Surgem assim, ferramentas e metodologias que verificam a aptidão dos arquivos à realização da preservação confiável a longo prazo, que vão desde a aplicação de modelos de maturidade, análises de riscos, auditorias internas e externas, até a conquista de certificações e selos de comprovação.

Esses processos permitem aos Arquivos a verificação do cumprimento de critérios que se relacionam aos objetivos da preservação dos documentos, tornando-se importantes para que o Arquivo tome ciência do seu estado atual. O mapeamento de critérios não atendidos identifica as fragilidades e riscos existentes fornecendo material para as suas resoluções ou mitigações, de acordo com cada caso.

A comunidade arquivística internacional elaborou diversas certificações de repositórios digitais, que podem ser aplicadas pelas próprias equipes, realizando o processo de autoavaliação ou através da contratação de um corpo avaliativo, que faz a verificação do atendimento dos critérios.

Nesse universo, pode-se citar a matriz de preservação NDSA que apresenta níveis de preservação que evoluem em sofisticação, culminando em um estado em que a preservação digital de uma organização apresenta sólidos indícios de confiabilidade (NATIONAL DIGITAL STEWARDSHIP ALLIANCE, 2022). Apresentando uma metodologia voltada à análise de riscos, o DRAMBORA possui uma ferramenta para a avaliação da probabilidade e impacto da ocorrência de eventos que possam comprometer a preservação digital dos repositórios, cobrindo desde os aspectos organizacionais até os tecnológicos, passando pelo tratamento dispensado aos pacotes de informação preservados (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007).

A nível de certificação formal de repositórios, encontra-se a norma ISO 16363:2012 que estabelece uma metodologia e critérios para a realização de auditorias que verificam se um determinado arquivo está em conformidade com o padrão de referência OAIS (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012). A obtenção da certificação, nesse caso, ocorre após a auditoria de um órgão externo que analisa todos os critérios detalhadamente e culmina

na obtenção da certificação. Similar à essa norma, a ISO 17068:2012 analisa o cumprimento de critérios no caso de repositórios externos à organização, ditos repositórios de terceiros, certificando que esses repositórios podem ser considerados confiáveis (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012).

Há ainda a opção de o repositório obter o Nestor Seal, selo que comprova que o repositório realiza preservação digital confiável através de auditorias internas que são posteriormente submetidas à organização que o controla. O selo possui características semelhantes à ISO 16363:2012, uma vez que o seu grupo de trabalho participou da elaboração da norma (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009). Destinada a ser o primeiro passo na verificação de confiabilidade dos repositórios, encontra-se o CoreTrust Seal, que requer uma auditoria interna e posterior revisão por especialistas da área de preservação de documentos, sem a necessidade de haver um corpo de auditoria formalizado. A obtenção desse selo ocorre após o envio da documentação comprobatória à organização que o controla, apresentando-se como um esforço inicial e comprometimento com a preservação confiável dos repositórios (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019).

Na Espanha, foi criado um guia que avalia critérios dos repositórios institucionais que verifica a presença de boas práticas e aplicações de padrões de metadados e serviços de coleta de informações. O guia foi elaborado no âmbito da Rede de Bibliotecas Universitárias (REBIUN) e dentre os seus objetivos, analisa se os repositórios institucionais espanhóis contribuem para a causa da Ciência Aberta, que amplia e difunde o acesso ao conhecimento científico. Ainda que o guia tenha sido elaborado com base na realidade espanhola, seus critérios podem ser utilizados para a avaliação de repositórios de outros países (AZORÍN, BERNAL, *et al.*, 2021).

0.2 DEFINIÇÃO DO TEMA, ALCANCE E LIMITAÇÕES

A partir do contexto descrito anteriormente, caracterizado pelos desafios implicados pela preservação digital e da necessidade de os repositórios digitais responderem a esses desafios, o presente estudo se baseia na seguinte problemática: **Como afirmar que a preservação digital dos repositórios brasileiros é confiável e de qualidade? Quais mecanismos são utilizados para avaliar os repositórios? Quais áreas funcionais de um repositório devem ser avaliadas para determinar a sua confiabilidade?**

Como referência e alcance, utiliza um subconjunto dos repositórios brasileiros, representados pelos repositórios digitais pertencentes à Rede Cariniana, com abrangência em todas as regiões geográficas do território brasileiro. Dentre os elementos que mais impactaram a pesquisa

foram aqueles vinculados à parte exploratória, dada a realidade brasileira, com uma ampla diversidade de regiões e divisões estatais e federais, a abrangência da totalidade dos repositórios se torna impossível. Neste sentido, foi necessária a busca de uma mostra representativa que explorasse a realidade, de onde advém a escolha de uma rede com múltiplas instituições com alcance em todo o país.

Trata-se de uma pesquisa de natureza descritiva e analítica quali-quantitativa que incorpora um estudo de caso e permite a elaboração de ferramentas adaptadas e propostas de melhoria.

Para o referencial teórico, empregam-se técnicas descritivas analítico-sintéticas baseadas em fontes bibliográficas, normativas internacionais e nacionais, assim como documentais, as quais facilitaram a análise e sistematização dos fatores mediadores na preservação a longo prazo dos objetos digitais que, por consequência, permitiram compreender o grau de complexidade funcional e tecnológico que envolve os repositórios digitais.

Ademais, consultaram-se as ferramentas mais relevantes de avaliação em uso no contexto internacional, as quais foram objeto de análise para a escolha de critérios essenciais para a avaliação de repositórios. Assim, os critérios de avaliação se basearam nas seguintes ferramentas e normas: ISO 14721:2012, ISO 16363:2012, matriz de níveis NDSA, Nestor Seal, CoreTrust Seal, DRAMBORA e o Guia para Avaliação de Repositórios REBIUN. Os seus conjuntos de critérios foram a base para a elaboração dos critérios e áreas utilizados no presente estudo.

Para a realização da pesquisa, de natureza aplicada, como já citado, utilizou-se os repositórios da Rede Cariniana, com as devidas permissões para a exploração dos dados. A Rede é integrada por 31 instituições, das quais 22 são instituições públicas federais, 4 são instituições estaduais e 5 são de natureza privada, de onde se obteve um total de 19 instituições respondentes, representando 61% do total de repositórios. Das instituições que não responderam, 7 são privadas ou estaduais e 5 são federais. A todos os repositórios foi solicitado o preenchimento do formulário de modo a se obter um panorama geral sobre a forma como a preservação digital é realizada nas suas respectivas instituições.

Por outro lado, uma vez feita a seleção, os problemas derivaram-se das ausências de respostas de alguns repositórios da Rede Cariniana. Assim mesmo, resulta limitante a proteção dos dados obtidos para não que não difundam os resultados daqueles repositórios que mostram debilidades manifestas. As demais limitações de caráter metodológico são apresentadas na seção 4 de metodologia.

Os resultados esperados na pesquisa centram-se, em primeiro lugar, na criação de uma ferramenta de avaliação dos repositórios que permita avaliar fatores relevantes para a qualidade da

preservação desses repositórios, em segundo lugar, em dispor de um mapeamento sobre o grau de confiabilidade da amostra selecionada de repositórios institucionais e da realidade brasileira. Ademais, vinculam-se os principais desafios da preservação em suas áreas a partir de enfoques heterogêneos: objeto, ciclo da vida e custódia e processos específicos. A seguir, são apresentados os objetivos planejados para a realização do presente trabalho.

0.3 OBJETIVOS

Os objetivos deste trabalho de pesquisa são apresentados a seguir, sendo divididos em dois níveis, para uma melhor compreensão. No primeiro nível, são declarados os objetivos gerais amplos (**OG**), cada um deles dividido em uma série de objetivos específicos (**OE**):

- **OG1:** definir um modelo geral de análise e avaliação de repositórios digitais com base em fontes normativas internacionais. Embora, como mencionado anteriormente, existam várias ferramentas relacionadas à auditoria, avaliação, análise de risco ou certificação de repositórios digitais, decidiu-se projetar um modelo geral porque as ferramentas existentes são vinculadas à avaliação de repositórios de acordo com uma determinada norma, o que condiciona o próprio processo de avaliação, ou são focadas em determinados aspectos específicos (avaliação de repositórios de terceiros, avaliação de risco, avaliação do impacto da ocorrência de eventos que comprometem a preservação ou a realização de auditorias internas). Como se pode observar, esse é um objetivo metodológico, para o qual foram estabelecidos os seguintes objetivos específicos:
 - **OE1.1:** identificar e analisar as ferramentas de avaliação de repositórios digitais usadas no contexto internacional, a fim de obter uma visão geral dos elementos comuns e das diferenças mais significativas entre elas;
 - **OE1.2:** identificar, selecionar e classificar os critérios de avaliação para a certificação de repositórios confiáveis identificados nas ferramentas mencionadas no ponto anterior, a fim de definir as áreas de gestão de repositórios digitais nas quais um processo de avaliação deve se concentrar. Definir as áreas de gestão nas quais os critérios de avaliação são classificados é uma etapa necessária para estabelecer um modelo de avaliação;
 - **OE1.3:** elaborar um questionário estruturado para a avaliação de repositórios digitais que possa ser aplicado em qualquer ambiente institucional, tecnológico e normativo;
 - **OE1.4:** projetar uma ferramenta de avaliação quantitativa, associada ao questionário indicado no ponto anterior, que permita: 1) avaliar de forma objetiva o nível de

qualidade de um repositório; 2) fazer comparações equitativas entre as diferentes áreas de gestão de um repositório; 3) estabelecer comparações entre repositórios diferentes, e; 4) elaborar um *ranking* de repositórios em relação às suas qualidades;

- **OE1.5:** testar o modelo de avaliação em um ambiente delimitado no contexto brasileiro, a fim de comprovar a aplicação do modelo e validar sua utilidade.
- **OG2:** analisar o estado da preservação digital dos repositórios digitais da Rede Cariniana usando o modelo definido anteriormente. Esta análise permitirá construir uma visão dos elementos internos dos repositórios que se relacionam com a confiabilidade. Pelas razões que serão explicadas no capítulo de metodologia, a Rede Cariniana é um bom exemplo de repositórios brasileiros, de modo que as conclusões tiradas dessa análise podem ser extrapoladas para outros sistemas de repositórios digitais do país. Para atingir esse objetivo geral, foram definidos os seguintes objetivos específicos:
 - **OE2.1:** avaliar aspectos relacionados ao planejamento e gestão dos repositórios digitais da Rede Cariniana;
 - **OE2.2:** avaliar os aspectos técnicos dos repositórios e a utilização de metadados de preservação;
 - **OE2.3:** avaliar as técnicas e mecanismos utilizados de modo a garantir a segurança e integridade dos dados;
 - **OE2.4:** avaliar de que forma os repositórios difundem seus materiais e se fazem acessíveis ao público;
 - **OE2.5:** identificar nos repositórios as áreas funcionais mais desenvolvidas (pontos fortes) e quais necessitam melhorias.
- **OG3:** determinar a existência de fatores externos que podem influenciar na qualidade dos repositórios digitais, estudando o caso da Rede Cariniana. Independentemente dos fatores de caráter interno, tratados no objetivo anterior, podem haver elementos externos que influenciam na qualidade dos repositórios, como o entorno geográfico, características do seu contexto institucional ou até mesmo o tempo decorrido desde a criação do repositório. Para atingir esse objetivo geral, foram definidos os seguintes objetivos específicos:
 - **OE3.1:** determinar a influência do entorno geográfico na qualidade dos repositórios. Se estabelece esse objetivo com o fim de avaliar a incidência de políticas gerais desenvolvidas nas diferentes regiões geográficas sobre o objeto de estudo;
 - **OE3.2:** determinar a influência dos aspectos relacionados com o entorno institucional na qualidade dos repositórios. Nesse sentido, se pretende analisar aspectos como o tipo de instituição, o tipo de parceria dos repositórios com a Rede

Cariniana, o orçamento, o tamanho das comunidades dos repositórios e determinados índices de qualidade institucionais.

- **OE3.3:** determinar como o tempo de criação do repositório (idade) influencia na sua qualidade.

Na próxima seção do texto, são apresentadas as hipóteses que guiaram a realização do presente estudo e que buscaram serem respondidas ao final do estudo.

0.4 HIPÓTESES

A partir da problematização, iniciou-se o trabalho intelectual de elaboração de hipóteses que, para Gil, se caracterizam como “[...] uma suposta resposta ao problema a ser investigado. É uma proposição que se forma e que será aceita ou rejeitada somente depois de devidamente testada” (2008, p. 41). Rampazzo, caracteriza hipótese como “a suposição de uma causa ou de uma lei destinada a explicar provisoriamente um fenômeno, até que os fatos venham a contradizê-la ou confirmá-la” (2005, p. 36). Sendo assim, foram elaboradas as seis hipóteses apresentadas a seguir.

A hipótese 1 (**H1**) é definida como **“os repositórios institucionais da Rede Cariniana possuem estruturas organizacionais, físicas e tecnológicas para a realização da preservação digital a longo prazo”**. Uma vez que tenham sido apresentados os desafios da preservação digital, torna-se imprescindível que existam as devidas estruturas para apoiar e executar a preservação digital em si. As estruturas organizacionais definem papéis, cargos e responsabilidades aos agentes envolvidos, enquanto a estrutura física é necessária tanto para a alocação das equipes quanto para sustentar a estrutura tecnológica necessária e fundamental para a preservação digital. A falta, debilidade ou ineficiência de alguma dessas estruturas pode vir a comprometer a preservação dos materiais digitais.

A hipótese 2 (**H2**) se define como **“os repositórios institucionais da Rede Cariniana utilizam diretrizes estabelecidas para a realização da preservação de seus materiais digitais”**. É a partir das diretrizes que os repositórios definem os requisitos de preservação, fornecendo subsídios para a implementação de regras para o depósito e acesso aos materiais preservados. Ainda, as diretrizes estabelecem como o repositório pode se prevenir de eventos que possam cessar e/ou comprometer as atividades dos repositórios, sem que haja prejuízos aos materiais preservados.

A hipótese 3 (**H3**) se define como **“os repositórios institucionais da Rede Cariniana não aplicam ferramentas de avaliação de confiabilidade”**. Visto que não são encontradas documentações referentes a processos de certificação, auditorias ou selos de confiança nos repositórios, tem-se o fato de que não são aplicados quaisquer métodos para a verificação de

confiança nos repositórios. Essa situação pode se originar do fato de que não existe uma ferramenta adaptada à realidade brasileira que atenda a fatores relevantes da preservação e facilite a avaliação dos repositórios. No âmbito da gestão documental brasileira, pode-se citar o e-ARQ (Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos) (CONSELHO NACIONAL DE ARQUIVOS, 2022) que possui bem estabelecidos e detalhados os critérios a serem cumpridos pelos sistemas de forma a implementarem uma gestão arquivística adequada. No âmbito da preservação, a Resolução N° 43 do CONARQ estabelece as diretrizes a serem seguidas para que um repositório possa ser denominado como Repositório Arquivístico Digital Confiável (RDC-Arq) (CONSELHO NACIONAL DE ARQUIVOS, 2015). No entanto, as diretrizes se apresentam de maneira ampla e não estabelecem um instrumento que possa ser utilizado efetivamente na comprovação de atendimento ou não.

A hipótese 4 (H4) é definida como **“a elaboração de um conjunto de critérios de confiabilidade mínimos, adaptados de padrões internacionais, é útil para a avaliação dos repositórios brasileiros e para a formulação de planos de melhoria”**. Uma vez que não existam ferramentas de avaliação voltadas às especificidades da realidade brasileira, a criação de um instrumento que englobe a avaliação de diferentes áreas funcionais dos repositórios com critérios que possuam a cobertura de instrumentos, normas, auditorias e métodos já consolidados poderia auxiliar na execução de avaliações nos repositórios. Uma vez aplicada a avaliação a partir desse conjunto de critérios, os repositórios poderiam comprovar seus aspectos de confiabilidade ou detectar pontos a serem evoluídos.

A hipótese 5 (H5) é definida como **“a qualidade dos repositórios é influenciada por fatores externos, relacionados aos seus entornos”**. Dado que os repositórios se tornam parte da sociedade, estão sujeitos a influências e interferências externas à sua gestão, que podem afetar diretamente as suas operações e qualidade. Dentre esses fatores externos, pode-se citar a localização geográfica, o contexto institucional e orçamento, por exemplo.

A hipótese 6 (H6) é definida como **“os repositórios mais antigos, criados há mais tempo, devem possuir mais qualidade que os repositórios mais recentes”**. Por estarem em operação há mais tempo, os repositórios antigos devem possuir mais qualidade, contando com políticas e diretrizes aplicadas, além de contarem com mais experiência e conhecimentos para a correção e realinhamento de fatores que tenham influenciado em suas operações.

0.5 ESTRUTURA DO TEXTO

O estudo se estrutura em duas partes, marco teórico e estudo aplicado. Na primeira parte, apresenta-se o objeto de preservação, a partir da análise de sua concepção e dos desafios próprios, condicionantes na sua preservação durante todo o seu ciclo de vida. Ademais, observam-se os meios e modelos de se preservar objetos digitais e as principais ferramentas para assegurar e demonstrar confiabilidades nos processos preservadores.

A segunda parte foca no estudo aplicado, o qual inclui o detalhe metodológico para criar o modelo de avaliação e o estudo exploratório vinculado à avaliação da Rede Cariniana de repositórios digitais, que facilita a comprovação do modelo avaliativo proposto.

O marco teórico apresenta de início as particularidades do objeto a preservar, quais são os desafios e onde preservar, para em seguida analisar o modo de assegurar a preservação. Articula-se em três capítulos. O Capítulo 1 detalha o objeto de preservação, os principais desafios, normas e ferramentas de preservação. O Capítulo 2 apresenta as particularidades de um Repositório Digital, evoluindo desde os Arquivos tradicionais, além de apresentar o modelo referência para repositórios digitais em uso atualmente, exemplos de ferramentas de preservação, a inserção da realidade brasileira nesse contexto e a contribuição dos repositórios ao Acesso Aberto. O Capítulo 3 aborda os temas relacionados com a avaliação de confiabilidade dos repositórios, a linha do tempo de métodos e ferramentas de avaliação, auditoria e certificação e conclui com a apresentação de algumas ferramentas destacadas.

A segunda parte da pesquisa se concentra na criação de uma ferramenta de avaliação e na verificação de sua validade mediante o estudo exploratório da Rede Cariniana. Neste sentido, o Capítulo 4 aborda em detalhes a metodologia utilizada no presente estudo, apresentando o universo de repositórios analisados, forma de comunicação empregada, além das perguntas e respostas utilizadas para avaliação de cada critério. O Capítulo 5 registra os resultados do estudo exploratório, apresentando os resultados por categorias considerando os itens integrantes e de modo geral.

Finalmente, apresentam-se as conclusões e considerações finais do presente estudo no Capítulo 6 e as referências bibliográficas no Capítulo Referências. A formatação dos títulos das seções do trabalho, as epígrafes e referências bibliográficas seguem as devidas normas ABNT

Parte I
Marco Teórico

A Preservação dos Objetos Digitais

“O tempo não existe, só existe o passar do tempo.”

Millôr Fernandes

1 A PRESERVAÇÃO DOS OBJETOS DIGITAIS: DESAFIOS E FERRAMENTAS

1.1 INTRODUÇÃO

Uma vez que se faça necessário preservar a produção intelectual da humanidade, seja por motivos históricos, legais ou jurídicos, há a necessidade de se estabelecer critérios e processos para que isto ocorra. A preservação, tanto no âmbito bibliográfico, como no âmbito documental, arquivístico e mesmo museológico, requer a elaboração de um rol teórico e, posteriormente, de processos que os apliquem na prática, com o objetivo de, no futuro, ser possível acessar e usar esses materiais para os mais variados fins.

Enquanto a humanidade lida com os desafios de preservação de suportes físicos desde muito tempo, utilizando tábuas de argila, pergaminhos ou papeis, por exemplo, a utilização de meios eletrônicos e digitais se insere na história recente da humanidade. A utilização de documentos digitais passou a fazer parte do cotidiano das pessoas com a ascensão dos computadores pessoais, facilitando a editoração de texto e utilização de planilhas de cálculo, além de diversas outras tarefas. Mais recentemente, no início do século XXI, com a disponibilização em larga escala de internet no meio doméstico e comercial, observa-se uma alta produção de materiais digitais e, com isso, trazem necessidades de preservação para que esses materiais permaneçam acessíveis e utilizáveis pelo tempo que for necessário.

A evolução tecnológica, enquanto permite grandes feitos e inovações, ao mesmo tempo traz desafios a preservação digital, uma vez que os suportes de armazenamento, *hardware* e *software* evoluem e, invariavelmente, se tornam obsoletos em pouco tempo. Térmens (2014, p. 5, tradução nossa), define:

No médio e longo prazo, os dispositivos computacionais tornam-se expostos em museus por serem artefatos que representam a tecnologia de uma civilização, possivelmente inutilizáveis após seu uso. Mas os dados digitais não sofrem o mesmo destino, podem sobreviver e continuar a ser utilizados para além da vida útil do *software* e *hardware* utilizados na sua criação. É por isso que não dizemos que os conservamos, mas que os preservamos para que possam ser reutilizados.

Não obstante, a preservação digital se utiliza de conceitos e preocupações da preservação analógica, como a conservação e segurança, por exemplo. A conservação visa evitar ou

mitigar ameaças aos materiais preservados, preocupando-se com a umidade do ar, riscos de enchentes ou incêndios, por exemplo, possuindo técnicas aplicáveis tanto para os suportes analógicos como para os suportes digitais, respeitando, obviamente as suas diferenças e sutilezas. Uma das formas de conservação do material digital é a sua replicação utilizando-se diferentes localizações geográficas e tecnologias, aproveitando-se da possibilidade da sua natureza que permite cópias infinitas, o que não pode ser realizado com os materiais analógicos. Não obstante, ainda há a necessidade de se antecipar à obsolescência do *hardware*, evitando que eventuais falhas comprometam o acesso aos materiais digitais ou os corrompam definitivamente.

No quesito de segurança, deve-se evitar o extravio e adulterações dos materiais, por exemplo, aplicando-se as devidas barreiras para cada tipo de material. Enquanto a utilização de barreiras físicas pode garantir a segurança dos materiais analógicos, os materiais digitais necessitam de maiores preocupações devido a sua natureza frágil e suscetível a modificações. Nesse caso, além das barreiras físicas, aplicam-se ainda técnicas de segurança digital que buscam impedir ataques, vazamentos de dados e adulterações indevidas, necessitando de mecanismos de monitoramento e rastreabilidade, utilizando-se a ISO 27000¹¹ para o tratamento da segurança das informações, por exemplo.

Ademais, a preservação necessita de técnicas específicas que garantam que os materiais já produzidos permaneçam acessíveis em um período futuro muitas vezes inalcançável para a imaginação humana. Uma vez que se deve manter a possibilidade de uso dos documentos, utilizam-se as técnicas de refrescamento de mídias, migração e conversão de formatos de arquivo, por exemplo, que efetivamente acabam por alterar os bits e bytes armazenados, buscando migrá-los para outras mídias, atualizá-los para versões mais atualizadas de formatos ou para outros formatos mais apropriados ao uso e reuso. Nesses casos, admitem-se mudanças periódicas controladas, definidas a partir de planos de preservação para que a usabilidade dos documentos seja mantida. Térmens faz um paralelo entre mundo analógico e o digital (2014, p. 12, tradução nossa):

Ao nos depararmos com documentos analógicos, podemos optar por não fazer nada, deixá-los trancados em um depósito e esperar para que depois de anos a falta de interação humana não tenha causado nenhum tipo de prejuízo, de tal forma que os documentos possivelmente ainda se encontrarão em um bom estado de conservação. Se também formos negligentes com os dados digitais, por exemplo deixando um disquete dentro de um arquivo por 40 anos, perderemos o seu conteúdo, pois após esse período este meio não será mais legível ou os meios técnicos adequados para lê-lo não existirão mais.

No presente capítulo pretende-se abordar como é realizada a preservação digital, buscando-se apresentar os desafios envolvidos e as ferramentas disponíveis para esse fim. São apresentados os desafios identificados na preservação digital, que vão desde a preservação das

¹¹ (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2018)

propriedades dos documentos, passam pelas estratégias, processos, cadeia de custódia e preservação, até a particularidade de se preservar determinados materiais especiais. Por fim, apresentam-se um conjunto de normas existentes relacionadas à gestão documental como um todo.

1.2 O DOCUMENTO DIGITAL COMO OBJETO DE PRESERVAÇÃO

1.2.1 A Perspectiva Tradicional

Na publicação *Manual of Archive Administration*, Sir Hillary Jenkinson, britânico, escolhe o termo *Archive* para se referir aos documentos que possuem valor histórico e que devem ser preservados por instituições nomeadas *Archives* (Arquivos). O autor registra a preferência por esse termo ao contrário de *Records*¹², pois a definição desse último é “[...] altamente técnica e restrita em seu sentido correto e extremamente flexível em seu uso comum” (JENKINSON, 1937, p. 2, tradução nossa), além desse fato, o termo *Archive* é bastante comum em outros idiomas além da língua inglesa. Ao tentar cobrir o máximo de contextos os quais os documentos estão ou podem estar inseridos, o autor define o termo documento como (JENKINSON, 1937, p. 11, tradução nossa):

Um documento que pode ser considerado como pertencente à classe de Arquivos é aquele que foi elaborado ou utilizado no curso de uma transação administrativa ou executiva (pública ou privada) da qual ele mesmo fez parte; e posteriormente preservado sob sua própria custódia para sua própria informação pela pessoa ou pessoas responsáveis por essa transação e seus legítimos sucessores.

Posterior à Sir Hillary Jenkinson, o pesquisador em gerenciamento e preservação de documentos, Theodore Schellenberg, norte-americano, traz em seu livro, originalmente publicado em 1956, uma definição bastante ampla do termo *record* (SCHELLENBERG, 2003, p. 16, tradução nossa):

Todos os livros, papéis, mapas, fotografias ou outros materiais documentais, independentemente de sua forma física ou características, feitos ou recebidos por qualquer instituição pública ou privada em cumprimento de suas obrigações legais ou em conexão com a transação de seus negócios apropriados e preservados ou apropriados para preservação por essa instituição ou seu legítimo sucessor como prova de suas funções, políticas, decisões, procedimentos, operações ou outras atividades ou por causa do valor informativo dos dados nele contidos.

Há que se notar exatamente a distinção entre os termos originais escolhidos para representar o mesmo conceito. Schellenberg faz exatamente uso do termo desprezado por Jenkinson, *Record*, adicionando fortemente a ideia de comprovação de fatos acontecidos e atividades

¹² A tradução literal do termo *record* para o português é registro, que também possui um sentido bastante amplo assim como na língua inglesa.

executadas por instituições carregando valor suficiente para que esses materiais sejam preservados. Admitindo-se que o termo instituições pode ser referir a órgãos governamentais, políticos, empresas, um indivíduo ou uma família, por exemplo.

Outro fato observado na definição do termo é que as informações devem sempre estar inscritas em um suporte, independente da sua forma, seja ela um livro, um mapa ou uma fotografia, por exemplo, porém sempre agregando valor probatório a essas informações e se tornando o foco principal da preservação.

Ao diferenciar as informações que devem ser preservadas, Schellenberg define o termo *archives* como “os documentos¹³ de qualquer instituição pública ou privada considerados dignos de preservação permanente para fins de referência e pesquisa e que tenham sido depositados ou selecionados para depósito em uma instituição arquivística” (SCHELLENBERG, 2003, p. 16, tradução nossa). Nessa definição, o autor se refere aos documentos que são alvos de preservação permanente, que devem ser guardados para sempre sem, no entanto, especificar quais critérios devem ser utilizados para a seleção desses documentos, uma vez que cada instituição pode estabelecer um rol diferente de critérios para essa seleção. Complementarmente, o autor acredita que a seleção para a preservação permanente dos documentos se faz por eles possuírem um valor adicional ao valor original pelo qual o documento foi originado, conhecido como valor secundário.

Para Vicenta Cortés Alonso, os documentos são “os resultados da atividade humana e os testemunhos e informações sobre a vida dos homens [...]” (ALONSO, 1981, p. 365, tradução nossa) reafirmando o caráter testemunhal dos documentos ao registrar atividades desenvolvidas por pessoas, independente das atividades profissionais desenvolvidas. Ainda, a autora ressalta a diferença dos documentos, ao que se deduz equivaler aos documentos arquivísticos, em relação a outras espécies como os documentos bibliográficos, arqueológicos e artísticos, necessitando tratamento adequado baseado nos elementos que os constituem. O tratamento nesse caso se refere aos processos arquivísticos pelos quais os documentos arquivísticos percorrem durante a sua vida, além dos métodos de análise diplomática específicos.

Na Espanha, a Lei 16/1985 de 25 de junho, do Patrimônio Histórico Espanhol em seu artigo 49, do capítulo I que trata do patrimônio documental e bibliográfico, define documento como (ESPANHA, 1985, p. 21, tradução nossa):

Toda expressão ou linguagem natural ou convencional e qualquer outra expressão gráfica sonora ou em imagem, registradas em qualquer tipo de suporte material, incluindo os suportes informáticos. São excluídos os exemplares não originais de edições.

¹³ O termo original utilizado é *Record*, sendo aqui traduzido para documento.

A definição da lei permite notar o cuidado por incluir diferentes gêneros documentais, registrados sob algum método em um determinado suporte, já admitindo a utilização de meios computacionais para esse fim. Ainda, faz questão de diferenciar os documentos de materiais bibliográficos.

Para Heredia (1991, p. 125, tradução nossa) “[...] os documentos são testemunhos e informação e assim podem ser usados como prova e como fonte de dados”. A autora ressalta que os documentos de arquivo são únicos, sem múltiplos exemplares, constituindo fontes primárias de informação, daí o perigo da ocorrência de suas perdas. Ainda, os documentos são produzidos dentro das estruturas das organizações e são agrupados em séries, sendo esse o caráter mais notório dos documentos de arquivo, criando-se assim organicidade. Ao se isolar um documento de arquivo desse contexto que o rodeia, corre-se o risco de perder o sentido da sua existência.

A segunda edição do *Diccionario de Terminología Archivística*, elaborado pela *Subdirección General de los Archivos Estatales*, órgão do Ministério da Cultura da Espanha apresenta a definição de documento (MINISTERIO DE CULTURA, 1995, p. tradução nossa):

Qualquer expressão em linguagem natural ou convencional e qualquer outra expressão gráfica, sonora ou imagética, recolhida em qualquer tipo de suporte material, incluindo suportes informáticos. São excluídos os exemplares não originais de edições. Documento de arquivo é o testemunho material de fato ou ato elaborado segundo características materiais e formais.

Essa definição engloba a apresentada na Lei de Patrimônio Histórico Espanhol, e ainda contempla o termo *documento de archivo* que registra o acontecimento de um fato ou ato e que deve seguir regras materiais e formais durante a sua elaboração. Essas regras definem os tipos documentais que ditam a forma e o conjunto de elementos que devem estar presentes nos documentos, construindo a base para o estabelecimento da autenticidade dos documentos digitais que é a necessidade de uma forma fixa e conteúdo estável (DURANTI, MACNEIL, *et al.*, 2005, p. 1).

Ainda ao final do século passado, em 1997, o *International Council on Archives* (ICA) elaborou um guia para a gestão de documentos eletrônicos sob o ponto de vista arquivístico, contendo a contextualização da situação trazida pelos documentos eletrônicos, os impactos da tecnologia, a definição de termos da área e possíveis estratégias a serem adotadas na gestão dos documentos eletrônicos. Dentre os primeiros termos definidos no documento, está a definição de *Record* (INTERNATIONAL COUNCIL ON ARCHIVES, 1997, p. 7, tradução nossa):

Um documento é a informação registrada produzida ou recebida no início, condução ou conclusão de uma atividade institucional ou individual e que compreende conteúdo, contexto e estrutura suficientes para fornecer evidências da atividade, independentemente da forma ou meio.

Novamente o caráter evidencial do documento é apresentado, inserindo-se três importantes características que o formam: conteúdo, contexto e estrutura. O conteúdo representa a própria informação e mensagem transmitida no documento. A estrutura se refere à forma como o documento se apresenta, o uso de fontes, símbolos e a disposição adotada para esses elementos. O contexto diz respeito ao meio em que o documento foi produzido, subdividindo-se em pelo menos três aspectos: informações extraídas do próprio conteúdo do documento, como o seu signatário, por exemplo; as relações existentes com outros documentos, e; a atividade e funções desempenhadas pelos agentes na produção do documento (INTERNATIONAL COUNCIL ON ARCHIVES, 1997, p. 22). O que diferencia o documento eletrônico é que o seu conteúdo é codificado em símbolos especiais e requer a utilização de um equipamento eletrônico para a sua leitura e interpretação¹⁴ (INTERNATIONAL COUNCIL ON ARCHIVES, 1997, p. 7).

Complementar à definição do ICA, Terry Cook esclarece as condições necessárias para que se tenha um documento (COOK, 1997, p. 24, tradução nossa):

Primeiro, dados e informações se tornam “documentos” apenas se a estrutura da informação (sua forma, apresentação, aparência, design) e seu contexto (sua ligação funcional de elementos de dados a um evento ou transação, com outras ligações a seus criadores, usuários, funções de animação, mandatos e outros registros relacionados, etc.) são adicionados ou integrados ao conteúdo real (seus fatos, palavras, números, imagens). Sem essa contextualidade ou “documentabilidade”, os criadores de informações eletrônicas não têm responsabilidade por sua precisão ou, mais seriamente, pelos eventos e transações dos quais as informações pretendem fornecer evidências.

Assim, os dados e informações somente se tornam documentos quando o seu conteúdo é integrado a uma estrutura documental formal e a um contexto devidamente registrado. Para Cook, sem essas características, que o autor chama de documentabilidade¹⁵, não há como se certificar da veracidade dos atos e transações que o documento pretende evidenciar. Sob um certo ponto, a diplomática do documento eletrônico se assemelha à do documento físico, ao analisar o conteúdo, forma, signos utilizados e a extração dos agentes envolvidos e seus papéis, porém, devido à volatilidade dos documentos eletrônicos, há que se extrair mais informações do contexto ao qual está inserido, buscando-se uma série de informações a respeito do contexto computacional utilizado e informações que previnam e detectem o corrompimento dos dados, por exemplo. Esse contexto, além de apoiar a autenticidade do documento, também serve de alimento para a manutenção da sua interpretabilidade, de onde se extrai as condições necessárias para a sua reprodução e migração, sendo característica fundamental para a preservação digital.

¹⁴ No caso dos documentos digitais, utiliza-se símbolos binários codificados que são lidos e interpretados por um computador ou outro dispositivo equivalente.

¹⁵ Tradução livre para o termo *recordness* cunhado pelo autor em sua publicação.

Hedstrom e Montgomery relatam que, de um modo geral, os desafios da preservação digital já estão bem estabelecidos (1998, p. 1, tradução nossa):

Os materiais digitais são especialmente vulneráveis à perda e destruição porque são armazenados em mídias magnéticas e ópticas frágeis que se deterioram rapidamente e podem falhar repentinamente devido à exposição ao calor, umidade, contaminantes transportados pelo ar ou dispositivos de leitura e escrita defeituosos. Mesmo que a mídia seja preservada intacta, os materiais digitais tornam-se ilegíveis se os dispositivos de reprodução necessários para recuperar as informações da mídia se tornarem obsoletos ou se o *software* que traduz a informação digital do formato legível por máquina para legível por humano não estiver mais disponível.

1.2.2 A Perspectiva Digital

Em 1999, o projeto InterPARES foi lançado voltando-se às preocupações inerentes à criação, uso e intercâmbio de documentos digitais, tais como a rápida obsolescência dos *softwares* e *hardwares* e a fragilidade das mídias digitais. A *autenticidade* dos documentos digitais foi uma das preocupações propulsoras do projeto, que se voltou à descoberta de mecanismos que pudessem comprovar que um determinado documento digital foi criado de acordo com as autorizações permitidas e permaneceu sendo o mesmo depois de um certo período de tempo, contemplando as qualidades de identidade e integridade dos documentos.

À medida em que o projeto InterPARES teve seu início, houve a necessidade de se alinhar, esclarecer e estabelecer os conceitos envolvidos na preservação digital. Assim, tanto os membros internos quanto o mundo externo ao projeto tendem a serem uníssonos nas interpretações dos documentos gerados pelo projeto.

Com essa finalidade, foi criado o glossário InterPARES (INTERPARES, 2001) com a finalidade de ser um documento a ser utilizado para o estudo dos demais documentos produzidos pelo projeto, buscando criar assim, uma consistência no uso dos termos. Ainda, tendo a noção de que o uso dos documentos do projeto são diversos e abrangem a comunidade mundial, adiciona-se como objetivo ao glossário “[...] garantir que os significados dos termos sejam comunicados eficazmente ao mundo fora do projeto” (HANNIGAN, GUERCIO, *et al.*, 2001).

O glossário se declara como não sendo compreensivo, no sentido de não somente existirem aqueles termos no universo arquivístico ou de aqueles serem as únicas e verdadeiras interpretações dos termos definidos, mas sim, um ponto de início para a compreensão do vocabulário arquivístico existente e utilizado no âmbito do projeto. Ademais, o glossário tem o intuito de contribuir para o diálogo entre as diferentes áreas do conhecimento que por ventura estarão colaborando entre si para a realização de uma efetiva preservação a longo prazo. Essa contribuição mostra-se importante na formação de equipes de trabalho interdisciplinares, tornando-

se um poderoso material para a dissolução de dúvidas e más interpretações de conceitos que podem estar presentes em mais de um campo de conhecimento, daí o esforço dispendido pelos autores buscando a máxima precisão de termos-chave.

O glossário InterPARES define o termo *Record* como sendo “um documento criado ou recebido e separado no decorrer de uma atividade prática” (INTERPARES, 2001, p. 6, tradução nossa), enquanto o termo *Electronic Record* é definido como “um documento que é criado (produzido ou recebido e separado) em formato eletrônico” (INTERPARES, 2001, p. 3, tradução nossa). É possível notar que a definição do termo *Electronic Record* engloba a definição do termo *Record*, diferenciando-se aquele pelo meio no qual foi criado. Ainda, cabe salientar que o *Electronic Record* de acordo com o estabelecido por InterPARES engloba qualquer dispositivo eletrônico que permita a geração de um documento, não restringindo-se apenas aos documentos digitais originados em computadores.

Adicionalmente, é definido o termo *Digital Component* como “um objeto digital que faz parte de um documento eletrônico, ou de um documento eletrônico reproduzido, ou que contém um ou mais documentos eletrônicos, ou documentos eletrônicos reproduzidos, e que requer métodos específicos de preservação” (INTERPARES, 2001, p. 3, tradução nossa). Essa definição já faz menção ao meio digital, que remete ao mundo binário, dos computadores, assumindo que um *Digital Component* só existe como pertencendo a um determinado Documento Eletrônico e que, por sua vez, deve empregar métodos, processos e atividades dedicados à realidade digital.

Jones e Beagrie definem o termo *Electronic Record* como “documentos criados digitalmente no dia-a-dia da organização e com status formal atribuído pela organização. Eles podem incluir, por exemplo, documentos de processamento de texto, e-mails, bancos de dados ou páginas da intranet” (2001, p. 25, tradução nossa). Destaca-se que essa definição datada do início do século XXI já abarcava documentos produzidos em processadores de texto e outros gêneros como as páginas *web*.

Thibodeau (2002, p. 6, tradução nossa), define o termo *digital object* como “um objeto de informação, de qualquer tipo de informação ou qualquer formato, expresso em formato digital”, assumindo que essa definição pode ser muito ampla ou genérica demais para que se trabalhe a preservação digital. Assim, o autor compreende que todo o objeto digital é um objeto físico, lógico e conceitual ao mesmo tempo, possuindo propriedades específicas em cada um desses níveis. No nível mais básico, o objeto físico é a inscrição de sinais em um meio físico, equivalendo aos *bits* e *bytes* gravados em uma mídia de armazenamento ou ainda em uma memória não persistente, estando esse objeto intimamente relacionado ao *hardware* utilizado. O objeto lógico é a forma como os *softwares* organizam as sequências de bits (*bitstreams*) para que consigam realizar a leitura e

permitir a interpretação humana. Por fim, o objeto conceitual é o objeto com que os humanos lidam no mundo real ao escrever ou ler um documento digital, por exemplo, abstraindo as duas camadas anteriores. Para o autor, a preservação de um objeto digital deve utilizar diferentes estratégias e mecanismos para cada um dos objetos que o compõe, assim como as relações entre os objetos devem ser mantidas.

A Sociedade Americana de Arquivistas (*Society of American Archivists* – SAA), em seu glossário elaborado em 2005, possui entre os significados para o termo *Record*: “Um trabalho escrito ou impresso de natureza legal ou oficial que possa ser utilizado como evidência ou prova; um documento”, ou ainda: “Dados ou informações em um formato fixo que são criados ou recebidos no curso da atividade individual ou institucional e separados (preservados) como evidência dessa atividade para referência futura” (PEARCE-MOSES, 2005, p. 326, tradução nossa). Pode-se extrair dessas definições a natureza de fonte de prova que o documento carrega, uma vez que ele pode ser utilizado para a comprovação da execução de atividades, sejam elas profissionais ou pessoais.

O glossário da SAA, possui uma entrada para o termo *Digital Record*, porém ela remete ao termo *Electronic Record*, definindo-o como “Dados ou informações que foram capturados e fixados para armazenamento e manipulação em um sistema automatizado e que requerem o uso do sistema para torná-lo inteligível por uma pessoa.” (PEARCE-MOSES, 2005, p. 141, tradução nossa).

O Dicionário do Livro (FARIA e PERICÃO, 2008, p. 406), possui entre as definições de documento, “qualquer elemento de conhecimento ou fonte de informação fixado materialmente, que possa ser utilizado para estudo, consulta ou prova, isto é informação”. É possível notar o foco das autoras em explicitar a necessidade de o documento possuir um elemento de suporte, assim como ter a finalidade de ser a comprovação de fatos, definidos como informações ou conhecimentos nesse caso. Essa definição vai de encontro àquela cunhada por Schellenberg (2003, p. 16).

A mesma obra ainda define documento eletrônico como “aquele que é elaborado mediante processamento eletrônico de dados. Suporte informático com informação computorizada que pode ser decodificada por meio de uma máquina.” (FARIA e PERICÃO, 2008, p. 408). Aqui, as autoras alargam o conceito de documento abarcando qualquer processamento eletrônico de dados, mesmo que flertando com algum processamento através de computadores. Por fim, tem-se a definição de documento digital como sendo “modalidade de documento no qual a informação é representada através de séries de bits; contém informação textual, visual e sonora codificada sob forma de dígitos, manipuláveis em sistemas informáticos” (FARIA e PERICÃO, 2008, p. 408). Essa definição já admite a codificação em formato de dígitos binários que é tratada, de forma geral, por

computadores. Um aspecto interessante dessa definição é a inclusão gêneros não textuais, característica pouco comum de outros autores.

Na Espanha, a Lei 11/2007, de 22 de junho, que trata sobre o acesso eletrônico dos cidadãos aos serviços públicos, traz em seu anexo, a definição de documento eletrônico: “informação de qualquer natureza em suporte eletrônico, arquivada em suporte eletrônico de acordo com formato específico e suscetível de identificação e tratamento diferenciado” (ESPANHA, 2007, p. 32, tradução nossa). A publicação primeiramente reconhece o direito aos cidadãos espanhóis em se comunicar com os órgãos do governo através de meio eletrônico, estabelecendo requisitos e delineando uma estrutura a ser adotada para que haja eficiência na comunicação e que os documentos eletrônicos gerados sejam dignos de confiança por todas as partes interessadas. Essa lei foi revogada posteriormente pela Lei 39/2005, de 1 de outubro.

Após dois anos, é publicado o Real Decreto 1671/2009, de 6 de novembro, o qual buscou desenvolver e especificar o acesso eletrônico de que tratava Lei 11/2007, especificando as características e requisitos do documento eletrônico em seu artigo 41, parágrafo 1 (ESPANHA, 2009, p. 20, tradução nossa):

Os documentos eletrônicos devem preencher os seguintes requisitos para sua validade: a) Conter informações de qualquer natureza; b) A informação deve estar arquivada em suporte eletrônico segundo um formato determinado e suscetível de identificação e tratamento diferenciado, e; c) Dispor de dados de identificação que permitam a sua individualização, sem prejuízo de sua possível incorporação a um processo eletrônico.

Essa definição já demonstra preocupação com o suporte e formatos utilizados nos documentos eletrônicos espanhóis, especificando a necessidade de tratamento diferenciado, que se pode compreender como procedimentos que apoiem a posterior preservação e acesso. Outra característica a ser ressaltada é a capacidade de identificação única dos documentos, mesmo quando inseridos em processos eletrônicos. Adiante, no artigo 42, parágrafo 1, admite-se a inclusão de metadados aos documentos, sob a seguinte definição (ESPANHA, 2009, p. 20, tradução nossa):

Entende-se por metadados, para efeitos do presente Decreto Real, qualquer tipo de informação em formato eletrônico associada a documentos eletrônicos, de natureza instrumental e independente do seu conteúdo, destinada ao conhecimento imediato e automatizável de qualquer das suas características, com a finalidade de garantir a disponibilidade, acesso, conservação e interoperabilidade do próprio documento.

O parágrafo 4 do artigo 42, esclarece que a sistematização dos metadados mínimos obrigatórios a estarem presentes nos documentos públicos espanhóis seria definido posteriormente pelo Esquema Nacional de Interoperabilidade. Essa Lei seria revogada posteriormente pela Lei 39/2015, de 1 de outubro.

O Esquema Nacional de Interoperabilidade foi instituído e regulamentado pelo Real Decreto 4/2010, de 8 de janeiro, tendo como princípios (ESPANHA, 2010, p. 6, tradução nossa): a) A interoperabilidade como **qualidade integral**; b) Caráter **multidimensional** da interoperabilidade, e; c) Enfoque em soluções **multilaterais**. Ter a interoperabilidade como qualidade integral significa que ela deve ser pensada desde o projeto de sistemas e serviços do governo. O caráter multidimensional busca envolver a interoperabilidade nas dimensões organizacionais, semânticas e técnicas, incluindo desde os acordos entre os órgãos administrativos até os sistemas informacionais, utilizando-se normas e padrões aplicáveis, perpassando pela dimensão temporal ao longo do tempo. A característica de soluções multilaterais foca em buscar por serviços que possam ser escalados, compostos e reutilizáveis para a construção de uma estrutura de solução modular e multiplataforma. Em seu anexo, a Lei traz um glossário de termos, dentre os quais possui a definição de documento eletrônico, que se mantém a mesma estabelecida na Lei 11/2007 de 22 de junho.

No que se refere à preservação dos documentos eletrônicos, o Real Decreto, no artigo 21, parágrafo 1, alínea h, especifica a necessidade de medidas para (ESPANHA, 2010, p. 13, tradução nossa):

[...] assegurar sua recuperação de acordo com o prazo mínimo de preservação determinado pelas normas administrativas e obrigações jurídicas, que se garanta a sua preservação a longo prazo, que se assegure seu valor probatório e sua confiabilidade como evidência eletrônica das atividades e procedimentos, assim como a transparência, memória e identificação dos órgãos das Administrações públicas e das Entidades de Direito Público vinculadas ou dependentes daquelas que exercem a competência sobre o documento ou processo.

Em julho de 2011, é publicada a Resolução que estabelece as Normas Técnicas de Interoperabilidade Documentos Eletrônicos para a Administração pública espanhola. A Resolução busca concretizar o Esquema Nacional de Interoperabilidade apresentado ainda na Lei 11/2007 de 22 de junho, e tem como objetivo geral (MINISTERIO DE POLÍTICA TERRITORIAL Y ADMINISTRACIÓN PÚBLICA, 2011, p. 87.094, tradução nossa):

Estabelecer os componentes do documento eletrônico, incluindo conteúdo, assinatura eletrônica e mínimo obrigatório e metadados, e seu formato, assim como as condições para seu intercâmbio e reprodução.

A Norma Técnica especifica detalhadamente os metadados mínimos obrigatórios dos documentos, assim como os esquemas XML para o intercâmbio de documentos e as informações básicas sobre as assinaturas de documentos eletrônicos. Define-se como componentes do documento eletrônico (MINISTERIO DE POLÍTICA TERRITORIAL Y ADMINISTRACIÓN PÚBLICA, 2011, p. 87.095, tradução nossa): a) Conteúdo, entendido como o conjunto de dados ou informação do documento; b) Quando for o caso, assinatura eletrônica, e; c) Metadados do

documento eletrônico. A cada um desses componentes, a Resolução atribui uma entidade que é representada através de atributos no esquema XML apresentado em seu Anexo II.

Cabe reconhecer-se a importância do estabelecimento da Norma Técnica no sentido de padronizar a gestão de documentos na administração pública espanhola, definindo o conjunto mínimo de metadados a estar presente junto aos documentos, favorecendo a uniformização de acesso, além de registrar informações de custódia. O intercâmbio de documentos entre sistemas heterogêneos torna-se facilitado e padronizado, uma vez que se estabelece a estrutura de dados e informações mínimos requeridas para a realização das comunicações. Como reflexo, esse arcabouço de metadados associados aos documentos apoia não somente a gestão documental como também tende a facilitar o processo de seleção para preservação dos documentos.

O Real Decreto 1708/2011, de 18 de novembro, estabelece o Sistema Espanhol de Arquivos e regulamenta o Sistema de Arquivos da Administração Geral do Estado e órgãos públicos e seu regime de acesso, definindo como sistema de arquivos (ESPANHA, 2011, p. 8, tradução nossa):

conjunto de sistemas arquivísticos departamentais, órgãos, arquivos e centros da Administração Geral do Estado e seus órgãos públicos, que atuam de forma coordenada de forma a garantir a correta gestão dos fundos, acervos, documentos e demais conjuntos documentais produzidos. Recolhidos no exercício das suas competências e facilitar o acesso dos cidadãos aos mesmos.

O Decreto ainda define o comportamento do Sistema em relação ao ciclo vital dos documentos, garantindo o acesso aos documentos públicos e estabelecendo a obrigatoriedade de haver tratamento arquivístico adequado aos documentos custodiados, obedecendo-se as normas nacionais e internacionais e instruções emanadas pela *Comisión de Archivos de la Administración General del Estado*. Em seu capítulo I, artigo 2, apresenta diversas definições sobre documentos, sua gestão e preservação, dentre as quais se destaca na alínea C a definição de Documento Eletrônico já trazida na Lei 11/2007, de 22 de junho (ESPANHA, 2011, p. 6, tradução nossa):

[...] informação de qualquer natureza em formato eletrônico, arquivada e em suporte eletrônico, segundo um formato determinado e suscetível à identificação e tratamento diferenciado.

De forma análoga, no Brasil, o Decreto N. 4.915 de 12 de setembro de 2003, estabeleceu o Sistema de Gestão de Documentos de Arquivo (SIGA) da administração federal, tendo uma de suas finalidades definida no artigo 2, alínea I (BRASIL, 2019):

garantir ao cidadão e aos órgãos e entidades da administração pública federal o acesso aos arquivos e às informações neles contidas, de forma ágil e segura, resguardados os aspectos de sigilo e as restrições legais.

Compõem o SIGA, o Arquivo Nacional do Brasil e os órgãos setoriais e unidades responsáveis pela gestão de documentos e arquivos das entidades da administração pública federal brasileira, sendo responsável por definir, elaborar e divulgar as diretrizes e normas relacionadas à gestão documental desses órgãos.

Na Espanha, a Lei 39/2015, de 1 de outubro, estabelece o Procedimento Administrativo Comum das Administrações Públicas e regula os requisitos de validade e eficácia dos atos administrativos e procedimentos administrativos públicos nacionais. Ademais, a Lei revoga diversos artigos da Lei 11/2007 de 22 de junho e do Real Decreto 1671/2009 de 6 de novembro, no que se refere ao acesso eletrônico de cidadãos espanhóis aos serviços públicos nacionais. No que cabe aos documentos eletrônicos, estabelece em seu artigo 17, alíneas 2 e 3 (ESPANHA, 2015, p. 89.358, tradução nossa):

2. Os documentos eletrônicos devem ser armazenados em um formato que garanta a autenticidade, integridade e conservação do documento, assim como sua consulta, independentemente do tempo decorrido desde sua emissão. Em qualquer caso, deve ser garantida a possibilidade de transferência dos dados para outros formatos e suportes que garantam o acesso a partir de diferentes aplicações. A eliminação de tais documentos será autorizada de acordo com as disposições dos regulamentos aplicáveis.

3. Os meios ou suportes sobre os quais os documentos são armazenados devem ter medidas de segurança, de acordo com as disposições do Esquema Nacional de Segurança, que garantam a integridade, autenticidade, confidencialidade, qualidade, proteção e conservação dos documentos armazenados. Em particular, eles devem assegurar a identificação de usuários e controle de acesso, assim como o cumprimento das garantias previstas na legislação de proteção de dados.

Nota-se a ênfase em garantir a autenticidade e preservação e acesso dos documentos pelo tempo que for definido, além de já prever o procedimento de atualização e conversão de formatos de arquivos, visando manter os documentos utilizáveis. A seguir, é tratado dos suportes de armazenamento dos documentos, especificando que se deve utilizar mecanismos que garantam a autenticidade, segurança e proteção dos dados, registrando-se e controlando-se os acessos aos documentos. Em seu artigo 26, alínea 2, a Lei 39/2015 define os requisitos para que um documento administrativo eletrônico seja considerado válido (ESPANHA, 2015, p. 89.363, tradução nossa):

Para serem considerados válidos, os documentos administrativos eletrônicos devem: a) Conter informação de qualquer natureza arquivada em meio eletrônico em formato específico suscetível de identificação e tratamento diferenciado; b) Dispor dos dados de identificação que permitam a sua individualização, sem prejuízo da sua eventual incorporação em um processo eletrônico; c) Incorporar a referência temporal do momento em que foram emitidos; d) Incorporar os metadados mínimos necessários; e) Incorporar as respectivas assinaturas eletrônicas de acordo com o disposto na regulamentação aplicável.

É possível notar uma complementação à definição presente no Real Decreto 1671/2009 de 6 de novembro, adicionando a dimensão temporal, através do momento de emissão do documento, além da noção de metadados mínimos e assinaturas eletrônicas. O artigo 27 da mesma

lei trata sobre a validade das cópias de documentos da administração pública e também exige a presença de metadados que indiquem que o documento se trata de uma cópia, mesmo que essa cópia seja fruto de uma digitalização de um documento em físico.

A Lei 40/2015, de 1 de outubro, do Regime Jurídico do Setor Público, em seu artigo 46, alínea 1, declara (ESPANHA, 2015, p. 89.442, tradução nossa): “Todos os documentos utilizados em ações administrativas serão armazenados eletronicamente, exceto quando isso não for possível”. A Lei estabelece a obrigatoriedade de uso de documentos eletrônicos na administração pública espanhola como padrão e não como exceção, o que auxilia a impulsionar o desenvolvimento de arcabouço de soluções tecnológicas que apoiem os sistemas do governo, tanto para a gestão quanto para a preservação e acesso dos documentos. Ademais, a fim de que seja possível garantir a integridade, autenticidade, confidencialidade, qualidade, proteção e preservação dos documentos, a alínea 3 do mesmo artigo 46 especifica as características que os meios e suportes devem possuir (ESPANHA, 2015, p. 89.442, tradução nossa):

Em particular, assegurar a identificação dos usuários e o controle de acesso, o cumprimento das garantias previstas na legislação de proteção de dados, bem como a recuperação e preservação a longo prazo dos documentos eletrônicos produzidos pelas Administrações Públicas que assim o exijam, de acordo com as especificações sobre o ciclo de vida dos serviços e sistemas utilizados.

A mesma Lei ainda estabelece no artigo 156, alíneas 1 e 2, o Esquema Nacional de Interoperabilidade e o Esquema Nacional de Segurança, substituindo o disposto na Lei 11/2007 de 22 de junho, que trata sobre o acesso eletrônico de cidadãos aos serviços públicos espanhóis (ESPANHA, 2015, p. 89.489, tradução nossa):

1. O Esquema Nacional de Interoperabilidade inclui o conjunto de critérios e recomendações em matéria de segurança, conservação e normalização da informação, formatos e aplicações que devem ser tidos em conta pelas Administrações Públicas na tomada de decisões tecnológicas que garantam a interoperabilidade.

2. O Esquema Nacional de Segurança tem por finalidade estabelecer a política de segurança na utilização dos meios eletrônicos no âmbito da presente Lei, sendo constituído pelos princípios básicos e requisitos mínimos que garantam adequadamente a segurança da informação tratada.

Em 2016, é publicada a segunda edição do *Guía de aplicación de la Norma Técnica de Interoperabilidad de Documento Electrónico* com o objetivo de apoiar a compreensão, aplicação e implementação da Norma Técnica de Interoperabilidade (NTI). No que se refere à definição de documento eletrônico, a guia resgata as definições trazidas no Real Decreto 4/2010 de 8 de janeiro e da Lei 39/2015 de 1 de outubro, trazendo uma nova definição para documento administrativo eletrônico (MINISTERIO DE HACIENDA Y ADMINISTRACIONES PÚBLICAS, 2016, p. 13,

tradução nossa): “o objeto digital administrativo que contém a informação objeto (dados y assinatura) e os dados associados à esta (metadados)”.

Sobre os metadados, a guia define (MINISTERIO DE HACIENDA Y ADMINISTRACIONES PÚBLICAS, 2016, p. 14, tradução nossa): “elemento que contextualiza o conteúdo, a estrutura e a assinatura de um documento, contribuindo para seu valor probatório e confiabilidade ao longo do tempo como prova eletrônica de atividades e procedimentos”. Essa definição realça a importância dos metadados como meio de garantir a autenticidade dos documentos produzidos pela administração pública espanhola tornando-os passíveis de serem usados como fontes de prova, além de apoiarem a preservação, uma vez que fornecem detalhes sobre os contextos e estrutura dos arquivos.

O conjunto de marcos legais espanhóis sobre documento eletrônico, formado pelos decretos reais, leis específicas e a própria NTI, cria um arcabouço seguro para a produção de documentos na administração pública ao definir uma estrutura para o intercâmbio de documentos e um conjunto mínimo de metadados, o que permite a integração de sistemas heterogêneos e propicia a preservação documental confiável.

Avançando para as definições de objetos de preservação, Ferreira (2006, p. 22) define o objeto físico como “[...] um conjunto de símbolos ou sinais inscritos num suporte físico” que necessitam de um processamento que decifre tais símbolos gravados para que a utilização pelos seres humanos aconteça. A maneira como os símbolos inscritos são organizados para serem interpretados pelos softwares forma o objeto lógico. A interpretação pelos seres humanos da apresentação das camadas montada pelos computadores é o que pode ser definido como o objeto conceitual.

As definições apresentadas pelo autor se aproximam bastante dos conceitos de Thibodeau (2002), porém Ferreira adiciona aos três objetos anteriores, o objeto experimentado que expressa a forma como cada ser humano interpreta um determinado objeto digital, ao que o autor não considera ser possível de ser preservada. Por fim, Ferreira (2006, p. 23) considera que “do ponto de vista do ser humano, o objeto conceitual constitui aquilo que deve ser preservado”, porém para que a preservação aconteça e a interpretabilidade dos objetos possa ser mantida, deve-se empregar diferentes técnicas e métodos para as demais camadas.

A segunda edição da Norma Geral Internacional de Descrição Arquivística (ISAD(G)) (INTERNATIONAL COUNCIL ON ARCHIVES, 2000), vigente até os dias atuais, adota as mesmas definições dos termos documento e documento arquivístico adotados pelo projeto InterPARES, refletindo a sua importância e alcance frente à comunidade arquivística. O rol de

padrões de descrição elaborados pelo ICA que sucederam ao ISAD(G) também adotou a mesma definição.

Em 2021, com o lançamento de uma nova proposta de padrão para a descrição arquivística, o ICA volta a definir *record* com um certo grau de abstração como “informação inscrita pelo menos uma vez por qualquer método, em qualquer suporte, em qualquer forma persistente e recuperável por um agente no curso da vida ou atividade profissional” (INTERNATIONAL COUNCIL ON ARCHIVES, 2019, p. 24, tradução nossa). É possível extrair dessa última definição a mesma natureza daquela definição cunhada pelo projeto InterPARES, notando-se a escolha por deixar genérico o meio em que o documento se encontra, desde que ele seja persistente e interpretável posteriormente. Ainda, o termo incorpora agora a definição de agente, que manipula o documento: “uma pessoa, ou grupo, ou uma entidade criada por uma pessoa ou grupo (mecanismo), ou um cargo, que age no mundo” (INTERNATIONAL COUNCIL ON ARCHIVES, 2019, p. 27, tradução nossa).

As normas ISO que tratam da gestão documental estabelecem um vocabulário contendo termos e definições acerca do domínio de conhecimento relacionado à própria gestão documental e são compartilhados entre as normas que se sucedem. A sua utilização permite a construção de normas claras e consistentes que uniformizam o entendimento e minimizam interpretações distorcidas ou ambíguas que podem comprometer o estudo e a aplicação das próprias normas.

A norma ISO 15489, que trata sobre a gestão de documentos, define o termo Documento(s) de Arquivo como sendo “informações produzidas, recebidas e mantidas como prova e como um ativo por uma organização ou pessoa para atender as obrigações legais ou transações de negócio” (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2018, p. 3). Esse termo definido pela versão brasileira da norma torna-se equivalente ao termo *Record(s)* definido na versão original da norma revisada em 2016 e ainda vigente (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016). É possível notar que a definição do termo não contempla o meio eletrônico ou digital, no entanto, torna ampla a sua utilização e passível de ser empregada nas normas de gestão e preservação de documentos que se sucederam.

O Dicionário Brasileiro de Terminologia Arquivística define o termo documento como sendo uma “Unidade de registro de informações, qualquer que seja o suporte ou formato” (ARQUIVO NACIONAL, 2005, p. 73), sem citar a natureza física do documento, seja ele analógico ou digital. O mesmo dicionário define documento digital como “Documento codificado em dígitos binários, acessível por meio de sistema computacional” (ARQUIVO NACIONAL, 2005, p. 75), agora sim delimitando o universo para os documentos gerados e/ou acessíveis através de computadores ou sistemas similares.

Complementarmente, o dicionário define o termo documento eletrônico como “Gênero documental integrado por documentos em meio eletrônico ou somente acessíveis por equipamentos eletrônicos, como cartões perfurados, disquetes e documentos digitais” (ARQUIVO NACIONAL, 2005, p. 75). Esse último termo se apresenta como um grande englobador de diversas naturezas de documentos, incluindo os documentos gerados e interpretados por computadores ou aqueles gerados analogicamente, mas interpretados por um computador, como é o caso dos cartões perfurados.

O Glossário dos Documentos Arquivísticos Digitais, elaborado pela Câmara Técnica de Documentos Eletrônicos (CTDE), órgão do Conselho Nacional de Arquivos (Conarq) define o termo Documento Arquivístico como “documento produzido (elaborado ou recebido), no curso de uma atividade prática, como instrumento ou resultado de tal atividade, e retido para ação ou referência” (CONSELHO NACIONAL DE ARQUIVOS, 2020, p. 24). A publicação deixa ressaltado que esse é o termo a ser usado equivalente ao termo *record* da língua inglesa. A CTDE ainda define o termo Documento Arquivístico Digital como “documento digital reconhecido e tratado como um documento arquivístico” (CONSELHO NACIONAL DE ARQUIVOS, 2020, p. 25).

Como forma a auxiliar uma padronização de sistemas gerenciadores de documentos, em 2001 foi criado o MoReq, em uma parceria entre o DLM Forum e a Comissão Europeia, na época em que poucos países europeus possuíam padrões nacionais para sistemas de gerenciamento de documentos. Dentre as suas características, desde início, sempre estiveram a busca por uma aplicação em escopo universal, buscando que a sua aplicação ocorresse em diversos países, inclusive fora da Europa. Outra característica buscada foi a disponibilização do conjunto de requisitos em diversos idiomas, procurando aumentar a sua adoção e promovendo uma uniformização entre as traduções disponibilizadas. Além dessas características, o modelo de requisitos se tornou um padrão *de facto*, uma vez que a sua grande adoção mundial acabou por se reconhecida como um padrão, mesmo que o modelo não tenha sido criado por um órgão regulador com essa ambição (DLM FORUM FOUNDATION, 2011, p. 16).

Após um processo de revisão dos requisitos iniciais, em 2008 houve a publicação da segunda versão do conjunto de requisitos, nomeada MoReq2 (EUROPEAN COMMISSION, 2008). Além da especificação dos requisitos, foi disponibilizada uma versão do *schema* XML a ser utilizado por desenvolvedores para adaptar os seus sistemas para a exportação de dados com informações padronizadas e que permitiam o intercâmbio desses dados entre sistemas que adotassem o MoReq2. Adicionalmente, foi lançado um *framework* de testes com o objetivo de verificar se um determinado sistema de informações adotava o conjunto de requisitos projetado, servindo também como uma espécie de certificação de sistemas de gerenciamento de documentos.

Essa característica permitia que fornecedores de software pudessem afirmar que as soluções desenvolvidas estavam de acordo com os requisitos MoReq2.

O MoReq2 utiliza a mesma definição de documento, ainda que bastante genérica, que a primeira versão da norma ISO 15489: “informação ou objeto registrado que pode ser tratado como uma unidade” (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2001) apud (EUROPEAN COMMISSION, 2008, p. 16, tradução nossa), no entanto, com a diferenciação de que o documento ainda não possui classificação e proteção contra modificações. A definição assumida para o termo *record*, equivale a da ISO 15489 que, por sua vez, equivale ao termo documento arquivístico da versão brasileira dessa mesma norma, enfatizando que ele “[...] deve incluir informações contextuais e, se aplicável, informações estruturais” (EUROPEAN COMMISSION, 2008, p. 17, tradução nossa). Por consequência, o termo documento eletrônico (*electronic record*) incorpora a definição de *record* adicionando uma nota de que ele “pode estar em formato eletrônico como resultado de ter sido criado por *software* de aplicativo ou como resultado de digitalização” (EUROPEAN COMMISSION, 2008, p. 17, tradução nossa).

Em 2010 foi lançada a versão mais recente do modelo de requisitos, recebendo a nomenclatura de MoReq2010 (DLM FORUM FOUNDATION, 2011) fruto de uma revisão dos modelos anteriores que incluiu rodadas de consultas à comunidade arquivística e especialistas na área. A versão elaborada substituiu o conceito de um modelo de requisitos para o novo conceito de requisitos modulares, contendo módulos que podem ser agrupados e rearranjados de acordo com a necessidade e realidade de cada organização que lide com sistemas de gerenciamento de documentos. Essa decisão se deve ao fato de que “[...] a nova abordagem modular permite que os consumidores especifiquem facilmente um conjunto flexível, mas abrangente e coeso de requisitos organizacionais simplesmente escolhendo uma combinação adequada de uma seleção de módulos que correspondem às suas necessidades organizacionais” (DLM FORUM FOUNDATION, 2011, p. 20, tradução nossa).

A característica de certificar sistemas que adotam o modelo foi mantida, no entanto, diferentemente da versão anterior na qual os sistemas tinham certa flexibilidade para a adaptação dos requisitos, a nova versão não permite alteração dos requisitos, porém, essa modificação que a certificação dos sistemas ocorra em módulos de acordo com o grau de maturidade dos sistemas e interesse das administrações das respectivas organizações.

A última versão disponibilizada do MoReq2010 assume a definição de que os documentos (*records*) são “aqueles pedaços de informação que têm um valor intrínseco que os torna importantes o suficiente para serem salvos e mantidos em segurança por seu valor probatório” (DLM FORUM FOUNDATION, 2011, p. 20, tradução nossa). É possível notar a ênfase no valor de

prova que a definição traz consigo, servindo o documento como evidência, seja para fins históricos, legais ou financeiros, por exemplo. Assim como a definição usada desde a primeira versão do MoReq, essa também se alinha à definição de documento de arquivo (*record*) especificada na norma ISO 15489. Complementarmente, o MoReq2010 define ainda um parâmetro para diferenciar informações, de um modo geral, de um documento de arquivo: “para decidir se uma informação é um documento ou não, seu contexto de negócios deve ser entendido, bem como sua relevância e significado para a organização” (DLM FORUM FOUNDATION, 2011, p. 22, tradução nossa).

A norma ISO 30300, que trata sobre informação e documentação, especificamente sobre a gestão de documentos, estabelece um vocabulário e os conceitos fundamentais sobre o assunto, dentre os quais, se destacam (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2021, p. 13, tradução nossa, grifo nosso):

Documento: Informação criada ou recebida e mantida como evidência e como ativo por uma organização, no cumprimento de obrigações legais ou decorrente de suas atividades;

Evidência: Informação que pode ser usada por si só ou em combinação com outra informação, para a estabelecer provas sobre um acontecimento ou ação. Nota 1: a evidência pode ser mostrada como inviolável e completa. Não se limita ao sentido jurídico do termo;

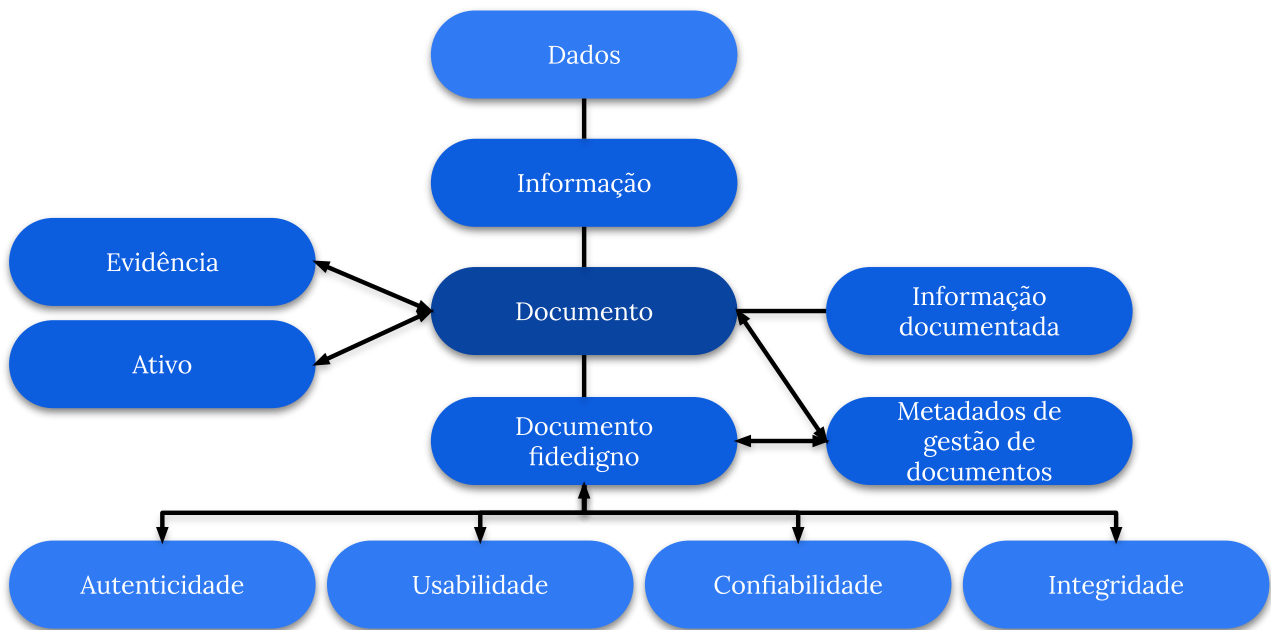
Ativo: Objeto, coisa ou entidade que tem um valor potencial ou real. Nota 1: através do tempo, o ativo pode ter um valor diferente para a organização.

Documento fidedigno: Documento que possui as características de autenticidade, confiabilidade, integridade e usabilidade.

Nota-se a necessidade de complementar a definição do termo documento com os termos evidência e ativo, caracterizando que o documento é usado como fonte de prova e possui valor para a organização. A definição de ativo, em sua nota, traz a noção que o valor dos documentos pode mudar ao longo do tempo, relacionando-se às fases do ciclo de vida documental. Uma vez que as fases corrente e intermediária do documento se encerram, existe a avaliação desse valor que pode destinar esse documento à eliminação ou à preservação permanente. A definição de documento fidedigno, por natureza, engloba os requisitos da definição de documento e ainda inclui que devem estar presentes características que permitam verificar a sua autenticidade.

A Figura 1 exibe os relacionamentos de conceitos da ISO 30300 que se ligam ao **Documento**. Inicialmente, tem-se os **Dados** que são os componentes de menor granularidade e que isolados não possuem um significado ou semântica, somente quando são combinados a partir de uma determinada lógica, os dados passam a representar uma Informação dentro de um contexto. Os conceitos de **Evidência** e **Ativo**, citados anteriormente, suportam a comprovação das atividades às quais o Documento se refere enquanto esse carrega um valor para a sua organização.

Figura 1 – Conceitos da ISO 30300 relacionados com documento.



Fonte: a partir de (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2021, p. 19)

O **Documento fidedigno** se relaciona com os termos: a) **Autenticidade**: quando é possível provar que o documento é o que pretende ser, que foi criado ou recebido por quem diz e sob as devidas autorizações, e foi criado ou recebido quando diz; b) **Usabilidade**: quando é possível acessar o documento, compreendê-lo e utilizá-lo para quaisquer fins; c) **Confiabilidade**: quando o documento reúne um conjunto mínimo de propriedades que permite o considerar uma representação fidedigna das atividades que testemunha, podendo ser considerado fonte de prova, e; d) **Integridade**: quando é possível atestar que o documento é completo e não sofreu alterações. Os **Metadados de gestão de documentos** são informações que apoiam a gestão dos documentos na organização assim como formam o conjunto de informações que vão garantir as características necessárias a um Documento fidedigno. O conceito de **Informação documentada** representa as informações que a organização gera e trata e que apoiam os próprios documentos, como determinadas evidências de resultados obtidos.

Para autores como Geoffrey Yeo, o documento é munido de caráter fluido e mutável e se altera à medida em que o tempo passa e os seus usos se modificam (YEO, 2022, p. 31, 34, tradução nossa):

[...] os documentos – sejam digitais ou analógicos – podem ser adequadamente considerados mutáveis ou fluidos. [...] Com o tempo, os documentos podem ser arquivados e re-arquivados, organizados e reorganizados, e armazenados e realocados em diferentes lugares; os vínculos entre um documento e outro podem ser criados e quebrados. [...] eles também estão sujeitos à fluência da interpretação. Ao longo do tempo, à medida que adquirem novos usuários e novos usos, geram uma série de novas interpretações e significados simbólicos. A fluidez de uso e fluidez de interpretação são fenômenos que

pertencem tanto ao mundo analógico quanto ao digital, mas a tecnologia digital permite fluidez de agrupamento em uma escala previamente desconhecida.

Mesmo com a característica da fluidez, Yeo, conclui que os arquivos tradicionais e digitais devem ser tratados da mesma forma (YEO, 2022, p. 35, tradução nossa): “As funções tradicionais dos arquivos não foram substituídas e não precisamos ver os arquivos digitais como algo diferente daqueles criados em épocas anteriores”.

Os documentos por estarem imersos nessa realidade fluida, ainda mais aguçada pelo mundo digital, trazem desafios à preservação, evidenciando o papel fundamental dos arquivos, destacado por Alonso López (2023, p. 116, tradução nossa):

Contra a fadiga causada pela mudança contínua e indiscriminada, o grande desafio é “estabilizar” os documentos superando a obsolescência, para garantir um “*continuum*”. Por isso mesmo que Duranti enfatizava a importância de um espaço físico. Quem, senão os arquivos, pode garantir a integridade e a arquitetura onde os documentos foram criados”.

1.2.3 Análise Desde uma Perspectiva Integral

O Quadro 1 exibe um resumo das definições do termo Documento e demais termos relacionados de acordo com cada autor ao longo do tempo, tanto na perspectiva tradicional quanto na perspectiva digital. Em geral, as definições carregam a noção de que o documento é a materialização de atividade que pode ser usado como evidência ou fonte de prova e, para isso, deve ser preservado adequadamente. Ao longo do tempo, a noção de documentos eletrônicos foi sendo introduzida, inicialmente, admitindo a utilização de aparelhos eletrônicos para a produção e uso dos mesmos, porém à medida em que os documentos surgiram e os seus usos se expandiram, as publicações incorporaram uma definição mais específica, utilizando computadores para a produção, armazenamento e acesso. Ressalta-se ainda, o surgimento do conceito de componentes digitais como sendo partes do documento, arquivos em geral, que são o alvo das estratégias de preservação.

A partir das definições, pode-se notar uma mudança de foco ao longo dos tempos. O conceito inicial de Jenkinson (1937) se alicerça fundamentalmente na custódia do documento, quando as qualidades, reputações e métodos dos custodiadores atribuíam confiabilidade aos seus documentos. Schellenberg (2003) adiciona a importância do processo de avaliação dos documentos, admitindo que os documentos podem possuir valor secundário, histórico e possuírem diferentes destinações. Alonso (1981) apresenta o aspecto das relações orgânicas dos documentos formando uma rede que permite conhecer um panorama dos processos executados por uma determinada organização.

Em meio ao crescente uso de documentos eletrônicos, em 1997 o ICA (1997) define o conceito de documento eletrônico e dá indícios sobre os critérios a serem utilizados na diplomática:

conteúdo, contexto e estrutura dos documentos. Cook (1997), corrobora e complementa essa visão, enfatizando a importância de haver um contexto que comprove que os fatos registrados nos documentados. A noção de confiabilidade, a partir daí, se desloca para focar no documento e a presença de elementos, como o contexto, que atestam que os fatos registrados nos documentos são de fato verídicos, foram produzidos de acordo com os processos previstos nas organizações e por agentes com as devidas responsabilidades.

Quadro 1 – Definições dos termos relacionados a documentos

Ano/Autor/Publicação	Características da definição	Elementos diferenciadores
1937 Sir Hillary Jenkinson (1937)	Documento elaborado ou utilizado no curso de uma transação administrativa. Preservado sob a custódia do autor ou pelos responsáveis pela transação.	
1956 Theodore R. Schellenberg (2003)	Documento produzido ou recebido por uma organização no cumprimento de obrigações legais ou relacionados às transações realizadas. Preservado como fonte de prova ou devido ao seu valor para a organização.	Já admitia vários gêneros documentais. Noção de valor do documento
1981 Vicenta Cortés Alonso (1981)	Documentos como testemunhos das atividades humanas. Unicidade e organicidade.	Os documentos são únicos e sua perda acarreta prejuízos. Se relacionam com outros documentos formando séries. Requerem a execução de processos arquivísticos.
1985 Lei 16/1985, de 25 de junho. (ESPANHA, 1985) (Espanha)	Documento registrado em suporte como expressão em linguagem natural ou convencional em qualquer expressão gráfica, sonora ou imagem. Admite o uso de suportes computacionais.	Inclusão de outros gêneros além do textual e uso de suportes computacionais.
1991 Antonia Heredia Herrera (1991)	Documento de arquivo são testemunhos e fontes primárias de informação. Únicos, sem múltiplos exemplares, produzidos na estrutura da organização. Agrupam-se em séries.	Agrupamento em séries, sugerindo a estrutura da organização como modelo.
1995 <i>Diccionario de Terminología Archivística</i> (MINISTERIO DE CULTURA, 1995)	O documento registra o acontecimento de um fato ou ato e segue regras materiais e formais.	Noção de utilização de regras formais para a produção de tipos documentais.
1997 Guia para a gestão de documentos eletrônicos. ICA (1997)	Documento é informação registrada de uma atividade institucional ou individual. Deve possuir conteúdo, contexto e estrutura suficientes para servir de evidência. Documento eletrônico tem o seu conteúdo codificado em símbolos especiais que requerem a utilização de equipamento para o seu uso.	Necessidade de o conteúdo, contexto e estrutura estarem alinhados para que o documento sirva de evidência.
1997 Terry Cook (1997)	Documento eletrônico deve possuir estrutura e contexto integrados ao seu conteúdo para serem evidências.	Forte ênfase na necessidade de haver contexto que suporte a comprovação dos fatos que o documento registra.
2001 Interpares (2001)	Documento criado ou recebido no decorrer de uma atividade prática. Documento eletrônico é um documento em formato eletrônico. Componente digital é um objeto digital que faz parte de um documento eletrônico e requer	Noção de documento eletrônico ser formado por componentes digitais que são o foco da preservação digital.

	métodos específicos de preservação.	
2002 Thibodeau (2002)	Objeto digital composto por objeto físico (bits), objeto lógico (arquivo) e objeto conceitual (documento).	Utilização de estratégias diferentes para preservar cada um dos objetos.
2005 <i>Society of American Archivists – SAA</i> (PEARCE-MOSES, 2005)	Documento com formato fixo produzido ou recebido no decurso de uma atividade e preservado como evidência para referência no futuro. Documento eletrônico é um documento que utiliza um sistema para armazená-lo e interpretá-lo.	Foco no formato do documento que é usado como evidência e, no caso do documento eletrônico, utiliza-se um sistema específico para o seu uso.
2005 Dicionário Brasileiro de Terminologia Arquivística (ARQUIVO NACIONAL, 2005)	Documento como unidade de registro de informações, independente do suporte. Documento digital é o documento codificado em dígitos binários e acessível por computadores. Documento eletrônico é o documento acessível somente através da utilização de equipamentos eletrônicos.	Definição genérica de documento. Definição de documento eletrônico engloba a de documento digital. Não são citados métodos de preservação ou de manutenção da autenticidade dos documentos.
2006 Miguel Ferreira (2006)	Objeto digital composto por diferentes objetos. O objeto conceitual, com o qual o ser humano interage é o que deve ser preservado.	Noção de composição do objeto digital assim como Thibodeau (2002), assumindo a utilização de diferentes técnicas para cada objeto.
2007 Lei 11/2007, de 22 de junho (Espanha) (ESPANHA, 2007)	Documento eletrônico possui informações de qualquer natureza e necessita de suporte específico que requer tratamento diferenciado.	Noção de tratamento específico dedicado para a preservação de documentos eletrônicos.
2008 Dicionário do Livro (FARIA e PERICÃO, 2008, p. 406)	Documento como fonte de informação e prova, inscrito em um suporte. Documento eletrônico utiliza processamento eletrônico dos dados. Documento digital é formado por bits, representa diversos gêneros documentais e necessita de sistemas informáticos para a sua utilização.	Documento fixado em um suporte a ser usado para fonte de estudo ou como prova.
2009 Real Decreto 1671/2009, de 6 de novembro (Espanha) (ESPANHA, 2009)	Documento eletrônico deve estar em suporte eletrônico adequado à preservação e deve ser possível acessá-lo mesmo quando não estiver contido em um processo maior. A preservação deve manter o valor probatório do documento e sua confiabilidade para ser usado como evidência das atividades e procedimentos pelo tempo em que for necessário.	Garantias de utilização de métodos de preservação que assegurem o valor probatório do documento (autenticidade).
2011 Moreq2010	Documentos são informações que têm valor intrínseco e devem ser mantidos por seu valor probatório.	Noção de valor probatório do documento e, por isso, deve ser preservado.
2011 Resolução 2011, de 19 de julho (Espanha) (ESPANHA, 2011)	O documento eletrônico é o conjunto de: conteúdo, assinatura eletrônica e metadados. Definição de um esquema XML para intercâmbio de documentos.	Apresentado o esquema XML para que a administração pública espanhola possa trocar documentos entre si.
2015 Lei 39/2015 (Espanha)	Os documentos eletrônicos devem ser armazenados em um formato que garanta a autenticidade, integridade e conservação do documento, assim como sua consulta, independentemente do tempo decorrido desde sua emissão. É admitida a migração e conversão dos formatos de arquivo visando a preservação.	Utilização de formatos de arquivos apropriados à preservação e modificação desses formatos quando for necessário.

(ESPANHA, 2015)		
2016 ISO 15489 (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2018)	Documento de Arquivo: informações produzidas, recebidas e mantidas como prova e como um ativo por uma organização ou pessoa para atender as obrigações legais ou transações de negócio” (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2018, p. 3)	Documento como prova e ativo, possuindo valor.
2016 Guia de aplicação da Norma Técnica de Interoperabilidade de Documentos Eletrônicos (MINISTERIO DE HACIENDA Y ADMINISTRACIONES PÚBLICAS, 2016)	Documento eletrônico formado por informações, assinaturas e metadados.	Noção de documento formado por partes.
2020 Glossário dos Documentos Arquivísticos Digitais (CONSELHO NACIONAL DE ARQUIVOS, 2020, p. 24)	Documento decorrente de uma atividade prática retido para a alguma ação futura ou para referência. Documento arquivístico digital: documento digital que necessita tratamento arquivístico.	Documento com resultado de uma atividade e que necessita de tratamento arquivístico.
2021 ISO 30300 (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2021)	Documento é informação criada ou recebida e mantida como evidência e como ativo por uma organização.	Noção de documento como evidência que deve ser preservado autêntico e possui valor para a organização.
2021 <i>Records In Contexts</i> (INTERNATIONAL COUNCIL ON ARCHIVES, 2021)	Documento é a informação inscrita em um suporte persistente, de forma recuperável, por um agente no decorrer de sua vida ou trabalho.	Definição concisa de documento, mas que engloba suporte, agente e acesso.

Fonte: elaboração própria.

Ao final do século XX, devido às transformações da sociedade, e impulsionada pelas tecnologias emergentes, houve uma mudança nos processos arquivísticos como um todo, a fim de se adaptar à realidade da época. Terry Cook (2001, p. 4, tradução nossa) fala sobre uma mudança de paradigma introduzida no mundo pós-moderno e computadorizado:

No coração do novo paradigma, está um afastamento de ver os documentos como objetos físicos estáticos, em direção a compreendê-los como conceitos virtuais dinâmicos; deixar de olhar os documentos como produtos passivos das atividades humanas ou administrativas para considerá-los como agentes ativos na formação da memória humana e organizacional [...]. Para os arquivistas, a mudança de paradigma requer afastar-se de se identificar como guardiões passivos de um legado herdado para celebrar seu papel em moldar ativamente a memória coletiva (ou social).

Em 2001, o Interpares (2001) cria a noção de que os documentos eletrônicos são formados por componentes digitais e estes requerem métodos de preservação específicos. Thibodeau (2002), subdivide o objeto digital em diferentes níveis de abstração: físico (bits), lógico (arquivo) e conceitual (documento). Ferreira (2006), se baseando em Thibodeau, admite a utilização de diferentes técnicas para cada nível de abstração dos objetos digitais.

A partir desse momento a definição de documento vai agregando detalhes relativos às suas estruturas e como a preservação pode ocorrer para que se mantenha a autenticidade. A verificação de autenticidade passa a necessitar de mais elementos para ser comprovada do que apenas a reputação da instituição custodiadora, verificando-se informações registradas desde o momento da produção dos documentos até as operações e interações realizadas.

Destaca-se a postura da Espanha na criação de leis e Decretos Reais voltados a tornar o documento digital como regra para o registro das atividades da administração pública, exigindo a existência de suportes e demais componentes necessários à preservação dos documentos autênticos.

Diante das definições de termos apresentados sobre os documentos e documentos digitais, é possível se observar a necessidade de preservá-los adequadamente de forma a tornar possível os seus usos como evidências ou fontes de prova. A preservação digital deve sobretudo se preocupar em manter os documentos autênticos e utilizáveis pelo tempo que for necessário, utilizando técnicas que permitam a comprovação das evidências, fatos e atividades das quais os documentos são testemunhas. Admite-se, portanto, um certo grau de mutabilidade dos arquivos preservados na busca de mantê-los utilizáveis ao longo do tempo, no entanto, as operações realizadas sempre devem ser documentadas e executadas por pessoas autorizadas a esses procedimentos. A seguir, são tratados os aspectos envolvidos na preservação digital bem como os desafios envolvidos em realizá-la.

1.3 OS DESAFIOS DA PRESERVAÇÃO DE DOCUMENTOS DIGITAIS

Uma vez que o objeto da preservação tenha sido apresentado em diferentes perspectivas, cabe à essa seção apresentar os desafios que o entorno digital traz à preservação dos documentos digitais. Dado que os bits e bytes de um documento digital podem mudar de suporte, serem transmitidos entre pessoas e sistemas, além de serem utilizados em diferentes dispositivos, a dinâmica implícita à sua natureza requer a definição de quais propriedades podem e devem ser preservadas para que permaneçam autênticos e interpretáveis ao longo do tempo. O desafio de se definir essas propriedades é apresentado na seção 1.3.1. Logo em seguida, faz-se necessário decidir quais estratégias serão usadas para preservar os documentos digitais e suas propriedades, de onde são identificados diversos desafios como os apresentados na seção 1.3.2. A operacionalização da preservação dos materiais digitais de uma empresa ou órgão institucional requer todo um aparato organizacional que possua cargos, funções e responsabilidades bem definidos, apresentando desafios como os da seção 1.3.3. A preservação dos documentos digitais se liga diretamente com os processos da gestão documental, requerendo uma pensar amplo que contemple todo o ciclo de vida documental, que se inicia com a produção dos documentos, contempla todas as operações e tramitações realizadas sobre os documentos e culmina na destinação final da preservação, de onde se identificam os desafios apresentados na seção 1.3.4. Complementarmente, a seção 1.3.5 trata sobre os desafios de se manter as Cadeias de Custódia e Preservação de modo a não comprometer a autenticidade dos documentos que podem levar à perda de confiança de todo um repositório. Por fim, a seção 1.3.6 apresenta os desafios identificados para a preservação de materiais especiais como os artísticos e audiovisuais, por exemplo.

1.3.1 O Desafio de Preservar as Propriedades dos Documentos

Dentre os aspectos comuns à preservação digital e analógica está a necessidade de manter os documentos autênticos durante o tempo necessário. Para se alcançar esse objetivo, há que se cercar de diversos processos e atividades iniciados já no período da gênese documental perpassando todo o ciclo de vida dos documentos até a preservação em si e a disponibilização de acesso aos usuários. Conforme esclarece Duranti (1999, p. 154, tradução nossa):

Uma diferença fundamental entre documentos eletrônicos e não eletrônicos é que os últimos são mantidos autênticos mantendo-os na mesma forma e estado de transmissão em que foram feitos ou recebidos e separados, enquanto os primeiros são mantidos autênticos por atualização contínua e migração periódica.

Para Beagrie *et al* (2008, p. 10, tradução nossa), “em contraste com os materiais impressos, a informação digital não vai sobreviver e permanecer acessível por acidente: requer uma gestão ativa contínua”. Enquanto a preservação do documento não eletrônico envolve técnicas para mantê-lo exatamente igual ao momento da sua criação, concentrando esforços principalmente para a manutenção do seu suporte, a preservação do documento eletrônico requer obrigatoriamente a migração periódica dos suportes, evitando a efemeridade dos armazenamentos. Portanto, a comprovação de autenticidade dos documentos eletrônicos não pode se basear exclusivamente na análise dos seus suportes.

Já no final do século passado, diversos estudos foram desenvolvidos a fim de se compreender como definir e manter a autenticidade dos documentos digitais, em contraposição aos métodos dos documentos analógicos. Duranti (1995) tratou de trazer os conceitos de Confiabilidade e Autenticidade¹⁶ para o universo dos documentos digitais. No documento em suporte papel, a sua confiabilidade em geral é baseada na confiabilidade dos seus produtores ou respaldada por pessoas com confiabilidade, seja por elas possuírem alguma reputação reconhecida ou por ocuparem cargos específicos como, por exemplo, escrivães e tabeliões. Em geral, os procedimentos utilizados para a comprovação da confiabilidade dos documentos fazem uso de selos, carimbos e assinaturas. Esses elementos quando corrompidos ou danificados fazem com que a confiabilidade dos documentos seja quebrada. Já o conceito de autenticidade desses documentos se refere à comprovação de que eles não sofreram manipulações, falsificações e outras adulterações, sejam elas propositalis ou acidentais, e que eles se referem realmente ao fato que eles atestam. Resumidamente, “um documento é confiável quando pode ser tratado como o fato do qual é prova. Em contraste, um documento é autêntico quando ele é o que afirma ser” (DURANTI, 1995, p. 7, tradução nossa).

Jones e Beagrie definem a autenticidade dos documentos digitais, dando pistas quanto à presença de metadados e a possível dificuldade de se aferir a autenticidade devido à facilidade de adulteração dos materiais digitais (2001, p. 24, tradução nossa):

O material digital é o que pretende ser. No caso de registros eletrônicos, refere-se à confiabilidade do documento eletrônico como um registro. No caso de materiais nato digitais e digitalizados, refere-se ao fato de que tudo o que está sendo citado é o mesmo de quando foi criado, a menos que os metadados que o acompanham indiquem quaisquer alterações. A confiança na autenticidade dos materiais digitais ao longo do tempo é particularmente crucial devido à facilidade com que as alterações podem ser feitas.

Para Brown, a autenticidade é derivada de três características essenciais: **confiabilidade** (*reliability*), **integridade** (*integrity*) e **usabilidade** (*usability*). A **confiabilidade** se refere a objeto “[...] ser uma representação completa e precisa da atividade cultural ou de negócio da qual ele atesta”

¹⁶ Os termos originais, em língua inglesa, são *Reliability* e *Authenticity*.

(BROWN, 2013, p. 194, tradução nossa), necessitando para isso, confiança no tratamento e manipulação desse objeto durante todo o seu ciclo de vida, seja no ambiente de gestão ou no ambiente de preservação. O alcance dessa característica se dá através da presença processos bem estabelecidos e transparentes na organização, além de requisitar a existência de metadados sobre o conteúdo, contexto e proveniência desses objetos. A **integridade** se refere a “proteger o objeto de alterações não autorizadas ou acidentais” (BROWN, 2013, p. 194, tradução nossa), se relacionando aos processos e métodos utilizados para esse fim. No âmbito da preservação digital, a integridade pode ser garantida através de processos de verificação dos bits dos arquivos, além do registro de todas as operações realizadas nos objetos. A **usabilidade** significa que “o objeto deve ser capaz de permanecer acessível por usuários autorizados, ao longo do tempo e das mudanças do ambiente tecnológico” (BROWN, 2013, p. 194, tradução nossa), referindo-se a manter a capacidade de acesso, interpretação e uso dos objetos pelos usuários ao longo do tempo, independente das mudanças tecnológicas ocorridas desde a sua produção. A preservação dos objetos deve-se utilizar de funções de preservação lógica, com as quais admite-se um grau de mutabilidade dos arquivos, desde que a sua usabilidade seja garantida, utilizando-se metadados que servem como guia para a realização dessas funções.

Enquanto os conceitos e técnicas para os documentos físicos já são vastamente utilizados e sacramentados, no meio digital, há que reformular a preservação, conforme exemplifica Brown (BROWN, 2013, p. 195, tradução nossa):

Nós temos vários séculos de experiência na preservação de artefatos materiais, sejam pinturas, livros impressos, esculturas ou manuscritos em pergaminhos. Então por que a preservação digital é tão frequentemente vista com sendo fundamentalmente mais desafiadora que a preservação tradicional? A resposta está na natureza da informação digital. Se colocarmos um livro em uma prateleira, ou um arquivo de papel em uma caixa, e armazená-lo mesmo em condições rudimentares sem nenhuma conservação adicional, podemos ser bem confiantes de que em 50 anos ele permanecerá utilizável. Armazene um CD ou HD portátil em uma prateleira por 50 anos, e a história vai ser bem diferente: mesmo que os dados sobrevivam intactos, a tecnologia necessária para ler o meio de armazenamento terá desaparecido há muito tempo.

Nesse contexto, Duranti, complementa e ainda define precisamente a preocupação que deve permear o universo dos documentos digitais (1995, p. 9, tradução nossa):

Os usuários precisam saber que o documento foi feito em circunstâncias controladas como parte do fluxo de trabalho regular, que foi feito dentro de um prazo razoável após a ocorrência dos fatos de que se trata e que foi gerado por alguém competente para fazer esse registro específico, com o dever ou o interesse direto de fazê-lo exato.

Como forma de nortear os aspectos da preservação digital, todos os agentes envolvidos devem sempre manter em mente o lema “documentos autênticos e não confiáveis não têm utilidade

para os usuários presentes e futuros, e isso significa que a principal condição que precisamos garantir é a confiabilidade dos registros” (DURANTI, 1995, p. 8, tradução nossa).

Para Waters e Garret (1996, p. 10, tradução nossa), o objetivo principal da preservação da informação digital é manter a sua integridade:

Qualquer que seja o método de preservação aplicado, no entanto, o objetivo central deve ser preservar a integridade da informação; ou seja, definir e preservar as características de um objeto de informação que o distinguem como um todo e um trabalho singular. No ambiente digital, as características que determinam a integridade da informação e merecem atenção especial para fins de arquivamento incluem o seguinte: **conteúdo**, **fixidade**, **referência**, **proveniência** e **contexto**.

Preservar o **conteúdo** (*content*) da informação digital envolve preservar, em nível mais baixo, os bits inscritos em suporte executando-se processos de migração e refrescamento. Em níveis mais elevados, o conteúdo é preservado através da migração e atualização de formatos de arquivos, entre outros processos, que buscam manter o objeto conceitual tal como foi produzido.

A **fixidade** (*fixity*) é preservada através do monitoramento constante de adulterações, intencionais ou não, do conteúdo preservado. A natureza dos objetos digitais permite a edição das informações de um documento deixando com poucos resquícios dessa execução, fazendo com que sejam necessários o emprego de técnicas que possam contrastar um objeto em um determinado instante de tempo com o objeto originalmente produzido a fim de se identificar essas modificações. O fato de se verificar que determinadas adulterações no objeto digital não são registradas e identificadas pode tornar a integridade de todo o acervo de um arquivo digital sob suspeita, tornando-o sem valor (WATERS e GARRET, 1996, p. 14).

A característica de **referência** (*reference*) se refere à capacidade de se identificar única e inequivocamente um objeto digital, permitindo que as pessoas possam acessar o material sempre que seja necessária a verificação de sua autenticidade. Com o uso da internet, pode-se adotar mecanismos de URI (*Uniform Resource Identifier*) como o DOI (*Digital Object Identifier*) que, quando acessados, respondem com uma cópia do objeto original, permitindo que a verificação de integridade. À medida em que uns objetos referenciam outros objetos internamente, a própria rede de referência fortalece a integridade desses objetos (WATERS e GARRET, 1996, p. 16).

A quarta característica relacionada à integridade se é a **proveniência** (*provenance*) que documenta a linha do tempo do objeto desde a sua criação até o momento atual. Para Waters e Garret (1996, p. 16, tradução nossa), “para preservar a integridade de um objeto de informação, os arquivos digitais devem preservar um registro de sua origem e cadeia de custódia”. Os dados de proveniência permitem construir uma trilha desde o momento da produção do objeto, identificando os agentes envolvidos, como autores, redatores, editores, etc., até as posteriores modificações necessárias, documentando-se todas as ações pertinentes à preservação desse objeto. Da mesma

forma, deve-se manter o registro dos custodiadores dos objetos de modo a compreender o papel que cada um deles desempenhou durante a vida dos objetos.

David Bearman (1993, p. 22, tradução nossa) resume a importância da proveniência dos documentos: “a proveniência, indiscutivelmente o conceito mais importante arquivologia, dita que os documentos devem ser entendidos com referência às suas origens na atividade”. O autor ainda esclarece o equívoco causado por alguns arquivistas ao definirem que a proveniência dos documentos é igual ao órgão que o produziu, sem levarem em consideração a função, a atividade evidenciada pelo documento e o sistema de gerenciamento de documentos utilizado para a sua produção, armazenamento e acesso. Assim, Bearman afirma (1993, p. 22, tradução nossa):

Basta dizer aqui que o que os analistas de sistemas chamariam função de negócio sendo conduzida, não o “órgão de origem”, determina a forma e o conteúdo dos registros e dita os procedimentos para sua criação e disseminação. [...] Por outro lado, se uma nova função for atribuída a um escritório, ela geralmente exigirá novos procedimentos acompanhados de novos tipos documentais, novas séries de arquivos.

A quinta e última característica necessária para a manutenção da integridade é o **contexto** (*context*) dos objetos digitais. De acordo com Waters e Garret (1996, p. 18, tradução nossa):

O contexto da informação digital inclui uma dimensão técnica, uma dimensão de ligação a outros objetos, uma dimensão de comunicação e uma dimensão social mais ampla. Os arquivos digitais devem estar atentos às maneiras pelas quais cada uma dessas dimensões contextuais afeta a integridade dos objetos sob seus cuidados.

A dimensão técnica se relaciona especificamente com os *softwares* e *hardware* utilizados para a produção e reprodução dos objetos digitais, documentando os requisitos necessários para que, no futuro, os objetos possam ser acessados e interpretados. A dimensão de ligação a outros objetos se refere ao meio em que um determinado objeto foi produzido, ligando-se a outros objetos correlatos que permitem a compreensão do meio em que estavam inseridos. No caso dos documentos digitais, essas ligações permitem conhecer quais documentos originaram um outro documento, por exemplo. A dimensão da comunicação se refere ao meio em que os objetos foram produzidos, registrando-se os requisitos necessários para que eles possam ser acessados ou reproduzidos. A dimensão social documenta o impacto ou uso que o objeto digital pode representar na sociedade, registrando informações acerca do ambiente externo do objeto ao qual ele está inserido. Resumidamente, “o contexto social dos objetos de informação digital é uma expressão do que se poderia caracterizar em outros termos como os interesses das partes interessadas nesses objetos” (WATERS e GARRET, 1996, p. 19, tradução nossa).

Para o registro dessas informações de contexto, faz-se uso da descrição arquivística, que nos documentos digitais é complementada pelos requisitos técnicos necessários para a sua leitura e

interpretação. Essas informações, em forma de metadados, devem ser recolhidas no instante mais próximo da produção do documento digital para que não se perca o contexto da sua gênese, que apoiará a autenticidade, um dos pilares da preservação digital. Cabe aos arquivistas das organizações definir quais modelos de descrição e metadados serão utilizados, arrecadando-se o máximo de metadados pertinentes para a formação desse contexto. Essa tarefa requer dos profissionais o exercício de antecipar quais informações poderão ser úteis e requeridas no futuro, muitas vezes intangível, para que os materiais preservados possam ser corretamente acessados. As consequências de eventuais falhas nesse processo podem fazer com que no futuro existam acumulados bytes que não poderão ser interpretados, resultando em documentos inúteis.

Já em 1993, Hedstrom (1993) alertava para as transformações necessárias que os profissionais arquivistas teriam que passar ao trabalharem com documentos eletrônicos. A autora previa uma concentração maior do trabalho desses profissionais no projeto dos sistemas de produção e criação de documentos, no intuito de se identificar como e quais metadados capturar para que os seus contextos e proveniências também fossem preservados. Já naquela época tornava-se nítida a dependência do documento com os seus metadados, sem os quais tanto a conservação do conteúdo do documento como a garantia de autenticidade estariam em perigo.

No início do século XXI, Terry Cook reformula conceitos e fazeres vindos da prática dos documentos analógicos, evidenciando o novo paradigma do documento digital, que não existe sem o seu contexto e é dinâmico, agregando mais e mais informações (metadados) à medida em o tempo avança (COOK, 2001, p. 22, tradução nossa):

O documento, portanto, deixa de ser um objeto físico para se tornar um “objeto” conceitual de dados, controlado por metadados, que virtualmente combina conteúdo, contexto e estrutura para fornecer evidências de alguma atividade ou função do criador. Além disso, como o contexto e os usos de um documento mudam ao longo do tempo (incluindo os usos arquivísticos), os metadados mudam e o documento e seu contexto são continuamente renovados. Os documentos não são mais fixos, mas dinâmicos. O documento não é mais um objeto passivo, um “registro” de evidências, mas um agente ativo desempenhando um papel contínuo na vida dos indivíduos, organizações e sociedade.

O conjunto de características da integridade dos objetos digitais vai se formando à medida em que os objetos perpassam o seu ciclo de vida, incrementando-se os registros de informações de acordo com as operações realizadas sobre eles. O nível de detalhamento e a qualidade dessas informações é o que permitirá a presunção de autenticidade dos objetos em maior ou menor grau.

Em 2001, a Força Tarefa de Preservação (*Preservation Task Force*) do projeto InterPARES refina a definição de autenticidade dos documentos (2001, p. 3, tradução nossa):

Um documento autêntico é aquele que é o que pretende ser e que está livre de adulteração ou corrompimento. Determinar que é o que pretende significa confirmar sua identidade. Determinar que está livre de adulteração ou corrompimento significa demonstrar que sua integridade permanece intacta através do espaço e do tempo.

Ferreira (2006, p. 20), também ressalta o aspecto indissociável da preservação digital com a manutenção da autenticidade dos documentos:

A preservação digital consiste na capacidade de garantir que a informação digital permanece acessível e com qualidades de autenticidade suficientes para que possa ser interpretada no futuro recorrendo a uma plataforma tecnológica diferente da utilizada no momento da sua criação.

Complementarmente, Ferreira (2006, p. 49) ainda resume a autenticidade como “um objeto é autêntico se estiver conforme o original e se a sua história custodial tiver sido devidamente documentada ao longo do tempo”. Ferreira expressa a mesma ideia de autenticidade definida pelo projeto InterPARES, atestando que o objeto deve estar conforme o original e possuir uma história custodial, características que compõem a integridade e identidade, respectivamente.

Enquanto a preservação do papel requer métodos e técnicas capazes de preservar o suporte (papel) e manter as informações legíveis (interpretáveis) pelo tempo que for necessário, a natureza do documento digital já traz embutida desafios à preservação. Na gênese do documento, a escolha do suporte e a própria forma com que os bits serão inscritos nesse suporte influenciam nas decisões a serem tomadas para a preservação desse documento. Por consequência, a leitura desse documento também vai depender de mecanismos que sejam capazes de acessar o conteúdo, interpretá-lo e disponibilizá-lo às pessoas interessadas. Ainda assim, a variedade de *hardwares* e *softwares* utilizados pode fazer com que haja diferenças entre o documento gravado e o documento lido, como o relatório do projeto InterPARES (2001, p. 3, tradução nossa) aponta:

As diferenças entre representações de armazenamento e apresentações de uso não ocorrem apenas entre diferentes registros e arquivos digitais. A mesma apresentação de um registro pode ser produzida a partir de uma variedade de representações de armazenamento. Por outro lado, uma única representação de armazenamento pode ser processada para gerar uma variedade de apresentações.

Adrian Brown, baseado no modelo lógico de informação OAIS, que especifica que um Objeto de Informação (*Information Object*) é composto de um ou mais Objetos de Dados (*Data Object*) que são interpretados através de uma ou mais Informações de Representação (*Representation Information*) (CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS, 2012, p. 4-20), resume os objetivos da preservação digital a partir de três ações (BROWN, 2013, p. 202, tradução nossa):

- 1) Manter a usabilidade do objeto de informação, garantindo que se tem sempre alguma combinação do objeto de dados a rede de representação na qual pode ser acessada

usando-se tecnologias e conhecimentos disponíveis; 2) manter a integridade de cada objeto de dados, garantindo que está protegido de alterações e perdas, e; 3) manter a confiabilidade do objeto de informação, através de documentações e processos de curadoria apropriados.

Assim, a preservação dos documentos digitais necessita de técnicas para preservar o aspecto físico dos documentos, os *bits*, porém necessita que mais informações sejam preservadas para que seja possível a interpretação, leitura e uso dos documentos.

Para preservar a integridade dos documentos, primeiro deve-se preocupar com a mídia que armazena o seu conteúdo, buscando-se técnicas que evitem o corrompimento dos dados como a CRCs e *checksums* por exemplo. Em geral, os próprios Sistemas Operacionais já implementam mecanismos e utilizam soluções para evitar a alteração de dados, assim como para recuperar possíveis perdas ou rupturas decorrentes de erros nos dispositivos de armazenamento. Essa primeira camada de proteção auxilia a diminuir a complexidade dos sistemas de preservação como um todo.

A manutenção da integridade dos documentos ainda pode ser tratada com replicação dos conteúdos utilizando tecnologias e localizações diferentes entre as cópias. Essas estratégias permitem um certo grau de resiliência dos sistemas de preservação em relação a catástrofes, por exemplo, além de não o tornar dependente de um único tipo de tecnologia.

A preservação da identidade dos documentos requer o recolhimento de informações que permitam atestar que o documento foi criado de acordo com as devidas autorizações, por agentes devidamente autorizados garantindo que a mensagem contida no documento está de acordo com as regras que o regulamentam. Para realizar-se essa missão, há que se recolher informações, mais especificamente metadados, desde a criação do documento buscando formar o contexto que será necessário à determinação da sua identidade que, por consequência, deverá ser preservado junto ao documento.

Tarefas como a transmissão dos documentos em uma rede de computadores, por exemplo, podem apresentar falhas que alteram o conteúdo do documento, mesmo que determinadas sejam contornadas pelos softwares e não sejam percebidas pelas pessoas interessadas. De outro lado, certas ações necessárias para a preservação realizam a alteração proposital dos conteúdos dos documentos, porém, nesses casos deve-se fazer o registro detalhado dessas ações garantindo a autenticidade que pode vir a ser contestada futuramente. É possível afirmar que “se houver um ponto em que não possamos afirmar razoavelmente que o registro continua a carregar sua mensagem original intacta, nunca mais poderemos afirmar que ele é autêntico” (INTERPARES PRESERVATION TASK FORCE, 2001, p. 4, tradução nossa).

No meio analógico, o conteúdo e o suporte de um documento são intimamente relacionados, para não dizer inseparáveis, e o processo de presunção de autenticidade se debruça,

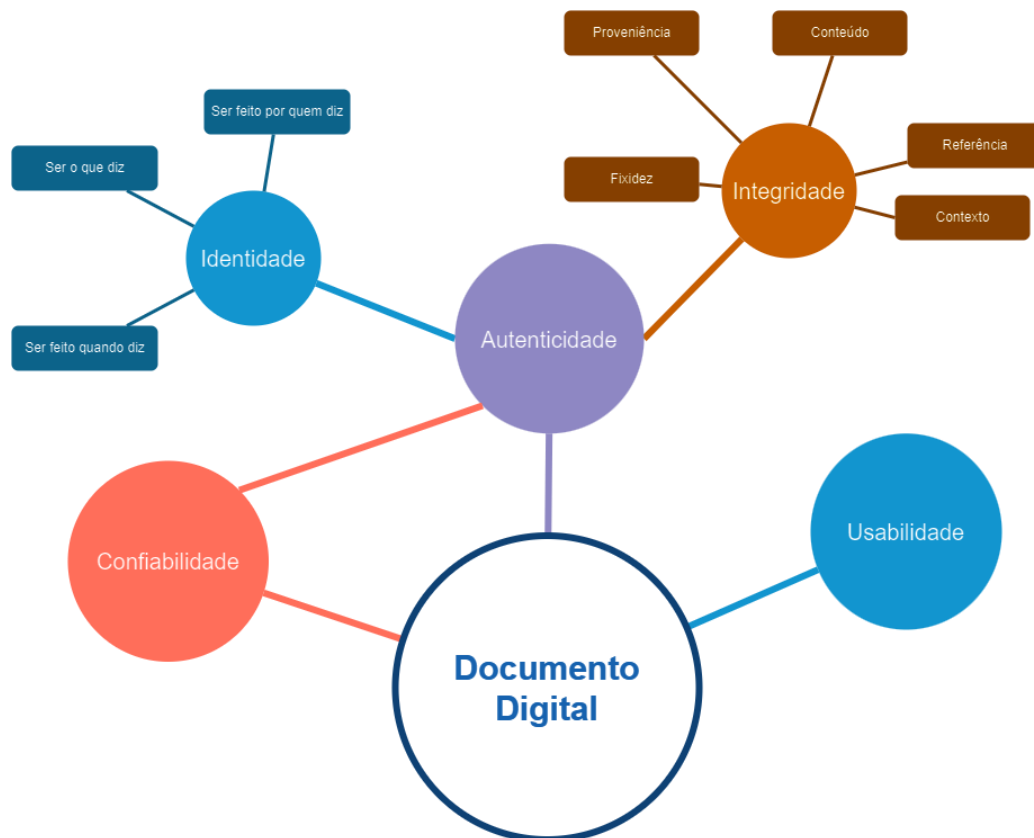
em grande parte, em analisar as características físicas do suporte (danificações, rasuras, tipo de papel usado, etc.), assim como os elementos intrínsecos (conteúdo, mensagem, autor, etc.) e extrínsecos (formato, tipologia, etc.). No meio digital, conteúdo e suporte nem sempre estão relacionados.

Ao se criar um documento em texto digital, por exemplo, as informações são inseridas através de um software e permanecem apenas na memória RAM do computador, que é efêmera e não durável. Ao se utilizar de uma funcionalidade específica desse software, os dados são então transferidos para um armazenamento durável para serem persistidos, utilizando-se muitas vezes um HD para isso. A transferência desse documento para outra pessoa ou máquina desacopla o seu conteúdo do suporte anterior. Conforme necessário, esporadicamente, o próprio Sistema Operacional utilizado pelo redator do documento pode escolher transferir os bits entre seções diferentes do HD a fim de executar otimizações internas. Assim, a metodologia para a comprovação da autenticidade dos documentos digitais deve se debruçar em analisar outros aspectos em detrimento do suporte.

Ferreira (2006, p. 51) declara que “o ambiente tecnológico é propício à introdução de modificações provocando um clima generalizado de desconfiança em relação à autenticidade deste tipo de material”. Muitas dessas modificações citadas podem ser fruto da aplicação de métodos de preservação dos documentos digitais ao se converter um formato de arquivo antigo para um mais atualizado, visando a interpretabilidade desses documentos, porém, quando essas ações são corretamente executadas e registradas, a autenticidade existente no formato de arquivo anterior do documento não pode ser refutada.

A Figura 2 faz um apanhado das principais propriedades relacionadas ao Documento Digital a serem levadas em consideração na sua preservação.

Figura 2 – Diagrama de propriedades relacionadas ao Documento Digital.



Fonte: elaboração própria

A propriedade “Identidade” é composta por elementos que permitam atestar que o documento é verídico (“Ser o que diz”), foi produzido sob a devida autorização (“Ser feito por quem diz”) e no momento atestado por ele (“Ser feito quando diz”). A propriedade “Integridade” é composta por elementos que se relacionam com o seu arquivo em si (“Conteúdo”), elementos para identificar corrompimentos (“Fixidez”), à sua identificação única e inequívoca (“Referência”), elementos que permitam rastrear a sua criação e ciclo de vida (“Proveniência”) e elementos que permitam conhecer o ambiente no qual o documento foi produzido e como ele se insere na organização (“Contexto”). A propriedade “Autenticidade” se forma pelo conjunto das propriedades “Identidade” e “Integridade”, sendo fruto da união das suas características, ao mesmo tempo em que se liga à propriedade “Confiabilidade”, que se refere à possibilidade de se dizer que o documento pode ser usado como fonte de prova. Por fim, tem-se a “Usabilidade” que refere a manter a possibilidade de uso documento digital ao longo do tempo, empregando-se as técnicas mais adequadas.

A seguir, são apresentados os desafios identificados na escolha das estratégias de preservação que um repositório deve realizar.

1.3.2 O Desafio de Escolha das Estratégias de Preservação

Nos Estados Unidos, em 1960, o *Census Bureau*, órgão responsável pela realização do censo nacional, armazenou os registros em formato eletrônico utilizando fitas UNIVAC tipo II-A como suporte, considerado então um padrão para armazenamento permanente (WATERS e GARRET, 1996). Após 16 anos, em 1976, o órgão norte-americano NARS (*National Archives and Records Services*) considerou que algumas séries dos dados recolhidos pelo censo possuíam caráter histórico e necessitavam ser preservados permanentemente, porém à essa altura, os leitores das fitas UNIVAC tipo II-A já não eram facilmente encontrados. Após esforços, foram localizados os equipamentos adequados para a leitura das fitas e os dados foram migrados, com para novas fitas, agora em uma tecnologia de padrão de indústria vastamente utilizado. Esse episódio, embora possua controvérsias a respeito do tratamento e a custódia dos dados¹⁷, explicita os perigos aos quais os dados eletrônicos, e por consequência, também os documentos eletrônicos, estão sujeitos.

Jones e Beagrie definem de maneira ampla a preservação digital como “a série de atividades gerenciadas necessárias para garantir o acesso contínuo a materiais digitais pelo tempo que for necessário” (2001, p. 24, tradução nossa). Os autores ainda definem o que seria a preservação a longo prazo: “acesso contínuo a materiais digitais, ou pelo menos às informações neles contidas, por tempo **indeterminado**” (2001, p. 25, tradução nossa, grifo nosso).

Beagrie *et al* (2008, p. 10, tradução nossa) resumem o grau de preocupação quanto à produção digital:

A preservação digital é o processo de gerenciamento ativo pelo qual garantimos que um objeto digital será acessível no futuro. O intervalo de tempo é potencialmente muito curto e, devido à rápida mudança na tecnologia e nos sistemas, isso nos afetará diretamente; e potencialmente, já que não temos ideia de até que ponto outros desejarão continuar acessando nossa produção digital do século XXI. Quando o fazem, é possível que grande parte da infraestrutura técnica que usamos para criar e ler nossos dados esteja indisponível.

A constante evolução tecnológica pode tornar obsoletos equipamentos até então tidos como o melhor padrão para a preservação, colocando em risco informações preciosas que podem desaparecer sem chances de serem recuperadas. Entre as estratégias usadas para se prevenir dos problemas de armazenamento, sejam eles decorridos pela fragilidade das mídias ou pela obsolescência dos equipamentos, é o refrescamento. Essa estratégia, que pode utilizar diversas técnicas, tem por objetivo copiar os dados de uma mídia para outra, em prazos estabelecidos, antecipando o tempo máximo de vida dos equipamentos, recomendado pelos fabricantes, visando

¹⁷ Os autores Donal Waters e John Garret (1996, p. 42) detalham, a partir de uma correspondência do *National Archives and Records Administration* (NARA), que pode não ter havido perda de dados, mas sim uma compreensão inadequada entre o órgão que possuía os dados, *Census Bureau*, e o NARS, que tinha interesse histórico nos dados, baseada no conhecimento limitado sobre os dados eletrônicos desse último.

otimizar o uso das mídias ao mesmo tempo em que diminui os riscos de ocorrências de falhas das mesmas.

Ferreira (2006, p. 33) considera que o refrescamento puro e simples não constitui uma estratégia de preservação digital, mas sim um pré-requisito fundamental: “a frequente verificação da integridade dos suportes físicos, assim como o seu refrescamento periódico, são consideradas atividades vitais num contexto de preservação digital”. Heslop, Davis e Wilson (2002, p. 11, tradução nossa) são categóricos e afirmam:

a curta vida útil das mídias de armazenamento contemporâneas significa que um programa de atualização constante das mídias é a única maneira de garantir a sobrevivência do material digital.

Há que se notar que os bits dos arquivos são transferidos de um lugar para outro, podendo sofrer modificações nesse processo de refrescamento das mídias, quando, por exemplo, a nova mídia utiliza uma codificação diferente da anterior. No mundo analógico, uma estratégia com esse mesmo intuito, de atualização do suporte, não pode ser realizada, gerando sempre novas cópias do documento original. No mundo digital, os bits originais do documento relativos ao instante da sua criação inscritos na mídia de suporte vão sendo migrados, evidenciando a forte diferença entre as metodologias de preservação dos dois mundos. Não obstante, cabe ressaltar que a realização dos refrescamentos deve ser realizada por pessoas capacitadas para que não haja perda de dados e documentando-se os procedimentos realizados.

Mais preocupante que a fragilidade e o tempo curto de vida das mídias de armazenamento, a obsolescência tecnológica traz desafios que a preservação digital deve lidar para conservar os documentos durante o tempo em que for necessário. Assim, os riscos da obsolescência digital não podem ser atacados apenas com a adoção de novos equipamentos o armazenamento dos documentos, necessitando atenção quanto aos formatos de arquivos que os documentos alvo de preservação utilizam.

A atualização de versões dos documentos ou a migração para novos formatos estão entre os procedimentos que a preservação pode lançar mão para manter os documentos acessíveis e interpretáveis ao longo do tempo. A atualização de versão em geral utiliza uma versão mais recente do *software* que produziu o documento original, gerando-se um novo documento, idêntico semanticamente ao original, porém agora acessível por versões mais atualizadas do *software*. Há que se notar que os bits do arquivo são efetivamente modificados para refletir a nova versão, no entanto, a preservação do objeto conceitual deve ser mantida, buscando aproximar o novo produto daquele produto original. Para Ferreira (2006, p. 38), “a atualização de versões é, possivelmente, a estratégia de preservação mais vulgarmente utilizada pela generalidade dos utilizadores”.

Outra estratégia que visa preservar a interpretabilidade dos documentos é a migração de formatos dos arquivos, convertendo os formatos originalmente fornecidos pelos produtores em outros formatos mais adequados à preservação. Em determinados casos, são identificados formatos de arquivos que podem ser descontinuados, suscetíveis à obsolescência, optando-se por migrar os documentos para um formato que suporte melhor a preservação a longo prazo. As empresas responsáveis por formatos de arquivos proprietários podem escolher não dar continuidade aos formatos antigos ou ainda deixar de suportar uma retrocompatibilidade dos formatos necessitando a migração dos documentos para outro formato.

A escolha dos formatos de migração deve considerar o menor grau de perda de informações relevantes sem que haja prejuízo ao objeto conceitual preservado, uma vez que os bits serão modificados. No caso da preservação de imagens, por exemplo, os responsáveis pela preservação devem especificar os parâmetros mínimos aceitáveis de qualidade para que a migração do documento original não sofra perdas significantes. Como definem Heslop, Davis e Wilson (2002, p. 12, tradução nossa):

Alguns atributos do objeto digital podem ser perdidos durante o processo de conversão [...]. O nível de perda de dados do processo de migração depende do número de tratamentos de preservação aplicados ao documento, da escolha do processo, do novo formato dos dados, do nível de intervenção humana e do trabalho descritivo pós-migração.

Alternativamente, pode-se utilizar o processo de normalização de formatos de arquivos, aceitando somente a ingestão de documentos que pertençam a um rol de formatos pré-determinados ou executando procedimentos de migração para esses formatos. A escolha por qual método adotar pode levar em consideração a complexidade e ferramentas necessárias para a migração para os formatos adequados ou demandar ao produtor do conteúdo o poder de decidir como realizar a migração de forma que o produto final tenha uma qualidade aceitável. A normalização dos formatos de arquivos tende a diminuir a complexidade de determinados processos realizados durante a preservação tornando-os mais facilitados e, por consequência, mais econômicos (FERREIRA, 2006, p. 38).

A definição do processo de normalização deve buscar por formatos de arquivo sustentáveis, em geral, padrões abertos, disponíveis à comunidade e independentes de plataforma (LEVELS OF PRESERVATION REVISIONS WORKING GROUP, 2019, p. 8). Um reflexo positivo da normalização é a uniformização do acervo, que pode facilitar a comunicação e interoperabilidade com outros repositórios que possuam essa mesma visão, mesmo que utilizem tecnologias diferentes.

Há ainda os casos em que a preservação utiliza estratégias de emulação, com as quais equipamentos atuais emulam um ambiente já obsoleto e de difícil reprodução. Como exemplo,

pode-se citar os casos de jogos de videogame de plataformas já em desuso ou raras como Atari, que requerem programas de computadores, chamados emuladores, para fornecer a mesma arquitetura de instruções que os equipamentos antigos possuíam. O mesmo caso se aplica a programas de *software* escritos em linguagens de programação descontinuadas, necessitando que haja um emulador que simulando a plataforma original para que eles sejam compilados e executados. Voutssás e Amozorrutia (2014, p. 101, tradução nossa), definem a emulação como: “A reprodução do comportamento e resultados de *softwares* e *hardwares* por meio do desenvolvimento de novos *softwares* e *hardwares* para serem executados em computadores atuais”. As técnicas de emulação ainda podem ser usadas nos casos de preservação de *websites*, construindo-se um contexto com as versões do navegador e sistema operacional utilizados, para que se mantenha a utilização do material mais fiel possível à experiência original. As demais características da preservação de *websites* são apresentadas na seção 1.3.6.

Em determinadas situações e de acordo com os interesses, pode ser necessária a preservação de todo o contexto tecnológico necessário para o acesso e uso dos materiais, preservando-se desde o sistema de arquivos, o sistema operacional, até os *softwares* que interpretam os formatos de arquivos e os disponibilizam para uso final. Busca-se, com isso, manter ao máximo a experiência, apresentação, renderização e interatividade dos materiais à época das suas produções. Essa estratégia específica se denomina como encapsulamento (BARNARD, DELGADO e VOUTSSÁS, 2014). Ferreira (2006, p. 43) complementa: “A estratégia de encapsulamento consiste em preservar, juntamente com o objeto digital, toda a informação necessária e suficiente para permitir o futuro desenvolvimento de conversores, visualizadores ou emuladores”. Nesse sentido, o encapsulamento preservaria o material digital para o uso futuro, na presença de ferramentas que permitam o seu uso, apoiando a própria emulação, quando for o caso. Um cenário em que as técnicas de encapsulamento podem ser úteis e recomendadas é o das artes digitais, no qual, se preserva todo o entorno tecnológico para que se mantenha a experiência interativa e audiovisual que autor projetou para as suas obras. A seção 1.3.6 aborda complementarmente a preservação de obras de arte digitais. Cabe referenciar o trabalho de Flores-Fernandez *et al.* (2022) que aborda outras estratégias de preservação complementares às apresentadas nessa seção do trabalho.

A Figura 3 retoma as principais estratégias de preservação abordadas nessa seção do trabalho. Inicialmente, cita-se o “Refrescamento” que significa migrar os bits e bytes dos arquivos para novas mídias. A “Emulação” visa preservar o material e usar emuladores daqueles *softwares* e *hardwares* que eram utilizados na produção desse material para a sua reprodução futura. O “Encapsulamento” requer, além da preservação do material, a preservação de todo o entorno de *hardware* e *software* tal e qual os utilizados à época da sua produção para que a sua reprodução seja

mais fiel possível ao passado. A “Normalização de formatos” é a decisão de o repositório utilizar e exigir somente determinados formatos de arquivo, visando facilitar a preservação e diminuir os riscos de não se conseguir utilizar os materiais futuramente. A “Migração de formatos” é a conversão dos formatos de arquivo recebidos para outros formatos mais adequados à preservação. Por fim, “Atualização de versões” é a utilização dos mesmos formatos de arquivo recebidos, porém em versões mais atuais, visando garantir o uso dos materiais com os *softwares* mais recentes.

Figura 3 – Estratégias de preservação.



Fonte: elaboração própria.

A próxima seção apresenta os desafios identificados em relação à criação e manutenção das infraestruturas das instituições responsáveis pela preservação de materiais.

1.3.3 O Desafio de Dispor de uma Infraestrutura Institucional de Preservação

Ao final do século passado, diversos estudos e iniciativas foram desenvolvidos com o intuito de compreender a forma como os documentos digitais são usados e como eles deveriam ser preservados. Deanna Marcum (1996, p. 451, tradução nossa), então presidente da Comissão de Preservação e Acesso do órgão norte-americano *Council on Library Resources*, alertava para alguns desafios da preservação digital:

Podemos ter certeza de que as informações armazenadas digitalmente estarão acessíveis no futuro? As vantagens da tecnologia digital são atenuadas por alguns riscos alarmantes. A leitura e a compreensão de informações em formato digital requerem equipamentos e *softwares*, que estão em constante mudança e podem desaparecer completamente do mercado dentro de alguns anos.

Portanto, o conjunto de ações a serem executadas na preservação dos documentos digitais requer uma estrutura com as devidas responsabilidades que seja respaldada por políticas bem definidas buscando atender as necessidades dos seus usuários. Waters e Garret definem o conceito de arquivos digitais como (1996, p. 8, tradução nossa):

[...] repositórios de informações digitais que são coletivamente responsáveis por garantir, por meio do exercício de várias estratégias de migração, a integridade e acessibilidade de longo prazo da herança social, econômica, cultural e intelectual da nação instanciada em forma digital.

Aqui, o termo nação é extrapolado para os arquivos digitais nacionais, mas a sua utilização pode se dar em âmbitos como instituições governamentais ou não, assim como empresas. Os autores ainda diferenciam arquivos digitais de bibliotecas digitais (WATERS e GARRET, 1996, p. 8, tradução nossa):

os arquivos digitais são distintos das bibliotecas digitais no sentido de que as bibliotecas digitais são repositórios que coletam e fornecem acesso a informações digitais, mas podem ou não fornecer armazenamento e acesso a longo prazo a essas informações.

Nesse sentido, os autores delimitam as responsabilidades das duas entidades, definindo que os arquivos digitais sempre são responsáveis pela preservação a longo prazo dos seus materiais, enquanto as bibliotecas digitais podem, opcionalmente, incorporar essa responsabilidade. Reconhece-se, no entanto, que ambas compartilham das mesmas funções como recebimento, seleção, avaliação, armazenamento e disponibilização de acesso.

1.3.4 O Desafio de Adaptar os Processos de Gestão Documental

Assumindo-se o fato de que a tecnologia evolui e os materiais digitais vão se tornando obsoletos, há a necessidade de se aplicar técnicas que modificam controladamente os materiais originais a fim de preservá-los e mantê-los usáveis. Logo, deve-se definir quais características terão preferência para serem preservadas em relação a outras, identificadas por Hedstrom e Lee, como propriedades significativas (*significant properties*), que afetam a **qualidade** (*quality*), **usabilidade** (*usability*), **renderização** (*rendering*) e **comportamento** (*behaviour*) dos objetos digitais. A decisão sobre quais propriedades devem ser consideradas leva em consideração diversos fatores (2002, p. 221, tradução nossa):

Nenhum conjunto único de regras de tomada de decisão pode ou deve ser aplicado a todas essas escolhas. Em vez disso, as decisões sobre quais propriedades significativas manter dependerão das prioridades institucionais, uso antecipado, conhecimento da comunidade designada, tipos de materiais envolvidos e recursos financeiros e técnicos disponíveis para o repositório.

Para cada tipo de material preservado, o repositório deve realizar uma análise para definição das propriedades significativas e, a partir disso, definir os métodos a serem empregados na preservação desses materiais. Como exemplo, pode-se citar a preservação de imagens, na qual o repositório decide não preservar formatos de arquivo que requerem muito espaço de armazenamento, por questões de custos e, por isso, define que todos os depósitos serão normalizados para um determinado formato de arquivo. Ao tomar essa decisão, o repositório deve avaliar se as propriedades significativas das imagens se mantêm, atendem às demandas da sua comunidade e permitem uma preservação sustentável ao longo do tempo. No caso de documentos textuais, o repositório pode optar por formatos de arquivo sustentáveis para a preservação como o PDF/A, por exemplo, desde que todas as propriedades significativas sejam preservadas e estejam de acordo com as necessidades da sua comunidade. No caso da preservação de websites, por exemplo, o repositório pode optar pela estratégia de emulação, preservando o conteúdo dos websites, navegador e sistema operacional, buscando manter a experiência visual, de comportamento e interação do usuário mais próxima possível do momento da produção dos documentos, caso essas sejam as demandas da comunidade.

Ao mesmo tempo em que se reconhece a importância da preservação dos materiais digitais, mais especificamente dos documentos digitais, ela somente pode ser realizada após a execução das funções arquivísticas de avaliação e seleção. Essas funções combinadas determinam os valores primários e secundários dos documentos analisando os seus elementos e relações com outros documentos, assim como determinam as suas destinações finais como a eliminação ou preservação permanente.

O resultado dessas funções é a determinação do que deve ser preservado ou esquecido para sempre, incumbindo um grau de importância e responsabilidade aos arquivistas que as realizam. Atribui-se a esses profissionais um papel social fundamental que permite ou nega, em determinados casos, o exercício da cidadania das pessoas, de onde se faz necessária a existência de critérios bem estabelecidos e transparentes para a avaliação e seleção dos documentos. Sobre a avaliação, Corujo (2020, p. 44), declara:

Para que este processo seja transparente, a avaliação deve inserir-se num quadro analítico sustentado em teorias, metodologias, políticas e programas explícitos, para os arquivistas prestarem contas do que fazem e de como o fazem na realidade, para que assim os resultados do seu trabalho possam ser mais bem contextualizados e percebidos pela sociedade para a qual trabalham. Tal exige que a avaliação deva ser exercida de maneira democrática, transparente, reflexiva e com base de evidência e justificativa (identificação e explicitação das decisões tomadas), já que ela é prova do poder exercido pelos arquivistas, cujas decisões têm reflexo no futuro, nomeadamente sobre as histórias que serão contadas (informação conservada) e as que não o serão (informação eliminada).

O autor dá aos arquivistas o papel central no processo de avaliação ao mesmo tempo em que ressalta a quem esses profissionais servem: a sociedade. Daí a necessidade de estabelecer transparência no processo, uma vez que a consequência desse trabalho será a construção do patrimônio documental contendo somente as informações selecionadas para preservação.

Como processo que apoia a avaliação e seleção, a classificação dos documentos desempenha papel de extrema importância na gestão dos documentos, ao passo em os agrupa, cria relações orgânicas entre eles e, dá a qualidade de organicidade ao arquivo. Schellenberg (2003, p. 39, tradução nossa) descreve a importância da organização das instituições, podendo ser traduzida pela classificação dos documentos: “a gestão eficiente de documentos públicos é de grande importância para o governo, e a eficiência de um governo pode frequentemente ser medida pela eficiência com que seus documentos são gerenciados.”

No Brasil, os órgãos públicos ligados ao Poder Executivo Federal utilizam o Código de Classificação de Documentos (PCD) e a Tabela de Temporalidade de Documentos (TTD) elaborados pela administração federal, sob a responsabilidade do Arquivo Nacional. No caso das atividades-meio dos órgãos, foi estabelecido um único CCD e TTD a ser seguido por todos os órgãos (ARQUIVO NACIONAL, 2020), já no caso das atividades-fim, conforme a natureza das atividades de cada órgão, foram desenvolvidos instrumentos específicos, como é o caso das IFES (Instituições Federais de Ensino Superior) (ARQUIVO NACIONAL, 2011). A critério de cada instituição, o Arquivo Nacional permite a inserção de tipos documentais nos CCDs para refletir a realidade dos seus documentos produzidos, respeitando a temporalidade e destinação final previstas no instrumento original ou estendendo-as, nunca diminuindo os tempos de guarda ou optando pela eliminação de documentos os quais estava prevista a preservação permanente.

A implantação dos CCD e TTD permite uma uniformização da gestão documental no âmbito do governo e, por consequência, estabelece critérios que garantem a preservação dos conjuntos de documentos necessários. Ao mesmo tempo em que essa metodologia retira dos arquivistas das instituições a responsabilidade por definir a destinação final dos documentos, o que deve ser preservado ou eliminado, ela acaba por servir aos governos, podendo ser tendenciosa de acordo com os interesses dos governantes em mandato. Sobre esse aspecto, Corujo (2020, p. 51) faz a seguinte reflexão: “os arquivistas devem questionar-se para quem trabalham: para o Estado e o seu aparelho burocrático centralista que se concilia com os interesses da elite, ou para a população, as comunidades – os ‘indígenas’ ou ‘autóctones’”.

Cabe aos arquivistas dos órgãos do governo que elaboram os CCDs e TTDs a intermediação entre os interesses do próprio governo e os interesses identificados na sociedade a fim de que se incluam, por exemplo, minorias que não encontram representatividade, mas que

poderão fazer uso de documentos preservados para a comprovação ou a reivindicação de direitos no futuro. Não obstante o conjunto de habilidades e conhecimentos técnicos requeridos pelos arquivistas, esse processo ainda requer habilidades políticas para que os interesses identificados na comunidade encontrem representação na gestão documental. Corujo (2020, p. 51) resume apropriadamente o desafio dos profissionais arquivistas:

A informação digital pode servir para ampliar o espectro de alcance e disseminação, mas também é preciso ter em conta como matéria em que se encontram os vestígios de memória dos não representados oficialmente. Ao avaliar esta informação, o arquivista será a ponte entre as instituições da Administração Pública e, mais do que a chamada “Sociedade Civil”, deverá ir mais além, e procurar a memória de quem não está representado, de quem é considerado marginal. Só assim é possível transmitir uma visão plural da Sociedade para o futuro.

Aplicados os devidos processos para a seleção dos documentos a serem preservados, tem-se a necessidade de criação de uma entidade que se responsabilize pelo recebimento dos documentos e aplique os processos necessários que vão garantir a preservação e acesso a longo prazo, construindo a memória documental das instituições. O relatório de Waters e Garret (1996, p. 9) apresenta o termo *repositories of digital information* (repositório de informações digitais) como um sistema nacional de repositórios interligados que garantam que nenhuma informação valiosa se perca e que elas estejam acessíveis no futuro. Para essa realização, os autores citam dois mecanismos, sendo o primeiro o cumprimento de critérios normalizados de certificações independentes, demonstrando aos usuários, administradores e financiadores dos repositórios as garantias de possuem a responsabilidade de preservar as informações digitais a longo prazo. O segundo mecanismo é o estabelecimento de uma estrutura à prova de falhas que suporte os próprios repositórios, respaldada por uma estrutura organizacional, apoiada em uma sustentabilidade financeira e respaldada pelos meios legais vigentes de cada realidade.

Para os autores, os repositórios digitais e as certificações são indissociáveis (WATERS e GARRET, 1996, p. 8, tradução nossa):

Sem a operação de um programa formal de certificação e um mecanismo à prova de falhas, a preservação do patrimônio cultural da nação em formato digital provavelmente dependerá excessivamente das forças do mercado, que podem valorizar as informações por um período muito curto e sem aplicar critérios de interesse público mais amplos.

A forma como essas garantias são expressadas e reconhecidas pelos usuários, administradores e financiadores dos repositórios se dá através do cumprimento de critérios normalizados de certificações independentes. O aspecto independente, nesse caso, busca tornar a avaliação de critérios o mais imparcial possível, livrando-se de possíveis visões enviesadas.

1.3.5 O Desafio de Garantir a Cadeia de Custódia e a Cadeia de Preservação

Enquanto as organizações tratavam apenas de documentos analógicos, a garantia de autenticidade se dava através do cuidado com a posse dos documentos, garantindo que somente as pessoas autorizadas tivessem acesso aos documentos. Entre a produção, usos dos documentos e a transferência para guarda em uma instituição arquivística custodiadora, o principal cuidado era para que a Cadeia de Custódia (*Chain of Custody* – CoC) com responsabilidades bem definidas não fosse interrompida a fim de se quebrar a confiança dos documentos em transferência (HIRTLE, 2000). Assume-se aqui o conceito de custódia (ARQUIVO NACIONAL, 2005, p. 62): “responsabilidade jurídica de guarda e proteção de arquivos, independentemente de vínculo de propriedade.”

O princípio da não interrupção da cadeia de custódia estabelece que “[...] através dos seus ciclos de vida, os documentos deveriam estar sob custódia de partes conhecidas e confiáveis para preservá-los intactos” (INTERPARES PRESERVATION TASK FORCE, 2001, p. 3, tradução nossa). Por consequência, a ruptura da Cadeia de Custódia cria brechas que facilitam adulterações nos documentos, assim como “[...] podem tornar impossível demonstrar que um documento não foi alterado” (INTERPARES PRESERVATION TASK FORCE, 2001, p. 3, tradução nossa).

O fato de se identificar que em algum dado momento durante o ciclo de vida do documento não foram adotadas medidas protetivas para a manter a integridade dos documentos já pode lançar dúvidas significativas sobre a autenticidade dos documentos armazenados (HACKETT, 2008, p. 704). Portanto, além de não haver a ruptura da Cadeia de Custódia, deve-se observar se todas as medidas de proteção dos documentos foram adotadas para que seja possível presumir as suas autenticidades. Higashi *et al* (2020, p. 505) destacam: “a escolha do custodiador confiável torna-se ainda mais proeminente quando se trata de documentos arquivísticos em ambiente digital, tendo em vista a facilidade de alterá-los sem deixar vestígios”.

A partir do crescente uso dos documentos digitais nas organizações e diante das ameaças que eles sofrem pelo fácil corrompimento ou pela interrupção de acesso pela obsolescência tecnológica, por exemplo, outras preocupações foram agregadas à Cadeia de Custódia, buscando minimizar os fatores que dificultem a presunção de autenticidade dos documentos. Criou-se assim o conceito de Cadeia de Preservação (*Chain of Preservation* – CoP) que é definida como “o conjunto de controles que se estende por todo o ciclo de vida dos documentos de forma a garantir as suas identidades e integridades ao longo do tempo” (DURANTI, MACNEIL, *et al.*, 2005, p. 4, tradução nossa). Complementarmente, o armazenamento do documento deve ser realizado de maneira a se garantir a sua correta interpretação e apresentação tal qual foi criado, utilizando-se informações para esse intuito além de outras que sejam necessárias para se presumir a autenticidade desde o momento de sua criação (INTERPARES PRESERVATION TASK FORCE, 2001).

Assim, entendeu-se a necessidade de haver uma Cadeia de Custódia Digital Arquivística (CCDA) que além de permanecer ininterrupta durante o uso dentro do sistema de gestão e na transferência para depósito no repositório, deve também incorporar informações que permitam a presunção de autenticidade desde a produção dos documentos a sua destinação final (FLORES, ROCCO e SANTOS, 2016) (GAVA e FLORES, 2021).

As políticas arquivísticas da organização devem prever os contratos e condições a serem cumpridas entre o sistema de gestão de documentos e a entidade responsável pela preservação dos documentos, de forma a não haver rupturas da CCDA, ainda que a produção e preservação ocorram dentro de uma mesma organização. A política de arquivos deve especificar a forma como a transferência dos documentos para a preservação ocorre, incluindo os requisitos mínimos e as responsabilidades inerentes a cada parte, sendo aconselhado que essa comunicação seja feita através de ferramentas automatizadas diminuindo-se interferências humanas. Deve-se buscar uma interoperabilidade entre os sistemas que diminua os riscos de perda e adulterações de informações, garantindo-se assim, a continuidade e fluidez da CCDA. Flores, Rocco e Santos (2016, p. 122) detalham de que forma o documento deve sair do ambiente de gestão e ser transferido para o ambiente de preservação:

A interoperabilidade entre o Sigad e o RDC-Arq deve ser feita de forma automatizada, de modo que não existam manipulações não autorizadas, e não se permitam inserções ou importações de documentos ou de pacotes de documentos que não estejam contemplados pela cadeia de custódia digital, que documenta todas as atividades, rotinas, recepções, conversões, normalizações etc., em metadados que ficam encapsulados nos pacotes de interoperabilidade de documentos. É importante destacar que esta interoperabilidade consiste basicamente em um empacotamento de acordo com os padrões da área e da adoção dos metadados, também de acordo com os padrões da área.

A Cadeia de Custódia se relaciona mais a um princípio arquivístico que trata de **restrições, permissões de acesso** e escolha de parceiros, por exemplo, enquanto a Cadeia de Preservação apresenta principalmente problemas a serem enfrentados através de **recursos computacionais para o armazenamento** dos documentos e seus elementos, assim como a comunicação entre os participantes da cadeia. A “CoP por sua vez, é a implementação tecnológica da CoC, que agora, precisa ser ressignificada em um ambiente digital, já que a CoP não substitui a CoC enquanto princípio” (GAVA e FLORES, 2021, p. 89). Para Gava e Flores (2021, p. 432), os dois conceitos representam “[...] processos que devem ocorrer concomitantemente, ambos buscando a manutenção da autenticidade dos documentos”.

Enquanto a preservação digital dos documentos sofre com as características intrínsecas de sua própria natureza como a facilidade de duplicação, renomeação, conversão, corrompimento e

a adulteração sem deixar vestígios, a preocupação com a autenticidade é constante e deve apresentar meios que a comprovem.

[...] a presunção da sua autenticidade deve se apoiar na evidência de que eles foram mantidos com uso de tecnologias e procedimentos administrativos que garantiram a sua identidade e integridade (componentes da autenticidade); ou que, pelo menos, minimizaram os riscos de modificações dos documentos a partir do momento em que foram salvos pela primeira vez e em todos os acessos subsequentes (CONSELHO NACIONAL DE ARQUIVOS, 2012, p. 1).

Enquanto Terry Cook (1993) discutiu o paradigma custodial e abordou o paradigma da arquivologia pós-custodial, na qual os documentos nem sempre são custodiados por seus produtores, mas são repassados a outras entidades com essa responsabilidade, Gava e Flores (2022) rediscutem essa visão, trazendo à tona as questões envolvidas Preservação Digital Sistêmica das instituições que acabam por requer uma CCDA compartilhada.

Partindo-se da CCDA, faz necessário compartilhamento da custódia dos documentos, necessitando a presença de, no mínimo, dois papéis para a realização da Preservação Digital Sistêmica de uma instituição: o administrador de preservação e o administrador de TI. O primeiro papel, geralmente desempenhado por arquivistas, lida com as questões arquivísticas atuando a partir das políticas arquivísticas de gerenciamento, preservação e acesso de documentos definidas, assim como interagindo ativamente nas suas revisões e reestruturações, buscando aplicar as melhores práticas disponíveis. O segundo papel, de administrador de TI, interage com o ambiente tecnológico sobre o qual a preservação digital se dá, utilizando e monitorando os *softwares* e *hardwares* envolvidos garantindo que os requisitos funcionais e não-funcionais sejam cumpridos (GAVA e FLORES, 2022). Faz-se necessário ressaltar que a custódia não pertence a um ou outro papel, mas é compartilhada entre eles, tendo ambos as suas responsabilidades definidas nas políticas institucionais.

1.3.6 O Desafio de Preservar Recursos Especiais

Ainda que os repositórios sejam estabelecidos, implantados e operacionalizados de maneira adequada à preservação digital, a natureza dos materiais sob custódia traz desafios a serem superados em maior ou menor grau de dificuldade. A preservação de *softwares* e jogos digitais, por exemplo, requer estratégias mais elaboradas e sofisticadas como a emulação para a preservação do objeto lógico. Nesse caso, utiliza-se um emulador de sistema operacional ou de *hardware* que permitirá a execução futura do *software* de maneira mais fiel ao método nativo ao qual o objeto foi criado. No entanto, a técnica de emulação requer o registro em detalhes de todo o ambiente tecnológico necessário para a execução do *software*, necessitando em alguns casos da preservação

dos próprios *softwares* de suporte, evoluindo para a técnica de encapsulamento. A alta complexidade dessa técnica, que pode requerer diversas combinações de *softwares* e arquiteturas de *hardware* é aconselhada quando se deseja preservar a forma de funcionamento do objeto preservado tal qual ele foi criado, caso contrário, quando possível, recomenda-se a migração de formatos (FERREIRA, 2006, p. 33).

Um outro gênero que lança desafios à preservação digital é o das obras de arte digitais, incluindo desde quadros digitais, imersões em realidade virtual até intervenções artísticas que interagem com o público. Mesmo uma sutil mudança tecnológica, como o aumento de resolução da tela, a descalibragem do perfil de cores de um monitor ou a ausência de conectividade da rede, pode prejudicar a percepção da obra de arte, desvirtuando a intenção inicial do artista. O mesmo pode ser aplicado ao caso de uma música preservada e reproduzida em um futuro distante, quando pode ser impossível reproduzir as sensações e nuances criadas pelo autor. Há que se analisar ainda o propósito da obra de arte digital e a forma como o seu envelhecimento ou obsolescência a afeta na medida em que o desejo do artista seja propositadamente a mudança ou adaptação ao longo do tempo, ou ainda, o desejo de que a obra cesse o seu funcionamento e pare de existir.

Christiane (2019) cita possíveis métodos e processos para a preservação da arte, partindo-se da estratégia mais “inelegante e impraticável” que seria coletar exatamente os *software* e *hardware* usados para a exibição das obras de arte, porém acabaria por tonar as instituições em museus de computadores. O segundo método seria a emulação dos equipamentos utilizados para a exibição da obra de arte, recriando o ambiente, porém se utilizar os equipamentos originais. O terceiro método seria a migração da obra de arte para formatos mais atualizados, o que nem sempre seria possível para todas as obras de arte. A nível global, existem iniciativas dedicadas a pesquisar e desenvolver métodos adequados voltados especificamente para a arte digital, dentre as quais, pode-se citar a INCCA (*International Network for the Conservation of Contemporary Art*), fundada em 1999 que tem entre os seus membros, arquivistas, historiadores da arte e profissionais ligados à curadoria e preservação da arte, tendo como missão “Coletar, compartilhar e preservar o conhecimento necessário para a conservação da arte moderna e contemporânea” (INTERNATIONAL NETWORK FOR THE CONSERVATION OF CONTEMPORARY ART, 2015).

A preservação de *websites* é outro gênero de documentos que traz desafios, ao que Pennock descreve (2013, p. 3, tradução nossa):

Uma das principais razões para o arquivamento de sites, particularmente na comunidade de patrimônio cultural, é o prazo relativamente curto dentro do qual o conteúdo pode ser, e tem sido, “perdido”: a web pode ser onipresente, mas os sites são transitórios.

Ainda que os *websites* estejam presentes na realidade e cotidiano das pessoas, a sua natureza sempre mutante pode tornar os conteúdos inacessíveis ou perdidos devido a reformulações, mudanças de projeto ou devido aos algoritmos que definem quais conteúdos são visíveis aos usuários. Em determinados casos, as referências usadas para os conteúdos (*links*) deixam de existir, impactando o acesso e localização desses conteúdos, que podem ser desde notícias, curiosidades, documentos oficiais governamentais até publicações científicas.

Partes restritas dos sites, que requerem autenticação com usuário e senha, por exemplo, também podem ser um obstáculo à preservação dos websites, necessitando tratamento adequado. A utilização de ferramentas que bloqueiam a utilização de robôs, conhecidos como *captchas*¹⁸, que requerem a solução de desafios por humanos pode prejudicar ou inviabilizar em certos casos a extração dos conteúdos dos sites.

Pennock oferece uma noção sobre os desafios do arquivamento de *websites* (2013, p. 4, tradução nossa):

Capturar e arquivar um site individual, pequeno e simples pode ser relativamente simples; ele pode ser copiado ou baixado do servidor e armazenado *offline* com relativamente pouco conhecimento técnico. O arquivamento da Web em escala e a longo prazo é um assunto mais complicado. Capturar sites grandes e complexos de forma recorrente, mantendo e identificando claramente as relações entre diferentes versões de um site e gerenciando simultaneamente os limites artificiais que inevitavelmente ocorrem em uma coleção ‘extraída’, requer uma solução mais complicada. Acrescente a escala absoluta de uma coleção de “domínio nacional” e o problema de longo prazo da obsolescência tecnológica, e o desafio aumenta novamente.

Dentre as técnicas usadas para a realização do arquivamento, pode-se a utilização dos *webcrawlers* que acessam o *site*, navegam pelos *links* e armazenam os conteúdos acessados, numa perspectiva ligada ao cliente do *site*. Dentre os desafios dessa técnica está na definição da quantidade de níveis que a ferramenta deve seguir e se a ferramenta deve navegar por outros sites externos aos do domínio inicial, por exemplo, o que pode levar a uma navegação infinita, assim como a necessidade de armazenamento infinito. Outras técnicas realizam o arquivamento pelo lado do servidor dos conteúdos, armazenando todas as versões disponibilizadas ao público.

Preservar a *web* ainda levanta preocupações sobre direitos de propriedade intelectual. Os sites podem conter materiais protegidos, incluindo imagens, vídeos e conteúdo escrito, que requerem permissões para preservação a longo prazo e acesso público. Navegar pelas complexidades legais e obter as permissões necessárias pode ser um desafio significativo. A escala absoluta da *web* representa um desafio substancial para os esforços de preservação. Com bilhões de sites e uma enorme quantidade de dados, é um desafio identificar e capturar tudo, necessitando

¹⁸ Programa ou sistema destinado a distinguir a entrada humana da máquina, geralmente como uma forma de impedir o spam e a extração automatizada de dados de sites.

abordagens seletivas para priorizar conteúdos valiosos, de interesse das instituições, e em risco para preservação.

O estabelecimento do escopo do conteúdo a ser preservado deve ser definido de acordo com os objetivos de cada instituição, optando-se por restringir a coleta a um determinado conjunto de sites ou até mesmo a todos os sites de um determinado país, por exemplo. Em ambos os casos, existem dificuldades em se definir os limites dessa coleta, por mais que um site pertença a um determinado país, ele pode conter conteúdos provenientes de sites de outros países. Cada estratégia definida vai possuir pontos fortes e fraquezas, cabendo à instituição determinar quais características serão priorizadas.

Há que se analisar os aspectos financeiros relacionados à preservação *web*, que requer recursos significativos, incluindo armazenamento, capacidade computacional e experiência. O financiamento para iniciativas de preservação da web pode ser limitado, dificultando a manutenção de esforços de preservação de longo prazo. A colaboração entre instituições, governos e organizações privadas pode representar um meio para garantir apoio e investimento contínuos.

Por mais que seja possível preservar as informações extraídas, a experiência do usuário na usabilidade e navegabilidade dos sites também pode impor desafios à preservação, uma vez que determinadas informações podem ser apresentadas como figuras, animações e vídeos, por exemplo. Outro aspecto a ser levado em consideração para a preservação, são os navegadores utilizados para acesso aos sites, uma vez que cada ferramenta pode implementar mecanismos diferentes de renderização das informações, trazendo experiências de uso diferentes.

Há ainda o desafio de se determinar a autenticidade dos websites arquivados, uma vez que pode ser impossível garantir a mesma forma como o conteúdo foi apresentado originalmente aos usuários, seja pela dinamicidade dessas informações, seja pelos métodos nativos que cada navegador pode utilizar para renderizar o conteúdo. Os sites estão sujeitos a constantes mudanças, atualizações e modificações, tornando difícil determinar o estado original de uma página da web. Garantir a precisão e a confiabilidade do conteúdo da *web* preservado exige documentação e metadados robustos.

Ball (2010, p. 10) sugere um conjunto de quatro propriedades significantes dos websites que devem ser mantidas durante a preservação e que vão subsidiar a autenticidade: texto, aparência, interatividade e dinamismo. A propriedade **texto** se refere ao conteúdo textual do site, que geralmente é incluído através de marcações HTML, imagens e demais conteúdos requisitados dinamicamente através de *javascripts*. A propriedade **aparência** se refere à disposição dos elementos e às cores, folhas de estilo (CSS), tamanhos e larguras dos elementos para o desenho da página. A **interatividade** se refere à navegação que o usuário executa no site através de *links* e

botões, por exemplo, utilizando somente os recursos do navegador, mas também inclui a interatividade na qual o usuário envia dados para o servidor e o conteúdo retornado altera o próprio conteúdo ou estilo visual, o que requer adicionalmente a preservação do servidor. A propriedade **dinamismo** se refere aos conteúdos gerados sob demanda da interação do usuário com o site, podendo ser fruto da execução de *Javascript* ou de conteúdos gerados pelo servidor, o que pode requerer a preservação de *snapshots* do servidor, incluindo-se os dados de bancos de dados.

1.4 AS NORMAS COMO FERRAMENTAS PARA ENFRENTAR OS DESAFIOS DA PRESERVAÇÃO

Uma vez que a preservação de materiais ultrapassa diferentes áreas e nichos que vão desde documentos arquivísticos, materiais museológicos até obras de arte e que se funde com a história da humanidade, é normal que se façam normas, padrões e guias de boas práticas para a sua realização. Esses instrumentos buscam estabelecer os melhores meios aplicáveis aos diferentes contextos para que os materiais permaneçam acessíveis, utilizáveis e confiáveis pelo maior tempo possível.

Nesse âmbito, podem-se citar normas ligadas à descrição arquivística, gestão documental e à preservação digital em si, formando um corpus regulatório, ou recomendatório, de extrema importância para que se alcance a excelência quanto à utilização de materiais num futuro muitas vezes inalcançável para o pensamento humano atual. A seguir são apresentados alguns dos marcos normativos que se relacionam direta ou indiretamente com a preservação de documentos.

1.4.1 Um Marco Normativo Complexo

A criação de normas a nível mundial constitui um grau de importância no sentido de permitir um entendimento comum nas mais variadas áreas, criando estruturas que permitem o desenvolvimento de processos, ferramentas e sistemas padronizados.

No âmbito da gestão documental, tem-se a ISO 15489 com a sua primeira parte publicada no início do século XXI, em 2001, modificada em 2016 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016), revisada em 2021 e ainda em vigor. A norma detalha os processos os quais os Sistemas de Gestão de Documentos de Arquivo (SGDA) (*Management Systems for Records – MSR*) possuem responsabilidade como a criação e captura de documentos, classificação e indexação, controle de acesso, armazenamento, uso e reuso, migração e conversão e destinação final. Sobretudo, a norma se debruça na gestão nas fases ativas, com prazos

curtos ou médios e não avança na preservação a longo prazo. A ISO 15489 é tratada com mais detalhes na seção 1.4.2.

Com o escopo voltado à criação, tratamento e uso de metadados dos documentos arquivísticos, foi criada a norma ISO 23081 em 2006 apoiando a ISO 15489 na gestão documental. A norma define que “a gestão de metadados é uma parte indissociável da gestão de documentos de arquivo, servindo a uma diversidade de funções e propósitos” (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2019, p. 3) tendo os metadados o papel de registrar contexto, conteúdo, estrutura dos documentos, além de registrar o histórico de gestão dos próprios documentos. A norma possui três partes, sendo a terceira, ISO 23081-3 voltada à realização de autoavaliações por parte das organizações, permitindo identificar a qualidade e estrutura dos metadados recolhidos pelos sistemas. A norma ISO 23081 é abordada da seção 1.4.5

Em 2005, se estabeleceu a norma ISO/TR 18492 preocupada com a preservação a longo prazo de documentos eletrônicos, definindo longo prazo como “[...] quando o período de retenção excede a vida útil esperada da tecnologia (*hardware* e *software*) usada para criar e manter as informações” (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2005). A norma reconhece o desafio de preservar documentos eletrônicos confiáveis por longos períodos de tempo e busca esclarecer e unificar as iniciativas utilizadas com esse propósito que, se acaso forem realizadas indiscriminadamente, podem levar a perda de dados ou da autenticidade dos documentos. TENTAR BAIXAR.

Ainda em apoio à ISO 15489, em 2012 foi elaborada a norma ISO 13008 para regulamentar os processos de conversão e migração de documentos de arquivo de forma a mantê-los autênticos e confiáveis (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2022). A norma define os requisitos que os sistemas de gestão documental devem atender para evitar problemas relacionados à obsolescência dos arquivos, diante das constantes mudanças tecnológicas, detalhando planos de migração e controles para a verificação de conformidade. Assim, busca-se manter a autenticidade dos documentos através de mudanças controladas, documentadas e regulamentadas.

No âmbito da gestão documental, surge a série de normas ISO 30300, voltada à implementação e governança de Sistemas de Gestão de Documentos de Arquivo (SGDA), se juntado a outras normas de sistemas de gestão como a ISO 9000, ISO 14001 e ISO/IEC 27001. A norma apoia o desenvolvimento de uma política de documentos de arquivo e prevê mecanismos para a avaliação de eficiência e monitoramento de desempenho do SGDA implantado. A seção 1.4.3 aborda a ISO 30300 mais detalhadamente.

No que se refere à descrição arquivística, tida como um dos principais processos executados em um arquivo, podem-se citar diversas normas tanto a nível internacional como iniciativas nacionais. Na década de 1990, o ICA (*International Council on Archives* - Conselho Internacional de Arquivos), iniciou a criação da *General International Standard Archival Description* (ISAD (G)) (Norma Geral Internacional de Descrição Arquivística), através de uma comissão inicialmente *Ad Hoc*, entregando em 1994 a primeira versão da norma. Em 2000, a segunda versão da norma foi publicada, incorporando revisões identificadas a partir de consultas à comunidade arquivística (CONSELHO INTERNACIONAL DE ARQUIVOS, 2000). Essa é a versão final da norma utilizada até os dias atuais. Posteriormente, o ICA ainda elaborou as normas ISAAR(CPF)¹⁹ (*International Standard Archival Authority Record for Corporate Bodies, Persons and Families* – Norma Internacional de Autoridade Arquivística para Entidades Coletivas, Pessoas e Famílias) para a descrição dos agentes que interagem com as entidades do arquivo, ISDF²⁰ (*International Standard for Describing Functions* – Norma Internacional para Descrição de Funções) com o objetivo de apresentar elementos para a descrição das funções desempenhadas pelos agentes e a ISDIAH²¹ (*International Standard for Describing Institutions with Archival Holdings* – Norma Internacional para Descrição de Instituições com Acervo Arquivístico) para fornecer elementos para a descrição das entidades custodiadoras em si.

Em 2006, no Brasil, é lançada a Norma Brasileira de Descrição Arquivística (NOBRADE) que foi elaborada pela Câmara Técnica de Descrição Arquivística (CTNDA) do Conselho Nacional de Arquivos (Conarq) através de um processo de elaboração que incluiu rodadas de consultas à comunidade arquivística de diversas regiões brasileiras. Ao se posicionar como não sendo uma tradução da ISAD(G) e ISAAR(CPF), tem como objetivo “a adaptação das normas internacionais à realidade brasileira, incorporando preocupações que o Comitê de Normas de Descrição do Conselho Internacional de Arquivos (CDS/CIA) considerava importantes, porém, de foro nacional” (CONSELHO NACIONAL DE ARQUIVOS, 2006, p. 9).

Em 2007, o governo espanhol criou a *Comisión de Normas Españolas de Descripción Archivística* (Comissão de Normas Espanholas de Descrição Arquivística – CNEDA) ligada ao Ministério da Cultura com o objetivo de desenvolver e atualizar as Normas Espanholas de Descrição Arquivística (NEDA), consideradas “[...] uma ferramenta aplicável e orientada à melhora contínua do acesso aos documentos de arquivo” (ESPANHA, 2007, p. 23.899). Após a identificação da necessidade atualizar o padrão de descrição arquivística espanhol, a CNEDA iniciou os seus trabalhos, publicando em 2017, após um trabalho de dez anos, o Modelo Conceitual de Descrição

¹⁹ (INTERNATIONAL COUNCIL ON ARCHIVES, 2004)

²⁰ (INTERNATIONAL COUNCIL ON ARCHIVES, 2007)

²¹ (INTERNATIONAL COUNCIL ON ARCHIVES, 2008)

Arquivística (NEDA-MC) (COMISIÓN DE NORMAS ESPAÑOLAS DE DESCRIPCIÓN ARCHIVÍSTICA, 2017).

Em 2016, o ICA, através do EGAD (*Experts Group on Archival Description* – Grupo de Especialistas em Descrição Arquivística) publicou a primeira proposta do modelo de descrição: “*Records in Contexts: a conceptual model for archival description*” (INTERNATIONAL COUNCIL ON ARCHIVES, 2016). Essa primeira publicação foi lançada em forma de rascunho de consulta (*consultation draft*) com o objetivo de difundi-lo na comunidade arquivística e fomentar pesquisas e contribuições para a elaboração final de um padrão que atenda às diversas realidades existentes nos arquivos mundo a fora. Além dos arquivistas, profissionais da gestão de documentos e da TI também foram instigados a analisar o modelo para contribuir com as suas colaborações pertinentes às suas respectivas áreas. Em 2022, houve uma nova rodada de consulta à comunidade para receber as contribuições da versão 0.2 e, após a compilação dos recebimentos, a versão foi disponibilizada para apreciação e identificação de pontos de melhoria que possam integrar uma nova futura versão do modelo (INTERNATIONAL COUNCIL ON ARCHIVES, 2022).

Esse apanhado de normas ligadas à gestão documental, incluindo a descrição arquivística, permite perceber a complexidade desses assuntos quando se trata de normalização e padronização. Um estudo aprofundado de todas as normas e suas interrelações acrescentaria uma complexidade ao presente tudo, tangenciando os seus objetivos propostos. Assim, extraiu-se um subconjunto para abordagem mais detalhada nas seguintes seções, baseando-se na importância e impacto para a gestão documental como um todo.

1.4.2 ISO 15489

Lançada com o intuito de criar um arcabouço teórico sobre o gerenciamento de documentos, a *International Organization for Standardization* (ISO) publicou a norma ISO 15489-1:2001, posteriormente revisada e substituída pela ISO 15489-1:2016, estabelecendo os conceitos e principais diretrizes sobre gestão de documentos. Dentre as definições iniciais está o conceito de processo de trabalho, definido como “a sequência de uma ou mais ações necessárias para produzir um resultado que cumpre com as regras que o regem” (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016, p. 3, tradução nossa). A norma explicita os processos envolvidos na criação, captura e gerenciamento de documentos: a) criação de documentos; b) captura de documentos; c) classificação e indexação; d) controle de acesso; e) **armazenamento**; f) **uso e reuso**; g) **migração** ou **conversão**; h) **destinação final**. Para cada um dos processos identificados, a norma estabelece os requisitos mínimos a serem atendidos pelos Sistemas de Gestão de Documentos de

Arquivo (SGDA) (*Management Systems for Records – MSR*) e cujos controles são encontrados no Anexo A da norma ISO 30301:2019. Destacam-se os processos mais comprometidos com o armazenamento e preservação a curto prazo (armazenamento durante um período de tempo definido, menos propenso a mudanças tecnológicas) e médio prazo (período de tempo mais ampliado, mas não indefinido, com possíveis mudanças tecnológicas,) pois a norma, como foi citado, tem uma orientação para os sistemas de gestão documental e não tanto para documentos permanentes, acessíveis em um período de tempo indefinido.

No Brasil, a sua tradução é representada pela norma ABNT NBR ISO 15489-1:2018 (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2018), já na Espanha a sua equivalente é a norma UNE ISO 15489-1:2016 (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2016). O escopo da norma torna claro que a sua aplicação engloba a preservação de documentos independente do seu suporte, sendo útil aos documentos analógicos, eletrônicos e digitais, uma vez que trata da construção dos pilares de uma instituição de arquivo. De uma forma geral, a norma pode ser interpretada como um guia e um manual de boas práticas a ser aplicado nos arquivos para atender os requisitos necessários de gerenciamento de documentos (MORO-CABERO, MARTÍN-POZUELO e BONAL-ZAZO, 2011, p. 108).

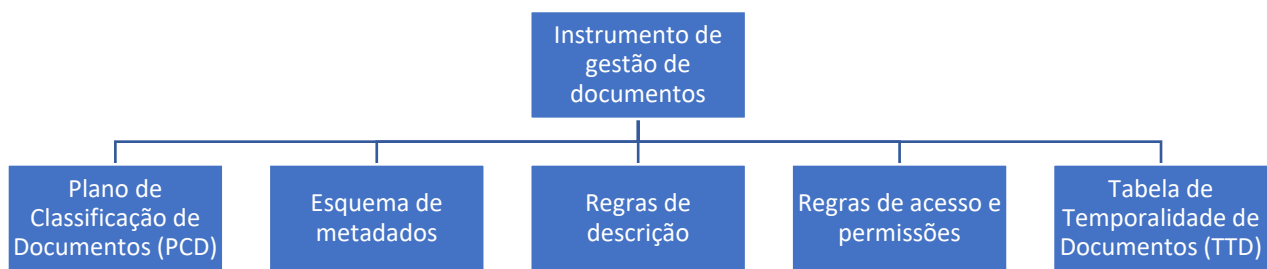
Entre as primeiras conceituações apresentadas está a que define o documento de arquivo como sendo “[...] tanto evidência da atividade de negócios como ativos de informação” (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016, p. v, tradução nossa). Assim, os documentos de arquivo se tornam testemunhos da execução de atividades e processos das instituições, apoiados pelo uso de metadados que detalham o contexto no qual o documento estava inserido no momento da sua criação ou captura e auxiliam a gestão do arquivo em si. O estabelecimento de um conjunto de metadados também corrobora para a presunção de autenticidade dos documentos sob custódia do arquivo, uma vez que provam as suas proveniências.

A criação de uma política de documentos que englobe os documentos de arquivo, compreendendo o seu ciclo de vida, assim como o estabelecimento de responsabilidades sobre os agentes, deve ser a pedra angular que norteará a inclusão do tratamento arquivístico nos processos e atividades das organizações. A elaboração da política deve ser alinhada à missão e aos valores da organização, identificando os processos que se apoiam ou geram documentos de arquivo juntamente com a legislação vigente aplicável ao negócio.

Os conceitos e princípios contidos na norma ISO 15489-1:2016, apresentados na Figura 4, cabe ressaltar, são independentes de modelos de negócios específicos, o que torna a norma passível de aplicação nos mais variados domínios. Do ponto de vista de preservação, a primeira versão da norma, de 2001, foca em documentos de suporte físico tratando dos cuidados necessários

em todos os processos de arquivos para que não ocorram perdas, destruições ou acessos indevidos aos documentos. A preservação de documentos digitais é abordada na norma de uma forma geral, refletindo o contexto de surgimento de sistemas computacionais para documentos digitais o qual estava inserida.

Figura 4 – Conceitos relacionados aos instrumentos de gestão de documentos.



Fonte: (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2021, p. 20, tradução nossa).

A norma prevê a necessidade de as instituições traçarem estratégias de tratamento dos sistemas de armazenamento digital, que contenham rotinas de backup de dados, verificações periódicas de integridade física e lógica dos documentos, conversão e migração de formatos de arquivo e eliminação de documentos. Ao mesmo tempo em que as rotinas de preservação envolvem os documentos, elas devem tratar sobre a preservação dos seus respectivos metadados que registram o contexto em que estão inseridos e vão sendo incrementados à medida em que os procedimentos arquivísticos são executados, documentando incrementalmente a sua proveniência. Os metadados devem abranger o contexto de negócio da instituição, os relacionamentos com outros documentos, o contexto social e legal e o registro dos agentes envolvidos nas operações de criação, gerenciamento e uso dos documentos (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016, p. 5).

A estruturação de um SGDA deve prever mecanismos que tornem os documentos de arquivo confiáveis, implementando métodos que garantam dois aspectos primordiais, que possibilitarão o uso dos documentos como fontes de prova: integridade e autenticidade. A integridade trata mais especificamente dos aspectos físicos dos documentos, garantindo que os documentos e seus componentes não sofreram e não estão vulneráveis a alterações intencionais ou não intencionais. Quando são necessárias ações que modifiquem os documentos, como, por exemplo, a migração de um formato de arquivo para preservação, estas devem ser bem

documentadas e estarem previstas nos procedimentos da organização. A autenticidade, apoiada na integridade, trata sobre a presença de informações e metadados que permitem a verificação do teor do documento, permitindo-se identificar que ele representa o que deveria ser e foi criado pelos agentes autorizados dentro de um processo existente na organização.

Como forma de complementação da norma, a sua segunda parte, ISO 15489-2:2001, consistia em um relatório técnico contendo uma metodologia para a aplicação da primeira parte. Foi apresentado um *workflow* a ser aplicado na análise da organização para a materialização e implementação da sua política de arquivo, iniciando com uma análise do contexto de funcionamento da organização, culminando no projeto de um sistema documental. Após um processo de revisão da norma, em 2016, essa segunda parte foi inativada, não estando mais vigente.

Dentre os principais produtos resultados da aplicação do *workflow* estavam o Plano de Classificação de Documentos (PCD) e a Tabela de Temporalidade de Documentos (TTD). O PCD reflete os principais processos da organização contendo subdivisões que os detalham até o nível de atividades, com os seus respectivos documentos. Conforme o interesse do arquivo, alinhado aos seus objetivos e/ou a legislação vigente, no PCD são configurados os prazos exigidos para cada fase do ciclo de vida dos documentos, culminando com a sua destinação final, que pode ser a preservação permanente ou a eliminação. Ambos instrumentos, em conjunto com as regras de acesso e permissão, configuram a base para a efetiva gestão documental mantendo os documentos organizados, com suas relações orgânicas e com as estratégias de difusão estabelecidas.

A segunda edição da norma (2016) inclui, como não podia ser de outro modo, além do PCD, TTD, regras de acesso e permissão, o esquema de metadados. Configuram-se deste jeito as quatro ferramentas relevantes para uma gestão operacional: 1) organização lógica dos conteúdos auxiliando-se do PCD e atendendo às atividades; 2) delimitação dos prazos das fases dos documentos e definição das suas destinações finais, mediante a TTD; 3) determinação dos documentos de livre acesso e o seu controle mediante permissões, e finalmente; 4) o esquema de metadados, o qual permitirá traçar os perfis mais ótimos a serem aplicados.

Destaca-se que os dois primeiros instrumentos, são obrigatórios no marco legal brasileiro para todo o tipo de gestão dos documentos, enquanto que os demais se mostram relevantes nos entornos eletrônicos.

1.4.3 ISO 30300

Conforme a demanda por Sistemas Gerenciadores de Documentos aumentou com o crescente uso de documentos digitais, foi criada a série de normas ISO 30300, que se estabelece

como uma Norma de Sistemas de Gestão (*Management System Standards* – MSS) compondo as demais séries existentes: a) ISO 9000, que trata de Sistemas de Gestão e Garantia da Qualidade; b) ISO 14000 que trata de Sistemas de Gestão Ambiental, e; c) ISO 27000 que rege os Sistemas de Gestão da Segurança da Informação. Enquanto as MSS possuem seções dedicadas aos documentos de arquivo que apoiam os seus processos ou resultam da execução deles, elas tratam de alguns procedimentos de gestão documental, porém carecem de um olhar mais específico sobre essa questão, que então é aprofundada e regulamentada pela série ISO 30300. Assim, a ISO 30300 se alinha e se integra às demais normas, estabelecendo processos padronizados para a gestão, armazenamento e preservação dos documentos, já que formam evidências do cumprimento dos seus processos (BUSTELO-RUESTA, 2015).

De uma maneira geral, as MSS permitem às organizações a sistematização de seus processos através de instrumentos normatizados que as auxiliam no cumprimento de suas missões. Como a execução das atividades dos processos deve ser materializada em documentos, a utilização de Sistemas Gerenciadores de Documentos em acordo com a norma ISO 30300 permite que os seus documentos se tornem ativos de informação confiáveis pelo tempo que for necessário, além de tornar a organização transparente e auditável.

Estas normas foram editadas no ano de 2011 e construídas como família MSR. Na sua segunda edição, a norma de fundamentos e terminologia constitui-se como norma de escopo terminológico geral para a gestão dos documentos.

Assim, a ISO 30300:2020, contém o vocabulário de termos e definições a respeito do gerenciamento de documentos, tornando-se a referência para as demais normas desse assunto. Os termos apresentados são organizados em seis seções que abordam: a) organização; b) documentos; c) classificação e avaliação; d) gerenciamento de documentos e processos de documentos; e) controle de documentos (instrumentos de auxílio à gestão de documentos), e; f) sistemas e sistemas de documentos. Do ponto de vista de preservação, o termo é definido como “o conjunto de medidas tomadas para manter a **usabilidade, autenticidade, confiabilidade e integridade** dos documentos através do tempo” (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2021, p. 15, tradução nossa), tais medidas incluem princípios, políticas, regras, estratégias, processos e operações. Na Espanha, a segunda versão da norma foi traduzida e resultou na norma UNE-ISO 30300:2021, enquanto o Brasil conta com a tradução da primeira versão, a ABNT NBR ISO 30300:2016.

A segunda norma da série, a ISO 30301:2019, tem como principal objetivo estabelecer os critérios básicos a serem cumpridos no projeto e implementação ou melhoria de um SGD de uma organização para os documentos, verificando a necessidade da análise contextual, a aplicação da sua

política de arquivos, avaliação dos riscos e as necessidades informacionais e documentais, assim como o estabelecimento e distribuição de responsabilidades de gestão e coordenação dos recursos para implementar processos, sistemas e instrumentos de gestão dos documentos. Ademais, a norma ainda fornece uma ferramenta de avaliação de conformidade com a própria norma. Para o planejamento do SGD, um dos primeiros passos a ser realizado é a compreensão do contexto organizacional, identificando fatores que a norteiam para o cumprimento da sua missão e entrega de resultados. Para tal, deve-se analisar desde os aspectos ligados à sua governança até o contexto tecnológico presente na organização projetando-se, assim, os requisitos fundamentais da sua gestão documental. Esse processo acaba se tornando constante com a avaliação, projeto e implementação de mudanças, formando um fluxo de melhorias contínuas de seu SGD.

A norma prevê que a organização disponha de atividades e processos que garantam que os seus documentos permaneçam preservados e acessíveis no seu SGD. Esses requisitos devem considerar a presença de mecanismos que identificam adulterações, intencionais ou não, e a constante avaliação da necessidade de migração de formatos dos arquivos, por exemplo. Adicionalmente, o SGD deve englobar processos de eliminação adequada dos documentos sob a sua guarda, cumprindo a política institucional ou atendendo a diretrizes legais.

O Anexo A da norma ISO 30301:2019 apresenta, em forma de tabela, os controles essenciais quanto ao cumprimento dos requisitos de gestão documental no que se refere a: a) processos de gestão de documentos; b) instrumentos de gestão de documentos, e; c) aplicações da gestão documental. Um extrato dos requisitos relativos à preservação de documentos presentes nas categorias de processos e instrumentos de gestão de documentos é apresentado no Quadro 2.

Os requisitos A.1.1 e A.1.2 se referem ao processo de criação de documentos no SGD. O requisito A.1.1 especifica que a produção dos documentos deve ocorrer o mais próximo, cronologicamente, das ações às quais eles se referem, visando capturar-se ao máximo as informações que vão compor os contextos de criação. O requisito A.1.2 se refere a mapear os tipos documentais necessários a cada processo de negócio das organizações estabelecendo-se as suas formas e regras de produção.

Os requisitos A.1.3 a A.1.6 tratam da captura de documentos no SGD. O requisito A.1.3 trata de se identificar quais processos de negócio requerem a captura de documentos dentro do SGD. O requisito A.1.4 especifica que o SGD deve reunir todas as informações possíveis de contexto, seja automaticamente ou manualmente, no momento da captura dos documentos. A.1.5 trata da unicidade dos documentos dentro da organização, atribuindo a eles um identificador único e inequívoco. O requisito A.1.6 trata sobre o registro de informações que permitam se conhecer os motivos e contexto envolvidos na ocasião de uma substituição de um documento por outro.

Quadro 2 – Requisitos do Anexo A da norma ISO30301:2019 referentes aos processos de gestão.

Identificador	Categoria	Requisito operacional
A.1.1	Produção de documentos	Deve-se criar os documentos no momento (ou logo em seguida) da atividade ou ação qual se referem, pelas pessoas que têm conhecimentos direto dos feitos ou pelos instrumentos normalmente utilizados pela organização para a realização da atividade.
A.1.2		Deve-se identificar e documentar a forma e estrutura de informação necessária como documento para cada processo de trabalho.
A.1.3	Captura de documentos	Deve-se decidir e documentar os métodos de integração da captura de documentos nos processos de trabalho.
A.1.4		No momento da captura, deve-se agregar as informações de contexto dos documentos.
A.1.5		Para os processos de negócio que requeiram evidências de sua captura, deve-se implementar um identificador único no momento da captura.
A.1.6		Quando um documento substitui um existente (atualização), como acontece com parte da informação documentada, a nova versão deve indicar a versão obsoleta, bem como as alterações efetuadas.
A.1.7	Classificação e indexação de documentos	Deve-se agrupar (classificar) os documentos de acordo com os processos de trabalhos com os quais estejam relacionados.
A.1.8	Armazenamento de documentos	Os meios de preservação/armazenamento dos documentos devem cumprir com as normas pertinentes para o suporte e tecnologias utilizado, com o objetivo de garantir que os documentos sejam acessíveis utilizáveis durante o tempo necessário.
A.1.9	Uso e reutilização	Os documentos digitais devem permanecer acessíveis e utilizáveis durante o tempo necessário.
A.1.10		Devem-se definir e implementar as ações sobre os documentos que devem ser registrados nos metadados.
A.1.11	Migração e conversão de documentos	Devem-se autorizar e documentar as migrações controladas de documentos para outra organização ou sistema.
A.1.12		Devem-se autorizar e documentar as conversões regulares dos documentos dos documentos, incluindo a conversão de formato analógico para digital (digitalização).
A.1.13		Durante os processos de migração e conversão, o aplicativo ou formato produzido deve ser mantido até que o processo seja concluído e a integridade e confiabilidade do formato ou aplicativo de destino seja confirmada.
A.1.14	Eliminação	Deve-se estabelecer e documentar os critérios para determinar os períodos de guarda dos documentos de acordo com os requisitos de cada processo de trabalho.
A.1.15		As decisões sobre recolhimento, transferência ou destruição dos documentos devem ser autorizadas e documentadas.
A.1.16		Os documentos devem ser destruídos mediante autorização e supervisão apropriada. A eliminação deve ser documentada.
A.1.17		Quando a natureza e complexidade do negócio e as necessidades formais exigirem, deve-se conservar informações de controle (metadados de registro, identificação e históricos) dos documentos que tenham sido eliminados.

Fonte: (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2019, p. 25, tradução nossa)

O requisito A.1.7 se refere à classificação e indexação dos documentos, especificando que o SGD deve possuir mecanismos de classificação de documentos que os agrupe de acordo com os processos de negócios. Trata-se de incorporar ao SGD um Código de Classificação de Documentos (CCD) que agrupe os documentos e mantenha as suas relações orgânicas.

O requisito A.1.8 trata do armazenamento dos documentos, focando na tecnologia adotada e ressaltando a necessidade de observar as normas pertinentes para esse fim. O atendimento desse requisito pelo SGD vai dar garantias à organização de que os métodos adotados para o armazenamento e acesso aos dados dos documentos estão em acordo.

Os requisitos A.1.9 e A.1.10 tratam do reuso e acesso aos documentos, relacionando-se fortemente aos princípios FAIR. O requisito A.1.9 trata de o SGD promover o acesso aos documentos, seja para a comprovação de fatos, prestações de contas ou para a extrações de informações úteis para a realização dos processos de negócio. O requisito A.1.10 trata de se definir quais os processos e ações sobre os documentos o SGD deve documentar em metadados, assim reunindo-se informações que podem ajudar a corroborar a autenticidade dos documentos.

Os requisitos A.1.11 a A.1.13 tratam da migração e conversão dos documentos através do SGD. O requisito A.1.11 trata de o SGD possuir procedimentos que autorizem e controlem a migração dos seus documentos para outras instituições ou sistemas. O requisito A.1.12 se refere ao SGD conter funcionalidades de conversões de formatos dos documentos incluindo as digitalizações. O SGD pode realizar conversões que visem a normalização de formatos para a utilização de formatos sustentáveis ou ainda para a disponibilização de formatos mais adequados ao acesso dos usuários. O requisito A.1.13 visa garantir que os documentos nos formatos originais sejam mantidos até que seja possível afirmar que a migração ocorreu com sucesso e o resultado seja autêntico.

Os requisitos A.1.14 a A.1.17 se referem aos processos que o SGD utiliza para a eliminação dos documentos e possuem relação direta com a TTD adotada pela organização. O requisito A.1.14 estabelece que o SGD deve possuir critérios para a eliminação dos documentos de acordo com os processos de negócio da organização. Nesse caso, o SGD deve respeitar as temporalidades definidas na TTD, disponibilizando funcionalidades que permitam à gestão o conhecimento de quais conjuntos de documentos estão em prazo para a realização das suas respectivas destinações finais. O requisito A.1.15 define que todas as decisões relacionadas às destinações finais dos documentos devem ser documentadas e autorizadas previamente, independente de se realizar o recolhimento, transferência a outro arquivo ou a própria eliminação. O requisito A.1.16 define que a eliminação dos documentos deve ocorrer sob supervisão adequada e sempre registrando-se os procedimentos realizados. O requisito A.1.17 trata sobre a possibilidade de o SGD manter metadados e demais informações necessárias dos documentos eliminados, de acordo com as necessidades do negócio ou visando o cumprimento de obrigações legais.

No Brasil, os órgãos e entidades pertencentes ao poder executivo federal seguem o fluxo de elaboração da Listagem de Eliminação de Documentos (LED) aprovada pela Comissão de Permanente de Avaliação de Documentos (CPAD) do órgão e depois autorizada pelo dirigente da instituição. A eliminação somente pode ser efetivada após esses trâmites, utilizando-se os procedimentos próprios para documentos digitais ou físicos (ARQUIVO NACIONAL, 2020).

Ainda no Anexo A da norma ISO 30301:2019, são detalhados 9 requisitos da seção A.2 que tratam dos instrumentos utilizados para a gestão documental, exibidos no Quadro 3. Os

requisitos A.2.1 a A.2.3 tratam sobre os esquemas de metadados dos documentos. O requisito A.2.1 trata de se projetar as informações que serão capturadas em metadados para a identificação dos documentos no SGD. O requisito A.2.2 se refere à identificação e registro das informações de descrição e controle dos documentos desde as suas criações até as demais operações realizadas, registrando-as em metadados. O requisito A.2.3 especifica explicitamente a necessidade de se implementar o registro de metadados para a identificação, gestão e controle dos documentos no SGD. O requisito A.2.4 trata sobre a classificação dos documentos, exigindo que o SGD permita a ligação entre o documento e as atividades às quais se referem. Na prática, significa incorporar um Código de Classificação de Documentos. Que reflete os processos e atividades de negócio, no SGD e disponibilizar funcionalidades para que os documentos possam ser atribuídos às devidas classes.

Quadro 3 – Requisitos do Anexo A da norma ISO30301:2019 referentes aos instrumentos utilizados na gestão documental.

Identificador	Categoria	Requisito operacional
A.2.1	Esquemas de metadados para a gestão de documentos	Deve-se definir e documentar a informação necessária para identificar os documentos de cada processo de trabalho, incluída a área da organização responsável por esses documentos e pelo processo de trabalho.
A.2.2		Deve-se identificar e documentar a informação de descrição e de controle (elementos de metadados) requerida na criação e controle dos documentos para cada processo de negócio.
A.2.3		Deve-se documentar e implementar as decisões sobre os metadados requeridos para identificar, gerenciar e controlar os documentos.
A.2.4	Quadros de classificação	Deve-se estabelecer e documentar um esquema de classificação para conectar as atividades de documentos
A.2.5	Regras e permissões de acesso	Deve-se estabelecer as regras de acesso aos documentos, documentá-las e mantê-las durante todo o tempo que os documentos necessitem.
A.2.6		As regras de acesso nas aplicações de gestão documental devem ser implementadas para se atribuir um status de acesso aos documentos e aos indivíduos.
A.2.7		As restrições de acesso, incluindo o uso de encriptação ou criptografia, devem ser eliminados depois de um período estabelecido.
A.2.8	Tabelas de temporalidade	As decisões sobre preservação ou eliminação dos documentos fundamentadas nos requisitos de negócio, legais ou de outra índole devem estar documentadas nas tabelas de temporalidade.
A.2.9		Deve-se autorizar e documentar as ações de preservação e eliminação das tabelas de temporalidade.

Fonte: (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2019, p. 26, tradução nossa)

Os requisitos A.2.5 a A.2.7 tratam das regras e permissões de acesso aos documentos, relacionando-se à política de arquivos da organização. O requisito A.2.5 especifica que as regras e permissões de acesso devem ser mantidas no SGD pelo tempo em que os seus documentos necessitarem. A ausência desse requisito pode fragilizar os acessos e usos dos documentos do SGD, possibilitando que documentos com teor restrito sejam disponibilizados indiscriminadamente, assim como documentos públicos e com informações pertinentes à sociedade não possam ser acessados. O requisito A.2.6 visa o tratamento de acesso e uso específico para determinados documentos e usuários, tratando possíveis exceções às regras previstas na política de arquivos da organização. O

requisito A.2.7 trata de o SGD possuir funcionalidades que removam as restrições de acesso dos documentos, inclusive as derivadas de encriptação, depois de decorrido um determinado prazo estabelecido.

Os requisitos A.2.8 e A.2.9 tratam das tabelas de temporalidade no SGD. O requisito A.2.8 trata explicitamente sobre o SGD possuir a tabela de temporalidade dos documentos documentada com os prazos definidos conforme as regras de negócio da organização e de acordo com a legislação vigente aplicável, para que a preservação ou eliminação dos documentos selecionados seja realizada. O requisito A.2.9 trata de se registrar as ações de preservação e eliminação e exigir que essas ações somente sejam realizadas mediante autorizações competentes, trazendo segurança à gestão documental.

A terceira parte do Anexo A trata dos requisitos das aplicações em si, da gestão documental e são apresentadas no Quadro 4.

Quadro 4 – Requisitos do Anexo A da norma ISO30301:2019 referentes às aplicações de gestão documental.

Identificador	Categoria	Requisito operacional
A.3.1	Integridade e Segurança	As aplicações de gestão documental devem garantir a integridade e segurança dos documentos para prevenir o uso, modificação, exclusão, distribuição, ocultação e/ou destruição não autorizadas dos mesmos.
A.3.2	Tecnologias	Deve-se selecionar as tecnologias para a criação e captura de documentos nos processos de trabalho (tanto automatizados como manuais). A escolha das tecnologias e quaisquer mudanças nas mesmas devem ser documentadas.
A.3.3	Inventário	As aplicações de gestão de documentos devem estar claramente identificadas e serem designadas a um proprietário responsável e documentadas em um inventário atualizado regularmente.
A.3.4	Documentação	As decisões de implementação nas aplicações de gestão documental devem ser documentadas, mantidas e postas à disposição de todos os usuários que necessitarem.
A.3.5	Disponibilidade	A disponibilidade das aplicações de gestão documental deve ser garantida e documentada.
A.3.6	Integridade	A integridade dos documentos não deve ser afetada por falhas atualizações ou manutenções das aplicações.
A.3.7	Monitoramento	Deve-se implementar e documentar o monitoramento regular do desempenho das aplicações de gestão documental em relação aos requisitos de negócio e objetivos da gestão de documentos.
A.3.8	Acesso	Deve-se estabelecer, documentar e manter as regras para o acesso às aplicações de gestão documental usadas no cumprimento das tarefas de administração do sistema.

Fonte: (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2019, p. 26, tradução nossa)

O requisito A.3.1 estabelece que as aplicações do SGD devem garantir a integridade e segurança dos documentos que gerencia, visando evitar acessos, adulterações ou destruições indevidas e desautorizadas. O requisito A.3.2 trata de se analisar as tecnologias mais adequadas para uso no SGD, exigindo que quaisquer mudanças sejam planejadas e documentadas apropriadamente. A escolha de quais tecnologias utilizar impacta diretamente em como os documentos são armazenados e acessados, por exemplo, adicionando melhorias ao SGD ou impactando

negativamente no uso e gerenciamento dos documentos, daí a importância de eventuais mudanças serem planejadas e documentadas.

O requisito A.3.3 estabelece que deve haver uma espécie de inventário, no qual conste quais aplicações são autorizadas aos usuários do SGD, trazendo como benefício, por exemplo, a possibilidade de se realizar auditorias que identifiquem o uso, ou o mal uso, do SGD. O requisito A.3.4 estabelece que devem existir documentações que detalhem a implementação do SGD, e estas devem ser mantidas e colocadas à disposição aos usuários que desejarem conhecê-las.

O requisito A.3.5 estabelece que o SGD deve se comprometer com a disponibilidade das suas aplicações, dessa forma, não impedindo que as ações necessárias à gestão documental sejam executadas. O requisito A.3.6 diz respeito à manutenção da integridade dos documentos gerenciados pelo SGD, ressaltando que ela não deve ser afetada pelas manutenções, preventivas ou reativas, realizadas no sistema.

O requisito A.3.7 estabelece que deve haver um monitoramento contínuo e documentado do desempenho do SGD a fim de que os requisitos de negócio sejam cumpridos, servindo como apoio, inclusive, para as decisões a serem tomadas quando da detecção do não cumprimento. O requisito A.3.8 trata de haver documentações que embasem os acessos dos usuários às aplicações do SGD para a execução dos seus processos de trabalho e administração do próprio sistema.

Após o estabelecimento de requisitos necessários para um SGD, a ISO 30302 oferece um guia prático de implantação de um SGD seguindo o disposto na norma ISO 30301. Mesmo que a instituição já possua algum sistema para a gestão de documentos, a norma pode ser utilizada para adequação dos processos de arquivo, favorecendo-se de uma normalização e alinhamento com as demais normas do assunto. As instituições que já possuam os seus processos organizados e desenhados de acordo com a norma ISO 15489-1 podem utilizar a norma ISO 30301 para aplicar os seus princípios e complementar os critérios e, com isso, desfrutar de um SGD padronizado e promover o seu cumprimento para se tornar auditável (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2015, p. 7). Assim, a norma ISO 30302:2022 mostra um excelente guia, incorporando no seu Anexo A exemplos de fontes de onde se verificar o controle dos processos, instrumentos e aplicações documentais.

Durante o processo de planejamento do SGD, deve-se estabelecer o alcance e escopo do sistema, definindo-se quais processos e órgãos da instituição ele abrangerá e a sua integração ou meio para se alcançar os objetivos de outras normas implantadas como as demais MSS (ISO 9000, ISO 14000 e ISO 27000, entre outros modelos de gestão padronizados). Dessa forma, os órgãos

diretores da instituição estarão cientes sobre o que esperar do SGD e conhecerão os fatores e impactos envolvidos para o alcance dos objetivos institucionais.

O planejamento da preservação de documentos deve levar em consideração o contexto da organização, identificando requisitos de negócio e legais que especificam, entre outros, os prazos em que os documentos devem ser conservados e as suas destinações finais. O instrumento que rege esses requisitos é denominado Tabela de Temporalidade de Documentos (TTD), elaborado através da análise das documentações geradas ou utilizadas nos processos da organização e as atividades que os compõem. O mesmo tratamento utilizado na preservação dos documentos deve ser aplicado aos seus metadados inseridos no SGD, fazendo com que ambos sempre existam em sintonia, compondo a proveniência dos documentos, garantindo as suas autenticidades, além de apoiar os processos de gestão documental. A gestão documental também deve estar prevista nos planos de continuidade da instituição, prevendo o tratamento adequado, possíveis transferências de custódia e recuperação de dados nos casos em que houver interrupção das atividades ou a ocorrência de catástrofes, por exemplo. No caso de arquivos governamentais, o plano de continuidade deve ser elaborado de maneira independente dos interesses dos diretores em exercício, garantindo que as mudanças de cargos não afetem o funcionamento do arquivo e a preservação dos documentos sob a sua custódia, conservando a memória das instituições públicas.

A implantação da série de normas ISO 30300 nas organizações traz consigo um conjunto de vantagens, tais como a compatibilidade e alinhamento às demais MSS, o estabelecimento formal de uma política de gestão documental, auxílio nas prestações de contas contábeis e demais obrigações legais vigentes e a gestão segura de suas informações. Como consequência, a administração torna-se mais transparente e auditável, melhorando a sua confiança junto aos seus colaboradores, fornecedores e demais órgãos externos.

1.4.4 e-Arq Brasil

Com o intuito de elaborar um conjunto de requisitos técnicos a serem aplicados nos sistemas de gestão de documentos, a Câmara Técnica de Documentos Eletrônicos, órgão vinculado ao Conselho Nacional de Arquivos, disponibilizou em 2006 a primeira versão do e-ARQ Brasil e em 2009 complementou a versão com o esquema de metadados a ser adotado (CONSELHO NACIONAL DE ARQUIVOS, 2011). Entre os anos 2017 e 2020 ocorreu o processo de elaboração de uma nova versão do e-ARQ Brasil para incorporar adaptações identificadas em estudos de caso e atualizações legais e de normas internacionais mais recentes, contando com diversas reuniões e uma

consulta à comunidade. Em 2022, a versão 2 do e-ARQ Brasil foi disponibilizada, autodeclarando-se como sendo (CONSELHO NACIONAL DE ARQUIVOS, 2022, p. 10):

Uma especificação de requisitos a serem cumpridos pela organização produtora/recebedora de documentos, pelo sistema de gestão arquivística e pelos próprios documentos, a fim de garantir sua confiabilidade e autenticidade, assim como seu acesso, pelo tempo que for necessário.

O conjunto de requisitos funcionais e não funcionais permite a adaptação ou a criação de sistemas que manipulam documentos, incorporando um tratamento arquivístico adequado à realidade brasileira. Esse conjunto de requisitos foi elaborado para ser independente de tecnologia e pode ser utilizado na gestão de documentos digitais e não digitais, formando um Sistema Informatizado de Gestão Arquivística de Documentos (SIGAD) (CONSELHO NACIONAL DE ARQUIVOS, 2022, p. 10). No caso do documento nato-digital, o SIGAD realiza toda a sua gestão e captura de metadados, já no caso do documento físico, o SIGAD somente faz o registro dos metadados apropriados para cada fase do seu ciclo de vida. Em complementação aos requisitos, o e-ARQ Brasil contém ainda um conjunto de metadados a ser adotado pelo SIGAD que registra o contexto dos documentos e ações executadas sobre eles e que permitem a verificação do cumprimento de determinados requisitos.

Um SIGAD se caracteriza por adicionar requisitos arquivísticos à gestão de documentos em um sistema informatizado, possuindo funcionalidades para a captura, armazenamento, indexação e recuperação de todos os documentos e seus componentes digitais (no caso dos documentos nato digitais) da instituição (CONSELHO NACIONAL DE ARQUIVOS, 2022, p. 21). Inseparáveis dos documentos, os seus metadados também devem ser capturados, mantidos e gerenciados pelo SIGAD, permitindo-se conhecer os contextos envolvidos em cada ação, desempenhando um papel central na presunção de autenticidade dos documentos da instituição

A garantia da autenticidade dos documentos, característica essencial para a preservação, é um dos requisitos de um SIGAD, fazendo-se uso de todos os mecanismos passíveis de serem aplicados de forma a não haver falhas de segurança e corrompimento de dados. A criação de diversas camadas de segurança no SIGAD se inicia com as instalações físicas e de *hardware* adotados, passa pela adoção de práticas seguras no desenvolvimento de *softwares*, e culmina na conscientização dos usuários no uso correto dos sistemas, que deve estar embasada na política de gestão de documentos da instituição. O conjunto dessas medidas traz completeza aos documentos do SIGAD, no sentido de ser possível examinar os seus elementos intrínsecos e extrínsecos, esses últimos registrados em metadados, que proporcionam graus de confiabilidade ao SIGAD, ao que o Conselho Nacional de Arquivos (2022, p. 29) resume:

Difícilmente, pode-se assegurar a veracidade do conteúdo de um documento; ela é inferida da completeza e dos procedimentos de produção. A confiabilidade é uma questão de grau, ou seja, um documento pode ser mais ou menos confiável.

No SIGAD, ainda, os documentos devem ser organizados de acordo com o Plano de Classificação de Documentos (PCD), estabelecendo-se assim uma relação orgânica entre eles, além de ser possível a aplicação da Tabela de Temporalidade de Documentos (TTD) em vigência permitindo que a destinação final seja executada nos documentos, eliminando-os ou preservando-os conforme cada caso.

Ao todo, o e-ARQ Brasil apresenta 391 requisitos que cobrem aspectos desde a criação dos documentos, tramitações, conformidade com a legislação, até a destinação final dos documentos, possuindo as seguintes obrigatoriedades (CONSELHO NACIONAL DE ARQUIVOS, 2022, p. 15):

- a) Obrigatório: quando o SIGAD deve obrigatoriamente cumprir com o requisito;
- b) Altamente desejável: quando o requisito não é obrigatório, mas o seu cumprimento melhora a gestão arquivística do sistema;
- c) Facultativo: quando cabe à instituição considerar a aplicação do requisito, conforme a sua realidade. Como resultado das análises dos gestores, determinados requisitos podem ser considerados obrigatórios para algumas instituições e para outras não.

O esquema de metadados do e-ARQ Brasil possui um total de 96 metadados divididos nas seguintes entidades (CONSELHO NACIONAL DE ARQUIVOS, 2022, p. 113):

- a) Documento: refere-se ao documento produzido ou capturado pelo SIGAD, podendo representar um processo ou dossiê. São definidos 37 metadados para a entidade;
- b) Evento de gestão do ciclo de vida: são as ações executadas durante o ciclo de vida dos documentos dentro do SIGAD, tais como a criação, captura, eliminação e transferência, por exemplo. São definidos 8 metadados para a entidade;
- c) Evento de gestão dos processos/dossiês: são as ações de protocolo que os processos e dossiês sofrem. São definidos 7 metadados para a entidade;
- d) Classe: se refere aos níveis do Plano de Classificação de Documentos (PCD) adotado pela instituição. São definidos 13 metadados para a entidade;
- e) Evento de gerenciamento de classe: são as ações que envolvem a manutenção do PCD, quando há a mudança de nomenclatura, nível, temporalidade ou destinação final, por exemplo. São definidos 6 metadados para a entidade;
- f) Componente digital: refere-se aos objetos digitais, arquivos de computador, que integram o documento, sendo possível que um determinado documento possua um ou mais componentes digitais. São definidos 15 metadados para a entidade;

- g) Evento de preservação: são as ações executadas sobre os documentos destinadas a preservá-los, tais como a migração, atualização ou conversão, por exemplo. São definidos 7 metadados para a entidade;
- h) Agente: entidade destinada aos atores que interagem com os documentos dentro do SIGAD. Todos os eventos executados sobre os documentos são executados por algum agente. São definidos 3 metadados para a entidade.

Da mesma forma que os requisitos, os metadados possuem obrigatoriedades diferentes entre eles (CONSELHO NACIONAL DE ARQUIVOS, 2022, p. 15):

- a) Obrigatório: quando o SIGAD deve registrar a informação referente ao metadado em questão;
- b) Obrigatório, se aplicável: deve-se avaliar a aplicabilidade do metadado. Caso seja aplicável, ele se tornar obrigatório;
- c) Facultativo: cabe à instituição definir se realizará o registro das informações referentes ao metadado.

Além de servir como base para a verificação do cumprimento e grau de adoção do e-ARQ Brasil, as obrigatoriedades dos requisitos e metadados permite o favorecimento de uma interoperabilidade entre diferentes sistemas. Durante a comunicação entre SIGADs, por exemplo, assume-se que um determinado metadado obrigatório estará presente em ambas as partes, assim como não é possível assumir que determinado metadado facultativo adotado por um SIGAD esteja presente em outro SIGAD.

Os papéis de diferentes profissionais são citados no e-ARQ Brasil, evidenciando a interdisciplinaridade necessária para a identificação, adoção, desenvolvimento e verificação de conformidade dos seus requisitos (CONSELHO NACIONAL DE ARQUIVOS, 2022, p. 26). Os profissionais da gestão de arquivos, como os arquivistas, auxiliam a incorporar a abordagem arquivística necessária aos SIGADs. Os profissionais de TI se comprometem com a orientação e desenvolvimento dos SIGADs no atendimento dos requisitos. Os usuários finais do SIGAD também devem ser envolvidos no processo a fim de se perceber as suas necessidades e demanda que, posteriormente, vão impactar diretamente na preservação dos documentos criados, capturados e tramitados no SIGAD. Essa interdisciplinaridade facilita o alinhamento de conceitos entre as áreas e proporciona uma visão holística do e-ARQ Brasil, ao mesmo tempo em que soma forças para a tomada de decisões e elaboração de políticas institucionais.

O e-ARQ Brasil ainda pode ser utilizado para se avaliar o grau de adesão de um sistema existente aos requisitos arquivísticos elaborados e, ainda que não possua aspecto de certificação, pode guiar as instituições nos processos de desenvolvimento ou aquisição de sistemas que façam a

gestão de documentos. Uma metodologia de planejamento e implantação do programa de gestão arquivística de documentos na instituição é sugerida pelo e-ARQ Brasil, através de um conjunto de passos que contém, entre eles, o monitoramento do SIGAD a ser executado periodicamente.

O último passo sugerido na implantação de um SIGAD é o monitoramento e a aplicação de ajustes necessários para a adequação à gestão arquivística de documentos da instituição. O monitoramento, nesse caso, se refere a analisar a conformidade da execução dos requisitos funcionais e não funcionais do SIGAD com o planejamento realizado, sendo recomendado que esse monitoramento seja executado periodicamente, no intuito de se identificar o mais cedo possível quaisquer inconformidades. O monitoramento visa assegurar o cumprimento dos requisitos do e-ARQ Brasil, ao que o Conselho Nacional de Arquivos (2022, p. 33) resume:

O monitoramento garantirá o retorno contínuo dos investimentos no programa de gestão arquivística de documentos, além de fornecer informação objetiva sobre a capacidade do órgão ou entidade em produzir e gerenciar documentos arquivísticos apropriados, garantindo o seu armazenamento de maneira segura.

Ainda que o e-ARQ Brasil não se caracterize como uma certificação, torna-se evidente a sua vocação de poder ser utilizado em auditorias que aferem os graus de conformidades de sistemas que se candidatem a se tornar SIGADs. Ao se cumprir com todos os requisitos obrigatórios e aplicáveis para os casos da instituição, poderia se afirmar que o sistema utilizado para a gestão de documentos possui os aspectos arquivísticos necessários para se intitular SIGAD. Complementarmente, o e-ARQ Brasil pode ser utilizado como ponto de partida para a adição de requisitos ao(s) sistema(s) informático(s) da instituição, englobando funcionalidades que vão cumprindo os requisitos de acordo com as demandas, prioridades e disponibilidade de recursos de cada caso.

1.4.5 ISO 23081

A norma ISO 23081 trata sobre o estabelecimento de um modelo para a gestão e utilização de metadados junto aos documentos de arquivo sem, no entanto, tratar um conjunto específico de metadados e as suas obrigatoriedades, uma vez que essas condições podem variar de acordo com o negócio da instituição e a legislação que a rege.

A norma é dividida em três partes: a) ISO 23081-1:2017 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2017): trata da definição teórica de metadados, seu projeto e incorporação nos arquivos. A sua primeira versão foi publicada em 2006; b) ISO 23081-2:2021 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2021): fornece meios para a aplicação prática dos conceitos abordados na primeira parte. A sua primeira versão foi

publicada em 2009 e passou por um processo de revisão em 2021, e; c) ISO 23081-3:2011 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2011): estabelece um método para a autoavaliação das instituições sobre a correta utilização dos metadados e seu gerenciamento, permitindo a análise da situação atual e o planejamento de melhorias. No Brasil, as três partes da norma foram traduzidas e originaram as normas ABNT NBR-ISO 23081-1:2019²², ABNT NBR-ISO 23081-2:2020²³ e ABNT NBR-ISO/TR 23081-3:2022²⁴.

Os metadados de documentos de arquivo são definidos como “informações estruturadas ou semiestruturadas que permitem a produção, gestão e uso de documentos de arquivo ao longo do tempo e dentro e entre domínios” (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2019, p. 2). Assim sendo, são porções de dados que dizem respeito sobre a estrutura do documento, seu teor, agentes envolvidos, destinatários, interessados e demais informações que detalham a política de arquivo da instituição e auxiliam a gestão documental. São os metadados associados aos documentos que auxiliam a presunção de autenticidade e por consequência, a sua confiabilidade, uma vez que detalham o contexto de sua gênese e as demais interações ao qual estiveram envolvidos no decorrer do seu ciclo de vida (TENNIS e ROGERS, 2012).

Outro aspecto fundamental no projeto adequado de utilização de metadados é permitir a usabilidade dos documentos, uma vez que eles descrevem as suas condições de acesso, direitos autorais envolvidos e presença de dados sensíveis, por exemplo. Em contraste com o suporte papel no qual muitos metadados não precisam ser declarados pois são implícitos, no suporte digital, o máximo de informações disponíveis sobre a estrutura dos documentos deve ser armazenada para permitir a preservação adequada bem como a sua utilização futura (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2017, p. 3). A especificação de formatos dos arquivos, estrutura lógica dos componentes, bem como os requisitos de mínimos de *software* e *hardware* para a sua utilização tornam-se essenciais para que a usabilidade dos documentos não seja comprometida ao longo dos anos, além de auxiliar os procedimentos de preservação adotados pelo arquivo.

A terceira parte da norma, ISO 23081-3:2011, contém uma ferramenta que auxilia o arquivo a realizar a sua autoavaliação para detectar o seu estado atual de utilização e armazenamento de metadados, além de apoiar o estabelecimento de metas de melhorias nos seus processos de captação de metadados. Um conjunto de listas de controle é apresentado na forma de checklists a serem utilizados por arquivistas, profissionais de TI e gestores da instituição, verificando o cumprimento ou a presença de cada requisito no sistema de arquivo. É essencial que a

²² (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2019)

²³ (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2020)

²⁴ (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2022)

instituição já tenha estabelecido e implantado a sua política de preservação, uma vez que o seu projeto de metadados do sistema de arquivo deve estar alinhado à essa política, fazendo com que a autoavaliação retrate o seu estado atual. De certa forma, a norma ISO 23081 se torna a materialização, através dos metadados no sistema de arquivo, da norma ISO 15489, uma vez que aquela se apoia nos objetivos e planos traçados por essa última.

Conforme o arquivo institucionaliza a descrição de seus documentos com o uso dos padrões de descrição arquivística do ICA (ISAD(G), ISAAR (CPF), ISDF e ISDIAH), nos seus diferentes níveis e especificidades, é possível afirmar o cumprimento da norma ISO 23081. A implantação das normas do ICA faz com que o contexto dos documentos em sua gênese já possua um padrão bem definido, tornando possível a interpretação sobre as circunstâncias envolvidas na criação e captura dos documentos da instituição. O uso de plataformas de difusão e descrição que suportam os padrões ICA como, por exemplo, o AtoM²⁵, proporciona uma grande riqueza de metadados no acervo institucional, podendo ser, inclusive, exportados para formatos de intercâmbio de informações entre sistemas, como EAD (*Encoded Archival Description*).

1.4.6 PREMIS

Em 2003, OCLC e RLG colaboraram para a formação do grupo de trabalho PREMIS (*Preservation Metadata: Implementation Strategies*) contendo profissionais e pesquisadores de diversos países, especialistas na área de metadados para a preservação digital. Os perfis desses membros incluíam profissionais de arquivo, de museus, bibliotecas, órgãos de governo e da iniciativa privada, buscando incluir participantes que tivessem trabalhado com preservação digital, uma vez que o foco do grupo era o trabalho prático. O propósito do grupo de trabalho foi tornar viável um protótipo realizado por outro grupo de trabalho apoiado pela OCLC e RLG, nomeado *The Framework*²⁶, que possuía um conjunto de metadados conceituais voltados à preservação digital de objetos. A realização desse trabalho resultou no Dicionário de Dados PREMIS, que é um conjunto de metadados, baseado no modelo OAIS, “[...] para objetos digitais arquivados, além de fornecer orientação e sugerir melhores práticas para criar, gerenciar e usar os metadados em sistemas de preservação” (PREMIS EDITORIAL COMMITTEE, 2015, p. 1, tradução nossa). Em 2005 foi disponibilizada a versão 1.0 do Dicionário de Dados PREMIS que recebeu atualizações e as versões 2.0, 2.1 e 2.2 foram publicadas em 2008, 2011 e 2012 respectivamente. A última versão publicada e atualmente em uso é a 3, de 2015.

²⁵ (ARTEFACTUAL SYSTEMS INC, 2022)

²⁶ (ONLINE COMPUTER LIBRARY CENTER, 2002)

Baseado na definição de que metadados de preservação são “informações que um repositório usa para apoiar o processo de preservação digital” (PREMIS EDITORIAL COMMITTEE, 2015, p. 2, tradução nossa), o grupo de trabalho elaborou metadados que permitissem a autenticidade e interpretabilidade dos objetos preservados ao longo do tempo, abrangendo as categorias normalmente usadas para a preservação digital: administrativa, técnica e estrutural. A categoria administrativa possui metadados voltados a registrar a proveniência dos objetos, incluindo informações relacionadas à propriedade intelectual e restrições de acesso. A categoria técnica possui detalhes sobre como as máquinas (computadores) devem utilizar os objetos e a categoria estrutural estabelece os relacionamentos entre os componentes do objeto digital preservado.

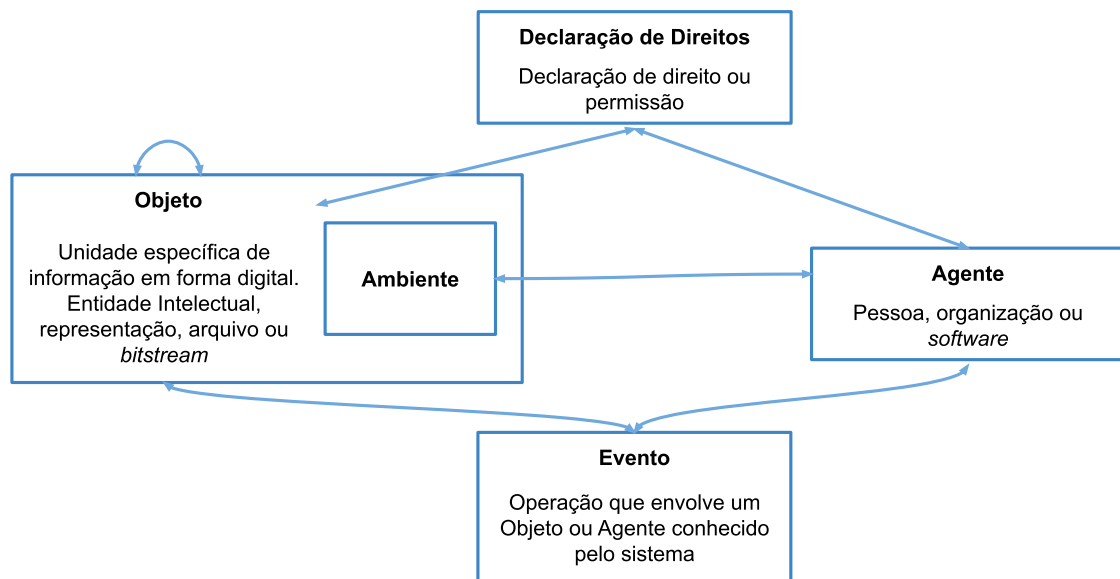
A descrição arquivística dos objetos preservados, muitas vezes realizada no ambiente de gestão, passa a ser representada através do Dicionário de Dados PREMIS no repositório, quando é complementada pelas informações dos processos de preservação executados. Daí a relevância do seu uso pelas plataformas de repositórios.

Dentre os princípios norteadores do grupo de trabalho estava a intenção de que os metadados fossem independentes de implementação, não especificando como ou quais tecnologias seriam utilizadas para armazená-los (se fossem armazenados), característica denominada neutralidade técnica. O Dicionário de Dados trata de Unidades Semânticas (*Semantic Units*), uma vez que esse termo se refere mais ao que deve ser registrado do que como deve ser registrado, nesse sentido, mais ligado ao termo elemento de metadado (*metadata element*) (PREMIS EDITORIAL COMMITTEE, 2015, p. 3). Dappert, Peyrard e Guenther fazem uma analogia do Dicionário de Dados PREMIS com a definição de um idioma, explicitando que os repositórios devem projetar como e quais informações registrar, de modo a melhor apoiar a preservação de seus objetos digitais, fazendo-se uso quando necessário de combinações de tecnologias e padrões (2016, p. 3, tradução nossa):

Compreender o dicionário sozinho não ensina como criar seus metadados de preservação. É qcomo uma definição de linguagem. É fornecida uma gramática e vocabulário básicos. Mas não lhe dá as frases que contam a sua história. Você usa suas construções para escrever sua própria história. Você tem considerável liberdade para fazer isso. Você não usará todas as palavras do idioma; você escolhe como estruturar o mundo que está descrevendo; você pode definir algum vocabulário personalizado para seu próprio domínio específico; você pode recorrer a línguas estrangeiras para expressar partes específicas de sua história, você pode criar seu próprio dialeto ou sotaque para fazer com que esse idioma atenda às suas necessidades.

As Unidades Semânticas são representadas como propriedades de entidades do modelo de dados PREMIS apresentado na Figura 5. Essas propriedades se traduzem em metadados que são registrados pelos repositórios e apoiam a realização da preservação digital.

Figura 5 – Modelo de dados PREMIS



Fonte: a partir de (CAPLAN, 2009, p. 8).

A entidade Objeto é o foco da preservação digital no repositório e pode ser do tipo Entidade Intelectual (*Intellectual Entity*), Representação (*Representation*), Arquivo (*File*) ou Sequência de Bits (*Bitstream*). Uma Entidade Intelectual “é uma criação intelectual ou artística distinta que é considerada relevante para uma determinada comunidade no contexto da preservação digital” (PREMIS EDITORIAL COMMITTEE, 2015, p. 8, tradução nossa), podendo ser compreendida como um documento digital e que pode ser formada por outras Entidades Intelectuais. A Representação é o conjunto de arquivos relativos a um objeto que são necessários para a sua renderização e para que seja possível a sua utilização. Um mesmo objeto pode conter várias representações. O Arquivo pode ser compreendido exatamente como um arquivo de computador, formado por bits que são agrupados de uma forma em que seja possível a sua leitura e interpretação por computadores.

A Sequência de Bits (*Bitstream*) representa uma porção do arquivo que pode ser alvo de estratégias específicas de preservação, porém, ela não existe de forma independente. Determinados formatos de arquivos de vídeo como por exemplo o AVI, possuem *bitstreams* separados para o áudio e para o vídeo, cabendo ao repositório decidir se trata a preservação dos *bitstreams* separadamente ou como parte de um arquivo único (CAPLAN, 2009, p. 9).

O Ambiente (*Environment*) é modelado como uma entidade para abrigar todas as informações necessárias sobre o ambiente necessário de *hardware* e *software* para a interpretação e utilização dos objetos preservados, podendo ser utilizada para a preservação dos próprios *softwares*. Assim, é possível preservar os objetos juntamente com os programas que os interpretam, favorecendo estratégias de preservação como a emulação, por exemplo.

A entidade Evento (*Event*) tem como função principal registrar os eventos de preservação executados pelo repositório sobre os objetos. Caplan (2009, p. 11) define a importância do registro dos eventos: “um registro preciso e confiável dos Eventos é essencial para manter a proveniência digital de um Objeto, que por sua vez é imprescindível para demonstrar a autenticidade deste”. A definição sobre quais eventos devem ser registrados é uma escolha de cada repositório, no entanto, o PREMIS recomenda que todas as ações que modificam o objeto preservado devam sempre ser registrados (PREMIS EDITORIAL COMMITTEE, 2015, p. 137) .

A entidade Declaração de Direitos (*Rights Statement*) é utilizada para registrar os aspectos legais e as permissões relacionadas ao objeto preservado, contendo detalhes sobre direitos autorais e permissões de uso, por exemplo. A presença dessas informações auxilia na tomada de decisões do repositório sobre a possibilidade de realizar migrações e conversões de formato ou os requisitos mínimos necessários para essas operações, assim como embasa a forma como a disseminação dos ocorrerá, de acordo com as permissões de uso de cada objeto.

A entidade Agente (*Agent*) registra informações sobre os responsáveis pela execução dos eventos ou dos detentores de direitos dos objetos (*Rights Statement*), representando pessoas físicas, organizações ou mesmo um *software*. Os agentes devem ser identificáveis inequivocamente no âmbito do repositório, associando-os aos eventos executados, o que embasa futuras verificações de autenticidade.

O Quadro 5 exibe um exemplo de definição de uma Unidade Semântica apresentada no Dicionário de Dados. A primeira entrada se refere ao nome da unidade, contendo um código que o identifica unicamente dentro do Dicionário de Dados, nesse caso, é apresentada Unidade Semântica “1.5 *objectCharacteristics*”, que agrega informações do objeto preservado. A propriedade Componentes semânticos apresenta subdivisões da unidade com os respectivos nomes e códigos. Nesse caso, são apresentadas as propriedades de nível de composição (1.5.1 *composition level*), fixidez (1.5.2 *fixity*), tamanho (1.5.3 *size*), formato (1.5.4 *format*), aplicativo de criação (1.5.5 *creatingApplication*), inibidores (1.5.6 *inhibitors*) e demais características (1.5.7 *objectCharacteristicsExtension*). Assim, o repositório pode utilizar essas propriedades para registrar adequadamente cada característica de um determinado arquivo.

Os campos Definição e Justificativa apresentam o propósito da propriedade e o motivo da sua existência, respectivamente. A Restrição de Dados, nesse caso, indica que a propriedade possui componentes semânticos, em forma de Contêiner, mas pode sinalizar que não há nenhuma restrição ou ainda que a informação deve vir de um vocabulário controlado. A Categoria de Objeto indica em quais casos a propriedade se aplica, a uma Entidade Intelectual, Representação, Arquivo ou *Bitstream*. Essa propriedade é complementada pela Repetibilidade e pela Obrigatoriedade que

indicam, respectivamente, se a propriedade pode ser expressada mais de uma vez e se ela é de uso obrigatório.

Quadro 5 – Trecho do Dicionário de Dados para a unidade semântica “size”.

Nome da Unidade Semântica	1.5 <i>objectCharacteristics</i>		
Componentes semânticos	1.5.1 <i>compositionLevel</i> 1.5.2 <i>fixity</i> 1.5.3 <i>size</i> 1.5.4 <i>format</i> 1.5.5 <i>creatingApplication</i> 1.5.6 <i>inhibitors</i> 1.5.7 <i>objectCharacteristicsExtension</i>		
Definição	Propriedades técnicas de um arquivo ou <i>bitstream</i> que são aplicáveis a todos os formatos ou à maioria deles.		
Justificativa	Existem algumas propriedades técnicas importantes que se aplicam a objetos de qualquer formato. A definição detalhada das propriedades de um formato específico está fora do escopo deste Dicionário de Dados, embora tais propriedades possam ser registradas em <i>objectCharacteristicsExtension</i> .		
Restrição de dados	Contêiner		
Categoria de objeto	Entidade Intelectual / Representação	Arquivo	Bitstream
Aplicabilidade	Não aplicável	Aplicável	Aplicável
Repetibilidade		Repetível	Repetível
Obrigatoriedade		Obrigatório	Obrigatório
Notas de uso	As unidades semânticas contidas em <i>objectCharacteristics</i> devem ser tratadas como informações pertinentes a um único objeto em uma única unidade <i>compositionLevel</i> . As características do objeto podem ser repetidas quando um objeto foi criado a partir de duas ou mais codificações, como compressão e criptografia. Nesse caso, cada repetição de <i>objectCharacteristics</i> terá uma unidade semântica <i>compositionLevel</i> incrementalmente maior. Quando se aplica criptografia, o bloco <i>objectCharacteristics</i> deve incluir a unidade semântica <i>inhibitors</i> . Um <i>bitstream</i> incorporado a um arquivo pode ter características de objeto diferentes das características do arquivo. Quando essas características forem relevantes para a preservação elas devem ser registradas.		

Fonte: adaptado de (CAPLAN, 2009, p. 15).

As Notas de uso apresentam informações sobre as possíveis intenções de uso da Unidade Semântica e/ou complementam a definição fornecida. Em alguns casos, o Dicionário de Dados especifica exemplos de valores da propriedade com o intuito de nortear a implementação da propriedade pelo repositório. Ainda, conforme necessário, o Dicionário de Dados inclui Notas de criação e manutenção especificando como o repositório pode obter as informações necessárias para a propriedade.

Com o intuito de permitir a inclusão de informações adicionais específicas à realidade de cada repositório, o Dicionário de Dados prevê a extensibilidade de determinadas propriedades.

Assim, é previsto um certo grau de adaptação do modelo, possibilitando a adição de características que auxiliem a realização da preservação e disponibilização de acesso aos objetos.

Mesmo que o Dicionário de Dados PREMIS se apresente como sendo independente de implementação, os autores disponibilizaram um *schema* XML a ser usado nos casos em que os repositórios optem por armazenar as informações dos objetos através de XML. De acordo com Caplan (2009, p. 16), a maioria dos repositórios já utiliza XML na importação e exportação de dados, o que favorece e facilita a interoperabilidade de plataformas. A utilização do Dicionário de Dados PREMIS em formato XML ainda pode ser combinada com o formato METS, que também armazena informações em formato XML. Várias informações são coincidentes entre os dois padrões, no entanto, algumas outras não encontram correspondências ou requerem decisões de projeto que podem dificultar a interoperabilidade, objetivo principal dessa combinação. Visando diminuir esses problemas, foi elaborado um guia de diretrizes e boas práticas para o uso combinado do Dicionário de Dados PREMIS com o METS (PREMIS EDITORIAL COMMITTEE, 2017).

A contribuição do Dicionário de Dados PREMIS com a preservação digital foi reconhecida através de diferentes premiações. Em 2005, recebeu o *Digital Preservation Award*, pela *British Conservation Awards*, em 2006 recebeu o *Preservation Publication Award* da *Society of American Archivists* e em 2012 foi finalista do *Digital Preservation Award* pela contribuição mais notável para a preservação digital na última década (PREMIS EDITORIAL COMMITTEE, 2015, p. 5). Em 2022, o Dicionário de Dados PREMIS recebeu o prêmio o *Digital Preservation Award* por ter feito o que os juízes definiram como “uma contribuição contínua, substancial e impactante para a comunidade de preservação digital nos últimos 20 anos” (DIGITAL PRESERVATION COALITION, 2022).

Dentre os repositórios digitais existentes, o Archivematica suporta os metadados da versão 3 do Dicionário de Dados PREMIS, utilizando-o para o registro de metadados técnicos, declarações de direito, eventos de preservação realizados sobre os objetos e de agentes. O repositório utiliza XML para o armazenamento das informações e faz uso da combinação PREMIS e METS (ARTEFACTUAL SYSTEMS INC, 2023). A Figura 6 exibe o trecho de um arquivo do Archivematica que registra os detalhes de um evento do tipo *filename change* (modificação do nome de arquivo), com um identificador único atribuído, realizado em 06/01/2020. Os detalhes da realização do procedimento foram registrados (*eventDetailInformation*), assim como o resultado da realização da operação (*eventOutcomeDetailNote*). O cabeçalho do código exibido permite notar a combinação dos padrões METS e PREMIS para a estruturação e armazenamento das informações.

Figura 6 – Registro do evento de modificação de nome de arquivo.

```

<mets:digiprovMD ID="digiprovMD_11">
  <mets:mdWrap MDTYPE="PREMIS:EVENT">
    <mets:xmlData>
      <premis:event xmlns:premis="http://www.loc.gov/premis/v3"
xsi:schemaLocation="http://www.loc.gov/premis/v3
http://www.loc.gov/standards/premis/v3/premis.xsd" version="3.0">
        <premis:eventIdentifier>
          <premis:eventIdentifierType>UUID</premis:eventIdentifierType>
          <premis:eventIdentifierValue>842774e2-7eeb-49a2-99ee-
eade80f6bdlf</premis:eventIdentifierValue>
        </premis:eventIdentifier>
        <premis:eventType>filename change</premis:eventType>
        <premis:eventDateTime>2020-01-06T15:01:52.881142+00:00</premis:eventDateTime>
        <premis:eventDetailInformation>
          <premis:eventDetail>prohibited characters removed: program="sanitize_names";
version="1.10.db4f6aca278e6daf2af160ed40349baf7c6f53af"</premis:eventDetail>
        </premis:eventDetailInformation>
        <premis:eventOutcomeInformation>
          <premis:eventOutcome></premis:eventOutcome>
          <premis:eventOutcomeDetail>
            <premis:eventOutcomeDetailNote>Original
name="%transferDirectory%objects/(brackets)/@at.txt"; new
name="%transferDirectory%objects/(_at.txt)"</premis:eventOutcomeDetailNote>
          </premis:eventOutcomeDetail>
        </premis:eventOutcomeInformation>
      </premis:event>
    </mets:xmlData>
  </mets:mdWrap>
</mets:digiprovMD>

```

Fonte: (ARTEFACTUAL SYSTEMS INC, 2023)

Análogo ao repositório Archivematica, o repositório RODA também é outra solução que utiliza o Dicionário de Dados PREMIS para o armazenamento de metadados que auxiliam na preservação confiável dos objetos (KEEP SOLUTIONS, 2021).

Ainda que o projeto do Dicionário de Dados PREMIS tenha tido como premissa ser flexível e independente de implementações, é desejável que haja conformidade entre as soluções que o implementam, garantindo consistência entre elas. Ao mesmo tempo em que implementar a conformidade pode trazer desafios aos repositórios, ganha-se vantagens nos seguintes aspectos (PREMIS EDITORIAL COMMITTEE, 2015, p. 1): intercâmbio de dados entre repositórios; certificação do repositório; compartilhamento de registros; uso de ferramentas de automação, e; suporte de fornecedores.

Dessa forma, foram definidos três níveis de conformidade a serem analisados pelos repositórios que implementam o Dicionário de Dados PREMIS, de acordo com o tratamento dos seus metadados (PREMIS EDITORIAL COMMITTEE, 2015, p. 4). No nível 1, o repositório possui alguma espécie de mapeamento que permite identificar a equivalência dos seus metadados com o PREMIS. No nível 2, o repositório implementa mecanismos, de preferência automáticos, que realizam a conversão e exportação dos seus metadados de preservação para o PREMIS. No nível 3, o mais alto, o repositório utiliza o próprio PREMIS para armazenar os seus dados de preservação, não requerendo mapeamentos, conversões e transformações.

Ainda, cada um dos níveis de conformidade se subdivide em outros dois níveis: A) analisam a conformidade somente da entidade Objeto, e; B) analisa a conformidade das entidades Objeto, Evento e Agente. Cabe notar que a entidade Declaração de Direitos não é contemplada na análise dos graus de conformidade, ainda que seja reconhecida a sua importância, pois muitas vezes as informações relativas aos direitos e propriedades intelectuais dos objetos já constam em outros metadados descritivos suportados pelos repositórios. O PREMIS Editorial Committee acredita que adicionar essa entidade na verificação de conformidade pode impor um requisito “desnecessariamente oneroso” para informações que já podem estar contempladas em outros metadados (PREMIS EDITORIAL COMMITTEE, 2015, p. 5).

A partir de um extrato de normas, modelos de requisitos e padrões estabelecidos, a presente seção buscou apresentar de que forma esses instrumentos podem contribuir para a superação dos desafios da preservação digital. Ao mesmo tempo, pode-se concluir que existe um vasto campo de instrumentos nesse sentido, que requer planejamento por parte dos repositórios e estudos de viabilidade para que possam ser implantados em suas estruturas e processos, analisando-se os ganhos e impactos causados em cada realidade.

2 O REPOSITÓRIO DIGITAL: NOVO DEPÓSITO PARA PROBLEMAS HETEROGÊNEOS NA TAREFA DE PRESERVAR CONTEÚDOS DIGITAIS

2.1 INTRODUÇÃO

Uma vez que se tenha definido o documento digital como objeto de preservação e se tenham apresentados os desafios envolvidos, a presente seção trata sobre a entidade que realiza efetivamente a preservação, o repositório digital. A tarefa de preservar documentos é para serem utilizados como fontes prova ou consulta é milenar e se funde com a própria história do homem, de onde se fez necessária a criação de estruturas próprias, os Arquivos, para a manutenção da custódia, centralização do acesso e aplicação de técnicas de conservação. Os Arquivos, nesse sentido, possuem o papel de serem guardiões da memória documental, ao mesmo tempo em que dão a chancela de autenticidade aos seus documentos custodiados.

A convivência com o mundo digital é, de certa forma, recente nessa linha do tempo. Faz-se necessária, portanto, a elaboração de processos adaptados a essa realidade, assim como a existência de estruturas, tanto organizacionais quanto tecnológicas, que deem conta dos desafios incumbidos, surgindo então os Repositórios Digitais ou Arquivos Digitais. Essas estruturas tomam a responsabilidade de estabelecer regulamentos próprios e de aplicar as devidas estratégias de preservação para que os materiais permaneçam autênticos e utilizáveis pelo tempo que for necessário, mesmo que esse tempo seja a eternidade. A evolução da realidade analógica para a inclusão dos Repositórios Digitais, com suas respectivas características e peculiaridades é tratada na seção 2.2.

Uma vez que os materiais digitais estejam preservados adequadamente, prossegue-se às suas difusões, propiciando ambientes de pesquisa e acesso mais abrangentes que os ambientes analógicos permitiam, devido às suas fisicalidades. Com o objetivo de fomentar o acesso e a transparência administrativa, em muitos casos, cria-se o ambiente ideal para a disponibilização de materiais que vão desde atos normativos, decisões administrativas, processos jurídicos até artigos científicos, teses e dissertações. Tem-se então a criação da Iniciativa de Acesso Aberto (*Open*

Access Initiative) e a contribuição dos Repositórios Digitais à essa iniciativa, sendo tratada em mais detalhes na seção 2.6.

Com o intuito de se estabelecer requisitos mínimos para a preservação de materiais digitais, de qualquer natureza, de forma confiável, criou-se o modelo de referência OAIS (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2021), que trata desde os aspectos organizacionais dos Repositórios Digitais até a infraestrutura tecnológica necessária. Dessa forma, cria-se um balizador, uma espécie de gabarito, pelo qual os repositórios se guiam para se estruturarem e definirem seus processos, sendo que, de acordo com o grau de aderência do repositório ao modelo OAIS, pode-se dizer que ele realiza a preservação digital confiável de seus materiais. O modelo OAIS é apresentado em detalhes na seção 2.3.

Buscando apresentar modelos de ferramentas de preservação utilizadas em repositórios, a seção 2.4 trata das principais alternativas existentes atualmente. Assim, é possível compreender as particularidades, formas de funcionamento e arquitetura de cada uma, bem como os seus pontos fortes e deficiências, que podem auxiliar em todas de decisão sobre a utilização de uma ou outra ferramenta.

De modo a adaptar o cenário mundial à realidade brasileira, criou-se a definição de um Repositório Arquivístico Digital Confiável (RDC-Arq) (CONSELHO NACIONAL DE ARQUIVOS, 2015), que se baseia principalmente no modelo de referência OAIS, mas inclui um viés de tratamento mais arquivístico de documentos em todas as suas fases. De forma explícita, a definição do RDC-Arq prevê uma forte colaboração entre as áreas de Arquivo e TI, buscando que ambas trabalhem em conjunto, porém com responsabilidades bem delimitadas, para que a preservação digital confiável a longo prazo possa se realizar. A seção 2.5 aborda o tema do RDC-Arq em detalhes.

2.2 DOS ARQUIVOS AOS REPOSITÓRIOS DIGITAIS

A partir da necessidade de se preservar a história e cultura da humanidade, surgem estruturas responsáveis por garantir o acesso a documentos históricos contendo leis, decisões, atos e manifestações culturais que afetam de alguma forma o mundo atual. Da mesma forma, há que se preservar os documentos sendo produzidos atualmente para acesso futuro, sob pena de não haver uma memória coletiva construída que possa comprovar a realização de atividades e processos que afetam indivíduos e até mesmo a sociedade como um todo.

É possível encontrar relatos de arquivos desde a antiguidade, contendo desde escrituras de propriedades de terras, até obras literárias, sob a custódia de órgãos incumbidos da

responsabilidade de preservar os documentos e fornecer acesso às pessoas interessadas. Assim como os cidadãos possuíam necessidade de comprovar determinados fatos, talvez os governos fossem os que mais se beneficiavam dessas fontes de informação, de onde extraíam dados de produção de agricultura, recolhimento de impostos e mapas cartográficos a respeito de seus territórios e possíveis novos interesses geográficos. Dentre as vantagens do estabelecimento dos arquivos está a organização dos seus materiais, utilizando-se de processos para a classificação e avaliação, o que traz agilidade e eficiência aos governos e, por consequência, requer a necessidade de preservação desses materiais. Existem ainda as razões culturais e históricas que levam à necessidade de preservação de obras literárias de grande importância regional, nacional ou global, por exemplo, ou ainda a preservação de atos normativos que estabeleceram e garantiram direitos a determinados grupos étnicos.

Para Schellenberg (2003), a experiência e trajetória na criação de arquivos nacionais na era moderna, em contraste com os arquivos antigos, da França, Inglaterra (Reino Unido) e Estados Unidos em particular merecem um destaque especial, sendo detalhados a seguir.

Na França, como resultado da Revolução Francesa, foi criado em 12 de setembro de 1790 o *Archives Nationales* (Arquivo Nacional) e em 25 de junho de 1794 teve os seus moldes definidos por decreto de lei com o objetivo de torná-lo um “depósito central para os arquivos nacionais”, no qual se estabeleceram três princípios que permanecem até os dias atuais: 1) centralização dos arquivos da nação; 2) acesso público gratuito, e; 3) criação de uma rede nacional de arquivos através da criação de serviços de arquivo em cada departamento (ARCHIVES NATIONALES, 2023). Ao longo do tempo, com o aumento na demanda por depósito de documentos e acessos ao acervo, o *Archives Nationales* passou por mudanças de localização, possuindo três instalações atualmente, sendo a última inaugurada em 2013 em Paris. Em 2006, através da publicação do Decreto de 24 de dezembro, *Archives Nationales* passa a ser tratado como um serviço nacional, com a missão de “coletar, classificar, inventariar, conservar, restaurar, prover acesso e promover os arquivos públicos das administrações centrais do Estado, os arquivos dos notários de Paris e grupos de registros privados de interesse nacional” (ARCHIVES NATIONALES, 2023). Para Schellenberg (2003, p. 5, tradução nossa), a Revolução Francesa que originou o *Archives Nationales*, causou um grande reconhecimento sobre o valor dos documentos na sociedade, levando a alcançar três importantes conquistas: 1) estabelecimento de uma administração arquivística nacional e independente; 2) reconhecimento do princípio de acesso público aos arquivos, e; 3) reconhecimento da responsabilidade do estado pelo cuidado dos documentos históricos valiosos.

Na Inglaterra, em 14 de agosto de 1838 foi criado o *Public Record Office* (PRO) após uma série de investigações identificarem mal uso e conservação dos arquivos custodiados pelo governo daquele país. Assim, originou-se o PRO, sob controle do cargo *Master of Rolls*, com a responsabilidade de preservar e dar acesso os documentos dos governos centrais da Inglaterra, País de Gales. Entre os anos 2003 e 2006 houve a formação do Arquivo Nacional (*The National Archives*), que uniu o *Public Record Office* e mais outros três órgãos de responsabilidades e atribuições semelhantes. Dentre os seus objetivos está produzir, prover acesso e preservar os documentos públicos sob sua guarda, além de utilizá-los como fonte de informação em consultorias do próprio governo (NATIONAL ARCHIVES, 2015).

Nos Estados Unidos, somente em 1934 foi criado o *National Archives* (Arquivo Nacional) com a missão de reunir e preservar os documentos de interesse do governo. Não muito diferente da situação da Inglaterra, antes da criação do Arquivo Nacional, os documentos do governo dos Estados Unidos sofreram diversas catástrofes como incêndios, inundações e deteriorações, situações essas que ao longo do tempo trouxeram luz à necessidade de haver uma estrutura dedicada à preservação desses documentos. Os documentos que até então eram armazenados em diferentes estruturas e localizações dispersas, como sótãos e porões sem o devido tratamento, a partir de então foram centralizados e passaram aos cuidados de equipes especializadas. Em 1985, a instituição passou a se chamar *National Archives and Records Administration* (NARA), enfatizando o seu papel desempenhado na gestão documental do país através da criação de atos e normas que regulamentam a produção e preservação dos documentos, além de administrar as diferentes localizações físicas criadas em todo o país para o armazenamento dos seus documentos (NATIONAL ARCHIVES, 2021).

Em comum entre as breves histórias relatadas dos arquivos dos três países, está a necessidade de haver uma organização responsável pelos documentos que seja confiável, disponibilize acesso e preserve os documentos sob sua responsabilidade, delimitando características do que vem a ser um repositório de documentos.

No mundo dos documentos digitais, Waters e Garret definiram arquivos digitais como (1996, p. 15, tradução nossa):

[...] repositórios de informações digitais que são coletivamente responsáveis por garantir, através do exercício de várias estratégias de migração, a integridade e acessibilidade a longo prazo da herança social, econômica e cultural da nação em formato digital.

Os autores ainda fazem distinção entre as funções de um arquivo digital e uma biblioteca digital, especificando que essa última recolhe e disponibiliza acesso aos seus materiais digitais, mas pode não conter processos de preservação a longo prazo. Os arquivos, no entanto,

englobam as funções das bibliotecas, como a seleção, obtenção, armazenamento e disponibilização de acesso, porém adicionando rotinas de preservação a longo prazo dos seus materiais. Outro requisito que diferencia essas estruturas se refere à natureza do material preservado, uma vez que a biblioteca armazena diversos exemplares de uma mesma obra, por exemplo, enquanto o arquivo armazena documentos únicos. As fronteiras que diferenciam ambas entidades podem nem sempre serem tão delimitadas e rígidas, a ponto de ambas possuírem processos, funções e finalidades quase idênticas e realizarem a preservação de seus materiais a longo prazo

O relatório da Força Tarefa de Waters e Garret posiciona os repositórios de informação digital como sendo partes de um sistema nacional de arquivos operando através de dois mecanismos, que são o cumprimento de padrões e normas de preservação pelos repositórios de maneira a garantir que não haja perda ou inacessibilidade dos materiais preservados ao longo do tempo e o estabelecimento de certificações formais para esses repositórios do sistema. As certificações estabeleceriam uma base comum sólida entre os repositórios que serviriam como um mecanismo a prova de falhas no sentido de que um repositório poderia englobar os materiais de um outro repositório o qual tenha sido identificado condições que coloquem em risco a preservação de seus materiais, seja por motivos administrativos, econômicos, tecnológicos ou advindos de catástrofes naturais, por exemplo (WATERS e GARRET, 1996).

Enquanto Waters e Garret definem que os repositórios são parte constituinte de um sistema de arquivos nacional, o modelo de referência OAIS, desde a sua versão de 2002, define o termo *Archive* (Arquivo) como “uma organização que pretende preservar informação para acesso e uso por uma Comunidade-Alvo” (CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS, 2002, p. 8, tradução nossa). Nesse mesmo ano, o relatório da RLG/OCLC sobre os atributos e responsabilidades de repositórios, define o termo repositório confiável como “Um repositório digital confiável é aquele cuja missão é prover acesso confiável a longo prazo aos recursos digitais gerenciados para a sua comunidade-alvo, agora e no futuro” (RESEARCH LIBRARIES GROUP, 2002, p. 5, tradução nossa).

Essa última definição de certa forma incorpora a definição de Arquivo da OAIS, no que se refere a preservar e dar acesso os seus materiais a longo prazo efetuando, para isso, a preservação digital, adicionando o aspecto da confiabilidade. Embora a definição OAIS não contemple explicitamente o caráter da confiabilidade, os seus requisitos e estrutura se voltam a esse fim, como será visto em seção posterior do corrente estudo.

A manutenção da autenticidade dos materiais é característica indispensável aos repositórios, de forma a construir um ambiente de confiança entre produtores e consumidores desses materiais. No entanto, é necessário atentar para as condições em que os materiais foram produzidos

e manipulados até o momento da ingestão no repositório, para que haja elementos que identifiquem o contexto de produção, proveniência e detalhamento sobre as operações realizadas até esse momento. No caso de documentos arquivísticos, faz-se necessário ainda existir as suas relações orgânicas com outros documentos e funções da organização produtora, normalmente, espelhando o seu Plano de Classificação de Documentos. Sem os devidos cuidados no âmbito do sistema de produção e gestão dos documentos, corre-se o risco de haver a ingestão de materiais que não podem ser comprovadamente autênticos, o que tornaria a preservação inútil.

Atendendo a uma demanda do *Joint Information Systems Committee* (JISC), Heery e Anderson (2005) elaboraram um relatório sobre repositórios digitais, onde identificaram um grande uso dos termos repositório e repositório digital, com significados variando de acordo com o entendimento de cada organização ou empresa. Dentre os primeiros objetivos, as autoras buscaram compreender o que distingue um repositório de outras coleções de materiais, diretórios, catálogos e bancos de dados, quais bases definem minimamente um repositório e qual a natureza dos materiais que eles preservam.

O relatório especifica os seguintes aspectos que diferenciam um repositório digital das demais coleções de materiais (HEERY e ANDERSON, 2005, p. 1, tradução nossa):

- 1) o conteúdo é depositado em um repositório pelo criador do conteúdo, proprietário ou terceiro; 2) a arquitetura do repositório conteúdo assim como metadados; 3) o repositório oferece um conjunto mínimo de serviços básicos como inserção, obtenção, pesquisa e controle de acesso aos seus materiais, e; 4) o repositório deve ser sustentável e confiável, bem apoiado e bem gerenciado.

A partir desse conjunto, pode-se ressaltar a característica de o repositório ser sustentável, no sentido de serem necessários mecanismos que garantam o seu funcionamento, indo desde a existência de uma estrutura adequada, passando pela sustentabilidade financeira e culminando no uso de tecnologias pertinentes aos seus objetivos. A necessidade de o repositório ser confiável se faz necessária e primordial a ponto de seus usuários terem garantias de que os materiais depositados são o que dizem ser sem sofrerem adulterações. Cabe ressaltar que, de acordo com esses aspectos, a preservação a longo prazo pode nem sempre ser um serviço disponibilizado pelo repositório, podendo estar presente de acordo com os objetivos, necessidades e principalmente as capacidades de cada repositório.

A melhoria na comunicação institucional é tida como um dos principais motivos para o estabelecimento de repositórios digitais, sejam eles institucionais ou de domínio específico, ao menos nos casos de eles serem de acesso aberto. Assim, quanto ao acesso, os repositórios devem prover acesso aos seus conteúdos, observando-se as questões legais e direitos de acesso, assim

como aos respectivos metadados, seja para consultas manuais ou para a coleta de informações (*harvesting*) (HEERY e ANDERSON, 2005, p. 2).

As razões que levam ao estabelecimento de um repositório, segundo Heery e Anderson, provêm de uma intersecção de interesses entre diferentes comunidades de uso, identificando-se entre elas as bibliotecas digitais, pesquisa, publicações, gerenciamento de documentos e preservação. Em comum, são identificados serviços mínimos que os repositórios devem oferecer (HEERY e ANDERSON, 2005, p. 2, tradução nossa):

- 1) Acesso aprimorado aos recursos; 2) Novos modos de publicação e revisão por pares; 3) Gerenciamento de informações corporativas (sistemas de gerenciamento de documentos e conteúdo); 4) Compartilhamento de dados (reutilização de dados de pesquisa, reutilização de objetos de aprendizagem), e; 5) Preservação de recursos digitais.

Em relação à arquitetura dos repositórios, tem-se o modelo de referência OAIS, que define um componente principal que gerencia, executa os processos de preservação e faz interface com os produtores e consumidores de materiais. Os detalhes, componentes e forma de comunicação utilizada pela arquitetura OAIS são detalhados na seção 2.3 do presente estudo.

Posterior ao modelo OAIS, Wheatley (2007) apresenta uma abstração do modelo de arquitetura de um repositório, apresentando três principais camadas:

1. Camada de armazenamento (*Storage layer*): gerencia o armazenamento físico dos materiais digitais e trata da integridade e segurança dos dados;
2. Camada de integração (*Middleware layer*): provê serviços e protocolos de interoperabilidade do repositório com outros sistemas e aplicações. A camada ainda inclui interfaces para acesso e pesquisa no repositório, gerenciamento de metadados e serviços de preservação;
3. Camada de aplicação (*Application layer*): possui serviços e interfaces que permitem que o usuário final pesquise e acesse o conteúdo preservado.

O autor ainda ressalta a importância de se utilizar padrões para a construção dos repositórios, de forma a garantir uma interoperabilidade e a realização plena da preservação a longo prazo, tal como o modelo OAIS.

Para dados de pesquisa de um domínio de conhecimento específico, ou mesmo para vários, em geral se utilizam repositórios de dados que contam com funcionalidades para o compartilhamento de dados entre pesquisadores e mecanismos para o referenciamento dos conjuntos de dados nos trabalhos científicos produzidos. Assim, combina-se a facilitação de acesso aos dados e compartilhamento entre pesquisadores, além de garantir a proveniência dos dados.

O compartilhamento de pesquisas e artigos entre pares costuma utilizar repositórios de domínio ou temáticos (*subject repositories*), definido por Café *et al* como (2003, p. 2):

Um repositório temático se constitui em um conjunto de trabalhos de pesquisa de uma determinada área do conhecimento, disponibilizados na Internet. Esses repositórios utilizam tecnologias abertas e seguem a filosofia da Iniciativa dos Arquivos Abertos, promovendo a maior acessibilidade à produção dos pesquisadores e à discussão entre seus pares.

Dentre os exemplos de repositórios temáticos, pode-se destacar o arXiv²⁷ como sendo um dos precursores nessa área. Surgido em 1991, não por acaso, juntamente com o surgimento da internet, inicialmente o repositório recebia e disponibiliza os arquivos via e-mail, quando em 1993 foi disponibilizada uma interface *web* para envio, consulta e acesso dos materiais. A demanda pelo compartilhamento de artigos da área física, principalmente *pre-prints*, foi a motivação inicial para a criação da plataforma, que fornece um certo grau de seleção, curadoria e avaliação dos trabalhos submetidos. O seu funcionamento se baseia na revisão por especialistas das áreas atendidas, sem que haja custos aos autores ou remuneração aos revisores, servindo como um grande serviço de compartilhamento de conhecimento e sendo um dos precursores da ciência aberta (GINSPARG, 2004). Atualmente, o repositório conta com mais de 2 milhões de artigos das áreas física, matemática, ciência da computação, biologia quantitativa, finanças quantitativas, estatísticas, engenharia elétrica, sistemas da ciência e economia, sendo mantido pela Cornell University (ARXIV, 2023).

Uma vez que os materiais tenham sido enviados e publicados pelo arXiv, eles permanecem acessíveis para sempre, o que faz com que a plataforma adote medidas internas para a preservação desses materiais. No processo de recebimento dos materiais são feitas verificações que nos arquivos enviados procurando preconizando formatos abertos e que não acarretem violações de direitos autorais.

Dentre os tipos de repositórios existentes, pode-se encontrar os repositórios institucionais. Crow define descreve o repositório institucional como “[...] coleções digitais que capturam e preservam a produção intelectual de uma comunidade única ou multiuniversitária” (2002, p. 4, tradução nossa). Para o autor, o repositório institucional atende às demandas de prover um meio de comunicação centralizado para a difusão do conhecimento acadêmico diferente dos meios de publicação tradicionais e servir como um meio de aumentar a visibilidade da instituição, trazendo valor e prestígio à instituição ao mesmo tempo em que adiciona os processos e responsabilidades da preservação a longo prazo dos seus materiais. A partir dessas demandas, o

²⁷ <https://arxiv.org/>

autor refina, detalha e sofisticada mais o conceito de repositório institucional (CROW, 2002, p. 16, tradução nossa):

um repositório institucional é um arquivo digital do produto intelectual criado pelo corpo docente, equipe de pesquisa e alunos de uma instituição e acessível a usuários finais dentro e fora da instituição, com poucas ou nenhuma barreira de acesso.

Para Café *et al*, o repositório institucional representa a união dos diversos repositórios temáticos de uma organização (2003, p. 4):

Um repositório institucional é a reunião de todos os repositórios temáticos hospedados em uma organização. No caso de uma universidade, cada departamento trata de uma área do conhecimento e, portanto, seu repositório temático será específico no assunto deste departamento. A união de todos os repositórios das diversas unidades de pesquisa comporá o repositório institucional, caracterizando-o como multidisciplinar.

Dessa forma, a instituição que controla o repositório tem o poder de definir os tipos de materiais depositados, requisitos para depósito, políticas de preservação e planos de acesso dos seus materiais, porém servindo como principal meio de consulta da sua produção intelectual e científica. Em certos casos, pode-se identificar uma intersecção das funções e alcances de repositórios institucionais e Arquivos universitários, por exemplo, no qual esse último se responsabiliza pelos documentos gerados pelas funções executadas na instituição e os preserva pelo tempo que for necessário, observando principalmente os prazos legais envolvidos.

Lynch (2003, p. 3, tradução nossa) traz uma definição sobre os repositórios institucionais universitários, abordando a responsabilidade por gerenciar, preservar e disseminar os materiais custodiados:

um repositório institucional universitário é um conjunto de serviços que uma universidade oferece aos membros de sua comunidade para o gerenciamento e disseminação de materiais digitais criados pela instituição e seus membros da comunidade. É essencialmente um compromisso organizacional com a administração desses materiais digitais, incluindo preservação de longo prazo, quando apropriado, bem como organização e acesso ou distribuição.

As responsabilidades do repositório trazidas pelo autor sempre devem ir de encontro aos interesses da sua comunidade, o que encontra confluência com a definição de Comunidade-Alvo definida na OAIS (CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS, 2012, p. 11, tradução nossa):

um grupo identificado de possíveis consumidores que deve ser capaz de entender um determinado conjunto de informações. A comunidade pode ser composta de várias comunidades de usuários. Uma Comunidade-Alvo é definida pelo Arquivo e essa definição pode mudar com o tempo.

Pode-se extrair diversas características dessa definição como a relação de consumidor, compreensão das informações e a mutabilidade no tempo. A relação de consumidor com o repositório se dá no sentido de possibilitar aos usuários o depósito de seus materiais para a preservação assim como o posterior acesso desses materiais, necessitando para isso, a execução de rotinas que garantam que as informações depositadas serão compreensíveis ou interpretáveis no futuro. Ressalta-se aqui que os consumidores do repositório podem ser uma pessoa, um sistema computacional ou até mesmo outro repositório que se relacionam entre si. O fato de admitir que a comunidade pode se modificar ao longo do tempo traz implícita a necessidade de monitoramento dos seus requisitos, tanto de produção quanto de acesso, assim como requer a capacidade de o repositório se adaptar a essas mudanças identificadas.

Um outro aspecto a ser levado em consideração na definição de repositórios é em relação à sua tipologia, que varia de acordo com o tipo de conteúdo, cobertura, função primária e grupo de usuários alvo. Quanto ao tipo de conteúdo deve-se analisar o material recebido e disponibilizado pelo repositório, podendo ser dados de pesquisa com ou sem tratamento, *pre-prints*²⁸, artigos revisados de revistas ou eventos, teses digitais, textos originais de publicações, objetos de aprendizagem e documentos institucionais como registros de alunos e de pessoal, por exemplo. Quanto à cobertura, variando do mais específico ao mais genérico, o repositório pode ser pessoal, de uma revista acadêmica, departamental, institucional, interinstitucional, nacional e internacional. Quanto à função primária, o repositório pode servir como acesso aprimorado a recursos, acesso a recursos de um domínio de conhecimento, preservação de recursos digitais, novas maneiras de disseminação, gerenciamento de ativos institucionais e compartilhamento e reuso de recursos. Quanto ao grupo de usuários alvo, o repositório pode se destinar a estudantes, professores e pesquisadores (HEERY e ANDERSON, 2005, p. 13).

Pode-se imaginar uma combinação bastante extensa das características citadas que levam a tipificar os repositórios, podendo-se concluir que essas definições não são estanques, fechadas e restritas. Um único repositório institucional pode, por exemplo, possuir documentos arquivísticos produzidos na execução de suas atividades, assim como materiais de teses digitais e artigos de seus pesquisados, abrangendo um grupo de usuários que engloba desde os funcionários, estudantes, professores e pesquisadores da instituição até a comunidade geral externa à instituição.

A instalação e uso de repositórios de domínio precedeu os repositórios institucionais, fato impulsionado pela busca das comunidades em atender as suas demandas e anseios. O surgimento crescente desses repositórios gerou a necessidade de criação de repositórios

²⁸ Artigos ou materiais de pesquisa ainda em estágio de revisão entre pares não submetidos à publicação ou submetidos, porém não publicados.

institucionais, de amplo alcance nas universidades, com o intuito de centralizar os esforços na manutenção, preservação e acesso desses materiais de pesquisa. Inicialmente, foram registradas resistências dessas comunidades de pesquisa em adotar os repositórios institucionais como meio para a publicação de seus materiais (XIA, 2008), porém como Björk demonstrou, houve uma mudança nesse cenário (2014, p. 705, tradução nossa):

atualmente, é muito mais fácil do ponto de vista gerencial lançar um repositório institucional, que geralmente é operado por funcionários dedicados de bibliotecas universitárias, do que um repositório de domínio, que pode carecer de financiamento inicial e pode exigir uma rede internacional de colaboradores para começar. [...] Os repositórios institucionais também são extensões naturais dos atuais sistemas de informação de pesquisa que quase todas as universidades agora usam para acompanhar sua produção de publicações, e está muito “na moda” iniciar repositórios institucionais.

Ao mesmo tempo em que Hockx-Yu, gerente do Programa de Preservação Digital do *Joint Information Systems Committee* (JISC), traz uma definição de repositório institucional, procurando delimitar o seu tipo de material em relação aos repositórios de domínio, a própria definição admite a inclusão de materiais que poderiam ser considerados destinados a repositórios de domínio, permitindo-se concluir que nem sempre essa fronteira entre ambos é fixa e linearmente delimitada (HOCKX-YU, 2006, p. 1, tradução nossa):

Enquanto alguns repositórios se concentram em domínios de assunto específicos, um repositório institucional armazena e torna acessíveis os ativos educacionais, de pesquisa e associados de uma instituição. Embora a maioria dos repositórios institucionais atualmente estabelecidos sejam repositórios de ‘e-prints’ que fornecem acesso aberto aos resultados de pesquisa de uma universidade ou instituição de pesquisa, o conteúdo não precisa ser limitado a e-prints, mas pode potencialmente incluir dados de pesquisa, material de aprendizagem, coleções de imagens e muitos outros tipos diferentes de conteúdo.

A existência de repositórios de domínio muito provavelmente irá perdurar por muitos anos, voltados às demandas específicas das comunidades atendidas, no entanto, a operacionalização desses repositórios pode se tornar inviável no sentido de necessitarem estruturas, equipes e verbas dedicadas à gestão, preservação e disponibilização de acesso a longo prazo dos materiais custodiados.

2.3 O MODELO DE REFERÊNCIA OAIS

Visando estabelecer uma estrutura mínima a ser seguida pelos repositórios digitais, foi criado o modelo de referência OAIS, que engloba elementos que vão desde a infraestrutura organizacional, passando pelo tratamento dos objetos digitais e se estende até a difusão dos materiais preservados. A importância do modelo OAIS para a preservação digital é tal, que o fato de um repositório ser aderente ao modelo o caracteriza como confiável e pressupõe que os seus

materiais estarão preservados corretamente pelo prazo que for necessário. Na presente seção do trabalho, o modelo de referência OAIS é detalhado, contemplando a sua gênese, apresentação dos conceitos-chave e requisitos e a sua arquitetura de componentes e pacotes.

2.3.1 A Origem do Modelo

O Comitê Consultivo para Sistemas de Dados Espaciais (*Consultative Committee for Space Data Systems – CCSDS*), composto por especialistas mundiais em sistemas de dados, elaborou um framework destinado, inicialmente, à preservação por tempo indefinido, dos dados provenientes de experimentos espaciais. O primeiro resultado desse esforço foi publicado como um Padrão Recomendado (*Recommended Standards*), conhecido como *Blue Book*, definindo os componentes e as interações entre si que devem estar presentes nos sistemas de preservação.

Após revisões do documento, o CCSDS modificou o documento e o tornou uma Prática Recomendada (*Recommended Practice*), conhecido como *Magenta Book* (Livro Magenta – 650.0-M-2) (CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS, 2012), retirando o seu aspecto impositivo e apresentando um guia geral para o tratamento de dados de missões espaciais. O trabalho do comitê resultou no modelo de referência para Sistemas Abertos de Arquivamento de Informações (SAAI), em inglês, *Open Archival Information System* (OAIS). Um acordo realizado entre o CCSDS e a *International Organization for Standardization* (ISO) transformou o Livro Magenta na norma ISO 14721:2012 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012), traduzida na Espanha para a norma UNE-ISO 14721:2015²⁹. No Brasil, a norma correspondente é a ABNT NBR 14721:2021³⁰, publicada em julho de 2021. Acerca do modelo OAIS, Ferreira destaca a sua maior contribuição para a preservação digital (2006, p. 28):

Um dos contributos mais notáveis desta iniciativa foi a definição de uma terminologia própria que viria a facilitar a comunicação entre os diversos intervenientes envolvidos na preservação de objetos digitais.

2.3.2 Conceitos e Requisitos Básicos do Modelo

Um **Arquivo OAIS** consiste em “[...] uma organização, que pode ser parte de uma organização maior, de pessoas e sistemas que aceitaram a responsabilidade de preservar informação e de torná-la disponível para uma Comunidade-Alvo.” (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2021, p. 1). Sobre o termo “aberto”, a norma ressalta que se relaciona ao

²⁹ (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2015)

³⁰ (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2021)

aspecto de a sua elaboração ser discutida em fóruns e comunidades de uma forma aberta, colaborativa e transparente e não implica em dizer que o acesso ao arquivo seja irrestrito.

O termo **Comunidade-Alvo** é definido como (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2021, p. 19):

Um grupo identificado de potenciais Consumidores, que convém que sejam capazes de entender, em particular, um conjunto de informações. A Comunidade-Alvo pode ser composta por múltiplas comunidades usuárias. Uma Comunidade-Alvo é definida pelo Arquivo e esta definição pode mudar com o tempo.

Cabe frisar que o termo original citado na norma ISO 16363 é “*Designated Community*” (CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS, 2012, p. 11), que teria como tradução literal o termo “Comunidade Designada”, da mesma forma, a norma AENOR-ISO 14721 traduziu o termo para “*Comunidad Específica*” (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2015, p. 21). Embora o termo tenha sido adaptado para os idiomas português e espanhol, as definições possuem os mesmos significados da norma original.

O conceito de Comunidade-Alvo, por sua vez se conecta ao conceito de **Base de Conhecimento** (*Knowledge Base*) (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2015, p. 20, tradução nossa): “Conjunto de informação, incorporada por uma pessoa ou sistema, que permite a essa pessoa ou sistema compreender a informação recebida”. Há que se ter consciência sobre a Base de Conhecimento da Comunidade-Alvo do repositório, uma vez que ela vai guiar as suas políticas de preservação e acesso.

Quanto mais específica for a Base de Conhecimento, mais especializado o seu público tende a ser, o que permite voltar as políticas do repositório à satisfação de demandas específicas. Normalmente, esse é o caso de repositórios de domínio ou de dados de pesquisa os quais se relacionam com comunidades especialistas em determinadas áreas que utilizam formatos de arquivos e conjuntos de dados específicos. Por outro lado, quanto mais ampla for a Base de Conhecimento, menos se pode focar em demandas específicas, levando o repositório a estabelecer políticas mais genéricas para acomodar todos os requisitos da sua comunidade ou, de maneira oposta, estabelecer políticas estritas no sentido de definir conjunto mínimo de requisitos comum a todos.

A organização disposta a se tornar um Arquivo OAIS deve aceitar cumprir um conjunto mínimo de responsabilidades, tais como negociar e aceitar a informação dos produtores, obter controle suficiente para garantir a preservação a longo prazo, definir a sua Comunidade-Alvo, garantir que os materiais preservados permaneçam compreensíveis pela Comunidade-Alvo, seguir políticas e procedimentos documentados de modo a garantir a preservação, mesmo que cesse o

funcionamento do Arquivo e garantir o acesso e autenticidade dos materiais preservados (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2021, p. 35).

A preocupação com a preservação dos documentos se inicia no estabelecimento dos requisitos mínimos a serem cumpridos pelos produtores que depositam seus documentos no repositório. Esse contrato acordado entre as partes determina quais formatos de arquivos são aceitos, o conjunto de metadados esperado, assim como as restrições de acesso referentes à presença de dados sensíveis ou direitos autorais. Tanto os sistemas de informação quanto as pessoas interessadas em preservar seus documentos devem estar cientes sobre as condições estabelecidas pelo arquivo e, caso necessário, adaptar suas rotinas de produção de documentos.

O funcionamento interno do repositório necessita realizar operações de manipulação nos objetos depositados a fim de realizar a preservação a longo prazo, o que requer estabelecer um contrato com os produtores de conteúdo no que se refere aos direitos autorais e de propriedade. O repositório deve ter controle dos materiais ingeridos e ter permissões dos produtores para a realização das rotinas de preservação, ao mesmo tempo em que deve observar os direitos de propriedade e acesso existentes. Cabe a ambos negociar os termos desse acordo, de modo que não haja infrações.

A definição da Comunidade-Alvo do repositório pode ser vista como um dos pilares do repositório, uma vez que é ela que vai determinar os requisitos de ingestão, procedimentos de preservação e requisitos de acesso aos materiais. Essa delimitação se torna fundamental e norteia as políticas de preservação a serem estabelecidas na instituição e cumpridas pelo repositório. Tão importante quanto a definição, é o monitoramento da Comunidade-Alvo, com vistas a detectar e reagir a mudanças de requisitos, adaptando o funcionamento e processos do repositório. Santos e Flores definem a importância do monitoramento da Comunidade-Alvo (2019, p. 125):

[...] será interessante que um repositório OAIS aplique ações para monitorar as comunidades designadas, e consequentemente, verificar se os documentos preservados são de fácil compreensão. Este é um processo complexo, pois as bases de conhecimento das comunidades designadas mudam o tempo todo, o que força uma verificação sistemática.

Garantir que os materiais permaneçam acessíveis ao longo do tempo, requer, primeiramente que se estabeleçam requisitos na ingestão dos materiais. Nos casos de materiais com formatos muito específicos, de softwares proprietários, por exemplo, pode ser necessário a exigência de inclusão de outros formatos que facilitem o acesso e a interpretação dos dados no futuro. Outra situação pode ser a ingestão de dados puros em texto, por exemplo, que não possuam instruções para a sua interpretação correta e que possam comprometer a interpretação desses dados no futuro.

Todo o funcionamento e processos de um repositório devem estar sustentados nas políticas que o regem, voltando-se à garantia de que a preservação a longo prazo é cumprida. Desde os processos realizados sobre os Pacotes de Informação de Arquivamento, até a utilização de tecnologias devem refletir as políticas estabelecidas de modo que não haja distorções entre a teoria e a prática realizada. Cabe ressaltar a importância das previsões das fontes de financiamento do repositório para o seu funcionamento e para as evoluções de tecnologia, estrutura e equipe que podem se tornar necessárias para a realização da preservação. Tão importante quanto garantir o funcionamento do repositório, é a garantia de que a preservação dos materiais nele depositados seguirá acontecendo, mesmo nos casos em que ele deixe de funcionar por quaisquer que sejam os motivos, o que requer um plano de sucessão que preveja a transferência de custódia a outra instituição, por exemplo.

A disponibilização de acesso aos materiais também é uma das responsabilidades primordiais do repositório, uma vez que ela dá o seu real sentido de existência. A partir dos materiais preservados, o repositório pode construir derivadas de acesso a serem acessadas pelos Consumidores, contendo cópias destinadas e otimizadas a esse fim, além dos metadados necessários para que seja possível presumir a autenticidade das informações prestadas. A adoção de processos bem estabelecidos e documentados pelo repositório permite à sua Comunidade-Alvo formar embasamento sobre a autenticidade dos materiais preservados, uma vez que essa deve ser uma das premissas principais dos repositórios.

Nos casos de preservação de fotografias ou documentos digitalizados, por exemplo, o repositório pode optar por utilizar formatos de arquivo mais populares e com menor tamanho que o original, facilitando o acesso dos usuários. O repositório deve, contudo, atentar às restrições de acesso dos seus materiais protegidos, disponibilizando-os somente quando os requisitos de acesso sejam cumpridos.

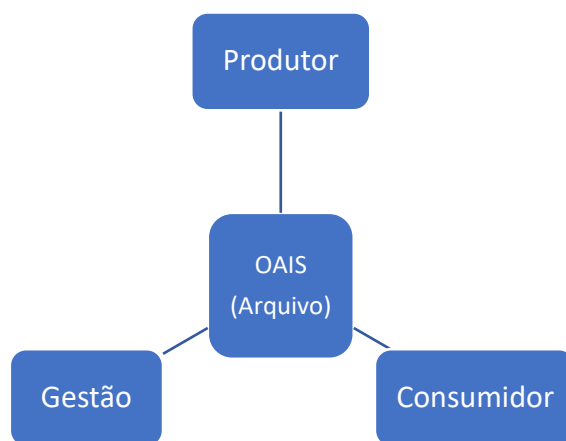
2.3.3 Arquitetura do Modelo

A arquitetura do sistema foi elaborada para ser utilizada por arquivos que recebem documentos para preservação, os gerenciam e disponibilizam acesso à sua comunidade. O tempo mínimo de preservação dos documentos recebidos pelo Arquivo pode variar conforme critérios da instituição, no entanto, o modelo desenvolvido não faz distinção quanto à temporalidade sendo possível utilizá-lo para documentos que serão eliminados em pouco tempo ou a documentos que devem ser preservados para sempre. O principal aspecto norteador do modelo é a preocupação na preservação de documentos que podem sofrer com mudanças de tecnologia ou mudanças nos

requisitos da comunidade atendida (CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS, 2012, p. 1-1).

A arquitetura do modelo define um componente principal que gerencia, executa os processos de preservação e faz interface com os produtores e consumidores de materiais, conforme exibe a Figura 7. O Produtor (*Producer*) representa o papel das pessoas ou sistemas que depositam os seus documentos no arquivo para preservação. A Gestão (*Management*) é responsável pelo gerenciamento dos documentos, aplicando a política de preservação estabelecida pela instituição. O Consumidor (*Consumer*) representa as pessoas que manifestam interesse em acessar os documentos sob a posse do Arquivo ou sistemas destinados à difusão dos documentos.

Figura 7 – Modelo de ambiente de um repositório OAIS.



Fonte: (CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS, 2012, p. 2-2, tradução nossa)

Internamente, a arquitetura OAIS trabalha com o conceito de Objetos de Informação (*Information Object – IO*) que é composto pelo(s) Objeto(s) de Dados (*Data Object – DO*) e pelas Informações de Representação (*Representation Information – RI*). O Objeto de Dados representa o conteúdo, os bits, de um arquivo digital que interage com o repositório. As Informações de Representação contêm as instruções para se interpretar corretamente o conteúdo do DO. Em determinados casos, podem haver várias RI que formam uma rede advinda do fato de que para a interpretar uma RI é necessária outra RI e assim, sucessivamente. O repositório pode optar por manter mais ou menos Informações de Representação, de acordo com as especificidades da Base de Conhecimento da sua Comunidade-Alvo. Assim, o IO se compõe a partir dos conteúdos de seus arquivos, juntamente com as instruções para as suas interpretações.

Toda a comunicação entre as entidades é realizada através de Pacotes de Informação (*Information Packages – IP*) que são definidos como (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2021, p. 22):

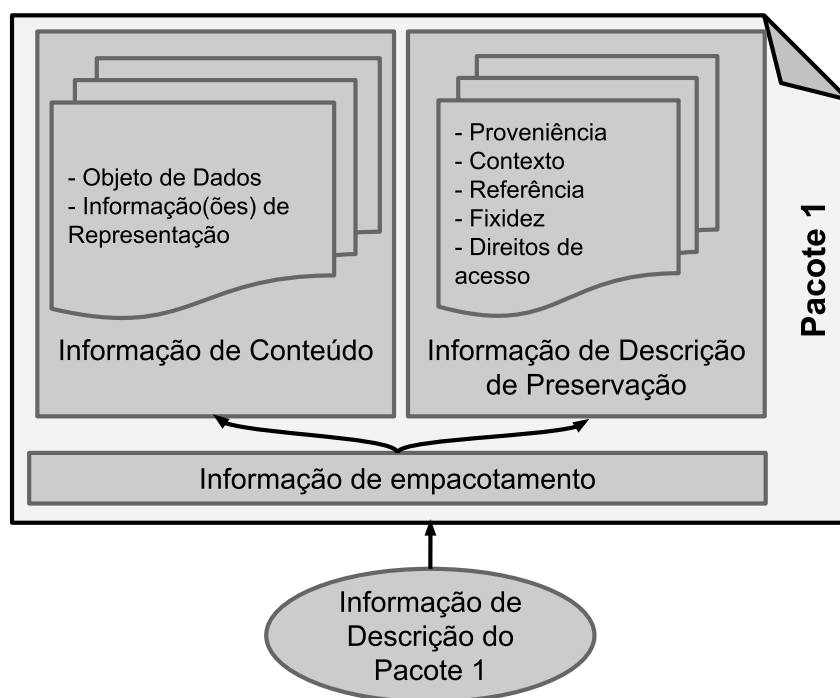
um container lógico composto de Informação de Conteúdo opcional e Descrição de Preservação opcional associada. Associada a esse Pacote de Informação, existe a Informação de Empacotamento utilizada para delimitar e identificar a Informação de Conteúdo e a informação de Descritor de Pacote, utilizada para as buscas da Informação de Conteúdo.

Santos e Flores resume a definição de Pacote de Informação, seu fluxo e as transformações realizadas no arquivo OAIS (2019, p. 119):

Os pacotes de informação consistem em estruturas de dados que o OAIS utiliza para transportar os documentos digitais e suas respectivas informações relacionadas, por meio de fluxos de informação definidos previamente no modelo funcional. Os pacotes de informação são utilizados na submissão realizada pelos produtores, na custódia realizada pelo próprio OAIS, e são disponibilizados para acesso ao consumidor. Logo, um mesmo pacote de informação será transformado para sua devida finalidade: submissão, preservação ou acesso.

A Figura 8 exibe o diagrama dos componentes de um Pacote de Informação e seus relacionamentos utilizando um pacote como exemplo. No interior do Pacote 1, encontram-se as entidades Informação de Conteúdo (*Content Information*) e Informação de Descrição de Preservação (*Preservation Description Information – PDI*). A Informação de Conteúdo representa o foco da preservação digital, e acomoda internamente o Objeto de Dados e sua(s) Informação(ões) de Representação. A PDI possui essencialmente metadados com as descrições do objeto a ser preservado, contendo cinco tipos de informações: proveniência, contexto, referência, fixidez e direitos de acesso.

Figura 8 – Estrutura do Pacote de Informação e seus relacionamentos.



Fonte: elaboração própria a partir de (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2021).

As informações de **proveniência** detalham a origem da Informação de Conteúdo, sua história e seus custodiadores. As informações de **contexto** descrevem o ambiente no qual a Informação de Conteúdo foi produzida e seus relacionamentos com outras informações externas ao próprio pacote. Utilizando-se como exemplo um documento arquivístico, o contexto pode conter detalhes sobre a sua posição no Plano de Classificação de Documentos institucional. As informações de **referência** possuem os identificadores únicos da Informação de Conteúdo, como por exemplo, DOI, ISBN, sendo utilizada para diferenciar Informações de Conteúdo entre si. A **fixidez** se responsabiliza por blindar a Informação de Conteúdo de alterações não autorizadas e não documentadas, utilizando cálculos como *checksums* para essas verificações. Os **direitos de acesso** contemplam informação a respeito das licenças, direitos autorais e permissões de uso e acesso da Informação de Conteúdo (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2021).

Sobre os pacotes de informação, Lavoie e Gartner, resumem as suas estruturas e papéis (2005, p. 9, tradução nossa):

O modelo de informações OAIS é uma conceituação dos objetos de informação recebidos, armazenados e disseminados por um repositório de preservação digital. O conceito central subjacente a esse modelo é o de um pacote de informações – uma combinação de algum conteúdo que é o foco da preservação, juntamente com seus metadados associados.

A Figura 9 exibe o modelo funcional de entidades de um arquivo OAIS. Sendo a comunicação realizada por pacotes, o **Produtor**, quando manifesta interesse em enviar um material

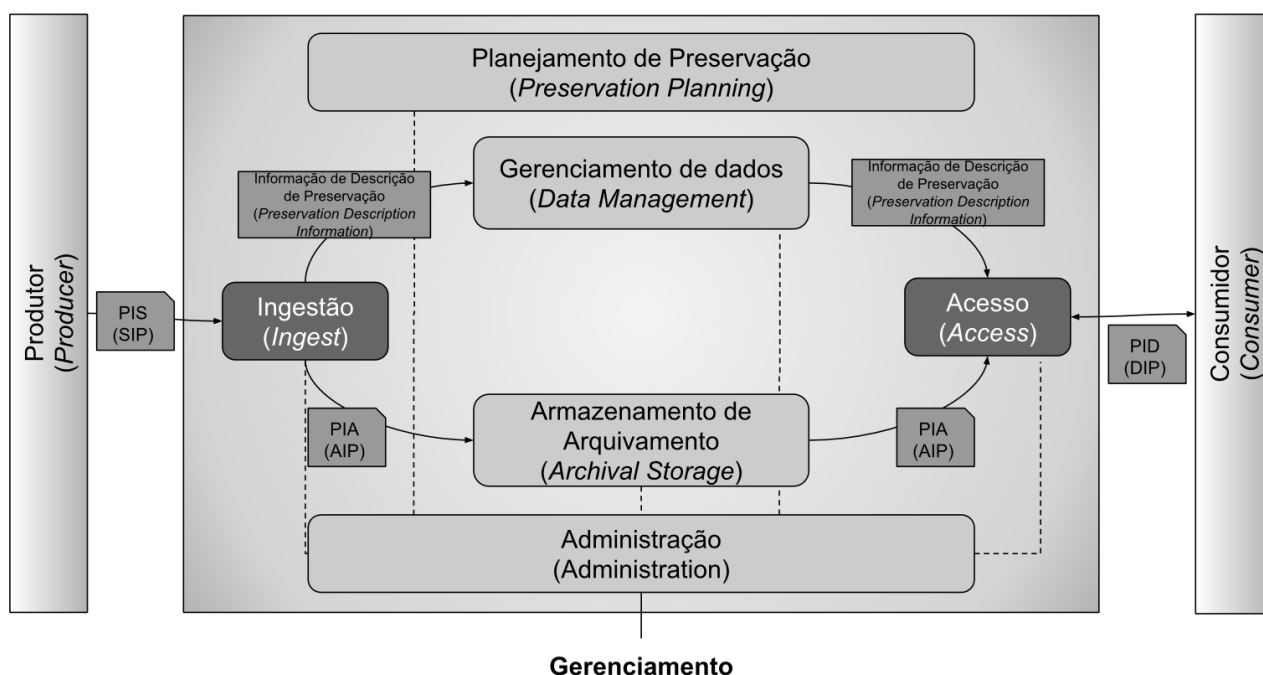
ao arquivo, o faz através do envio de um Pacote de Informação de Submissão (*Submission Information Package* – SIP), contendo o material e todas as partes que o compõe, assim como os metadados que descrevem a sua estrutura e a sua história de vida, formando os registros da sua proveniência.

A entidade funcional de **Ingestão** (*Ingest*) executa rotinas para verificação de integridade e conformidade do material com a política de preservação, identificando e validando os formatos de arquivo e os metadados recebidos. No caso de serem identificados problemas que não estejam de acordo com as regras especificadas, o SIP pode ser rejeitado e o Produtor é notificado.

Após serem realizados os procedimentos de ingestão, são enviadas as Informações de Descrição de Preservação (*Preservation Description Information* – PDI) para a entidade **Gerenciamento de Dados** (*Data Management*), que normalmente utiliza bancos de dados internos que permitem pesquisas, localizações e manutenções dos metadados dos materiais preservados. A entidade desempenha um papel crucial na organização do arquivo e preservação dos objetos digitais, gerenciando os seus metadados descritivos, administrativos e de preservação. Dentre essas informações, pode-se citar título, autor, data de criação, formato de arquivo, histórico de versões e direitos de acesso. Esses metadados são essenciais para identificar, descrever e contextualizar os objetos digitais.

Tão importante quanto a preservação dos objetos digitais é a preservação dos seus metadados para que eles permaneçam acessíveis, interpretáveis e utilizáveis ao longo do tempo. Para isso, a entidade se envolve na implementação de estratégias de backup e armazenamento redundante, bem como a adoção de práticas de migração e conversão de formatos para garantir a acessibilidade contínua dos metadados, independentemente das mudanças tecnológicas.

Figura 9 – Modelo funcional de entidades de um arquivo OAIS.



Fonte: elaboração própria a partir de (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2021).

A entidade ainda realiza a organização e indexação dos objetos digitais a partir dos seus metadados, facilitando a descoberta e recuperação eficiente quando necessário. A indexação pode envolver a criação de índices por palavras-chave, classificação em categorias ou a utilização de sistemas de gerenciamento de conteúdo, por exemplo. Cabe à entidade ainda o controle de versões dos objetos digitais, especialmente quando são atualizados ou modificados. Permitindo o rastreamento de mudanças ao longo do tempo e a recuperação de versões anteriores, se necessário. Esses históricos de modificações e controles abastecem a realização de auditorias e verificações de conformidade realizadas sobre o arquivo.

Um dos produtos gerados pela entidade Ingestão é o Pacote de Informação de Arquivamento (PIA) (*Archival Information Package – AIP*) que é enviado à entidade funcional **Armazenamento de Arquivamento** (*Archival Storage*), responsável por garantir o armazenamento seguro e confiável dos objetos digitais preservados no arquivo. Santos e Flores fornecem uma visão sobre a transformação do SIP (2020, p. 28):

Logo, os documentos arquivísticos digitais e os metadados contidos no SIP serão reorganizados na forma de um AIP, e receberão uma informação descritiva oriunda desse processo. Tal informação é essencial para manutenção da autenticidade visto que toda a alteração nos documentos digitais deverá ser registrada para garantir que não foi realizada de forma arbitrária.

A entidade Armazenamento de Arquivo trata do armazenamento dos AIP nas mídias utilizadas pelos servidores, criando hierarquias e estruturas que a permitam realizar operações

periódicas como varreduras de vírus e verificações de integridade. De um modo geral, são os AIP que viabilizam a preservação autêntica dos materiais digitais a longo prazo, garantindo a manutenção da proveniência, organicidade e registro dos seus históricos de custódia (SANTOS e FLORES, 2020).

Essa entidade tem ainda a responsabilidade de monitorar e prevenir desastres ligados à corrupção de dados, mídias, meios de armazenamento do arquivo e acessos indevidos, utilizando estratégias como replicação de dados (*clusters*, *storages* e RAID, por exemplo), *checksums* e permissões de usuários nos sistemas de arquivo. Para isso, requer a implementação de uma infraestrutura adequada para suportar o armazenamento dos objetos digitais, que pode incluir a utilização de sistemas de armazenamento em disco, fita magnética, servidores em nuvem ou outras tecnologias de armazenamento em massa. Essa infraestrutura deve ser dimensionada para atender aos requisitos de capacidade e desempenho do arquivo.

A segurança dos objetos digitais pode ainda requerer que se estabeleçam mecanismos de redundância e resiliência, tais como a replicação dos dados em locais geograficamente distintos, a utilização de sistemas de armazenamento redundantes e a implementação de técnicas de recuperação de falhas para garantir a continuidade do acesso aos objetos digitais, mesmo em caso de falhas do sistema. Em conjunto, o arquivo deve ainda possuir de mecanismos de recuperação de desastres que apoiem as atividades de entidade, além necessitar de realizações periódicas de testes de recuperação para garantir a eficácia desses procedimentos.

A entidade funcional de **Planejamento de Preservação** (*Preservation Planning*) é responsável por desenvolver estratégias e planos para garantir a preservação de longo prazo dos objetos digitais armazenados em no arquivo. A entidade tem a responsabilidade de identificar e avaliar os riscos que podem afetar a integridade, autenticidade e acessibilidade dos objetos digitais ao longo do tempo. Isso inclui riscos tecnológicos, como obsolescência de formatos e mídias, bem como riscos operacionais e ambientais, como desastres naturais e falhas de infraestrutura. A partir da avaliação de riscos, a entidade elabora estratégias adequadas para mitigar esses riscos e garantir a preservação contínua dos objetos digitais. Isso pode envolver a seleção de formatos de arquivo estáveis e abertos, a implementação de práticas de armazenamento redundantes e a definição de políticas de migração e conversão de formatos.

A entidade ainda é responsável por definir ações específicas como a realização de migrações periódicas de formatos, a criação de cópias de segurança em locais geograficamente distintos e a implementação de mecanismos de detecção e correção de erros. O monitoramento contínuo do ambiente de preservação é outra das suas responsabilidades, avaliando se as estratégias e ações definidas estão sendo eficazes. Esse processo envolve a revisão regular das políticas e

procedimentos de preservação, a realização de auditorias de conformidade e a atualização das estratégias de acordo com os avanços tecnológicos e as mudanças nas necessidades dos objetos digitais preservados.

A entidade funcional de **Administração** (*Administration*) é responsável por supervisionar e gerenciar as operações gerais do arquivo através da coordenação de atividades, desempenhando um papel central na sua governança. A entidade realiza a orquestração de funcionamento e comunicação de todas as demais entidades funcionais do repositório, coordenando-as para a realização da preservação dos seus materiais. A Administração também coordena a interação com Produtores e Consumidores de informação, garantindo a conformidade com as políticas estabelecidas.

A entidade ainda possui a atribuição de monitoramento do desempenho e avaliação da eficácia do repositório. Essa realização acontece através do monitoramento do cumprimento das políticas, avaliação da qualidade dos serviços prestados, análise de métricas e indicadores de desempenho e a identificação de áreas de melhoria. Com base nessa avaliação, a Administração pode tomar medidas corretivas e implementar aprimoramentos no arquivo.

A entidade funcional de **Acesso** é responsável por facilitar e fornecer acesso aos objetos digitais preservados no arquivo, desempenhando papel fundamental em garantir que os usuários finais possam encontrar, recuperar e usar os objetos digitais preservados. A interface de comunicação com os consumidores dos materiais pode oferecer métodos de pesquisa e navegação que permitam aos usuários a localização dos materiais de interesse. A forma com que os usuários se comunicam com essa entidade pode ser realizada através de uma interface gráfica amigável ou através de uma API, que fornece os métodos de acesso, pesquisa e localização para serem utilizados por ferramentas externas ao repositório.

A partir das requisições recebidas dos Consumidores, a entidade se comunica com a entidade Gerenciamento de Dados para a realização das pesquisas, que as registra para controle e retorna as PDI referentes à consulta. A entidade então recebe da entidade Arquivamento Arquivístico os AIP correspondentes às PDI. A partir dos AIP, ocorre a geração de Pacotes de Informação de Disseminação (*Dissemination Information Package* – DIP) para serem enviados aos Consumidores como resposta às requisições. O DIP possui uma cópia do documento preservado, geralmente otimizada para acesso e em formato compreensível para a comunidade, assim como um subconjunto de metadados. A entidade **Consumidor** pode representar tanto pessoas usuárias do sistema realizando pesquisas de interesse sobre os materiais, quanto plataformas de acesso e difusão que requisitam conjunto de dados para disponibilizados à Comunidade-Alvo do repositório.

Ainda, é possível arranjar arquivos entre si, referindo-se ao conceito de cooperação de arquivos. No cenário mais básico, os DIP de um são enviados a outro repositório para ingestão, que os transforma em SIP. No caso de os arquivos serem cópias fiéis entre si, a implementação desse cenário favorece a disponibilidade, garantindo que um arquivo funcione no lugar de outro no caso de acontecimento de falhas de cessação de funcionamento. Outros cenários de cooperação de arquivos ainda são possíveis como a federação, que pode melhorar o desempenho e resposta aos usuários, além de integrar os acervos entre si. Outro cenário existente é o compartilhamento de funções entre arquivos diferentes, que pode visar o aproveitamento de recursos e estruturas existentes entre arquivos ou que ainda podem ser consequência de hierarquias organizacionais (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2015, p. 99).

A entidade ainda se responsabiliza pelo controle de acesso aos materiais do repositório, garantindo a segurança e a privacidade dos objetos digitais e estabelecendo controles adequados de acesso. Essas finalidades podem envolver a implementação de autenticação de usuários, autorização de níveis de acesso, restrição de conteúdo sensível e registro de atividades de usuários para fins de auditoria e conformidade.

Como a entidade realiza a comunicação com os consumidores dos materiais preservados no arquivo, ela também desempenha um papel importante na divulgação e promoção do próprio repositório, uma vez que é através das suas interfaces que os materiais são disponibilizados. Assim, o seu funcionamento está diretamente ligado às políticas de acesso desenvolvidas e implantadas pelo repositório, podendo se interligar a plataformas de difusão específicas. Cabe destacar que em determinados repositórios, pode haver cobranças de taxas para acessos a materiais ou a formatos específicos, com maior qualidade por exemplo, desses materiais, adicionando essas atribuições de controle à entidade Acesso.

Exterior ao arquivo, tem-se o Gerenciamento, que representa o(s) órgão(s) que efetivamente operam, e administram o arquivo como um todo, relacionando-se à estrutura e equipes da organização mantenedora do arquivo.

O modelo de referência OAIS apresenta-se como um passo inicial para o projeto de um arquivo de uma organização, sendo totalmente independente de tecnologias, linguagens de programação e sistemas operacionais. A partir do modelo, arquivistas, bibliotecários, museólogos, analistas de TI e demais profissionais podem formar consenso sobre como disponibilizar as funcionalidades e interfaces previstas para cada um dos papéis envolvidos, de forma a desenvolver e executar a preservação digital a longo prazo.

Cabe ressaltar que um repositório OAIS representa muito mais que uma ferramenta tecnológica que se comunica por pacotes e realiza operações de preservação. Um dos passos iniciais

no estabelecimento de um repositório é a definição de sua missão e visão, que norteiam a criação das políticas institucionais que definem o que deve ser preservado, para quem preservar e quem vai preservar. Para isso, as instituições devem prever estruturas organizacional, física e tecnológica com definições de atribuições e responsabilidades dos papéis envolvidos, sendo guiados sempre pelas políticas estabelecidas.

Arellano e Oliveira, sintetizam o processo de estabelecimento organizacional de um repositório (2016, p. 469): “Os aspectos organizacionais da preservação digital são mais desafiadores que os assuntos técnicos, por isso muitos membros da comunidade têm se centrado nesses aspectos”. Granger cria a definição de *Deep Infrastructure* (Infraestrutura Profunda) que conecta e estabelece as relações entre a infraestrutura tecnológica, organização, princípios legais e contexto cultural dos usuários do repositório, relatando um certo temor de que se foque apenas na parte tecnológica e que as soluções adotadas sejam mais voltadas aos interesses de empresas e fornecedores de softwares, do que aos usuários do próprio repositório (2002).

Ainda que o modelo OAIS tenha sido projetado para a preservação digital a longo prazo, o caso dos documentos permanentes a serem preservados para sempre torna o horizonte de tempo muito amplo e distante, fazendo com que seja praticamente impossível prever todas as possíveis mudanças tecnológicas que ainda estão por vir. Haja vistas às transformações a nível tecnológico e de sociedade ocorridas nos últimos cinquenta anos, por exemplo. No entanto, o modelo OAIS possui entidades com responsabilidades bem definidas que devem monitorar constantemente o ambiente tecnológico visando se antecipar a problemas de obsolescência tecnológica, seja no âmbito de suportes de armazenamento ou de formatos de arquivo. Esse monitoramento ativo pode ainda levar a evoluções do próprio modelo OAIS, buscando se adaptar aos cenários que vão surgindo ao longo do tempo.

Nos casos de repositórios governamentais ou de órgãos ligados ao governo como as universidades públicas, por exemplo, as constantes mudanças de governos levantam a necessidade de se traçar estratégias de preservação que perpassem a duração dos mandatos políticos, fazendo com que os arquivos se tornem o mais distante possível das políticas envolvidas. No entanto, esses arquivos devem se manter vigilantes quanto ao surgimento de governos com regimes ou vieses ditatoriais que podem fazer com que documentos arquivísticos, publicações e dados científicos sejam adulterados a favor de seus governantes ou ainda eliminados ilegalmente. A preservação dos documentos é alicerce para o pleno exercício da cidadania e garantia de direitos, ao mesmo tempo em que forma o patrimônio histórico de uma nação.

Uma vez comprovado que o modelo de referência OAIS é adequado e se estabelece como um padrão para a realização da preservação digital, cria-se a necessidade de haver formas e

mecanismos de verificação de conformidade dos repositórios com o modelo de referência. O estabelecimento de padrões de certificações e auditorias permitem a determinação do grau de aderência de um determinado repositório ao modelo, o que traz confiança à sua Comunidade-Alvo, apoiadores e financiadores. É nesse contexto que surge a ISO 16363, também conhecida como ACTDR (*Audit and Certification of Trustworthy Digital Repositories*) como método e ferramenta de auditoria dos repositórios, normalizando essa tarefa a nível global (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017). A norma é tratada em detalhes em capítulo posterior do presente trabalho.

2.4 EXEMPLOS DE FERRAMENTAS PARA REPOSITÓRIOS

Nessa seção, são apresentadas ferramentas disponíveis para que os repositórios realizem a preservação digital, tendo como objetivo apresentar a forma como são estruturados.

2.4.1 FEDORA

Dentre os primeiros repositórios surgidos, está o FEDORA (*Flexible Extensible Digital Object Repository Architecture* – Arquitetura Flexível e Extensível de Repositório de Objetos Digitais) iniciado como um projeto de pesquisa da Cornell University em 1997 com o intuito de possuir uma arquitetura modular *open source* sobre a qual são adicionados serviços e módulos que combinados e integrados fornecem funções de um repositório de objetos digitais. O repositório se destina a ser usado por arquivos, bibliotecas, museus, instituições de memória, entre outros. Payette e Lagoze (1998, p. 41, tradução nossa), autores do projeto inicial, sintetizam a sua criação:

Um requisito fundamental de uma arquitetura aberta para bibliotecas digitais é um meio confiável e seguro para armazenar e acessar conteúdo digital. FEDORA é uma arquitetura de repositório e objeto digital projetado para atender a esses requisitos, ao mesmo tempo em que fornece extensibilidade e interoperabilidade. As suas principais características são: 1) suporte a tipos de dados heterogêneos; 2) acomodação de novos tipos à medida que surgem; 3) agregação de dados mistos, possivelmente distribuídos, em objetos complexos; 4) capacidade de especificar várias difusões de conteúdo desses objetos, e; 5) capacidade de associar esquemas de gerenciamento de direitos a essas difusões

Em 2009, o projeto passou a ser responsabilidade para a empresa DuraSpace, sob os cuidados da sua divisão que cuida de projetos apoiados pela comunidade. Atualmente, o repositório se encontra na versão 6.4.0 e, dentre as suas principais características, está a existência de uma interface HTTP Rest para integrações com outras ferramentas e plataformas de gestão de documentos, por exemplo. Internamente, o FEDORA utiliza o sistema de arquivos do sistema operacional para a armazenamento dos arquivos submetidos e um banco de dados para o

armazenamento dos metadados e para a realização de buscas. Assim, os objetos preservados tornam-se independentes da ferramenta, podendo ser acessados sem o seu uso em caso de falhas, por exemplo.

Ainda que o FEDORA possua diversas funcionalidades para suportar a preservação digital, os seus desenvolvedores assumem que essa responsabilidade pode ser compartilhada com diversas outras ferramentas e plataformas, daí a sua grande característica de ser modulável e extensível. O seu principal foco é a preservação dos objetos em suas mídias, possuindo mecanismos para a verificação de adulterações e geração de novas versões de arquivos no caso de migrações de formato, por exemplo. Assim, o repositório não se coloca como um concorrente de outras soluções de repositório, mas sim, uma plataforma que pode ser agregada para complementar funcionalidades na busca da preservação digital a longo prazo.

Internamente, o FEDORA possui serviços que se comunicam com o armazenamento persistente, suportando tanto sistemas de arquivo tradicionais como armazenamento em nuvem. Ainda, são disponibilizadas APIs de busca e gerenciamento do repositório que fornecem funcionalidades para o desenvolvimento de ferramentas de integração com o repositório, possibilitando que não somente pessoas operem o repositório, mas também outros *softwares* desenvolvidos. As escolhas de projeto utilizadas no repositório FEDORA permitem a combinação e uso de serviços personalizados de acordo com as necessidades de cada instituição, facilitando, inclusive integrações com outras ferramentas e sistemas de gestão documental.

2.4.2 Archivematica

Dentre as opções de ferramentas de repositório, ainda pode-se encontrar o *software* de código aberto Archivematica, com arquitetura modular baseada em microsserviços voltado a preservar de maneira confiável e autêntica materiais a longo prazo, estando em conformidade com o modelo funcional OAIS. Além do padrão OAIS, a ferramenta ainda implementa outros padrões de fato utilizados mundialmente como METS, PREMIS e Dublin Core. O seu funcionamento permite a ingestão de conteúdo digital, processando-o para garantir que atenda aos padrões de preservação e, em seguida, armazenando-o em um formato para preservação a longo prazo. O seu principal objetivo é (ARTEFACTUAL SYSTEMS INC, 2023, p. tradução nossa):

O objetivo do projeto Archivematica é fornecer aos arquivistas e bibliotecários com capacidade técnica e financeira limitada as ferramentas, metodologia e confiança para começar a preservar a informação digital hoje. O projeto conduziu um caso de uso completo do OAIS e uma análise de processo para sintetizar as etapas específicas e concretas que devem ser executadas para cumprir o modelo funcional OAIS, desde a Ingestão até o Acesso.

O Archivemática usa um conjunto de fluxos de trabalho automatizados (*workflows*) para gerenciar o processo de preservação podendo ser personalizados para atender a requisitos específicos de preservação. Esses *workflows* podem conter tarefas como identificação de formato de arquivo, conversão de formato, extração de metadados, controle de qualidade e armazenamento de arquivo, por exemplo.

A arquitetura interna do Archivemática possui elementos e serviços aderentes às arquiteturas definidas por Wheatley (2007) e pelo próprio modelo OAIS, ainda que com nomenclaturas diferentes. O Serviço de Armazenamento (*Storage Service*) provê uma infraestrutura de armazenamento para os objetos digitais, armazenando esses conteúdos em discos locais, em rede ou nuvem. Por tratar de questões bastante técnicas como armazenamento físico e algoritmos de checagem de integridade, por exemplo, esse serviço em geral é administrado e operacionalizado por profissionais de área de TI, devido à familiaridade com as ferramentas e termos técnicos utilizados.

O Painel de Controle (*Dashboard*) disponibiliza uma interface web para gerenciar as atividades de preservação, permitindo a ingestão de materiais, monitoramento das transferências e gerenciamento de metadados dos materiais. De maneira a facilitar e tornar mais amigável a ingestão de materiais, o Archivemática permite que sejam enviados os arquivos, indicando a localização dos respectivos metadados ou ainda preenchendo-os manualmente, iniciando os processos para a geração do SIP como, por exemplo, verificações de identificação de formato, compatibilidade de formatos, varredura por vírus e extração de metadados. A partir de então, o pacote gerado é de fato ingerido pelo repositório que inicia os processos de preservação e geração do AIP. Cabe notar que a interface é um meio de permitir as pessoas de realizarem essas tarefas manualmente, no entanto, a ferramenta disponibiliza uma API para integração de outros sistemas de informação como, por exemplo, sistemas de gestão de documentos das organizações.

O Armazenamento Arquivístico (*Archival Storage*) realiza as operações necessárias sobre o *Storage Service* de modo a garantir a integridade e dos arquivos e prover acesso aos mesmos para os outros serviços. Esse serviço permite a administração dos AIPs presentes na plataforma, disponibilizando ferramentas de verificação e gerenciamento de metadados.

O Plano de Preservação (*Preservation Planning*) permite aos administradores do repositório, usualmente arquivistas, configurar os planos e políticas de preservação, definindo requisitos de formatos de arquivos aceitos, além do comportamento a ser executado nas migrações e normalizações dos arquivos. O serviço de Administração (*Administration*) permite o gerenciamento do repositório como um todo, além de controlar os usuários e suas permissões de acesso.

O Sistema de Acesso (*Access System*) trata da geração dos DIPs a partir dos AIPs do repositório, além de realizar a conexão do repositório com ferramentas de difusão como o AtoM³¹ ou ArchivesSpace³², por exemplo. O Archivematica não disponibiliza aos usuários da comunidade uma interface para localização e acesso dos materiais preservados, porém disponibiliza integrações com ferramentas de difusão através da comunicação por DIPs, indo de encontro ao que o modelo OAIS define.

2.4.3 RODA

Como alternativa ao Archivematica, existe a ferramenta RODA (Repositório de Objetos Digitais Autênticos), originado de um projeto português entre a Universidade do Minho e o Instituto de Arquivos Nacionais/Torre do Tombo, mais tarde migrado para a Direção Geral do Livro, Arquivos e Bibliotecas (DGLAB), com a finalidade de desenvolver um repositório “[...] capaz de incorporar, descrever e dar acesso a todo o tipo de informação digital produzida no contexto da administração pública” (CORUJO, 2014).

O repositório foi projetado para ser totalmente aderente ao modelo de referência OAIS, além de agregar outros padrões como EAD³³, PREMIS³⁴ e METS³⁵ para o tratamento de metadados descritivos e técnicos. Sobre os aspectos técnicos, o repositório foi desenvolvido a partir de ferramentas de código aberto, baseando-se principalmente na estrutura provida pelo FEDORA devido a sua flexibilidade de possibilidade de combinação de módulos.

Internamente, o RODA se organiza através de serviços, dentre os quais pode-se citar o Serviço de Dados RODA (*RODA Data Service*) que utiliza serviços FEDORA para disponibilizar as funcionalidades de gestão de dados e armazenamento, ingestão e acesso aos objetos, pesquisa dos conteúdos armazenados e registro de todas as ações executadas dentro do repositório. Sobre esse serviço, criou-se a camada *RODA Core Services* (Serviços Básicos RODA) responsável por tratar o fluxo de ingestão, consultas ao repositório e demais funções administrativas do repositório. De forma a permitir a integração com outras ferramentas e plataformas, diversas funcionalidades dessa camada são disponibilizadas através de *webservices*, permitindo a interoperabilidade e ingestão de objetos, por exemplo. A interação com os usuários do repositório ocorre através da *RODA Web User Interface* (Interface Web de Usuário RODA), sendo através dela que administradores, arquivistas, produtores e consumidores realizam as operações desejadas.

³¹ Acrônimo para *Access to Memory*: <https://www.accesstomemory.org/pt-br/>

³² <https://archivesspace.org/>

³³ Encoded Archival Description

³⁴ PREservation Metadata: Implementation Strategies

³⁵ Metadata Encoding & Transmission Standard

A operacionalização da preservação dos materiais no RODA utiliza tarefas configuradas pelos administradores e arquivistas que desencadeiam ações específicas como varredura de vírus, conversão e migração de formatos de arquivos, por exemplo.

Quanto à implantação do RODA, destaca-se a possibilidade de instalação em um único servidor ou a sua distribuição em diferentes servidores, de modo a otimizar o seu desempenho e segurança, diminuindo pontos únicos de falha.

O RODA utiliza um modelo de dados baseado em três tipos de objetos: Objeto de Descrição (OD), Objeto de Representação (OR) e Objeto de Preservação (OP). Os Objetos de Descrição armazenam as descrições submetidas pelos produtores, além de formar uma hierarquia que representa o fundo de arquivo preservado. Cada OD de um documento mantém relação com um ou vários Objetos de Representação que é composto pelos arquivos que formam a representação digital do objeto. O terceiro objeto, Objeto de Representação, se relaciona aos outros objetos e tem como responsabilidade registrar as suas proveniências e histórico de ações de preservação realizadas dentro do repositório (CORUJO, 2014).

Uma das funcionalidades que o núcleo do RODA não possuía nativamente era a que se relacionava ao Plano de Preservação (*Preservation Planning*). Essa funcionalidade tem o objetivo principal de manter os objetos acessíveis e interpretáveis à comunidade do repositório, executando migrações e conversões dos formatos de arquivo de acordo com o monitoramento ativo. Para preencher essa lacuna, o RODA desenvolveu uma integração com o framework CriB (*Conversion and Recommendation of Digital Objects Format*), projeto desenvolvido na Universidade do Minho em Portugal, através de serviços para o tratamento dos arquivos ingeridos no repositório. Ramalho *et al* especificam essa integração (2008, p. 6):

Depois de ingerir os objetos, o RODA normalmente invoca o serviço de identificação de formato de arquivo fornecido pelo CriB para determinar se o objeto digital recentemente depositado está bem formado e reconhecido como estando em um dos formatos de preservação declarados na preservação política do arquivo. Se o objeto ingerido ainda não estiver em um formato de preservação aceitável, o CriB pode ser consultado para encontrar serviços de migração disponíveis capazes de realizar a normalização correspondente.

O modelo de objeto do RODA foi desenhado para garantir a integridade e autenticidade dos objetos digitais ao longo do tempo e para fornecer uma plataforma robusta e sustentável para a preservação a longo prazo do conteúdo digital. Assim, os objetos unem os conteúdos físicos dos materiais preservados, as suas descrições, suas representações e a trilha de operações de preservação realizadas pelo repositório.

2.4.4 DSpace

Outra ferramenta disponível para a preservação digital em repositórios é o DSpace, que suporta o armazenamento, gerenciamento e disseminação de conteúdo digital. O sistema é projetado com base no Modelo de Referência OAIS e fornece várias ferramentas e recursos para dar suporte à organização, descrição e descoberta de conteúdo digital. Ainda que a ferramenta se baseie no OAIS, a sua aderência e conformidade não se dá por completa, devido à ausência de uma entidade equivalente ao Plano de Preservação e pela forma como gerencia os seus objetos internamente, que não corresponde aos AIPs (DE GIUSTI, LIRA, *et al.*, 2012) (DE GIUSTI e VILLARREAL, 2018).

De natureza open source, ele foi desenvolvido a partir de um projeto em 2002 que contou com a participação do MIT Libraries e HP Labs, mais tarde, em 2007, houve a criação da DSpace Foundation como uma instituição sem fins lucrativos para o desenvolvimento do software a suporte aos usuários. Em 2009, DSpace Foundation e Fedora Commons, criaram, após diversas colaborações em conjunto, a organização DuraSpace. A partir de 2019, houve uma fusão entre DuraSpace e Lyris, tornando essa última a responsável pela liderança do desenvolvimento e da comunidade de usuários da ferramenta.

O DSpace ainda hoje é disponibilizado como um software de código aberto disponível a toda comunidade, tendo como missão “criar um *software* de código aberto superior aproveitando as habilidades de uma comunidade ativa de desenvolvedores, a energia e os insights de usuários engajados e ativos e o suporte financeiro de membros do projeto e provedores de serviços registrados” (DURASPACE, 2023).

O sistema utiliza uma interface *web* para acesso dos usuários, onde o produtor do material faz a submissão dos seus arquivos digitais. O DSpace então forma os Arquivos de Dados (*Data File*), também conhecidos como *bitstreams*, que possuem informações técnicas sobre os seus formatos que auxiliarão na preservação dos materiais. O próximo processo é a construção dos Itens (*Item*), tidos como um “átomo arquivístico”, que são compostos pelos arquivos enviados e seus metadados de descrição e técnicos. Internamente, os metadados são indexados para buscas e navegação no repositório. O agrupamento lógico dos itens forma Coleções (*Collections*). Os agrupamentos ou conjuntos de Coleções foram as Comunidades (*Community*), que são o nível mais alto da hierarquia de conteúdo do DSpace, podendo se estabelecer uma relação com as partes da organização como, por exemplo, os departamentos ou centros de ensino de uma universidade. A administração e curadoria do repositório é realizada pelos administradores do repositório que têm a responsabilidade pela organização dos materiais e executam rotinas de preservação como a migração e conversão de formatos dos arquivos (DURASPACE, 2023).

A forma modular como a arquitetura do DSpace é construída permite a criação e integração de diversos repositórios, ao mesmo tempo em que os consumidores realizam as buscas e localização dos materiais de maneira unificada, facilitando essa tarefa. Essa característica permite, por exemplo, a formação de um único repositório institucional a partir de diversos outros repositórios temáticos dentro de uma universidade. Em estudo publicado em 2017, Costa e Leite, identificaram que a DSpace é o *software* mais utilizado para repositórios institucionais na América Latina (2017, p. 161). Essa predominância do *software* também foi relatada na América do Norte, onde dentre os repositórios pesquisados da *American Research Libraries*, a grande maioria (70%) utiliza ou planeja utilizar o DSpace, segundo estudo conduzido por Bailey Jr. (2006, p. 16). Em 2018, outro estudo no âmbito dos repositórios presentes no Diretório de Repositórios de Acesso Aberto (OpenDOAR) também detectou a predominância do *software* DSpace, relatando o seu uso por 40% dos repositórios, sendo o restante dividido por outras ferramentas (ALI, LOAN e MUSHATQ, 2018, p. 216). Esse percentual se mantém praticamente o mesmo nos dias atuais: 39% (JOINT INFORMATION SYSTEMS COMMITTEE, 2023).

Ainda que muitas iniciativas individuais de instalações e uso de ferramentas de repositório tenham surgido e ainda venham a surgir, esse fato não necessariamente caracteriza o estabelecimento de um repositório, seja ele institucional, temático ou de dados de pesquisa. Há que se atentar para o fato de que um repositório requer uma política de preservação, estrutura formal, física e organizacional, operada por equipes dedicadas, qualificadas e empenhadas em realizar a preservação dos materiais. Essa estrutura requer financiamento tanto para a implantação inicial, quanto para a sua manutenção, necessitando um planejamento estruturado a fim de que seja possível garantir à Comunidade-Alvo do repositório que os materiais preservados recebem tratamentos adequados e estarão disponíveis daqui a muito tempo.

Anterior às ferramentas, é necessário desenhar processos com a definição dos papéis e responsabilidades dos agentes designados, caso contrário, corre-se o risco de se ter um *software* instalado que serve como um depósito de materiais. Sem um norte, representado pelas políticas de preservação, sem processos e sem estrutura adequada, a preservação dos materiais é levada ao sabor do vento, refletindo a (des)organização que o controla, baseada em vieses individuais. Ao mesmo tempo, o repositório como instituição, fonte de materiais autênticos, dados primários e voltado à preservação, muitas vezes eterna, deve ser superior aos governantes em mandato, subsidiando a sociedade no acesso aos materiais que permitam o exercício da cidadania.

2.5 A EXPERIÊNCIA DO BRASIL: REPOSITÓRIO ARQUIVÍSTICO DIGITAL CONFIÁVEL – RDC-ARQ

No Brasil, em 2015, o Conselho Nacional de Arquivos (CONARQ) publicou a Resolução N° 43, estabelecendo as diretrizes necessárias para a implementação de um Repositório Arquivístico Digital Confiável (RDC-Arq) (CONSELHO NACIONAL DE ARQUIVOS, 2015). Após a criação de um arcabouço teórico-prático sobre os Sistemas Informatizados de Gestão Arquivística de Documentos (SIGAD), criou-se a necessidade de definir parâmetros para a realização da preservação dos documentos arquivísticos produzidos nas organizações. Assim, o principal objetivo da Resolução N°43 é (CONSELHO NACIONAL DE ARQUIVOS, 2015, p. 5):

Indicar parâmetros para repositórios arquivísticos digitais confiáveis, de forma a garantir a autenticidade (identidade e integridade), a confidencialidade, a disponibilidade, o acesso e a preservação, tendo em vista a perspectiva da necessidade de manutenção dos acervos documentais por longos períodos de tempo ou, até mesmo, permanentemente.

A resolução se baseia fortemente na ISO 16363, que trata de certificação de repositórios baseados no modelo OAIS, porém adiciona requisitos para cobrir especificamente os documentos arquivísticos digitais, uma vez que eles possuem características próprias. Ainda que a publicação não tenha caráter de norma ou certificação, ela traz diversas concepções de outras normas internacionais já bastante alicerçadas, como a ISO 14721³⁶, ISO 16919³⁷ e a própria ISO 16363³⁸. Sendo assim, ela pode ser usada como norteadora para a implantação de sistemas que tenham o objetivo de preservar documentos e se tornarem RDC-Arqs.

A definição de um RDC-Arq se dá através da união três conceitos:

- a) Repositório digital: sistema informático utilizado para o armazenamento dos documentos, composto de *hardwares* e *softwares* específicos para o desempenho da preservação, apoiado por uma estrutura organizacional que a gerencia;
- b) Repositório arquivístico digital: gerencia os documentos e respectivos metadados aplicando procedimentos arquivísticos que visem a preservação do documento, protegendo impreterivelmente a sua identidade, integridade e a relação orgânica com demais documentos;
- c) Repositório digital confiável: tem a missão de manter os documentos sob a sua responsabilidade autênticos e acessíveis pelo tempo que for necessário, aplicando

³⁶ Modelo de Arquivo OAIS (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012)

³⁷ Requisitos para entidades certificadoras de Repositórios Digitais Confiáveis. (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2014)

³⁸ Auditoria e Certificação de Repositórios Digitais Confiáveis (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012)

todos os procedimentos adequados para esse fim. O aspecto da confiabilidade do repositório pode ser atestado através de auditorias e certificações que demonstrem à sua comunidade alvo a sua capacidade de preservação dos documentos.

Assim, tem-se que “um repositório arquivístico digital confiável deve ser capaz de atender aos procedimentos arquivísticos em suas diferentes fases e aos requisitos de um repositório digital confiável” (CONSELHO NACIONAL DE ARQUIVOS, 2015, p. 10). Para se alcançar os objetivos de preservação digital, sobre tudo à longo prazo, o RDC-Arq divide as suas responsabilidades tanto do projeto, quanto da execução e operacionalização entre profissionais de arquivos e profissionais da TI, ambos adicionando, respectivamente, a visão necessária para o tratamento arquivístico dos documentos e a disponibilização dos meios tecnológicos necessários ao repositório. Braga, Holanda e Canelhas analisam os papéis dos atores envolvidos na implementação de um RDC-Arq, enfatizando a parcela de participação da própria gestão institucional (2022, p. 6):

Ao se analisar os requisitos identificados na proposta do Conarq, a primeira conclusão obtida é que não é possível propor a implementação de um projeto de RDC-Arq sem a participação de atores provenientes de diversas áreas. Além da equipe responsável pela gestão documental e pela tecnologia da informação, a participação da alta gestão viabiliza o cumprimento de diretrizes estabelecidas no escopo do Conarq.

Cabe ressaltar que o período de preservação dos documentos pode variar de acordo com os critérios legais e definições próprias das organizações. Como o propósito da preservação é manter a autenticidade e acessibilidade dos documentos, sobretudo no caso dos documentos digitais, é fundamental manter-se um monitoramento desses aspectos que podem se corromper por problemas técnicos ou pela obsolescência tecnológica, podendo ser necessária a adoção de um RDC-Arq não só para documentos permanentes, mas também para documentos em fase corrente e intermediária. Enquanto em fase ativa ou semiativa, os documentos são usados e acessados, necessitando procedimentos de preservação, ainda que a destinação final desses documentos possa ser a eliminação.

O RDC-Arq deve ser independente do sistema de produção e captura dos documentos, muitas vezes um SIGAD, no sentido de os funcionamentos de ambos não possuírem dependência entre si para os seus funcionamentos. Há que se ter a possibilidade de acessar os documentos diretamente no RDC-Arq, criando-se assim camadas de isolamento que se comunicam através de interfaces, mas que não afetam as suas formas internas de funcionamento. A interoperabilidade do RDC-Arq, sobretudo das soluções informatizadas adotadas, é outro aspecto a ser levado em consideração, procurando-se softwares que adotem padrões estabelecidos reconhecidos em nível nacional ou internacional (CONSELHO NACIONAL DE ARQUIVOS, 2015, p. 11).

Análogo a um Arquivo OAIS, o RDC-Arq deve possuir toda a estrutura necessária ao seu funcionamento e, inclusive, possuir um planejamento no caso do cessamento do seu funcionamento. O RDC-Arq deve possuir uma estrutura organizacional, com equipes bem dimensionadas e capacitadas, utilizando-se de governança que demonstre transparência, capacidade para o cumprimento da preservação dos documentos sob sua custódia e sustentabilidade financeira. Há que se atentar aos aspectos legais em vigência, tanto internacionais quanto da esfera do RDC-Arq, de modo a não os infringir, o que poderia levar à interrupção do seu funcionamento ou ao comprometimento da autenticidade e acessibilidade dos documentos (CONSELHO NACIONAL DE ARQUIVOS, 2015, p. 12).

No que tange à comunicação com o RDC-Arq, é adotada o mesmo projeto do Arquivo OAIS, utilizando-se pacotes de informação. O envio de materiais ao repositório, tanto de um SIGAD como de outra fonte, deve utilizar o Pacote de Submissão de Informação (PSI ou SIP em inglês) que contém os documentos e todos os seus metadados. Após o recebimento do PSI, o repositório gera o Pacote de Arquivamento de Informação (PAI ou AIP em inglês) que vai registrar todas as operações executadas necessárias à preservação dos documentos.

Os aspectos práticos da preservação dos documentos, definidos pela política de preservação digital da organização, são executados pelo repositório e registrados no PAI, sendo possível se conhecer em determinado ponto do tempo quais operações foram efetuadas e quem as efetuou, corroborando para a manutenção da autenticidade dos documentos. As migrações de versões, conversões de formatos e varredura de antivírus são algumas das operações realizadas no repositório que devem ser registradas no PAI.

A externalização dos documentos do repositório utiliza o Pacote de Disseminação de Informação (PDI ou DIP em inglês) que empacota as derivadas de acesso dos documentos e seus metadados. Através da disponibilização dos pacotes PDI, o repositório cumpre a política de acesso aos documentos e permite que a Comunidade-Alvo tenha acesso ativo, ou através de requisições, aos materiais que tem interesse. Seja através de um serviço ou integrado a outras ferramentas, o repositório deve propiciar a divulgação dos documentos e seus respectivos metadados, sempre observando as restrições de acesso e direitos autorais cabíveis em cada caso.

Ainda que as diretrizes para o estabelecimento de um RDC-Arq não possuam caráter de certificação, é possível a sua utilização para a verificação de conformidade dos seus requisitos em repositórios das organizações, criando-se graus de maior ou menor aderência que podem balizar possíveis contratações ou monitorar a evolução das instituições.

Conforme observam Andrade e Chagas (2023), é possível encontrar similaridades entre os requisitos de um RDC-Arq, os critérios da certificação Nestor Seal e a norma ISO 16363. Devido

a certificação Nestor Seal possuir critérios mais amplos, uma vez que abarca diferentes contextos e áreas, o RDC-Arq possui forte visão arquivística, fazendo com que vários dos seus critérios se equivalham a um critério Nestor Seal, por exemplo. No caso da ISO 16363, ocorrem similaridades conceituais entre os critérios, porém, devido ao esforço de adaptação do RDC-Arq aos critérios da norma, sugere-se a obtenção da certificação da norma internacional, devido ao seu maior reconhecimento.

Complementarmente, é possível utilizar-se o estudo de Santos que elaborou o Manual para Auditoria de Repositórios Arquivísticos Digitais Confiáveis (2018), como ferramenta de auditoria para a verificação de conformidade dos repositórios das organizações. A obra se baseia nos requisitos do modelo OAIS³⁹ e ACTDR para comprovar que os requisitos de um RDC-Arq estão presentes e são plenamente alcançados em um determinado repositório. Cabe ressaltar que o presente estudo incorpora critérios presentes no modelo de referências OAIS e da ACTDR, da mesma forma que o Manual, porém vai além ao incluir critérios que cobrem outras ferramentas e instrumentos como o Nestor Seal⁴⁰, CoreTrust Seal⁴¹, DRAMBORA⁴² e o Guia de Avaliação de Repositórios REBIUN⁴³, conforme demonstrado na seção 4.2.2.

2.6 A CONTRIBUIÇÃO DOS REPOSITÓRIOS PARA O ACESSO ABERTO

Uma vez que os repositórios estejam estabelecidos e a preservação digital esteja ocorrendo a pleno, a difusão dos seus materiais contribui para a ampliação de acesso a esse conhecimento. No caso de repositórios que preservam e difundem materiais de pesquisa como artigos, teses e dissertações, por exemplo, a disponibilização de acesso contribui para que diversos pesquisadores, de diversas localizações, possam utilizá-los como referências e/ou bases para a elaboração de novas pesquisas, fomentando uma distribuição de conhecimento científico a nível global. Essa é a base da Iniciativa Acesso Aberto (*Open Access Initiative*), apresentada nessa seção do trabalho.

2.6.1 Os Fundamentos do Acesso Aberto

A importância da difusão de materiais pelos repositórios digitais em relação à disseminação de conhecimento científico fez surgir tratados como o *Budapest Open Access*

³⁹ (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012)

⁴⁰ (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009)

⁴¹ (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019)

⁴² (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007)

⁴³ (AZORÍN, BERNAL, *et al.*, 2021)

Initiative (BOAI) (BUDAPEST OPEN ACCESS INITIATIVE, 2002), *Bethesda Statement on Open Access Publishing* (BROWN, CABELL, *et al.*, 2003) e *Berlin Declaration on Open Access to Knowledge in the Science and Humanities* (OPEN ACCESS MOVEMENT, 2003) que buscam criar diretrizes para regulamentar e estabelecer requisitos mínimos para o acesso aberto (*open access*). Dentre as definições de acesso aberto, pode-se citar a encontrada na Declaração de Berlin (OPEN ACCESS MOVEMENT, 2003, p. 1, tradução nossa):

Estabelecer o acesso aberto como um procedimento válido requer, idealmente, o compromisso ativo de todo e qualquer indivíduo produtor de conhecimento científico e detentor de patrimônio cultural. As contribuições de acesso aberto incluem resultados de pesquisas científicas originais, dados brutos e metadados, materiais de origem, representações digitais de materiais pictóricos e gráficos e material multimídia acadêmico.

Essas iniciativas buscaram difundir o acesso livre das pesquisas realizadas a nível mundial, o que acabou gerando uma demanda pela preservação e acesso a longo prazo dos materiais através dos repositórios de dados de pesquisa (*research data repositories*). Em 2007, a OECD (*Organization for Economic Co-operation and Development*), fórum que reúne a participação de 30 países na busca por soluções para os desafios econômicos, sociais e ambientais da globalização mundial desenvolveu o *OECD Principles and Guidelines for Access to Research Data from Public Funding* (Princípios e Orientações para Acesso aos Dados de Pesquisa com Financiamento Público) no intuito de construir uma base sólida de dados de pesquisa provenientes de financiamentos públicos que possam ser compartilhados e reusados por diferentes pesquisadores ao longo do tempo. Inicialmente, o guia faz a definição de dados de pesquisa (ORGANIZATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT, 2007, p. 13, tradução nossa):

[...] são definidos como registros factuais (escalas numéricas, registros textuais, imagens e sons) usados como fontes primárias para pesquisas científicas e que são comumente aceitos na comunidade científica como necessários para validar os resultados da pesquisa. Um conjunto de dados de pesquisa constitui uma representação sistemática e parcial do assunto que está sendo investigado.

A ideia da OECD é tornar os dados de pesquisas financiadas por verbas públicas compartilháveis entre os países membros, buscando facilitar o estabelecimento de parcerias para a realização de pesquisas, além de permitir o reuso e interpretação desses dados sob óticas diferentes das que os produziram. Como reflexo imediato, o grupo espera um melhor retorno das pesquisas que utilizam verbas de seus governos, ou seja, investimento público.

Em termos de preservação dos dados, o guia estabelece que as instituições de pesquisa devem avaliar individualmente os prazos de preservação dos seus conjuntos de dados, selecionando quais conjuntos devem ser mantidos a longo prazo e quais podem ser eliminados, de acordo com grau de importância da pesquisa e potencial de utilidade para os demais pesquisadores do grupo.

Para Ray Uzwyshyn, os repositórios de dados de pesquisa podem ser definidos como (2016, p. 18, tradução nossa) :

Os repositórios de dados de pesquisa *online* são grandes infraestruturas de banco de dados criadas para gerenciar, compartilhar, acessar e arquivar conjuntos de dados de pesquisadores. Os repositórios podem ser especializados e destinados a agregar dados disciplinares ou mais gerais, coletando dados de áreas de conhecimento maiores, como ciências ou ciências sociais.

A ideia é permitir que pesquisadores depositem os dados relativos às suas pesquisas, juntamente com instruções e metadados que permitam o seu uso e interpretação para que outros pesquisadores possam utilizá-los em suas pesquisas. Adicionalmente, conforme a política do repositório, pode haver um certo grau de curadoria, revisão e validação dos dados, seja através de ferramentas automatizadas ou por processos de revisão entre pesquisadores. As ferramentas utilizadas para esses repositórios podem ainda disponibilizar uma forma de referenciar os conjuntos de dados inequivocamente nas publicações, através de identificadores persistentes (*persistent identifiers*) como o DOI ⁴⁴ ou UNF⁴⁵, garantindo o acesso diretamente aos dados usados nos estudos.

Conforme a área de conhecimento atendida pelo repositório, os arquivos submetidos aos repositórios de dados podem apresentar uma ampla gama de tipos, incluindo desde formatos de texto, planilhas de cálculo, dados geoespaciais, até formatos muito específicos de programas utilizados pelos pesquisadores. Esse grande leque de possibilidades de formatos trazido pela Comunidade-Alvo do repositório pode acarretar uma grande complexidade no projeto e implementação dos processos de preservação, uma vez que pode ser necessário e obrigatório tratar formatos proprietários, nem sempre apropriados à própria preservação.

Assim como a OECD, que justifica que os dados de pesquisas financiados por verbas públicas devem ser publicados, nos Estados Unidos, o *Office of Science and Technology Policy* (OSTP) (Gabinete de Política Científica e Tecnológica) publicou em 2013 um memorando com o objetivo de melhorar e aumentar o acesso aos resultados de pesquisas financiadas por verbas federais. Está entre os princípios do Memorando a seguinte declaração (OFFICE OF SCIENCE AND TECHNOLOGY POLICY, 2013, p. 1, tradução nossa):

A Administração está empenhada em garantir que, na maior extensão e com o menor número possível de restrições e de acordo com a lei e os objetivos estabelecidos abaixo, os resultados diretos da pesquisa científica financiada pelo governo federal sejam disponibilizados e úteis para o público, indústria e a comunidade científica. Tais resultados incluem publicações revisadas por pares e dados digitais. [...] O acesso a conjuntos de dados digitais resultantes de pesquisas financiadas pelo governo federal permite que as empresas concentrem recursos e esforços na compreensão e exploração de descobertas. [...] Além disso, a disponibilidade mais ampla de publicações revisadas por pares e de dados

⁴⁴ Digital Object Identifier

⁴⁵ Universal Numeric Fingerprint

científicos em formatos digitais criará mercados econômicos inovadores para serviços relacionados à curadoria, preservação, análise e visualização.

Essa decisão do governo teve como objetivo exigir das agências de pesquisa a criação de planos de acesso aos resultados de suas pesquisas bem como dos dados usados como fontes primárias de informação, no sentido de ampliar o acesso ao público dessas informações e maximizar o impacto das pesquisas financiadas por verbas federais. Para a implementação do acesso aberto das publicações e dados de pesquisa, as agências de pesquisa devem contar com repositórios para preservação a longo prazo das publicações e seus metadados, sem cobranças de taxas ou impostos, salvas condições protegidas por leis específicas.

Um aspecto interessante desse Memorando foi o estabelecimento de um período de 12 meses após a publicação dos resultados das pesquisas, denominado como embargo, para disponibilização de acesso livre a esses materiais. O prazo desse embargo poderia variar de acordo com cada plano de acesso das instituições podendo, inclusive, ser suspenso a partir de uma petição com justificativas aplicáveis a cada caso e domínio de pesquisa.

Em 2022, o governo americano, através do OSTP, publicou outro memorando com objetivo de reafirmar o compromisso das agências de pesquisa na garantia de acesso livre e imediato dos resultados de pesquisas financiadas por verbas federais. Dentre as primeiras recomendações do Memorando está a remoção do prazo de embargo estabelecido no Memorando antecessor de 2013, tornando as publicações imediatamente disponíveis após as suas finalizações (OFFICE OF SCIENCE AND TECHNOLOGY POLICY, 2022, p. 2, tradução nossa):

Anos de feedback público indicaram que a principal limitação do Memorando de 2013 é o embargo opcional de 12 meses ao acesso público de qualquer publicação resultante de pesquisa financiada pelo governo federal. Esta disposição limitou o acesso imediato aos resultados da pesquisa financiada pelo governo federal apenas para aqueles que podem pagar por ela ou que têm acesso privilegiado por meio de bibliotecas ou outras instituições. Meios financeiros e acesso privilegiado nunca devem ser os pré-requisitos para obter os benefícios da pesquisa financiada pelo governo federal que o público americano merece.

Como reforço à disponibilização livre e imediata das publicações, o Memorando cita as lições aprendidas com a epidemia de COVID-19 quando diversas organizações científicas a nível global se uniram na disponibilização pública e imediata para compartilhamento de seus achados. Essa decisão possibilitou grandes avanços da ciência, em especial na área da saúde, ao que o Memorando conclui (OFFICE OF SCIENCE AND TECHNOLOGY POLICY, 2022, p. 2, tradução nossa): “os *insights* de pesquisas novas e de ponta decorrentes do apoio de agências federais devem estar imediatamente disponíveis – não apenas em momentos de crise, mas em todos os momentos”

O Memorando reconhece o grande avanço promovido pelos planos de acesso das instituições desde 2013, contabilizando mais de 8 milhões de publicações acadêmicas acessíveis

publicamente e uma estatística de mais de 3 milhões de acesso a essas publicações diariamente, o que levou a uma mudança de paradigma na qual o conhecimento estava circunscrito a silos de pesquisa e então passou para uma cultura de colaboração e compartilhamento de dados de pesquisa na comunidade científica. Assim como o Memorando de 2013, o Memorando de 2022 decreta que as instituições devem possuir repositórios para depósito, preservação e acesso das publicações e dos dados que apoiaram as respectivas pesquisas, no intuito de promover o acesso e fortalecer a confiança pública na ciência financiada pelo governo federal.

Uzwyshyn entende que a disponibilização de acesso às pesquisas financiadas com verbas públicas representa uma obrigação dos pesquisadores e suas agências, no sentido de prestarem contas à sociedade pagadora de seus impostos (2016, p. 19, tradução nossa):

O raciocínio é que, se um pesquisador está tirando proveito dos contribuintes públicos, a pesquisa deve ser acessível ao público por meio do artigo e dos dados originais. Compartilhar esses dados ajuda a manter vital a economia em geral, facilitando uma competição saudável para a comercialização e disseminação da descoberta. Se os pesquisadores não tiverem planos de gerenciamento de dados, sua chance de obter uma bolsa diminui.

À medida em que os dados de pesquisa são depositados nos repositórios, cresce a responsabilidade por manter os materiais autênticos e interpretáveis ao longo do tempo, a fim de sustentar os resultados já encontrados, o que os caracteriza como evidência dos estudos. O caráter evidencial desses dados pode ser considerado correspondente aos documentos, que são considerados evidências dos fatos que os seus conteúdos sustentam, daí a necessidade de se utilizar repositórios com processos de preservação a longo prazo que garantam a autenticidade, acesso e interpretabilidade dos materiais custodiados.

Uma vez que sejam identificadas falhas que coloquem a autenticidade dos dados custodiados à prova, pode ocorrer o fenômeno de invalidação em cascata de todas as pesquisas tal como um castelo de cartas que desmorona, anulando os estudos baseados diretamente nesses dados comprometidos ou em estudos derivados. Os impactos podem ser imensuráveis sendo tanto materiais, quando as verbas destinadas aos estudos se tornam desperdiçadas, quanto imateriais, causando a perda de credibilidade em pesquisadores e instituições de pesquisa como um todo. Se faz necessária assim, a implementação de repositórios confiáveis que possuam estruturas, equipes e verbas adequadas aos seus funcionamentos, além de contarem com certificações formais ou informais, aplicações de modelos de maturidade e demais ferramentas que atestem as suas capacidades de realizar uma preservação confiável.

2.6.2 Iniciativas Institucionais

Dentre as estratégias para a promoção de acesso, impulsionamento e difusão dos conhecimentos preservados nos repositórios, foram criadas ao longo dos anos ferramentas que unificam os repositórios institucionais e facilitam a busca e localização de estudos realizados. Dentre essas iniciativas, está o *Directory of Open Access Repositories* (OpenDOAR) (Diretório de Repositórios de Acesso Aberto), criado a partir de um projeto iniciado na Universidade de Nottingham, Inglaterra.

A ferramenta busca criar um índice de repositórios que disponibilizam materiais científicos de acesso aberto a nível mundial, estabelecendo critérios mínimos para a inclusão dos repositórios, além de contar com um processo de revisão a partir de seus membros e da comunidade para a garantia de qualidade. Dentre esses critérios, pode-se citar a obrigatoriedade em disponibilizar os materiais de forma livre, sem cobranças de taxas e sem a necessidade de realização de registros e cadastros dos usuários. Os materiais dos repositórios listados no OpenDOAR podem ser de natureza científica como artigos, livros e teses, por exemplo, ou documentos arquivísticos das instituições, desde que possam ser reutilizados pelos usuários.

Quando iniciou o seu funcionamento, em 2005, o OpenDOAR possuía 75 repositórios listados. Atualmente, após 18 anos de atividades, são listados ao todo 6.027 repositórios, sendo os Estados Unidos da América o país com mais repositórios cadastrados na plataforma. Dentre os tipos de conteúdo fornecidos pelos repositórios, os artigos de revistas científicas são o material predominante (JOINT INFORMATION SYSTEMS COMMITTEE, 2023).

Com objetivos semelhantes à OpenDOAR, criou-se a COAR (*Confederation of Open Access Repositories* – Confederação de Repositórios de Acesso Aberto), fruto da evolução e ampliação do projeto DRIVER iniciado em 2009 e que buscava conectar somente os repositórios europeus. A missão do COAR é “aumentar a visibilidade e a aplicação dos resultados da pesquisa por meio da colaboração em toda a rede global de repositórios” (CONFEDERATION OF OPEN ACCESS REPOSITORIES, 2023, p. 3, tradução nossa). A organização sem fins lucrativos conta atualmente com mais de 150 participantes a nível mundial, sendo 1 do Brasil e 5 da Espanha, representando bibliotecas, universidades, instituições de pesquisa, entre outros tipos de repositórios, buscando agir como uma “voz global para a comunidade de repositórios” e em conjunto desenvolver as suas cinco direções estratégicas (CONFEDERATION OF OPEN ACCESS REPOSITORIES, 2023, p. 4, tradução nossa):

- 1) defender o importante papel de uma rede internacional distribuída de repositórios abertos no avanço da ciência aberta e da bibliodiversidade; 2) fornecer suporte para fortalecer e modernizar repositórios e redes de repositório; 3) promover alinhamento e

interoperabilidade entre repositórios, redes de repositórios e entre repositórios e outros sistemas e serviços; 4) definir e promover a adoção de novos comportamentos, tecnologias e funções para repositórios e sua interação com outros serviços de valor agregado, e; 5) aumentar a sustentabilidade e eficácia da organização e fortalecer a marca COAR.

A partir dessas estratégias, a COAR desenvolve ações de pesquisas sobre repositórios, desenvolve e publica boas práticas e promove eventos com o intuito de difundir o universo de conceitos, ferramentas e usos de repositórios entre os membros. A promoção da interoperabilidade entre os repositórios também é uma característica da COAR através do estabelecimento de vocabulários de termos comuns e padrão de metadados dos repositórios, possibilitando e facilitando a integração dos repositórios das organizações participantes da confederação. Essas ações, além de contribuir para o Acesso Aberto, também se relacionam com os princípios FAIR⁴⁶ tornando os materiais Localizáveis (*Findability*) e Interoperáveis (*Interoperability*).

Adicionalmente, a COAR disponibiliza um kit de ferramentas (*toolkit*) que apresenta aos administradores de repositórios um guia para a implementação das boas práticas sugeridas, tendo como seu principal objetivo (CONFEDERATION OF OPEN ACCESS REPOSITORIES, 2019, p. tradução nossa):

O objetivo do kit de ferramentas é fornecer aos gerentes de repositório as melhores práticas e recursos educacionais para apoiar a interoperabilidade, descoberta e desenvolvimento de serviços de valor agregado. O kit de ferramentas fornece acesso a recursos relacionados à função de repositórios, descoberta e interoperabilidade, repositórios de próxima geração e contém links para informações técnicas para implementação e gerenciamento de plataformas de repositório.

Na Espanha, a Fundação Espanhola para a Ciência e Tecnologia (FECYT) em conjunto com a Rede de Bibliotecas Universitárias (REBIUN), criou em 2007 o RECOLECTA (*Recolector de Ciencia Abierta*) como uma iniciativa que visa promover e apoiar práticas de ciência aberta no país. Ele atua como uma infraestrutura nacional para a coordenação, desenvolvimento e implementação de políticas e práticas de acesso aberto na pesquisa e inovação (FUNDACIÓN ESPAÑOLA PARA LA CIENCIA Y LA TECNOLOGÍA, 2023).

O RECOLECTA é uma plataforma que facilita a coleta e agregação de resultados de pesquisa de acesso aberto produzidos por instituições espanholas coletando metadados de repositórios institucionais, revistas e outras fontes e facilitando a descoberta e acesso de conteúdo de acesso aberto para pesquisadores, formuladores de políticas e o público em geral.

Além de promover e aumentar a visibilidade e o impacto da pesquisa nacional, a plataforma também apoia a conformidade com políticas e iniciativas nacionais e internacionais sobre a questão do Acesso Aberto e Ciência Aberta. Os repositórios participantes devem

⁴⁶ (WILKINSON, DUMONTIER, *et al.*, 2016)

implementar conjuntos de metadados padronizados, além de disponibilizar serviços de colheita (*harvesting*) de metadados utilizando o protocolo OAI-PMH. Assim, através de um ponto único de acesso, pode-se pesquisar publicações científicas e demais materiais que estejam inseridos nos repositórios individuais das organizações, agregando ao todo 176 repositórios nacionais, atualmente.

O protocolo OAI-PMH (*Open Archives Initiative Protocol for Metadata Harvesting* – Protocolo de Iniciativa de Arquivos Abertos para Colheita de Metadados) é amplamente adotado para compartilhar e colher metadados de repositórios de forma padronizada facilitando a interoperabilidade entre repositórios digitais ao permitir a troca de metadados entre eles (OPEN ARCHIVES INITIATIVE, 2023).

A sua operação se dá através de um cliente (coletor) que envia solicitações para um servidor (provedor de dados) usando protocolos HTTP. O servidor responde com registros de metadados que seguem um esquema XML específico, geralmente no formato Dublin Core ou outros formatos de metadados. Dentre as vantagens da sua utilização está a colheita eficiente de metadados, permitindo que os repositórios os compartilhem esforço mínimo, facilitando a agregação de várias fontes e criação de coleções abrangentes provenientes de vários repositórios. Como reflexo, o OAI-PMH promove a interoperabilidade de dados ao impor um formato padrão para os registros de metadados. Isso torna mais fácil para diferentes sistemas e serviços entender e processar os metadados colhidos de forma consistente.

De um modo geral, o OAI-PMH desempenha um papel crucial no campo de bibliotecas digitais e repositórios institucionais, tornando-se um padrão amplamente adotado para facilitar a troca de metadados, promover a interoperabilidade e apoiar o desenvolvimento de serviços inovadores baseados em coleções de metadados agregados.

Ainda que essas iniciativas de acesso aberto ou coletores não tratem especificamente sobre a preservação, elas dão conta sobre a forma como os repositórios se comunicam e interoperam, promovendo a difusão e ampliação dos materiais preservados nos repositórios. A operacionalização da preservação dos materiais de cada organização deve ser implementada internamente, porém ao incluir a difusão dos materiais na missão dos repositórios, vislumbra-se uma ampliação das suas comunidades-alvo, alcançando-se públicos de proporções mundiais, muitas vezes. Como reflexo, os repositórios ganham a uniformização dos seus metadados.

3 AVALIAR E CERTIFICAR A CONFIABILIDADE NOS REPOSITÓRIOS DIGITAIS

3.1 INTRODUÇÃO

Cada vez mais os requisitos de negócio, assim como as legislações, exigem que as organizações cumpram critérios que comprovem que a preservação digital é realizada plenamente. Desde requisitos dos sistemas de gestão de documentos até a capacidade de preservação dos repositórios digitais, tudo é avaliado, buscando-se fazer que os materiais digitais consigam vencer o tempo, sendo preservados autênticos, acessíveis e interpretáveis pelo tempo que for necessário.

Inicialmente, os esforços se concentravam em buscar soluções, técnicas e sistemas isolados para atender requisitos específicos da preservação digital, evoluindo-se para uma visão mais holística. Assim, compreende-se que a preservação digital perpassa diversas áreas de uma organização, necessitando, inclusive de interações entre diferentes profissionais, ao mesmo tempo em que surge a necessidade de se ter critérios para avaliar a real efetivação da preservação digital.

Maemura, Moles e Becker (2017, p. 1619, tradução nossa), declaram:

A avaliação organizacional é um componente-chave da preservação digital eficaz, pois é necessário demonstrar eficácia, um pré-requisito para uma melhoria efetiva e um mecanismo para ajudar as partes interessadas a estabelecer confiança em objetos, sistemas e repositórios digitais.

A comprovação da preservação digital, por consequência, permite o estabelecimento e demonstração de confiança na organização e seus sistemas utilizados. Essa demonstração pode-se realizar, por exemplo, através de realizações de auditorias, obtenção de certificações e aplicações de autoavaliações, demonstrando o grau de maturidade de preservação digital da organização.

Cada organização pode construir o seu caminho para a efetivação da preservação digital, como exemplificam Maemura, Moles e Becker (2017, p. 1620, tradução nossa):

Primeiramente, as organizações que iniciam uma nova iniciativa de preservação digital podem se beneficiar de estruturas que facilitam o **planejamento inicial**. Em segundo lugar, as organizações com processos e fluxos de trabalho existentes podem se beneficiar de avaliações que dão suporte à **melhoria**, identificando áreas específicas para investimento e desenvolvimento organizacional. Em terceiro lugar, organizações experientes com processos documentados e formalizados podem realizar avaliações para **certificação** externa para reconhecer a conformidade com os padrões.

O objetivo assim, é possibilitar que uma organização inicie aplicando ferramentas e instrumentos que auxiliam a preservação digital e servem para a verificação do seu grau de maturidade, servindo como parâmetro para verificação da sua própria evolução ou como comparação com outras organizações. Uma vez que a organização se volta aos processos da preservação digital, a segunda etapa é a realização de melhorias em pontos estratégicos identificados, buscando alcançar a excelência e demonstrando aos seus colaboradores, fornecedores e patrocinadores a sua vocação e comprometimento com a preservação digital. A realização de certificações formais, ao passo que pode ser a realização máxima de determinadas organizações, pode não ser o alvo principal de outras, uma vez que elas podem aplicar outras ferramentas e modelos de maturidade que, juntamente, tornam-se provas de que estão aptas para a realização da preservação digital plena.

De modo abordar o assunto proposto, o presente capítulo demonstra na seção 3.2 o caminho a ser percorrido pelos repositórios rumo à confiabilidade, partindo-se dos conceitos de confiança e como eles se encaixam na realidade dos repositórios, assim como os mecanismos de auditorias e certificações usados para se atestar a confiança. Em seguida, a seção 3.3 apresenta os principais métodos e ferramentas de auditorias e certificações em forma de linha do tempo, que percorre desde o ano 1996 até os dias atuais. Por fim, a seção 3.4 detalha algumas das ferramentas apresentadas na linha do tempo.

3.2 AVALIAR E ALCANÇAR A CONFIANÇA NA PRESERVAÇÃO DIGITAL

3.2.1 O Conceito de Confiança

No ambiente composto por documentos digitais, repositórios e preservação, surge a necessidade de se criar uma relação de confiança entre produtores, preservadores e consumidores dos documentos. A confiança forma a base de todo um sistema que utiliza documentos digitais para diversos processos em uma organização e para além das suas fronteiras. Opostamente, a quebra da confiança, ou ainda a falta de meios de se comprovar a confiabilidade estabelecida, pode por a própria organização em risco a ponto de colapsá-la.

Há, portanto, que se estabelecer confiança entre as partes envolvidas na preservação de materiais digitais como um todo, realizando-se esforços e utilizando-se todos os meios que forem aplicáveis. Corujo (2022, p. 125), define que a confiança se baseia em:

Expectativas em relação às intenções ou ao comportamento dos outros, seja fornecida pela interação continuada entre as partes, seja pelo sentimento de segurança derivado das

estruturas que sustentam uma instituição, seja ela derivada da informação credível que se constitui pela reputação demonstrável/certificação.

Niklas Luhmann abordou a relação entre confiança e risco em suas teorias sobre os sistemas sociais, identificando que a confiança desempenha um papel importante na redução do risco nas interações sociais. Sendo os sistemas sociais caracterizados por uma complexidade crescente, torna-se impossível para os indivíduos controlarem todas as variáveis e consequências de suas ações, emergindo a confiança como um mecanismo que permite que os indivíduos assumam riscos ao interagir com os sistemas sociais. A confiança, assim, reduziria o risco ao fornecer uma base de previsibilidade nas interações sociais, permitindo que os indivíduos ajam sem ter que calcular e verificar constantemente as consequências de cada ação. Nesse contexto, os indivíduos precisam confiar nos sistemas especializados (como sistemas jurídicos, sistemas financeiros, sistemas políticos, entre outros) para realizar suas funções de maneira adequada (LUHMANN, 1988).

No entanto, Luhmann também reconhece que a confiança está sujeita a riscos e pode ser abalada por eventos inesperados ou pela quebra de expectativas. Quando ocorre uma quebra de confiança, os indivíduos podem sentir-se expostos a riscos maiores, tornando-se mais cautelosos e desconfiados em relação aos sistemas e às instituições. Essa situação pode levar a um aumento da percepção de risco e a mudanças nas dinâmicas sociais.

O autor faz uma distinção dos termos “confiança” (*trust*) e “crença” (*confidence*) especificando que ambos se referem a expectativas criadas que podem ser quebradas, porém, a confiança é uma construção baseada na análise de alternativas e variáveis, enquanto a crença está mais associada a valores tidos como verdade. Nesse caso, a confiança requer um esforço e se faz necessária somente quando o resultado de uma ação pode trazer arrependimento ou desapontamento (LUHMANN, 1988, p. 98, tradução nossa):

A distinção entre confiança e crença depende de nossa capacidade de distinguir entre perigos e riscos, sejam eles remotos ou de interesse imediato. A distinção não se refere a questões de probabilidade ou improbabilidade. A questão é se a possibilidade de desapontamento depende ou não do seu próprio comportamento anterior.

Ao transportar esses conceitos para a preservação digital, a confiança de que uma determinada organização irá realizar a preservação de acordo é fundamentada pela análise da sua capacidade em realizar essa tarefa. A organização ao demonstrar o cumprimento de determinados critérios, através da aplicação de modelos de maturidade, auditorias ou certificações auxilia na criação do sentimento de confiança dos seus depositantes. A partir de análises criteriosas dos processos, métodos e ferramentas de uma organização pode-se definir que ela é confiável e adquire,

registra e constrói contextos de informações sobre os seus documentos que os tornam igualmente confiáveis.

Cabe fazer ressaltar uma distinção entre as instituições, no que se refere à natureza pública e privada. No caso de instituições públicas, os produtores (depositantes) podem construir a confiança embasada muito fortemente nas suas reputações, ou ainda, por medidas legais, serem obrigados a confiarem nessas instituições, ao que Luhmann associaria ao termo *crença*, e que elas vão realizar os seus papéis, mesmo sem qualquer meio de se atestar que são efetivamente confiáveis. Já no âmbito de instituições privadas, o interesse em receber e custodiar materiais pode ser o foco principal, levando à necessidade de comprovarem que atendem critérios para a conquista da confiança de seus produtores. Caso contrário, a falta de confiança pode inviabilizar a operação dessas instituições.

Márquez contextualiza o grau de necessidade de confiança no mundo atual, onde a produção de informação e documentos é constante, assim como a incerteza em crer nessas mesmas informações produzidas (2017, p. X, tradução nossa):

Da enorme produção e consumo de informação deriva, por sua vez, uma problemática cuja tendência também tem de aumentar: a questão da qualidade e credibilidade da informação que se consome. Cada vez que mais usuários se perguntam até onde podem confiar na informação a que acessam, obviamente existem problemas derivados de crer em informações que não são fidedignas, com as consequências que isso pode acarretar, assim como problemas derivados de duvidar da informação recebida. Assim, na medida em que a produção e o consumo de informações aumentam, também aumentam os problemas de confiança ou desconfiança nessas informações.

Márquez ainda analisa que o termo “confiança” é muito amplo e, portanto, difícil de se definir um único significado, podendo variar e se transformar de acordo com a passagem do tempo, localização geográfica, cultura e disciplina que o aplica. O termo acaba por se relacionar a outros termos com significados particulares e, de certa forma, entrelaçados como, por exemplo, verdade, veracidade, credibilidade, confiabilidade, risco, certeza e fé (2017, p. 5).

Ao analisar as definições da palavra “confiança” em outros idiomas, Márquez aponta para a tradução dos termos da língua inglesa para o espanhol, ao que se pode aplicar ao português da mesma forma. O autor apresenta a relação entre os termos “*trust*” e “*confidence*”, do inglês, que seguidamente são traduzidos como significados para o termo “confiança”, do espanhol e português, no entanto, ressalta que o segundo termo teria como melhor equivalente a palavra “certeza” (2017, p. 8). De maneira análoga, o termo “*trustworthiness*” seguidamente se traduz ao português como “confiabilidade” e ao espanhol como “*confiabilidad*” ou “*fiabilidad*”, referindo-se à qualidade de ser confiável, digno de confiança ou digno de crédito e, por consequência, o termo “*trustworthy*” se traduz como “confiável”. A importância dessas equivalências e traduções do inglês para outros

idiomas se deve ao fato da existência de uma vasta literatura publicada nessa língua, buscando diminuir erros e más interpretações.

No ambiente da informática, Márquez analisa que a confiança pode ser medida a partir de métodos cartesianos para atestar que um determinado programa computacional é ou não confiável, baseando-se na probabilidade que esse programa tem de fornecer uma saída de acordo com os resultados esperados. Assim a confiança seria “a medida da confiabilidade com que uma entidade de computador – sistema, programa ou equipamento – fornece informações corretas ou executa uma operação corretamente, sem desperdício de recursos” (MÁRQUEZ, 2017, p. 63, tradução nossa). O autor ainda embasa essa definição com a definição do termo “*reliability*” do Glossário Padrão IEEE de Terminologia da Engenharia de Software: “Confiabilidade. A capacidade de um sistema ou componente de executar suas funções necessárias sob condições estabelecidas por um período de tempo especificado” (THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, 1990, p. 62, tradução nossa).

Após analisar a confiança sobre diferentes prismas, Márquez identifica que a confiabilidade é a qualidade de algo (pessoa, sistema, instituição) ser fidedigno, íntegro e ter credibilidade, ao passo que a confiança é um resultado da análise dessa qualidade e seus elementos, definindo o termo, de forma ampla, como (2017, p. 97, tradução nossa):

Confiança é a disposição de uma parte (o confiante) de realizar uma ação com uma contraparte (o confidente), na qual o primeiro deve ceder o controle ao segundo, no qual coloca o confiante em uma certa situação de vulnerabilidade, já que existe um risco de o confidente possa não realizar no futuro a ação esperada, com dano ou prejuízo ao confiante, e essa situação cria uma certa incerteza nele. Para reduzi-la, o confiante deve, portanto, avaliar a confiabilidade do confidente, entendendo-a como a capacidade e disposição possíveis do confidente em realizar a ação desejada.

Ao analisar o mundo digital, sobretudo as informações disponibilizadas em websites, Márquez evolui e leva em consideração a confiança da informação digital, que analisa principalmente o seu conteúdo e fonte (2017, p. 101, tradução nossa):

A confiança na informação digital é o resultado de um processo de avaliação subjetiva de um determinado usuário sobre a medida de confiabilidade de uma determinada informação com base em sua apreciação da soma da qualidade da informação em si, a solvência da fonte e a credibilidade do site que a emite.

A preocupação com a confiança dos documentos se torna maior e ganha corpo no advento dos documentos eletrônicos, uma vez que se identificou que a preservação desses documentos requer a sua modificação para perpassar principalmente a obsolescência tecnológica, seja nas suas mídias e suportes, seja na migração e conversão de formatos. Essa mutabilidade não acontecia nos documentos físicos, imutáveis desde a produção, e que necessitavam de processos de conservação (MÁRQUEZ, 2017, p. 68).

Pode-se dizer ainda que a confiança ainda se relaciona com o grau de transparência de uma instituição, referindo-se à adoção de processos bem estabelecidos e públicos, assim como a publicação de seus documentos. Assim, nos casos de países democráticos, busca-se a implementação de leis e normativas de transparência buscando tornarem-se menos opacos quanto a disponibilização de informações e, principalmente, documentos que possam alterar a forma como se exerce a cidadania. No Brasil, cita-se a Lei nº12.527, conhecida como LAI (Lei de Acesso à Informação) que estabeleceu a transparência ativa como padrão e não exceção (BRASIL, 2011). Na Espanha, cita-se o conjunto de Leis e Decretos Reais que modificaram a administração pública ao implantar os documentos eletrônicos como padrão, preocupando-se com o registro de elementos que embasam a presunção de autenticidade (ESPANHA, 2007) (ESPANHA, 2009) (ESPANHA, 2010) (ESPANHA, 2011) (ESPANHA, 2015) (ESPANHA, 2015).

No âmbito dos documentos físicos, a confiança se relaciona fortemente aos signos e elementos que ele carrega, mas sobretudo à confiança atribuída aos seus custodiadores, criando uma espécie de dependência mútua. A confiança em determinadas instituições é muitas vezes atribuída baseada nos seus status ou reputação como, por exemplo, arquivos de igrejas e órgãos públicos, portanto, os documentos sob suas custódias também seriam confiáveis.

Com o passar do tempo, passou-se a se exigir mais elementos para se denominar que uma instituição é confiável, assim como os seus documentos. O universo volátil, frágil e mutante dos documentos digitais trouxe a necessidade de se rever a natureza desses elementos que permitem atestar a confiança dos próprios documentos e das instituições custodiadoras. Uma vez que a confiança envolve a crença do produtor de documentos de que a instituição custodiadora irá realizar a preservação adequada dos seus documentos, cabe à instituição demonstrar a presença de critérios de confiabilidade, conquistando assim a confiança dos produtores.

Duranti descreve o cenário em que os documentos arquivísticos se encontram e dá ideia sobre os elementos externos a si mesmos que compõem as suas respectivas autenticidades e que constroem a confiança (1994, p. 52):

As relações entre os documentos, e entre eles e as transações das quais são resultantes, estabelecem o axioma de que um único documento não pode se constituir em testemunho suficiente do curso de fatos e atos passados: os documentos são interdependentes no que toca a seu significado e sua capacidade comprobatória. Em outras palavras, os documentos estão ligados entre si por um elo criado no momento no qual são produzidos ou recebidos, que é determinado pela razão de sua produção e é necessário à sua própria existência, à sua capacidade de cumprir seu objetivo, ao seu significado, confiabilidade e autenticidade. Na verdade, os registros documentais são um conjunto indivisível de relações intelectuais permanentes tanto quanto de documentos.

Ao se analisar o relatório de Waters e Garret, encontra-se a presença dos atributos que compõem a confiabilidade dos arquivos e que levam à criação da confiança pelos usuários dos arquivos (1996, p. 21, tradução nossa):

Para garantir a longevidade da informação, talvez o papel mais importante na operação de um arquivo digital seja o gerenciamento da identidade, integridade e qualidade dos próprios arquivos como uma fonte confiável do registro cultural. Os usuários de informações arquivadas em formato eletrônico e de serviços de arquivamento relacionados a essas informações precisam ter a garantia de que um arquivo digital é o que diz ser e que as informações ali armazenadas são seguras a longo prazo.

3.2.2 Repositórios Digitais de Confiáveis

O documento elaborado pela RLG, que define os atributos e responsabilidades de um repositório confiável, traz o que deveria ser a missão principal de um repositório: “Um repositório digital confiável é aquele cuja missão é fornecer acesso confiável e de longo prazo a recursos digitais gerenciados para sua comunidade designada, agora e no futuro” (RESEARCH LIBRARIES GROUP, 2002, p. i, tradução nossa). Essa definição traz consigo o aspecto primordial da confiabilidade no acesso aos documentos do repositório, ao que se alicerça em todo um arcabouço de procedimentos que iniciam na ingestão e passam pela execução da preservação em si, garantindo que a comunidade do repositório consuma documentos fidedignos e que se crie uma relação de confiança com o repositório. Adiante, o documento identifica que a confiança no âmbito dos repositórios digitais é formada a partir de três níveis (2002, p. 9, tradução nossa):

- 1) Como as instituições culturais conquistam a confiança das suas comunidades designadas;
- 2) Como as instituições culturais confiam em seus fornecedores externos, e; 3) Como os usuários confiam nos documentos fornecidos a eles por um repositório.

No primeiro nível, aproveita-se a confiança já possui nas instituições existentes como bibliotecas, museus e arquivos, desfrutando do prestígio e reputação construídos ao longo dos tempos, armazenando e dando acesso aos seus materiais, ao que expressa da seguinte forma (2002, p. 9, tradução nossa):

Embora os desafios apresentados na preservação da informação digital sejam muito diferentes e exijam novas soluções, o público provavelmente terá pelo menos alguma confiança de que as instituições culturais terão sucesso com base no sucesso do passado.

No segundo nível, são os fornecedores externos que devem conquistar a confiança dos arquivos, através do cumprimento dos contratos firmados e a comprovação de atendimento de requisitos mínimos. Nesse caso, a existência de certificações dos fornecedores e apresentações de resultados de auditorias são aspectos que contribuem para a construção da confiança.

No terceiro nível, o arquivo deve possuir processos e mecanismos que respondam às perguntas “Como um usuário pode ter certeza de que o documento recebido é o solicitado? Como o documento pode ser verificado para ser o item exato depositado no repositório digital no passado?” (2002, p. 10, tradução nossa). Para que os documentos sejam confiáveis, o arquivo deve se precaver de adulterações dos documentos preservados, adotando medidas de segurança e níveis de acesso para evitar alterações intencionais e medidas de checagem de corrompimentos e monitoramento para evitar as alterações acidentais. Em conjunto, o repositório deve ainda registrar as rotinas executadas internamente e possuir elementos que permitam ao usuário final atestar a proveniência dos documentos acessados.

Ross e McHugh, no contexto da popularização e grande adoção de repositórios, institucionais relatam a mudança da percepção de confiança dos usuários, ao que definem ser uma “desconfiança saudável” (2006, p. tradução nossa):

Dentro de nossa comunidade, nossa perspectiva parece estar mudando de um comportamento de confiança otimista para um comportamento de confiança pessimista. Ou seja, surgiu um nível saudável de pessimismo que levou muitos a reconhecer que os repositórios digitais só são dignos de confiança se puderem demonstrar que possuem as propriedades de confiabilidade.

A partir desse ponto, os autores veem um cenário propício ao desenvolvimento de certificações e auditorias com critérios objetivos para se atestar a confiabilidade dos repositórios, ao mesmo tempo, fazem uma crítica aos métodos de certificação e auditoria em desenvolvimento à época, sobre a forma de se atestar o cumprimento dos critérios. Os autores sustentam que os critérios devem vir acompanhados de um conjunto de evidências que facilitem atestar os seus cumprimentos e que permita aos repositórios a visualização de um horizonte a ser perseguido (2006, p. tradução nossa):

A questão das categorias de evidências necessárias para facilitar as auditorias e possibilitar a certificação precisa receber consideração adequada: qualquer ferramenta que omita a descrição das evidências que contribuirão para o processo de auditoria é incompleta. Se um *checklist* de auditoria tiver aspirações de aplicabilidade prática, seus critérios devem detalhar não apenas os padrões esperados e exigidos, mas também os meios pelos quais sua obtenção pode ser demonstrada e avaliada. Da mesma forma, se essas ferramentas devem promover a autoavaliação de um tipo rigoroso e confiável e provavelmente fornecer uma boa previsão para o resultado de uma auditoria externa independente, elas devem ser abrangentes, independentemente ou em combinação com uma ou mais recursos mais vinculados.

O TRAC argumenta que a confiança de um repositório deve analisar não somente as suas tecnologias empregadas, mas sim observar a organização de uma maneira holística (ONLINE COMPUTER LIBRARY CENTER, 2007, p. 3, tradução nossa):

Ao determinar a confiabilidade, deve-se observar todo o sistema no qual a informação digital é gerenciada, incluindo a organização que administra o repositório: sua governança; estrutura organizacional e pessoal; políticas e procedimentos; aptidão financeira e sustentabilidade; os contratos, licenças e responsabilidades sob os quais deve operar; e sucessores confiáveis de dados, conforme aplicável. Além disso, as práticas de gerenciamento de objetos digitais, infraestrutura tecnológica e segurança dos dados em vigor deve ser razoável e adequada para cumprir a missão e os compromissos do repositório.

Após a análise do repositório, a confiança com a sua comunidade pode ser conquistada ou aumentada com a divulgação das formas e meios implantados para garantir a preservação e acesso confiáveis dos seus materiais (ONLINE COMPUTER LIBRARY CENTER, 2007, p. 4, tradução nossa):

A comunicação dos resultados da auditoria ao público – transparência – gerará mais confiança, e auditorias objetivas adicionais, potencialmente levando à certificação, promoverão mais confiança no repositório e no sistema que o suporta. Por fim, alcançar o status de confiança não é uma conquista única – conquistada e esquecida. Para manter o status confiável, um repositório precisará realizar um ciclo regular de auditoria e/ou certificação.

Para o Digital Curation Centre (DCC) e DigitalPreservationEurope (DPE), a confiança em um determinado repositório inicia com a identificação e avaliação dos riscos aos quais ele está exposto, culminando na demonstração e publicização dos esforços realizados para evitar e mitigar tais riscos. A avaliação dos riscos, nesse caso, analisa o repositório desde a sua estrutura organizacional, passando pelo tratamento dispensado aos objetos digitais, até a infraestrutura de tecnologia e segurança empregados (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007).

O Nestor Working Group ao criar o seu catálogo de critérios para avaliação e certificação de repositórios confiáveis, define a confiabilidade de uma maneira ampla e depois foca a definição mais em relação aos aspectos de TI (2009, p. 5, tradução nossa):

Confiabilidade é a capacidade de um sistema operar de acordo com seus objetivos e especificações (ou seja, faz exatamente o que se propõe a fazer). Do ponto de vista da segurança de TI, os valores básicos são integridade, autenticidade, confidencialidade e disponibilidade. A segurança de TI é, portanto, um pré-requisito importante para repositórios digitais confiáveis

Yoon analisa o significado do termo “confiança” a partir de usuários de repositórios de dados e a define em dois níveis: “Em um nível mais alto, os usuários percebem a confiança como sua vontade de depender ou contar com algo ou alguém (incluindo uma organização) em circunstâncias incertas” (2014, p. 30, tradução nossa). Especificamente tratando-se dos repositórios de dados, a confiança possui forte ligação com os dados, suas qualidades e integridades:

Os repositórios devem ser honestos e não enganar ninguém intencionalmente. Essa forte presença de verdade e honestidade na definição de confiança dos usuários em repositórios de dados reflete a importância da integridade e validade dos dados em suas pesquisas. A

aquisição de dados válidos e precisos é o primeiro passo para qualquer tipo de pesquisa que reutilize dados produzidos por terceiros. A questão da integridade dos repositórios também está relacionada à integridade e validade dos dados como fontes confiáveis de dados.

O estudo ainda relatou que a presença de documentações sobre os processos dos repositórios contribui para a construção da confiança. Em um caso específico, um participante relatou possuir mais confiança em um repositório que possui a documentação mais completa do que em outro repositório. Outro fator considerado para o estabelecimento da confiança, identificado no estudo, foi a influência da comunidade de usuários do repositório. Os participantes relataram que foram recomendados a usar os repositórios por outros usuários do repositório, criando uma espécie de cadeia de confiança que passa de pessoa para pessoa.

Yoo conclui trazendo a importância de o repositório ser percebido como confiável por sua comunidade, ao mesmo tempo em que ressalta a necessidade de métodos que permitam medir a confiança da sua comunidade (2014, p. 32, tradução nossa):

Ganhar a confiança dos usuários nos repositórios é importante porque uma das principais missões dos repositórios é servir suas distintas comunidades de usuários. Ao entender as perspectivas dos usuários sobre repositórios “confiáveis”, os repositórios podem melhorar seu status “confiável” porque as percepções de confiança dos usuários são frequentemente (embora nem sempre) relacionadas às práticas dos repositórios. [...] A confiança é um conceito complexo de medir, mas ter uma forma padronizada de medir a confiança dos usuários pode ajudar a demonstrar como os repositórios efetivamente ganharam a confiança dos usuários e foram percebidos como fontes confiáveis de informação.

Corujo faz um paralelo entre instituições que preservam patrimônio físico e digital, comparando a confiança nelas depositadas e o grau de exigência e desafios existentes no mundo digital (2022, p. 78):

Podemos considerar que as instituições ligadas ao patrimônio cultural são fidedignas, na medida em que são responsáveis pelos materiais e objetos que compõem esse patrimônio, e também pela sua guarda, pelas condições de acesso e de garantia de preservação a longo prazo. A confiança nestas instituições baseia-se no conhecimento, na experiência e na prática demonstrável pela salvaguarda de objetos físicos. Mas no âmbito do patrimônio digital, dadas as suas características e os desafios apresentados na sua preservação, esta confiança e a garantia de fidedignidade são mais difíceis de alcançar, e muitas das instituições culturais, como bibliotecas, museus e arquivos, não estão atualmente preparadas para cumprirem com as suas funções tradicionais de armazenamento, preservação e de garantir a acessibilidade, no âmbito do patrimônio digital, tendo de recorrer a serviços especializados de entidades externas.

Sendo assim, a confiança se torna um ativo de extrema importância para os repositórios, que é construído à medida em que as expectativas das partes envolvidas vão sendo atendidas, ao passo em que a perda desse ativo, a confiança, acarreta diretamente na credibilidade dos repositórios e, por consequência, nos materiais por ele preservados. Essa construção passa pelo repositório demonstrar a adequação organizacional, governança e transparência para a realização da preservação assim como a sua sustentabilidade financeira, além da utilização de recomendações da

área e atendimento às normas aplicáveis. A partir desse ambiente, a comunidade alvo do repositório pode perceber o seu valor e comprometimento com a preservação digital, alcançando-se a confiança. Adicional a essa percepção, mais ou menos subjetiva e demonstrável em graus diferentes, o repositório pode se submeter a auditorias e processos de certificação que deem ao repositório a chancela de cumprimento de critérios fundamentais para a preservação digital, aumentando o clima geral de confiança.

3.2.3 Auditorias e Certificações: Mecanismos de Avaliação para Assegurar a Confiança

Em geral, o processo de auditoria precede ao de certificação, uma vez que um determinado modelo de requisitos é escolhido para ser aplicado no repositório identificando as suas potencialidades e vulnerabilidades. A análise dessa auditoria permite ao repositório planejar ações de melhoria que podem levar aos gestores do repositório à decisão de se submeter a uma determinada certificação (SANTOS, 2018, p. 44).

A auditoria se caracteriza por ser um “processo sistemático, independente e para obter evidência objetiva e avalia-la objetivamente, para determinar a extensão na qual os critérios de auditoria são atendidos” (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2018, p. 1). A execução da auditoria pode ser realizada por uma equipe interna da instituição, podendo ser denominada como autoavaliação. Eventuais partes interessadas na instituição como fornecedores, por exemplo, também podem formar um corpo avaliativo com o intuito de realizar auditorias. No universo dos repositórios de documentos digitais, esse papel poderia ser desempenhado pelos produtores que depositam seus documentos e buscam verificar se a preservação deles ocorre conforme os padrões estabelecidos na auditoria. Uma outra forma de realizar a auditoria é através da contratação de uma organização de auditoria credenciada e capacitada para esse fim, tornando o processo independentes de todas as partes interessadas na instituição.

O processo de obtenção de uma certificação para o repositório envolve diferentes níveis da organização, sendo necessário avaliar desde a estrutura organizacional, pessoas envolvidas, até a estrutura técnica utilizada na preservação dos documentos. A certificação, com isso, afirma que os procedimentos e estrutura empregados no funcionamento do repositório como um todo, estão de acordo com requisitos estabelecidos definidos pela comunidade internacional. Pelo lado das pessoas e entidades que depositam os seus documentos, há a confiança de que a preservação segue padrões bem definidos, com a garantia de estarem armazenados a longo prazo. Ao mesmo tempo, os patrocinadores e financiadores do repositório também se resguardam compreenderem que os seus recursos são empregados da maneira correta e certificada.

Os processos de aplicação às certificações de confiança atualmente disponíveis requerem grandes esforços e geram determinados custos. No caso de realizar uma autoavaliação, o repositório deve investir na capacitação das pessoas envolvidas e alocá-las para a realização do trabalho necessário, que se inicia com a coleta de informações e documentos e termina com o parecer final de avaliação de cumprimento do conjunto de critérios. No caso das certificações de nível estendido e formais, há a necessidade de contratação de um corpo auditor habilitado a realizar a auditoria, além da necessidade de se destacar uma equipe para acompanhamento e fornecimento das informações necessárias.

Em todos os casos, o processo de auditoria e certificação gera determinados custos aos repositórios, sejam eles através de pagamentos efetuados ou de investimento de horas de trabalho do pessoal envolvido. No entanto, os benefícios gerados influenciam de diversas maneiras os processos, as equipes e os financiadores do repositório. Em determinados casos, a auditoria pode levar à otimização de processos, tornando-os mais eficientes e gerando resultados mensuráveis como a quantidade de recursos economizada, por exemplo. Em outros casos, os benefícios podem não ser facilmente tangíveis como a confiança gerada nos financiadores do repositório e o aumento da qualidade da documentação como um todo. Muito embora a métrica de Retorno do Investimento (*Return On Investment – ROI*) seja difícil de se calcular nos casos das certificações de confiança dos repositórios, acredita-se que os esforços despendidos valham a pena (LINDLAR e SCHWAB, 2022).

A obtenção de uma determinada certificação não representa um fim, mas sim, um processo contínuo de manutenção dos processos que garantem a confiabilidade do repositório, assim como as suas melhorias, sendo necessárias reavaliações periódicas conforme o tempo de validade de cada uma e conforme novas versões dos conjuntos de requisitos vão surgindo.

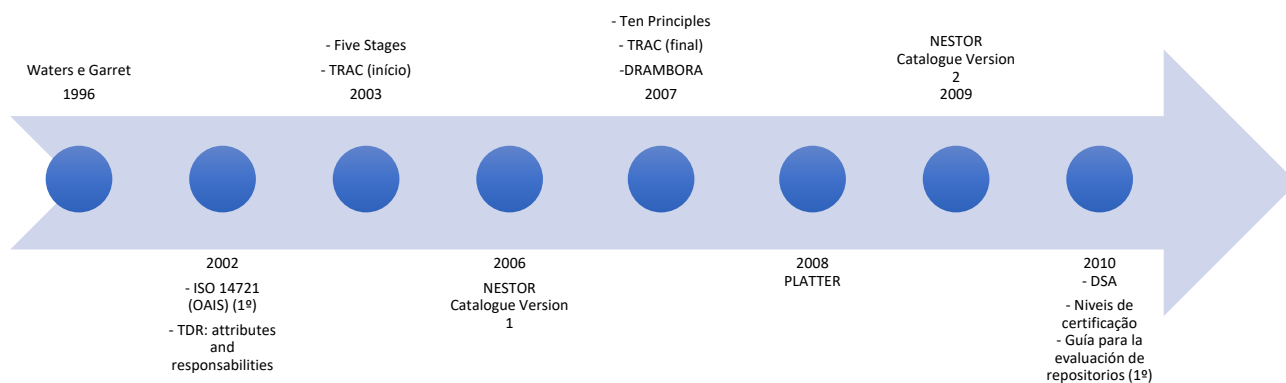
3.3 A LINHA DO TEMPO DOS MÉTODOS E FERRAMENTAS DE AUDITORAS E CERTIFICAÇÕES

Nessa seção é apresentada uma linha do tempo dividida em partes que aborda o surgimento e evolução de alguns métodos e ferramentas utilizados para a realização de auditorias (internas e externas) e obtenção de certificações. A fim de melhor apresentar a linha do tempo, ela foi dividida entre os acontecimentos até 2010 e os posteriores a esse ano até os dias atuais.

3.3.1 O Período de 1996 a 2010

O final do século XX e início do XXI é marcado pelo estabelecimento dos alicerces das definições sobre o que é um repositório digital, quais são as suas características e atribuições e como comprovar que os repositórios as cumprem, criando um ambiente confiável e seguro para a preservação de materiais digitais. A Figura 10 exibe uma linha do tempo até o ano 2010 com marcos importantes do tema de auditorias, certificações e recomendações relacionados à preservação digital.

Figura 10 – Linha do tempo da preservação digital de 1996 a 2010.



Fonte: elaboração própria

O primeiro marco exibe o relatório da Força Tarefa de Arquivamento de Informações Digitais elaborado em 1 de maio de 1996 e comandada por Donald Waters e John Garret (1996) que traça um panorama da preservação digital e seus desafios para manter as informações acessíveis e autênticas ao longo do tempo. Dentre as recomendações dos autores, está a criação de mecanismos para a certificação dos repositórios (WATERS e GARRET, 1996, p. 45, tradução nossa): “um processo de certificação para arquivos digitais é necessário para criar um clima geral de confiança sobre as perspectivas de preservação da informação digital”.

Em janeiro de 2002, o Comitê Consultivo para Sistemas de Dados Espaciais (*Consultative Committee for Space Data Systems – CCSDS*) publica a primeira versão do modelo OAIS (*Open Archival Information System*), como uma recomendação técnica com o intuito já de criar um padrão para repositórios digitais que permitisse as suas uniformizações e comparações de funcionamento. Essa recomendação “fornece uma base para uma maior padronização dentro de um

contexto de arquivamento e deve promover uma maior conscientização dos fornecedores e apoio aos requisitos de arquivamento” (CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS, 2002, p. iii, tradução nossa). O modelo foi construído para a preservação de informações digitais a serem preservadas por tempo indeterminado, que necessitam cuidados para superar as mudanças tecnológicas envolvidas, ainda que essas informações não sejam permanentes. Diversas funções ligadas à preservação são contempladas, como a ingestão, armazenamento, gerenciamento de dados, acesso e difusão, utilizando-se de pacotes de informação e entidades com responsabilidades específicas e bem delimitadas.

Em maio de 2002, a união de esforços entre RLG (*Research Libraries Group*) e OCLC (*Online Computer Library Center*) criou o documento *Trusted Digital Repositories: Attributes and Responsibilities* (TDR) que estabeleceu “[...] uma estrutura de atributos e responsabilidades para repositórios digitais confiáveis e sustentáveis, capazes de lidar com a variedade de materiais mantidos por grandes e pequenas instituições de pesquisa” (RESEARCH LIBRARIES GROUP, 2002, p. i). O relatório identifica atributos técnicos e de governança que um repositório digital precisa possuir para ser confiável, dentre os quais, pode-se citar a conformidade com o modelo de referência OAIS. Adicionalmente, é ressaltada a importância de existirem certificações para os repositórios digitais (RESEARCH LIBRARIES GROUP, 2002, p. 38, tradução nossa): “Uma estrutura e processo de certificação para repositórios digitais são cruciais e a sua ausência tem sido um impedimento para a atribuição de confiança”.

Em 2003, Anne Kenney e Nancy McGovern lançam o modelo de avaliação *The Five Organizational Stages of Digital Preservation* para que as organizações possam identificar os seus graus de maturidade em relação à preservação de materiais digitais (KENNEY e MCGOVERN, 2003): “descrevemos cinco etapas definíveis que os repositórios culturais atravessarão em seu caminho para desenvolver um programa de preservação digital totalmente maduro”. Os estágios foram elaborados em alto nível, sem se aprofundar em muitos detalhes, e possuem indicadores que auxiliam a verificar os seus cumprimentos.

No primeiro estágio, *Stage 1. Acknowledge: Understanding that digital preservation is a local concern*. (Reconhecimento: Compreender que a preservação digital é uma preocupação local), a organização deve reconhecer que precisa lidar com a preservação digital e quanto mais cedo houver essa consciência, menor é a probabilidade de se perder materiais que podem se mostrar essenciais no futuro, seja por questões históricas, jurídicas ou contábeis. “No estágio um, as instituições passam da dependência do conteúdo digital para o reconhecimento de que há um problema, para o reconhecimento do interesse próprio institucional e da necessidade de agir” (KENNEY e MCGOVERN, 2003).

No segundo estágio, *Stage 2. Act: Initiating digital preservation projects* (Agir: Iniciando projetos de preservação digital), a organização cria projetos e os executa a fim de atacar os seus problemas de preservação digital identificados no estágio anterior. Iniciativas de cooperação e grupos de estudos são encorajadas a fim de se conhecer e aplicar as boas práticas e padrões existentes, buscando-se definir os formatos de arquivos adequados, metadados a serem recolhidos e definindo a infraestrutura técnica necessária. Kenney e McGovern destacam a importância desse estágio (2003):

Esta fase é geralmente a de menor duração, pois a inadequação desta abordagem torna-se aparente quando os projetos terminam. No entanto, esta é uma fase muito importante pela qual as instituições devem passar, adquirindo habilidades e experiência necessárias na aplicação de conceitos e fazendo a ponte entre a pesquisa de preservação digital e as implementações reais. É também uma fase sóbria, durante a qual as inadequações das práticas anteriores se tornam mais aparentes.

No terceiro estágio, *Stage 3. Consolidate: Sequing from projects to programs* (Consolidação: passando de projetos para programas), à medida em que os projetos iniciados no estágio anterior vão sendo concluídos, a organização toma ciência de que as atividades que envolvem a preservação digital devem fazer parte da sua rotina, levando à criação de programas permanentes. Esse estágio possui um impacto econômico na organização, uma vez que os custos da preservação digital, antes limitados a projetos, agora devem cobrir toda a organização. “Há uma ampliação do escopo da preservação digital do nível do projeto para o institucional, processo que acaba levando à identificação de redundâncias, à redução de ineficiências e ao estabelecimento de prioridades.” (KENNEY e MCGOVERN, 2003).

No quarto estágio, *Stage 4. Institutionalize: Incorporating the larger environment* (Etapa 4. Institucionalizar: Incorporar o ambiente maior), existe a institucionalização de políticas de preservação digital, procurando unificar os processos necessários e otimizar setores e funções, procurando tornar a gestão da preservação digital mais alcançável. A transição do estágio 3 para o 4 requer um planejamento a nível institucional que pode afetar diversos setores, por isso, essa é a transição que tende a levar mais tempo para se iniciar nas organizações. Como balizador, utiliza-se nesse estágio a norma ISO 14721 para que o arquivo estabelecido esteja no padrão OAIS. “Um impulsionador para a passagem para o estágio 4 pode ser o fardo cada vez mais pesado de gerenciar coleções grandes e heterogêneas que estão hospedadas em vários pontos dentro da instituição e que exigem infraestruturas técnicas complexas e caras” (KENNEY e MCGOVERN, 2003).

No quinto estágio, *Stage 5. Externalize: Embracing inter-institutional collaboration and dependencies* (Etapa 5. Externalizar: Incorporando a colaboração interinstitucional e as dependências), a organização busca outras instituições, via parcerias ou consórcios, para oferecer serviços integrados e promover o compartilhamento de experiências. No aspecto econômico, muitas

vezes existe a possibilidade de divisão de custos entre as instituições beneficiando as partes envolvidas. Dentre os impactos positivos do alcance desse estágio, os autores destacam (KENNEY e MCGOVERN, 2003):

Tal união de recursos cria o potencial e a oportunidade para uma camada de serviços em cima dos repositórios que estará disponível para todos os membros e pode trazer benefícios significativos e inesperados para os participantes. [...] O tema nesta fase é que o todo pode ser maior do que a soma de suas partes.

Ainda no ano de 2003, foi criada a Força Tarefa em Certificação de Repositórios Digitais, fruto de uma parceria entre o RLG e o NARA (*National Archives and Records Administration*) com a missão de criar critérios para “desenvolver critérios para identificar repositórios digitais capazes de armazenar, migrar e fornecer acesso confiável a coleções digitais” (ONLINE COMPUTER LIBRARY CENTER, 2007, p. i, tradução nossa). Tal conjunto de critérios deveria ser aplicável a diferentes tipos de repositórios digitais, de bibliotecas a arquivos, de organizações acadêmicas a empresas privadas. Nascia assim, o esforço conjunto que resultaria mais tarde, em 2007, no TRAC.

Em 2006, o projeto alemão NESTOR, acrônimo para *Network of Expertise in Long-term Storage of Digital Resources*, através do *Working Group Trusted Repositories – Certification*, publicou a primeira versão do seu catálogo de critérios para a avaliação de repositórios digitais como um rascunho público para colaborações. O trabalho iniciado em 2004, fortemente embasado no TDR⁴⁷ e que passou por rodadas de aplicação em casos de uso e interações com profissionais da comunidade arquivística. O foco do catálogo de critérios foi a aplicação em repositórios digitais alemães, no entanto, já se aconselhava a sua aplicação a nível internacional, observando-se sempre as legislações locais e normas internacionais aplicáveis. A primeira intenção do projeto era, através dos critérios, fornecer um guia aos repositórios para que se tornassem confiáveis: “o catálogo pode ser utilizado como instrumento de autoavaliação em todas as etapas do desenvolvimento desde o conceito passando pela especificação até a implementação” (DOBRATZ, SCHOGER e STRATHMANN, 2007, p. 2, tradução nossa).

No entanto, a transformação do catálogo em uma ferramenta de auditoria e certificação já estava no horizonte do projeto, assim como a criação de um padrão alemão e/ou internacional. Os esforços do projeto culminaram, em 2009, na versão 2.0 do catálogo de critérios.

No ano de 2007, em Chicago, USA, juntaram-se ao *Center for Research Libraries* (CRL) com âmbito na América do Norte, outras três grandes organizações envolvidas na preservação digital, *The Digital Curation Centre* (DCC) do Reino Unido, *DigitalPreservationEurope* (DPE)

⁴⁷ (RESEARCH LIBRARIES GROUP, 2002)

representando a União Europeia, e NESTOR da Alemanha, para a elaboração de princípios básicos a serem atendidos por repositórios digitais, assim como para servirem como ponto inicial na construção de futuras certificações e auditorias. Sob a ressalva de que os repositórios devem adequar seus processos e atividades de preservação de acordo com os seus tamanhos e necessidades das suas comunidades-alvo, o grupo lançou os 10 princípios para a preservação digital (CENTER FOR RESEARCH LIBRARIES, 2007):

1. O repositório se compromete com a manutenção contínua dos objetos digitais para a(s) sua(s) comunidade(s);
2. Demonstra aptidão organizacional (incluindo financeira, de pessoal e processos) para cumprir o seu compromisso;
3. Adquire e mantém os direitos legais e contratuais necessários e cumpre com as suas responsabilidades;
4. Tem um quadro político eficaz e eficiente;
5. Adquire e ingesta objetos digitais com base em critérios declarados que correspondem a seus compromissos e capacidades;
6. Mantém/Assegura a integridade, autenticidade e usabilidade dos objetos digitais que mantém ao longo do tempo;
7. Cria e mantém os metadados necessários sobre as ações realizadas em objetos digitais durante a preservação, bem como sobre os contextos relevantes de produção, acesso e uso antes da preservação;
8. Cumpre os requisitos de difusão necessários;
9. Tem um programa estratégico de planejamento e ação de preservação;
10. Possui infraestrutura técnica adequada para a manutenção contínua e segurança de seus objetos digitais.

É possível notar que os princípios são bastante genéricos e não entram em detalhes sobre como os repositórios devem cumpri-los, porém norteiam as iniciativas que os repositórios devem tomar para a preservação adequada de objetos digitais. Entre os princípios, observa-se o foco nas necessidades da comunidade-alvo, a necessidade de haver uma estrutura organizacional e técnica adequada, assim como a necessidade de existirem políticas de ingestão, preservação e difusão dos materiais do repositório. Ainda, o princípio 6 trata da integridade, autenticidade e usabilidade do material preservado, características implícitas e essenciais de um repositório confiável. A publicação dos princípios se tornou um marco na preservação digital, principalmente pela grande influência e representatividade das organizações envolvidas, sendo utilizada como

ponto de partida para as publicações, normas e instrumentos de avaliação de repositórios que se seguiram.

Em fevereiro de 2007 foi publicada a versão 1.0 do documento *Trustworthy Repositories Audit & Certification: Criteria and Checklist* (TRAC), trabalho iniciado em 2003 pela RLG⁴⁸ e NARA, fruto da necessidade de se definir e comprovar o aspecto confiável de um repositório digital, definindo-se como (ONLINE COMPUTER LIBRARY CENTER, 2007, p. 2, tradução nossa):

[...] as melhores práticas e pensamentos atuais sobre a infraestrutura organizacional e técnica necessárias para (um repositório) ser considerado confiável e capaz de certificação. [...] Estabelece uma definição básica de um repositório digital confiável e estabelece os componentes que devem ser considerados e avaliados como parte dessa determinação. [...] Por fim, documenta os critérios que os repositórios confiáveis poderão atender, fornecendo explicações e exemplos.

A criação dos critérios utilizou como pedra angular a ISO 14721 (OAIS), assim como foi influenciada pelos trabalhos antecessores como os 10 Princípios da Preservação Digital (*Ten Principles*) e o catálogo de critérios NESTOR. O checklist de critérios é dividido em três principais seções: 1) Infraestrutura organizacional; 2) Gerenciamento do objeto digital, e; 3) Tecnologias, infraestrutura técnica e segurança. Outros padrões e ferramentas de avaliação de repositórios que se seguiram também utilizaram esse mesmo formato de seções para apresentarem os seus critérios.

A primeira seção avalia desde os aspectos das políticas de documentos digitais da organização, passando pela verificação da adequação da sua estrutura quanto às responsabilidades e processos existentes, a comprovação da sustentabilidade financeira do repositório e a verificação de existência de aspectos de governança e transparência.

A segunda seção se conecta diretamente às entidades funcionais do modelo OAIS e mescla aspectos organizacionais e técnicos, analisando os processos existentes na Ingestão (*Ingest*), Arquivamento (*Archival*) e Difusão (*Dissemination*). Em termos gerais, é avaliado o tratamento do repositório sobre os pacotes SIP, AIP e DIP, analisando-se a qualidade das informações, registradas em metadados, e como eles são aderentes às políticas de arquivo adotadas na organização.

A terceira seção trata dos aspectos técnicos que o repositório deve implementar a fim de garantir a preservação dos documentos a longo prazo, assim como as suas respectivas interpretabilidades. Nessa seção encontram-se critérios que verificam, por exemplo, a adequação dos *hardwares*, *softwares* e mecanismos que evitam o corrompimento de dados, assim como os meios utilizados para evitar acessos indevidos aos documentos ou que possam ferir as permissões de acesso.

⁴⁸ Em 2006, o RLG foi incorporado ao OCLC (*Online Computer Library Center*, Inc)

Juntamente com a elaboração do TRAC, foram iniciados os trabalhos para tornar os critérios um padrão internacional ISO, culminando na ISO 16363:2012. Conscientes do tempo necessário para esse processo da normalização, Estados Unidos, Alemanha e Reino Unido desencadearam processos paralelos para a implementação de certificações em seus países e territórios.

Ainda em fevereiro de 2007, como fruto de um trabalho colaborativo entre o DCC e DPE, foi lançado o DRAMBORA (*Digital Repository Audit Method Based on Risk Assessment*), como um instrumento (*toolkit*) de avaliação do repositório a partir da identificação dos riscos envolvidos, seus impactos, probabilidade de ocorrência e medidas projetadas para evitá-los ou mitigá-los. A aplicação do *toolkit* no repositório “[...] procura determinar se o repositório fez todos os esforços para evitar e conter os riscos que possam impedir sua capacidade de receber, curar e fornecer acesso a informações digitais autênticas e contextualmente, sintaticamente e semanticamente compreensíveis”. (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 11, tradução nossa).

A percepção dos riscos envolvidos na preservação digital dá aos administradores do repositório uma visão ampla dos seus pontos fortes e aspectos a serem trabalhados, fornecendo um panorama que pode ser utilizado como *benchmark* de evolução do repositório, em um processo contínuo de avaliação. Ainda, essa análise de riscos pode servir como embasamento na captação de recursos, frente aos patrocinadores do repositório, para a melhoria de pontos chave. Quanto ao gerenciamento de riscos, é dito (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 20, tradução nossa):

Hoje em dia, a preservação digital é muitas vezes definida como um exercício de gestão de risco em que o objetivo é converter a incerteza sobre a manutenção da usabilidade de objetos digitais autênticos em riscos quantificáveis. O objetivo de um repositório digital é fazer todo o possível para mitigar os riscos que impedem sua capacidade de fornecer acesso a informações digitais autênticas.

O DRAMBORA apresenta uma metodologia para a identificação dos riscos no repositório digital que são então catalogados e avaliados. A partir desse primeiro trabalho, os riscos são avaliados quanto ao seu grau de impacto no funcionamento das atividades do repositório, incluindo a manutenção da autenticidade dos materiais, assim como a probabilidade de ocorrência. Baseado nessa análise, são definidas as ações a serem tomadas para se evitar o risco e, na ocorrência deles, as medidas a serem tomadas para o retorno à situação normal de funcionamento. Ainda que o DRAMBORA torne claro que a análise de riscos deve levar em conta a realidade e contexto individuais de cada repositório, é fornecido no Apêndice 3 um catálogo de 78 exemplos de riscos de repositórios digitais que pode ser utilizado como ponto de partida para o trabalho de

avaliação (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 134).

Em 2008, o DPE elaborou o *Planning Tool for Trusted Electronic Repositories* (PLATTER) com o objetivo de fornecer uma ferramenta aos repositórios para a verificação de conformidade dos 10 princípios de preservação digital⁴⁹ elaborados no ano anterior com a participação do próprio DPE. O PLATTER não se denomina um instrumento de certificação, mas a sua aplicação favorece a aplicação de posteriores submissões a auditorias e certificações, dentre as (DIGITALPRESERVATIONEUROPE, 2008, p. 6, tradução nossa):

Um repositório planejado utilizando PLATTER ficará numa posição forte quando vier posteriormente a aplicar um dos instrumentos de auditoria existentes para confirmar a adequação dos seus procedimentos para manter a usabilidade a longo prazo do seu material e o acesso ao mesmo.

Dentre as ferramentas de auditoria que o PLATTER recomenda aos repositórios está o DRAMBORA, não coincidentemente, elaborado pelo próprio DPE. A primeira parte do PLATTER possui questões que buscam compreender o repositório em si, analisando o seu propósito, meios legais envolvidos, o seu tamanho em relação aos dados e equipes, métodos de aquisição de dados, interoperabilidade e critérios técnicos a respeito do armazenamento dos dados.

O PLATTER é formado por 9 grandes Planos de Objetivos Estratégicos (*Strategic Objective Plans – SOP*) que se subdividem em objetivos e que devem preferencialmente estar em conformidade com os requisitos SMART⁵⁰. Assim, o resultado da análise do repositório tende a se tornar mais palpável e concreto, facilitando a criação de *benchmarks* de progresso entre avaliações subsequentes do próprio repositório ou entre repositórios diferentes. A seguir são apresentados os SOP que compõem do PLATTER (DIGITALPRESERVATIONEUROPE, 2008, p. 20):

1. Plano de Negócios (*Business Plan*): se refere à forma como o repositório se posiciona como negócio ao realizar a preservação digital, definindo, principalmente as suas fontes de financiamento e mecanismos de monitoramento para a identificação de eventuais dificuldades financeiras;
2. Plano de Pessoal (*Staffing Plan*): plano que trata das equipes envolvidas, definindo as responsabilidades adequadamente, assim como buscando compreender as habilidades e conhecimentos necessários para as equipes realizarem os seus trabalhos. Deve possuir metas que possibilitem a verificação de cumprimento de participações em cursos e capacitações, por exemplo;

⁴⁹ (CENTER FOR RESEARCH LIBRARIES, 2007)

⁵⁰ Acrônimo em inglês: *Specific* (Específico), *Measurable* (Mensurável), *Attainable* (Atingível), *Realistic* (Realista) and *Time-related* (Temporal). Significa que o objetivo deve ser específico, possuir uma métrica que valida o sucesso, ser realista e possuir um período de tempo para ser executado e verificado.

3. Plano de Dados (*Data Plan*): estabelece o tratamento que o repositório realiza nos dados e metadados nos pacotes SIP, AIP e DIP. Pode definir a periodicidade para a execução de procedimentos de preservação, como a verificação de vírus, por exemplo;
4. Plano de Aquisição (*Acquisition Plan*): se refere à forma como os objetos digitais são depositados no repositório, estabelecendo padrões e formatos mínimos aceitáveis e aspectos legais envolvidos, assim como os procedimentos de monitoramento;
5. Plano de Acesso (*Access Plan*): o repositório deve possuir uma política de acesso aos objetos preservados refletindo as demandas identificadas da sua Comunidade-Alvo. Deve-se estabelecer um plano que reflita essa demanda e mantenha um monitoramento contínuo das necessidades de acesso da Comunidade-Alvo;
6. Plano de Preservação (*Preservation Plan*): deve existir um plano de preservação que guie todas as estratégias de preservação do repositório, buscando antecipar demandas e fragilidades que podem surgir no futuro, nem sempre tão próximo. Devem existir mecanismos de monitoramento das tecnologias envolvidas e formatos de arquivos utilizados na preservação, visando garantir o acesso aos objetos pelo tempo que for necessário;
7. Plano do Sistema Técnico (*Technical System Plan*): um dos planos mais importantes do repositório, visto a total dependência em tecnologia para realização da preservação dos objetos a longo prazo. O plano deve especificar os requisitos de hardware e software necessários para a manutenção das suas atividades, além de apresentar as medidas adotadas para prevenir acessos indevidos e corrompimento dos seus dados.
8. Plano de Sucessão (*Succession Plan*): o repositório deve contar com um plano que garanta a preservação dos seus objetos mesmo que ele mesmo deixe de existir. O plano deve conter os requisitos mínimos do repositório que vai receber os seus objetos, especificando, por exemplo, tipos de licenças necessárias, formatos de dados e metadados aceitos.
9. Plano de Desastre (*Disaster Plan*): cabe ao repositório elaborar um plano a ser posto em prática em casos de desastres que ameacem a continuidade do funcionamento do repositório. Como apoio à elaboração do plano, deve-se usar ferramentas de análises e gerenciamento de riscos. Salienta-se que além dos desastres naturais que podem acometer o repositório, deve-se prever aqueles relativos a drásticas mudanças tecnológicas e perdas de pessoas chave das equipes, por exemplo.

Um aspecto interessante a ser destacado é diferenciação entre os Planos de Desastre e de Sucessão. O PLATTER inclui como desastre quaisquer problemas financeiros, políticos ou até mesmo de mudança do ambiente de negócios, por exemplo, classificando-os como riscos a serem identificados e avaliados. Assim o Plano de Desastre deve incluir mecanismos de monitoramento desses riscos que permitam ao repositório evitá-los ou antecipar-se a eles tomando as decisões mais adequadas.

Em contra partida, o Plano de Sucessão é apresentado sob a ótica de que mesmo que se empreguem as melhores técnicas de preservação atuais, elas podem não ser suficientes para um horizonte de preservação a (muito) longo prazo. Assim, o plano deve ser revisado periodicamente buscando identificar parcerias com outros repositórios que podem herdar os dados em caso de cessamento das suas atividades, mesmo sabendo-se que as negociações desses acordos de sucessão possam ser difíceis (DIGITALPRESERVATIONEUROPE, 2008, p. 37).

Ao final da primeira década do século XXI, em novembro de 2009, o *Working Group Trusted Repositories-Certification* do NESTOR publicou a versão 2 do seu catálogo de critérios para repositórios digitais confiáveis. Essa foi a versão final dos critérios publicada que, embora tenha a mesma quantidade de critérios que a primeira versão, 54 ao todo, incorpora uma visão mais amadurecida dos conceitos que permeiam a preservação digital, fruto da participação dos membros do grupo de trabalho em outras iniciativas como o *Ten Principles*⁵¹, TRAC⁵² e DRAMBORA⁵³, por exemplo. Essa é a versão utilizada até os dias atuais para a auditoria de critérios e certificação.

Em 2010 foi lançado o *Data Seal of Approval* (DSA), como resultado do trabalho iniciado ainda em 2008 pela *Data Archiving and Networked Services* (DANS). Essa organização originou-se da demanda de duas outras organizações holandesas, a *Royal Netherlands Academy of Arts and Sciences* (KNAW) e a *Dutch Research Council* (NWO), com o objetivo de “[...] criar um selo de aprovação de dados que ajude a garantir que os dados arquivados ainda possam ser encontrados, compreendidos e utilizados no futuro” (CORETRUSTSEAL, 2023).

O DSA era composto por 16 diretrizes para a preservação digital e alinhava-se a outras diretrizes e certificações da época como o Nestor Seal, TRAC e DRAMBORA, por exemplo. Segundo L’Hours, Kleemola e Leeuw (2019, p. 3, tradução nossa), “Ao extrair o conjunto mínimo de requisitos dessas fontes, o *Data Seal of Approval* procurou garantir uma alta qualidade e uma gestão confiável dos dados para o futuro sem exigir a implementação de novas normas, regulamentos ou investimentos pesados”. Conforme o resultado da avaliação das diretrizes que perpassam a produção, o arquivamento e o consumo dos dados digitais, os repositórios obtinham ou

⁵¹ (CENTER FOR RESEARCH LIBRARIES, 2007)

⁵² (ONLINE COMPUTER LIBRARY CENTER, 2007)

⁵³ (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007)

não a certificação que poderia ser exibida em suas publicações e meios de comunicação. Como passo inicial, o DANS recomendava que o repositório realizasse uma autoavaliação de atendimento dos critérios a fim de perceber o seu grau de maturidade. A cada critério era dada uma resposta que indicava o grau de atendimento que variava desde o repositório ainda não ter considerado o requisito até o seu atendimento por completo na prática. Para uma melhor validação da auditoria, ela deveria ser submetida à avaliação de especialistas da área de preservação digital, buscando uma visão externa, porém não há a necessidade de submissão a um corpo formal de certificação. Assim que o repositório julgava adequado, poderia submeter a sua avaliação aos auditores do DANS que analisavam os dados e documentos que fundamentavam o atendimento de cada critério, culminando na obtenção do selo DSA. Em 2013, iniciou a parceria entre o DSA e o *World Data System* (WDS), órgão do Conselho Internacional para Ciência (*International Council for Science* – ICS) que culminou na primeira versão da certificação CoreTrust Seal em 2017 e substituiu o *Data Seal of Approval*.

Em julho de 2010, com o intuito de se criar um entendimento comum e colaborativo quanto às certificações de confiança dos repositórios, houve a criação do grupo *European Framework for Audit and Certification of Digital Repositories*, composto por membros do CCSDS⁵⁴, DSA Board⁵⁵ e DIN (*Deutsches Institut für Normung* – Instituto Alemão de Normalização)⁵⁶. O grupo elaborou e assinou o documento *Memorandum of Understanding* (MoU) (GIARETTA, HARMSSEN e KEITEL, 2010) criando três possíveis níveis de certificação: *Core Level* (Nível Básico ou Essencial), *Extended Level* (Nível Estendido) e *Formal Level* (Nível Formal).

No *Core Level*, os repositórios devem ter obtido a certificação *Data Seal of Approval* (DAS), que mais tarde se tornou CoreTrust Seal. No *Extended Level*, os repositórios devem possuir a certificação Core e ainda ter realizado uma autoavaliação baseada na ISO 16363 ou DIN 31644 revisada por membros externos e estar publicamente disponível. No *Formal Level*, os repositórios, além de terem cumprido os requisitos Core devem possuir uma certificação realizada por uma auditoria externa baseada na ISO 16363 ou DIN 31644.

O nível de certificação que cada repositório vai obter vai depender do seu grau de organização, disponibilidade e capacidade das equipes, qualidade da auditoria interna e disponibilidade financeira. Mesmo nas auditorias internas, o processo pode se tornar custoso devido ao tempo envolvido em capacitações para que aplicação da certificação ocorra corretamente. Como

⁵⁴ David Giaretta como presidente do CCSDS, comitê responsável pelo ACTDR.

⁵⁵ Henk Harmsen como presidente do DSA Board, responsável pela certificação Data Seal of Approval.

⁵⁶ Christian Keitel como presidente do grupo de trabalho “Trusted Archives – Certification” do DIN e membro do NESTOR.

consequência dessa abordagem, a auditoria pode ter resultados enviesados que não refletem a verdadeira realidade do repositório.

No caso das auditorias externas, existem os custos de contratação dos corpos auditores, além da necessidade de criação de equipes que vão reunir e fornecer os dados necessários para a realização auditoria. Essa abordagem se torna mais imparcial que a realizada internamente, uma vez que o corpo avaliador não integra a equipe do repositório, fornecendo uma visão desconectada e independente sobre os processos adotados. Esse nível de certificação estabelece um procedimento de avaliação comum entre instituições, permitindo a comparação entre elas e possíveis colaborações mútuas para o alcance ou melhoria de requisitos comuns.

Ainda em 2010, na Espanha, é lançada a primeira versão do *Guía para la evaluación de repositorios institucionales de investigación*, com a participação de autores de diversas universidades espanholas e fruto de uma parceria entre a FECYT (*Fundación Española para la Ciencia y la Tecnología*) e REBIUN-CRUE (*Red de Bibliotecas Universitarias – Conferencia de Rectores de las Universidades Españolas*) e tem como objetivo “[...] ser uma ferramenta de auditoria interna para melhorar a qualidade dos repositórios, quer se encontrem já em funcionamento (detectando os seus pontos fortes e fracos) quer em fase de planeamento (fornecendo diferentes aspetos a ter em conta)” (BARRUECO CRUZ, CABALLOS VILLAR, *et al.*, 2010, p. 6, tradução nossa).

O guia, contendo 32 critérios de avaliação, avalia aspectos do repositório em diferentes níveis, como a sua visibilidade perante a sua comunidade, suas políticas implantadas, aspectos legais cumpridos, metadados recolhidos e disponibilizados, diretrizes DRIVER⁵⁷, logs e estatísticas e segurança, autenticidade e integridade dos dados. Os critérios aceitam apenas respostas “Sim” ou “Não”, sinalizando o cumprimento ou não dos requisitos apresentados. Em caso de atendimento, o respondente deve anexar os documentos que comprovem esse fato.

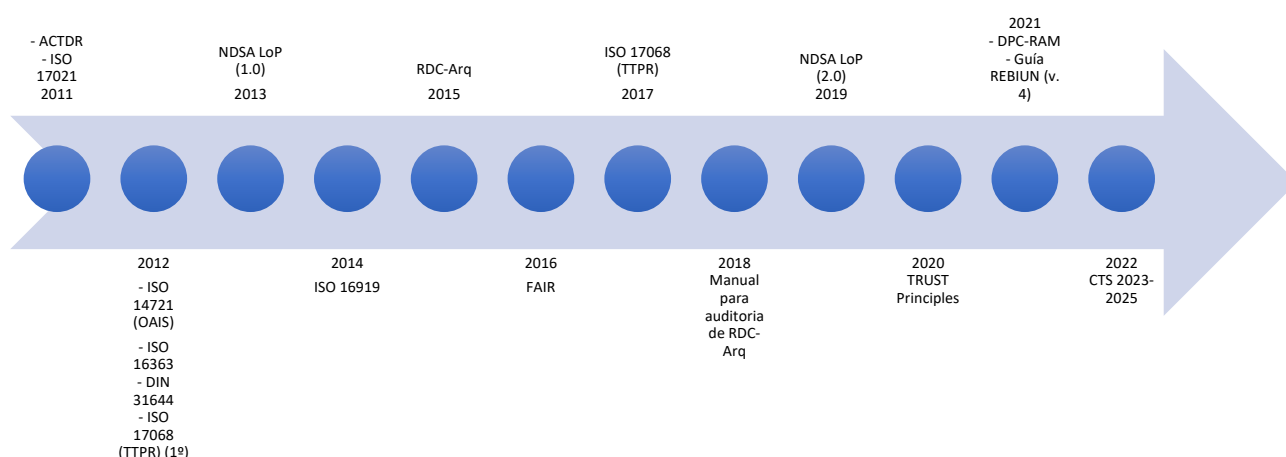
3.3.2 O Período de 2011 a 2025

A partir da segunda década dos anos 2000, surgem novas ferramentas e métodos relacionados às auditorias e certificações de repositórios, buscando reforçar os critérios de confiabilidade. A Figura 11 exibe os principais marcos dessa área a partir de 2011.

⁵⁷ *Digital Repository Infrastructure Vision for European Research*. Projeto criado para estabelecer uma rede comum de repositórios europeus de pesquisa para a criação de serviços de busca e localização de conteúdos globais. As diretrizes se baseiam principalmente na presença nos repositórios de serviços com o protocolo OAI-PMH, possuindo graus de conformidade de acordo com o atendimento de critérios obrigatórios ou recomendados. (BARRUECO CRUZ, CABALLOS VILLAR, *et al.*, 2010, p. 11).

Em setembro de 2011, o CCSDS lança a publicação *Audit and Certification of Trustworthy Digital Repositories* (ACTDR) (CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS, 2011) como prática recomendada para a auditoria e certificação de repositórios digitais, necessidade já levantada na primeira versão da ISO 14721 de 2002, buscando efetivamente certificar os repositórios em conformidade com o modelo OAIS e buscando terminar com as autodeclarações de conformidade dos repositórios. Não obstante, a publicação já possuía a intenção de se tornar um padrão ISO posteriormente e teve como base o TRAC (ONLINE COMPUTER LIBRARY CENTER, 2007) e o *TDR: Attributes and Responsibilities* (RESEARCH LIBRARIES GROUP, 2002).

Figura 11 – Linha do tempo da preservação digital (a partir de 2011).



Fonte: elaboração própria.

O ACTDR possui critérios que avaliam as três grandes áreas do repositório: 1) infraestrutura organizacional; 2) Gestão do objeto digital, e; 3) Gestão de riscos de infraestrutura e segurança. A primeira área possui requisitos que avaliam a organização como um todo, verificando a sua estrutura, governança, sustentabilidade financeira e cumprimento das obrigações legais aplicáveis. Essa área, termos gerais, valida a vocação e pré-disposição do repositório em realizar a preservação de materiais digitais.

A segunda área analisa o tratamento dos pacotes OAIS SIP, AIP e DIP, utilizados, respectivamente, na submissão, arquivamento e difusão dos materiais, verificando se as políticas arquivísticas estão sendo aplicadas na prática. São analisados requisitos técnicos como a utilização de formatos de arquivos sustentáveis e a presença de metadados que garantem a proveniência dos

materiais antes do arquivamento, assim como metadados ligados aos procedimentos de preservação em si.

A terceira área analisa a infraestrutura técnica do repositório e a gestão de risco empregada, procurando identificar se os *hardwares* e *softwares* utilizados de fato se adequam à preservação digital e possuindo monitoramentos que identificam as trocas e atualizações necessárias. Como aspecto central, é analisada a forma como o repositório se previne do corrompimento dos dados dos AIPs, o registro dos episódios detectados e as estratégias empregadas para reverter a situação ocorrida. Ainda nessa seção, é analisada a gestão de riscos quanto à segurança do repositório, verificando se os papéis e responsabilidades da equipe são bem definidos e executados, além de se certificar que os controles de acesso físico e dos sistemas é funcional e efetivo, evitando acessos indevidos e invasões.

Ainda no ano de 2011, foi publicada a primeira versão da norma ISO 17021, que especifica os requisitos necessários a serem cumpridos pelos órgãos que realizam auditorias e certificações. A norma se destina às auditorias e certificações de sistemas de gestão ambiental, de qualidade e segurança da informação, por exemplo e, mais tarde em 2009 serviu de base para a criação da norma ISO 16919, que fala especificamente dos órgãos de auditoria e certificação de repositórios digitais confiáveis.

Em 2012, houve o lançamento da segunda versão da norma ISO 14721, elaborada e lançada como prática recomendada pelo CCSDS e que contou com atualizações de termos, definições e diagramas. A própria definição de Arquivo OAIS sofreu modificação, assumindo que ele pode ser parte de uma organização maior (CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS, 2012, p. 1, tradução nossa, grifo nosso): “Um Arquivo OAIS, consiste em uma organização, que pode fazer parte de uma organização maior, de pessoas e sistemas que aceitou a responsabilidade de preservar a informação e torná-la disponível para uma Comunidade Designada”. Foi adicionado ainda o conceito de Propriedade de Informação Transformacional, que auxilia na determinação da autenticidade dos objetos preservados. Essa propriedade se refere às informações que devem ser preservadas na representação da informação quando são executadas operações de preservação irreversíveis, fornecendo meios para que seja possível constatar que o novo objeto corresponde semanticamente ao anterior.

Em fevereiro de 2012, a prática recomendada 652.0-M-1 do CCSDS, conhecida como ACTDR (CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS, 2011) foi adotada como padrão ISO 16363 utilizada para a auditoria e certificação de repositórios digitais confiáveis, mantendo a mesma estrutura, critérios e metodologia. Nesse mesmo ano, houve o lançamento da primeira versão da norma ISO 17068 (*Trusted Third Party Repository (TTPR) for Digital Records* –

Repositório de Terceiros Confiável para Documentos de Arquivo Digitais), buscando contemplar a realidade da terceirização de repositórios para a custódia de documentos digitais e “detalha os serviços de custódia autorizados de um TTPR para garantir integridade e autenticidade comprováveis dos registros digitais dos clientes e servir como fonte de evidência confiável” (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012).

Em abril de 2012 na Alemanha, foi publicada a norma DIN 31644 (DEUTSCHES INSTITUT FÜR NORMUNG, 2012), que estabelece critérios para repositórios digitais confiáveis e tem como base central o catálogo de critérios NESTOR (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009). Assim, o país passa a contar com uma norma a ser usada para certificações dos seus repositórios, utilizando o trabalho desenvolvido e amadurecido do NESTOR. Nos termos dos níveis de certificação, a norma DIN 31644 pode ser usada no Nível Estendido (*Extended Level*), quando o repositório realiza a autoavaliação ou no Nível Formal (*Formal Level*), quando o repositório submete a sua avaliação a um corpo externo avaliador.

Em 2013, houve o lançamento da versão 1 do *NDSA Levels of Digital Preservation* (NDSA LoP) (PHILLIPS, BAILEY, *et al.*, 2013). A publicação em forma de matriz possui 5 categorias de áreas da preservação digital: 1) Armazenamento e localização geográfica; 2) Fixidez dos arquivos e integridade dos dados; 3) Segurança da informação; 4) Metadados, e; 5) Formatos de arquivo. Cada categoria possui 4 níveis para soluções e processos repositórios que vão evoluindo: Nível 1 (Proteger os seus dados), Nível 2 (Conhecer os seus dados), Nível 3 (Monitorar os seus dados) e Nível 4 (Reparar os seus dados). A ferramenta se define como (PHILLIPS, BAILEY, *et al.*, 2013, p. 1, tradução nossa):

um conjunto de recomendações sobre como as organizações devem começar a construir ou aprimorar suas atividades de preservação digital. Um trabalho em andamento, pretende ser um conjunto de diretrizes relativamente fácil de usar, útil não apenas para aqueles que estão começando a pensar em preservar seus ativos digitais, mas também para instituições que planejam os próximos passos para aprimorar seus sistemas de preservação digital e fluxos de trabalho existentes.

Os níveis foram elaborados para serem independentes de tecnologias, tipo de conteúdo e formatos de arquivo, orientando ações práticas a serem tomadas pelos repositórios para agregar qualidade à preservação digital e diminuir os riscos de perdas de dados. A intenção foi construir uma matriz que aplicasse boas práticas de preservação e fosse de fácil entendimento, ao mesmo tempo em que é menos burocrática que as certificações para repositórios digitais confiáveis da época, como o TRAC e ACTDR, por exemplo.

Em março de 2014, a segunda versão da prática recomendada 652.1-M-2 do CCSDS, foi publicada, sob o nome *Requirements for Bodies Providing Audit and Certification of Candidate Trustworthy Digital Repositories* (Requisitos para Órgãos de Auditoria e Certificação de

Repositórios Digitais Confiáveis). Nesse mesmo ano, a prática recomendada foi adotada como a norma ISO 16919. A norma busca estabelecer as competências e posturas necessárias dos auditores de modo a garantir uma auditoria confiável, justa, transparente, imparcial e totalmente de acordo com a ISO 16363. De um modo geral, a norma se baseia na ISO/IEC 17021 que também trata de órgãos de auditoria para sistemas de qualidade, porém inclui requisitos específicos para os repositórios digitais. A norma chama a atenção para as auditorias e certificações de repositórios feitas por órgãos que não estejam em conformidade com os seus preceitos (CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS, 2014, p. B-2, tradução nossa):

Um repositório digital que seja auditado e certificado por um órgão não conforme com esta Prática Recomendada do CCSDS pode correr o risco de ter um certificado que não inspira confiança aos seus usuários. Também corre o risco de que quaisquer dados confidenciais revelados durante a auditoria possam estar sujeitos a uso indevido.

Em setembro de 2015, o Conselho Nacional de Arquivos do Brasil (CONARQ) publicou a Resolução Nº 43 estabelecendo diretrizes para a implementação de Repositórios Arquivísticos Digitais Confiáveis (RDC-Arq) no âmbito das entidades participantes do Sistema Nacional de Arquivos (SINAR), que inclui, entre outros, o Arquivo Nacional brasileiro e os arquivos dos poderes Executivo, Legislativo e Judiciário federais. O objetivo da resolução é: (CONSELHO NACIONAL DE ARQUIVOS, 2015)

Indicar parâmetros para repositórios arquivísticos digitais confiáveis, de forma a garantir a autenticidade (identidade e integridade), a confidencialidade, a disponibilidade, o acesso e a preservação, tendo em vista a perspectiva da necessidade de manutenção dos acervos documentais por longos períodos de tempo ou, até mesmo, permanentemente.

Dentre as referências utilizadas na elaboração do documento, pode-se citar o modelo OAIS (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012), o *TDR: attributes and responsibilities* (RESEARCH LIBRARIES GROUP, 2002), TRAC (ONLINE COMPUTER LIBRARY CENTER, 2007) e a ISO 16363 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012), por exemplo. De uma forma geral, a sua estrutura segue o mesmo modelo do TRAC e da ISO 16363, contendo uma seção para os requisitos da infraestrutura organizacional, uma seção para o gerenciamento dos documentos digitais e uma seção para a infraestrutura técnica, tecnológica e segurança do repositório. O trabalho do CONARQ engloba muitos dos conceitos das normas internacionais, porém busca se adaptar à realidade brasileira, integrando-se com padrões e definições utilizados em âmbito nacional como, por exemplo, NOBRADE (Norma Brasileira de Descrição Arquivística) (CONSELHO NACIONAL DE ARQUIVOS, 2006) e SIGAD (CONSELHO NACIONAL DE ARQUIVOS, 2022).

Em 2016, Wilkinson et al (WILKINSON, DUMONTIER, *et al.*, 2016), a partir de uma iniciativa entre diversos órgãos, instituições e empresas, publicaram os princípios FAIR, acrônimo de *Findability, Accessibility, Interoperability, and Reusability*, buscando estabelecer um conjunto de boas práticas para promover principalmente o reuso de dados científicos. Essas práticas visam tornar os dados e documentos Localizáveis (*Findability*), Acessíveis (*Accessibility*), Interoperáveis (*Interoperability*) e Reusáveis (*Reusability*) promovendo uma maior difusão e reuso das informações como um todo.

Em 2017, após um processo de revisão, foi lançada a versão vigente da norma ISO 17068 (*Trusted Third Party Repository (TTPR) for Digital Records* – Repositório de Terceiros Confiável para Documentos de Arquivo Digitais com os requisitos que fundamentam a contratação de repositórios e garantem que os serviços prestados e documentos fornecidos sejam considerados confiáveis entre todas as partes envolvidas.

Em 2018, como resultado da sua Tese de Doutorado, Henrique Machado dos Santos publicou o Manual para Auditoria de Repositórios Arquivísticos Digitais Confiáveis (SANTOS, 2018) fornecendo uma visão arquivística do modelo OAIS e dos critérios ACTDR ao mesmo tempo em que os aproxima do RDC-Arq e contextualiza com a cadeia de custódia. A publicação não tem como objetivo ser um instrumento de auditoria e/ou certificação, mas sim ser utilizado como apoio no planejamento e implementação de repositórios. Adicionalmente, pode-se usá-la para a melhoria ou adequação dos repositórios arquivísticos quanto aos critérios das normas OAIS e ACTDR e os requisitos para o estabelecimento de um RDC-Arq.

Em outubro de 2019 foi lançada a versão 2.0 do NDSA *Levels of Digital Preservation* com modificações pontuais nos nomes das áreas funcionais e dos níveis, porém mantendo a mesma intenção da versão anterior de ser uma ferramenta prática de análise e planejamento de ações do repositório (LEVELS OF PRESERVATION REVISIONS WORKING GROUP, 2019).

Em 2020, um grupo vasto de pesquisadores, incluindo aqueles envolvidos na criação do CoreTrustSeal, elaborou os princípios TRUST estabelecendo diretrizes que permitam aos repositórios adquirir confiança das suas comunidades demonstrando que possuem métodos adequados à preservação (LIN, CRABTREE, *et al.*, 2020). O nome TRUST vem do acrônimo em inglês para *Transparency* (Transparência), *Responsability* (Responsabilidade), *User focus* (Foco no usuário), *Sustainability* (Sustentabilidade) e *Technology* (Tecnologia).

O princípio da Transparência (*Transparency*) diz respeito ao repositório deixar claro os serviços que presta, tempo mínimo de preservação e requisitos a serem cumpridos pelos interessados em submeter seus documentos à preservação. A Responsabilidade (*Responsability*)

significa que o repositório deve estabelecer diretrizes e processos que garantam a autenticidade dos dados sob sua guarda.

Foco no usuário (*User focus*) significa projetar o repositório de acordo com os requisitos da sua Comunidade-Alvo, procurando compreender as suas necessidades de preservação, acesso e interpretação dos dados. O princípio da Sustentabilidade (*Sustainability*) trata sobre o repositório garantir que vai manter os seus serviços disponíveis para cumprir com a preservação no prazo que se comprometeu. Por último, o princípio da Tecnologia (*Technology*) é sobre o repositório possuir e se comprometer a prover toda a infraestrutura tecnológica necessária para realizar a tarefa da preservação e acesso digital dos documentos nele depositados. Em geral, as certificações de confiança de repositórios visam verificar o cumprimento dos princípios TRUST, sendo a quantidade de critérios e a profundidade de análise deles variável entre cada uma delas.

Em março de 2021, foi disponibilizada a segunda versão do *Digital Preservation Coalition Rapid Assessment Model* (DPC RAM) (DIGITAL PRESERVATION COALITION, 2021) com o objetivo de ser uma ferramenta de análise rápida que busca identificar a capacidade e maturidade dos repositórios em realizar a preservação digital. De acordo com DPC (2021, p. 2, tradução nossa), “o modelo fornece um conjunto de recursos organizacionais e de nível de serviço que são classificados em um conjunto simples e consistente de níveis de maturidade”. Dentre as referências usadas para a implementação do guia, estão o *CoreTrust Seal* e o *NDSA Levels of Digital Preservation*. A primeira versão da ferramenta havia sido lançada em setembro de 2019.

O DPC define o termo “Arquivo Digital” como “[...] uma instalação onde o conteúdo em formato digital com valor duradouro é armazenado e gerenciado para preservação a longo prazo” (DIGITAL PRESERVATION COALITION, 2021, p. 5, tradução nossa). Embora diferente, a definição se assemelha bastante com a de um Arquivo OAIS, explicitando a necessidade de haver uma instituição, órgão ou instalação que se comprometa em realizar efetivamente a preservação a longo prazo.

Quadro 6 – Elementos de avaliação de maturidade do DPC RAM.

a) Capacidades Organizacionais	
–	Viabilidade organizacional: analisa a governança, estrutura organizacional, equipe e meios de financiamento do repositório;
–	Política e estratégia: analisa as políticas, estratégias e procedimentos envolvidos na gestão do repositório;
–	Aspectos legais: verifica o cumprimento e adesão do repositório quanto à base legal envolvida na preservação digital, que geralmente varia conforme a localização do repositório;
–	Capacidade de TI: analisa o quanto a TI destinada ao repositório suporta os aspectos envolvidos na preservação digital;
–	Melhoria contínua: avalia os processos implantados no repositório e se eles apresentam metas e monitoramento do progresso de alcance dessas metas;
–	Comunidade: avalia o engajamento e a contribuição do repositório com a preservação digital na comunidade.
b) Capacidades de Serviço	
–	Aquisição, transferência e ingestão: avalia a forma com os processos de aquisição, transferência e ingestão estão implantado no repositório, verificando a curadoria realizada a uso de ferramentas automatizadas;

- Preservação dos bits: avalia os processos implantados que se referem efetivamente à preservação a longo prazo dos bits dos conteúdos digitais sob a custódia do repositório;
- Preservação do conteúdo: analisa os meios empregados pelo repositório no sentido de preservar a sua acessibilidade e uso ao longo do tempo;
- Gestão de metadados: verifica se o repositório recebe e mantém metadados suficientes para apoiar a preservação a longo prazo;
- Descoberta e acesso: avalia as iniciativas e mecanismos que o repositório adota para permitir a descoberta do conteúdo preservado e disponibilizá-lo aos seus usuários.

Fonte: (DIGITAL PRESERVATION COALITION, 2021, p. 7, tradução nossa)

A ideia do DPC foi criar um modelo de fácil aplicação, independente de tamanho do repositório, independente de tecnologia e das estratégias específicas de preservação utilizadas pelo repositório. Por se intitular um modelo rápido, os autores sugerem que o modelo pode ser aplicado várias vezes na instituição, tornando-se um instrumento de comparação de resultados e planejamento de ações para a evolução dos aspectos necessários. Assim, ao longo do tempo, o repositório tem condições de traçar a evolução do seu estado quanto à preservação digital. O DPC RAM apresenta onze elementos de avaliação de maturidades, apresentados no Quadro 6.

Os elementos de avaliação da primeira parte do modelo analisam em alto nível a forma como a estrutura organizacional do repositório se apresenta, buscando identificar a maturidade organizacional do repositório além do seu comprometimento e vocação para a preservação a longo prazo. Os elementos de avaliação da segunda parte possuem uma granularidade mais baixa, analisando os detalhes dos mecanismos de preservação utilizados, voltados à operacionalização do repositório.

A cada um dos elementos de avaliação, deve ser atribuído um nível de maturidade que utiliza uma escala de 0 a quatro. O nível zero significa que o repositório não possui nenhum conhecimento ou possui um mínimo conhecimento sobre o tema. Em oposição, o nível quatro de maturidade sinaliza que o repositório cumpre os requisitos e já os aplica de uma forma otimizada.

Em maio de 2021 é lançada a quarta versão do *Guía para la Evaluación de Repositorios Institucionales de Investigación*, em vigência atualmente, contando com 86 critérios que foram atualizados para englobar as últimas normas e diretrizes de boas práticas observadas na comunidade arquivística internacional (AZORÍN, BERNAL, *et al.*, 2021).

Em setembro de 2022, é publicada versão 1.0 dos requisitos CoreTrust Seal a serem válidos no período de 2023 a 2025 (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2022), como parte do processo periódico e programado de revisão. Complementarmente, foram publicados o guia estendido que possui comentários e discussões e um glossário de termos, ambas auxiliam na compreensão dos requisitos tanto por parte das equipes que se candidatam à certificação quanto para os revisores dos requisitos. Em termos gerais, a última versão ainda

preserva a mesma quantidade de critérios da primeira versão de 2010, 16 ao todo, no entanto, algumas modificações foram feitas.

Os níveis de conformidade dos critérios, que antes eram cinco, agora foram substituídos para duas possibilidades: 1) em progresso, e; 2) implementado. O primeiro nível significa que o repositório está trabalhando no requisito e deve apresentar um cronograma de trabalho para o seu cumprimento. O segundo nível sinaliza ao revisor que o repositório possui o requisito totalmente implementado e operacional, sendo necessário fornecer evidências que comprovem esse fato. Caso julgue pertinente, o revisor pode rebaixar um determinado requisito para o nível em progresso, constatando que não houve o seu cumprimento na totalidade. Nos casos de novas reaplicações da certificação de um repositório, é esperado que requisitos que antes estavam em progresso estejam implantados, sinalizando uma evolução do repositório.

Uma revisão da linha do tempo dos principais marcos das certificações e auditorias de repositórios digitais permite apreciar a preocupação com esse tema desde o final do século passado e, sobretudo, a diversidade de enfoques e métodos de avaliação e análise existentes nas fontes identificadas. Como exemplo dessa diversidade, em 1996, Waters e Garret (WATERS e GARRET, 1996), no relatório da *Task Force on Archiving of Digital Information* do *Commission on Preservation and Access* e do *The Research Libraries Group*, apontavam para a necessidade de haver métodos sólidos de certificações e organismos credenciados para a realização das auditorias necessárias a essas certificações. Em 2002, são publicados os atributos e responsabilidades necessários a um repositório digital confiável (RESEARCH LIBRARIES GROUP, 2002), no mesmo ano em que a primeira versão do modelo referencial OAIS é publicada pelo CCSDS e tornada norma internacional ISO 14721 (CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS, 2002).

Em 2007, são publicados os 10 princípios para um repositório confiável (CENTER FOR RESEARCH LIBRARIES, 2007), que serviu como base para todas as certificações, auditorias e ferramentas de medição de maturidade que se seguiram. Nesse mesmo ano, é publicado o *checklist* TRAC (ONLINE COMPUTER LIBRARY CENTER, 2007) como uma ferramenta objetiva para a realização de auditorias e emissão de certificação dos repositórios. Ainda em 2007, é publicado o DRAMBORA (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007), voltado à identificação e avaliação dos riscos inerentes à preservação digital pelos repositórios.

Em 2009, fruto do trabalho do grupo Nestor, na Alemanha, é lançada a certificação NestorSeal que serviria de base para a certificação alemã DIN 31644. Em 2010, é publicado o

Memorandum of Understanding que estabelece os níveis de certificação de repositórios. Nesse mesmo ano, é publicada a primeira versão do *Data Seal of Approval*.

Em 2011 é publicado o ACTDR pelo CCSDS (CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS, 2011) para auditoria e certificação de repositórios baseado na primeira versão do modelo OAIS. Em 2012 é publicada a última versão da ISO 14721 com o atual modelo OAIS, assim como a ISO 16363 para auditoria e certificação de repositórios baseada no ACTDR. Ainda em 2012, é publicada a primeira versão da ISO 17068 sobre *Trusted Third Party Repositories* (TTPR – Repositórios de Terceiros Confiável) e a norma alemã DIN 31644 que estabelece critérios para a determinação de um repositório digital confiável.

Em 2014, o CCSDS publica a prática recomendada 652.1-M-2 estabelecendo diretrizes e requisitos para os órgãos de auditoria e certificação de repositórios digitais (CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS, 2014), sendo nesse mesmo ano transformada na ISO 16919. Em 2017 é publicada a segunda versão, atualmente em vigência, da ISO 17068 que trata dos TTPR.

Em Bonal Zazo e De Lorenzo-Cáceres (2017) é realizada uma comparação de certificações entre o Catálogo de Critérios Nestor, TRAC e ACTDR (ISO 16363), identificando-se que, embora com nomenclaturas diferentes, as três apresentam seus critérios divididos em três grandes áreas, não coincidentemente, semelhantes às grandes áreas do modelo OAIS: a) Infraestrutura da organizacional; b) Gestão do objeto digital, e; c) Infraestrutura técnica e segurança. A área em que predominam mais critérios é a da Gestão do objeto digital (51,4%), seguida pela área da Infraestrutura organizacional (28,6%) e pela área da Infraestrutura técnica e segurança (19,8%).

Dentro da Gestão do objeto digital, o processo de transformação do SIP em AIP é o que mais concentra critérios nas três auditorias e certificações. No caso da ACTDR (ISO 16363), 44,2% dos critérios se concentram nesse processo, enquanto Nestor e TRAC apresentam número muito semelhantes entre si, 34,7% e 33,3%, respectivamente (BONAL ZAZO e DE LORENZO-CÁCERES, 2017). Essa concentração denota o tamanho da importância em se estabelecer critérios bem definidos para a criação do AIP, uma vez que ele será o alvo da preservação, sendo manipulado e modificado à medida em que os processos de preservação em si, vão ocorrendo.

Eventuais falhas na criação do AIP podem acarretar perdas irreversíveis de documentos permanentes, por exemplo, necessitando de uma nova ingestão dos materiais quando houver a possibilidade. Após o recolhimento ao repositório permanente, os produtores dos materiais podem optar por eliminá-los, uma vez que lhes foi garantida a preservação. Daí a importância de haver processos, ferramentas e capacitações adequadas no repositório para a criação e preservação do AIP.

De maneira geral, as três certificações e auditorias avaliam três âmbitos dos repositórios:

1) contexto organizacional e institucional; 2) atividades técnicas de tratamento dos objetos digitais custodiados, e; 3) entorno tecnológico e de segurança. Bonal Zazo e De Lorenzo-Cáceres (2017) ainda destacam: “a análise deste último (entorno tecnológico e de segurança) dos aspectos referidos deve ser abordada não só do ponto de vista da gestão documental, mas também, e sobretudo, do ponto de vista da segurança informática”.

O Quadro 7 resume as ferramentas e normas de auditoria e certificação com as respectivas metodologias de aplicação. É possível notar que em alguns casos, o processo de auditoria é realizado como uma autoavaliação do próprio repositório, servindo para identificar o estado da preservação digital, assim como pode ser um incentivo para processos mais profundos, que requerem a participação de corpos auditores, por exemplo.

Quadro 7 – Ferramentas de auditoria, certificação e maturidade de repositórios digitais.

Ferramenta	Criador	Metodologia
ISO 16363 (ACTDR)	CCSDS	Auditoria externa realizada pelo PTAB.
Nestor Seal (DIN 31644)	Nestor	Autoavaliação por checklist e submissão ao NESTOR para obtenção do selo.
TRAC	RLG (incorporado ao OCLC), NARA e CRL	Autoavaliação por checklist para <i>benchmark</i> ou auditoria do CRL para a formalização.
CoreTrust Seal	DANS e ICS-WDS	Autoavaliação revisada por especialistas e submissão ao CoreTrust Seal Board para obtenção do selo.
DRAMBORA	DCC e DPE	Autoavaliação de riscos. Pode complementar outras ferramentas.
PLATTER	DPE	Autoavaliação através de SOPs (<i>Strategic Objective Plans</i>)
NDSA LoP	NDSA	Autoavaliação e medição do progresso nos níveis.
DPC-RAM	DPC	Autoavaliação rápida do grau de maturidade da preservação digital do repositório através das capacidades organizacionais e de serviço.

Fonte: elaboração própria.

No caso específico da ISO 16363, é requerido que a auditoria seja realizada por um corpo de auditores específico para que a certificação seja alcançada, no entanto, a existência de outras avaliações, mesmo as autoavaliações, tende a facilitar esse processo, servindo como subsídio para os auditores. No caso do DRAMBORA, a avaliação de riscos pode ser utilizada por si só para a demonstração de que o repositório possui medidas para se prevenir, mitigar ou reagir aos riscos, assim como pode ser usado como ponto de partida e complementar outras auditorias e avaliações.

A seguir, alguns dos marcos apresentados na linha do tempo são detalhados, tratando-se de normas, auditorias e certificações relativas à preservação digital.

3.4 IDENTIFICAÇÃO E CARACTERIZAÇÃO DE FERRAMENTAS DE AVALIAÇÃO

Nessa seção do trabalho são apresentadas algumas ferramentas de avaliação e certificação destacadas da linha do tempo da seção anterior. Essas escolhas possuem notadamente grandes influências no desempenho da preservação digital e contribuem significativamente para o estabelecimento da confiabilidade dos repositórios. Inicialmente, na seção 3.4.1, são apresentados os princípios FAIR, que visam agregar, sobretudo, qualidade de uso dos materiais preservados. Na seção 3.4.2 é apresentada a matriz de níveis de preservação NDSA, tendo em uma dimensão, cinco áreas funcionais e, em outra dimensão, quatro níveis de evolução ou sofisticação das soluções implementadas pelos repositórios. A seção 3.4.3 apresenta a norma ISO 16363 que é utilizada, principalmente, para certificar que um determinado repositório é aderente à ISO 14721, que trata do modelo de referência OAIS, apresentado anteriormente na seção 2.3. A seção 3.4.4 aborda a ISO 17068 que se refere à utilização e contratação de repositórios de terceiros, externos às organizações. As seções 3.4.5 e 3.4.6 abordam, respectivamente os selos de confiança Nestor Seal e CoreTrust Seal, ambos com os seus devidos critérios de avaliação dos repositórios. A seção 3.4.7 apresenta o instrumento DRAMBORA que constitui de um método de auditoria que avalia os riscos pelos quais o repositório pode estar exposto, determinando os seus impactos e medidas preventivas ou de mitigação. Por fim, a seção 3.4.8 apresenta o Guia de Avaliação de Repositórios Institucionais de Investigação REBIUN, criado para avaliar os repositórios espanhóis quanto à disponibilização de acesso dos materiais preservados e como os repositórios se conectam e se integram à ferramenta RECOLECTA.

3.4.1 FAIR

Criados para serem usados como um norte de boas práticas no tratamento de dados científicos, os princípios FAIR (*Findability, Accessibility, Interoperability, and Reusability*) buscam a ampla difusão e uso desses dados podendo serem aplicados no âmbito dos documentos como um todo (WILKINSON, DUMONTIER, *et al.*, 2016). Essas práticas visam tornar os dados e documentos Localizáveis (*Findability*), Acessíveis (*Accessibility*), Interoperáveis (*Interoperability*) e Reusáveis (*Reusability*).

Os quatro princípios foram pensados para serem implementados desde os produtores de dados até os publicadores, frequentemente plataformas de difusão, buscando facilitar a sua localização, interpretação e reutilização em diversos estudos além dos que os originaram. Assim, o compartilhamento do conhecimento se torna otimizado e pode ser usado como ponto de partida por diversos pesquisadores.

Cada um dos atributos foi dividido em critérios a serem seguidos pelos próprios dados para que se alcance os princípios FAIR, sendo apresentados no Quadro 8. A especificação dos atributos tem como objetivo que os dados possam ser localizados, reutilizados e possuam a semântica apropriada, focando-se nos papéis dos produtores e custodiadores desses dados. É possível notar ainda que os princípios FAIR são agnósticos de tecnologias e plataformas, mas preza pela utilização de padrões estabelecidos ou padrões *de facto* que sejam amplamente utilizados, promovendo uma maior interoperabilidade entre sistemas. Assim, os princípios podem ser utilizados também na coleta automatizada de informações que utilizam ferramentas específicas para esse fim e subsidiam novas pesquisas.

Quadro 8 – Atributos dos princípios FAIR.

Princípio	Requisitos
Para serem Localizáveis: (<i>To be Findable</i>)	F1. Atribuir identificadores únicos globais e persistentes aos (meta)dados; F2. Descrever os dados com metadados de forma rica (definido no requisito R1); F3. Os metadados claramente e explicitamente incluem os identificadores dos dados que descrevem; F4. Os (meta)dados são registrados ou indexados através de um recurso pesquisável.
Para serem Acessíveis: (<i>To be Accessible</i>)	A1. Os (meta)dados são recuperáveis pelos seus identificadores usando protocolos de comunicação padronizados; A1.1 O protocolo é livre, aberto e universalmente implementável; A1.2 O protocolo permite procedimentos de autenticação e autorização, quando necessário; A2. Os metadados são acessíveis, mesmo quando os dados não são mais disponíveis.
Para serem Interoperáveis: (<i>To be Interoperable</i>)	I1. Os (meta)dados usam uma linguagem formal, acessível, compartilhada e amplamente aplicável para a representação do conhecimento; I2. Os (meta)dados usam vocabulários que seguem os princípios FAIR; I3. Os (meta)dados incluem referências qualificadas para outros metadados.
Para serem Reutilizáveis: (<i>To be Reusable</i>)	R1. Os (meta)dados são ricamente descritos com uma pluralidade de atributos relevantes e precisos; R1.1 Os (meta)dados são entregues com uma licença de uso clara e acessível; R1.2 Os (meta)dados são associados com uma detalhada proveniência; R1.3. Os (meta)dados seguem padrões relevantes de domínio da comunidade.

Fonte: (WILKINSON, DUMONTIER, *et al.*, 2016, p. 4, tradução nossa)

É importante observar que a implantação dos princípios FAIR nos repositórios não necessariamente requer a publicização livre dos dados, mas sim, deve respeitar as licenças de usos, direitos autorais e a proteção de dados quando for necessária. Outro fator importante é enriquecer os metadados com instruções e detalhes sobre o contexto da obtenção e/ou geração dos dados, adicionando-se as condições de reprodução que possam embasar futuros pesquisadores na reutilização dos mesmos. Para Winkinson *et al.* (2016, p. 1, tradução nossa), “o bom gerenciamento de dados não é um objetivo em si, mas é o principal canal que leva à descoberta e inovação do

conhecimento e à subsequente integração e reutilização de dados e conhecimento pela comunidade após o processo de publicação de dados”.

3.4.2 NDSA Levels of Digital Preservation

Sob a forma de um consórcio de instituições norte-americanas comprometidas com a preservação a longo prazo, foi criada a organização NDSA (*National Digital Stewardship Alliance*) em 2010, cuja missão é “estabelecer, manter e aumentar a capacidade de preservar os recursos digitais da nossa nação em benefício das gerações presentes e futuras” (LEVELS OF PRESERVATION REVISIONS WORKING GROUP, 2019, p. 3, tradução nossa). Como forma de fornecer um ponto de partida para as instituições com a intenção de realizar a preservação digital, foi criada a ferramenta chamada *Levels of Digital Preservation – LoP* (Níveis de Preservação Digital), com sua primeira versão publicada em 2013 que foi posteriormente revisada para o lançamento da versão 2.0 em 2019 (LEVELS OF PRESERVATION REVISIONS WORKING GROUP, 2019).

O projeto inicial de criação dos níveis iniciou quando alguns dos membros identificaram a necessidade de haver um guia de práticas acessível tanto para instituições que estavam iniciando o projeto de implementação de um programa de preservação assim como para instituições que já possuíam um programa implantado. A motivação do grupo foi criar uma ferramenta independente de formatos de arquivos utilizados ou plataformas de armazenamento, tornando-a agnóstica do ponto de vista das tecnologias empregadas, tamanho e natureza das instituições. Essas características ampliam o espectro de instituições aptas a utilizar a ferramenta.

A verificação do atendimento de cada um dos níveis permite que a instituição trace um diagnóstico da sua situação, identificando o seu nível atual ao mesmo tempo em que fornece um conjunto de ações que devem ser realizadas para que o próximo nível seja alcançado. À medida em que a instituição progride nos níveis, a sua solução de preservação digital se torna mais sofisticada e menos propensa à perda dos dados sob sua custódia. A Figura 12 exibe a matriz dos níveis, com a dimensão vertical representando as áreas funcionais e a dimensão horizontal representando o nível de conformidade para cada área funcional.

A equipe responsável buscou desenvolver a matriz de forma a possuir detalhes suficientes para a sua compreensão, ao mesmo tempo em que a tornou sucinta a ponto de ser possível colocá-la em uma única página, “[...] um objetivo requerido para melhorar a acessibilidade intelectual do gráfico e desmistificar o conjunto, por vezes desconcertante, de atividades envolvidas na preservação digital” (LEVELS OF PRESERVATION REVISIONS WORKING GROUP, 2019, p.

2, tradução nossa). A própria matriz fornece práticas acessíveis e diretas para o atendimento dos critérios dos níveis ao mesmo tempo em que apresenta exigências menos rígidas do que as apresentadas nas certificações de confiança de repositórios, por exemplo.

É possível notar que as áreas abrangem os aspectos práticos do funcionamento de um repositório digital de preservação, sendo que não há a presença de critérios de avaliação do contexto organizacional da instituição, tornando o modelo aplicável a diferentes tipos e tamanhos de organizações. Cada nível busca verificar diretamente se o repositório possui mecanismos ou processos que atacam diretamente aspectos críticos da preservação digital.

O avanço dos níveis de cada área visa a evolução do tratamento da preservação, iniciando na etapa de conhecimento do conteúdo do seu repositório, no Nível 1. Na área “**Armazenamento**”, são sugeridas ações em um nível de sofisticação inicial, buscando-se a proteção mínima do conteúdo através de duas cópias de *backup*, por exemplo, que não precisam estar necessariamente em localizações geográficas distintas. Na área “**Integridade**”, o conteúdo pode ser recebido sem a presença de informações que sustente a sua integridade, porém nesse caso, o repositório já deve contar com ferramentas que gerem e armazenem essas informações.

A área “**Controle**”, há a necessidade de o repositório mapear os agentes que interagem com o conteúdo conhecendo-se o papel de cada um a fim de se diminuir o conjunto de possibilidades de acessos não autorizados, por exemplo. É importante notar que deve haver o mapeamento tanto dos agentes humanos quanto dos agentes de software que executam ações sobre o conteúdo preservado.

A área “**Metadados**”, não há a necessidade de que o conteúdo recebido possua metadados próprios, mas o repositório deve possuir um inventário do conteúdo sob sua posse armazenando-o em localização diferente do conteúdo em si. A área “**Conteúdo**” é atendida através da verificação e registro da identificação dos formatos de arquivo presentes no repositório, atendendo ao objetivo do Nível 1 que é justamente o conhecer o conteúdo armazenado.

O Nível 2 tem como objetivo principal a proteção do conteúdo, adotando processos mais sofisticados que o nível anterior. Na área “Armazenamento”, é necessário ter no mínimo três cópias de *backup*, sendo uma das cópias em localização geográfica distinta das demais. Na área “Integridade”, há a preocupação com que as verificações de integridade sejam realizadas sempre que houver movimentação ou cópia do conteúdo, buscando-se identificar possíveis corrompimentos o mais cedo possível. Pode-se notar a recomendação para que os dados de integridade sejam armazenados separadamente do conteúdo, evitando-se assim, um ponto único de falha em caso de corrompimento dos dados.

Na área “**Controle**” deve haver o planejamento e execução prática das definições e restrições de uso dos agentes sobre o conteúdo armazenado.

Na área “Metadados”, é verificado se há o recolhimento de informações suficientes para saber do que trata o conteúdo, podendo-se para isso utilizar metadados de diversas naturezas. A área “Conteúdo” trata da verificação de atendimento dos formatos de arquivos recebidos, assim como a sugestão de haver um relacionamento com os produtores de conteúdo para que utilizem formatos sustentáveis, ou seja, formatos que tenham uma probabilidade maior de estarem acessíveis com o passar do tempo (LEVELS OF PRESERVATION REVISIONS WORKING GROUP, 2019, p. 8). Normalmente, adota-se formatos de arquivos abertos, em oposição a formatos de arquivos proprietários, que podem ser descontinuados ou se tornarem obsoletos, levando à inacessibilidade do conteúdo.

No nível 3, as ações da matriz de preservação são relacionadas ao monitoramento do conteúdo preservado. Na área “Armazenamento”, há a preocupação de haver ao menos um *backup* em localização geográfica diferente das demais e com ameaças de desastres diferentes. Mostra-se preocupação com o monitoramento da obsolescência das mídias de suportes observando o tempo de vida útil de cada elemento. Na área “Integridade”, a matriz avalia a existência de processos periódicos de verificação da integridade do conteúdo, documentando-se a execução dos processos em si, que serve como comprovação assim como integra a trilha de auditoria dos conteúdos armazenados.

A área “Controle” analisa a presença de registros (*logs*) que armazenam os detalhes das ações executadas pelos agentes sobre o conteúdo armazenado. A área “Metadados” verifica se o repositório tem estabelecido um padrão de metadados a ser seguido para o registro de informações pertinentes ao seu conteúdo, assim como verifica se existem ações que identificam metadados ausentes e os preenchem.

O Nível 4 representa o grau máximo de sofisticação da matriz de níveis de preservação, quando o repositório deve possuir processos que visam a sustentação do conteúdo preservado. Na área “Armazenamento”, busca-se a máxima diversificação dos mecanismos e tecnologias de suporte, recomendando-se três localizações de armazenamento com localizações geográficas distintas e com ameaças de desastres diferentes entre si. É verificado também se há diversificação quanto aos tipos de mídia utilizados assim como um planejamento que trate e evite a obsolescência das mídias, *hardware* e *software* utilizados na preservação.

A área “Integridade” avalia a presença de processos de verificação de integridade de acordo com a detecção de determinados eventos, sendo estes executados preferencialmente automaticamente. A reparação dos conteúdos corrompidos, sempre que necessária, também deve ser

possível, utilizando as técnicas mais apropriadas. A área “Controle” requer que existam processos periódicos de auditoria dos registros das atividades executadas pelos agentes a fim de que seja possível identificar eventuais ações não autorizadas ou mal intencionadas que possam por em risco o conteúdo preservado.

A área “Metadados” requer que os padrões de metadados escolhidos sejam completamente implementados e estejam presentes no conteúdo, assim, com a utilização desses padrões de metadados para armazenar as ações de preservação realizadas sobre o conteúdo. A área “Conteúdo” analisa se o repositório dispõe de estratégias suficientes que garantam que o conteúdo preservado possa ser acessado futuramente. É analisada a existência de estratégias de preservação como a migração de formatos e a emulação de plataformas, por exemplo.

A avaliação da matriz de níveis de preservação nas instituições pode, muitas vezes, apresentar diferentes níveis atingidos em cada área, refletindo as características únicas de cada repositório. A progressão para níveis mais avançados da matriz vai depender das escolhas das equipes que devem estar alinhadas às respectivas missões dos repositórios e, em muitos casos, pode não haver interesse para implementar as progressões.

De uma forma geral, os níveis de preservação devem ser usados como guia para a identificação do estado atual do repositório e projeto de melhorias na preservação, não representando um conjunto prescritivo de ações a serem seguidas à risca (LEVELS OF PRESERVATION REVISIONS WORKING GROUP, 2019, p. 3). Cada instituição deve avaliar como irá implementar na prática os níveis desejados, ressaltando-se que os níveis não levam em consideração o contexto holístico de um programa de preservação que deve estabelecer políticas de preservação e acesso, recursos, financiamento, etc.

Embora o processo de aplicação dos níveis de preservação em uma instituição não represente uma certificação, pode-se utilizá-lo como comparativo de evolução dos mecanismos de preservação tanto internos quanto com outras instituições, assim como pode servir como um ponto inicial de demonstração de comprometimento aos produtores e consumidores do repositório (BAUCOM, KUSSMANN, *et al.*, 2019, p. 7). Da mesma forma, o resultado da avaliação pode ser usado para a busca de recursos financeiros junto aos financiadores do repositório.

O *Levels of Preservation Revisions Workgroup* (2019, p. 3), grupo responsável pela elaboração da matriz de níveis de preservação e integrante da NDSA, deixa claro que os níveis de preservação devem seguir evoluindo através da observação das demandas da comunidade arquivística internacional. Em 2020, o *Levels of Digital Preservation Working Group* se tornou um vencedor do prêmio *Digital Preservation Award* ao receber o prêmio do Conselho Internacional de Arquivos de colaboração e cooperação (*International Council on Archives Award for Collaboration*

and Cooperation) como forma de reconhecimento ao trabalho realizado para a preservação digital de conteúdos (NATIONAL DIGITAL STEWARDSHIP ALLIANCE, 2022).

Figura 12 – Matriz de Níveis de Preservação NDSA.

Área funcional	Nível 1 (Conheça seu conteúdo)	Nível 2 (Proteja seu conteúdo)	Nível 3 (Monitore seu conteúdo)	Nível 4 (Sustente seu conteúdo)
Armazenamento	Ter duas cópias completas em localizações distintas. Documentar todos o suporte de armazenamento onde o conteúdo está armazenado. Inserir o conteúdo em um suporte de armazenamento estável	Ter três cópias completas com, no mínimo, uma cópia em uma localização geográfica distinta. Documentar o armazenamento e suporte de armazenamento indicando os recursos e dependências requeridas para seu funcionamento	Ter no mínimo uma cópia em uma localização geográfica com ameaça de desastre diferente das outras cópias. Ter no mínimo uma cópia em um tipo de suporte de armazenamento diferente. Monitorar a obsolescência do armazenamento e suas mídias.	Ter, no mínimo, três cópias em localizações geográficas diferentes, cada uma com diferentes ameaças de desastres. Maximizar a diversificação de armazenamento a fim de evitar pontos únicos de falha. Ter um planejamento e executar ações para tratar a obsolescência do <i>hardware</i> de armazenamento, <i>software</i> e mídias.
Integridade	Verificar a informação de integridade caso tenha sido disponibilizada juntamente com o conteúdo. Gerar a informação de integridade caso não tenha sido disponibilizada junto com o conteúdo. Executar varredura de vírus em todo o conteúdo; isolar o conteúdo em quarentena caso seja necessário.	Verificar informação de integridade ao mover ou copiar conteúdo. Usar dispositivos de bloqueio de escrita se trabalha com as mídias originais. Fazer backup da informação de integridade e armazenar a cópia em uma localização distinta do conteúdo.	Verificar a informação de integridade do conteúdo em intervalos fixos. Documentar os processos de verificação de informação de integridade do conteúdo e os respectivos resultados. Executar auditoria de informação de integridade sob demanda.	Verificar informação de integridade em respostas a eventos ou atividades específicas. Substituir ou reparar o conteúdo corrompido quando necessário.
Controle	Determinar os agentes humanos e de software autorizados a ler, gravar, mover e eliminar conteúdo.	Documentar e por em prática quais os agentes humanos e de software autorizados a ler, gravar, mover e eliminar conteúdo.	Manter logs e identificar os agentes humanos e softwares que executaram ações no conteúdo.	Executar revisão periódica dos logs de ações e acesso.
Metadados	Criar inventário do conteúdo, documentando sua localização de armazenamento atual. Fazer backup do inventário e armazenar no mínimo uma cópia separada do conteúdo.	Armazenar metadados suficientes para se conhecer do que se trata o conteúdo (pode incluir alguma combinação de metadados administrativos, técnicos, descritivos, de preservação e estruturais)	Determinar quais padrões de metadados utilizar. Encontrar e preencher as lacunas de metadados a fim de alcançar os padrões determinados	Gravar as ações de preservação associadas com o conteúdo e quando elas ocorreram. Implementar os padrões de metadados escolhidos
Conteúdo	Documentar os formatos de arquivos e outras características essenciais de conteúdo incluindo quando e como foram identificadas.	Verificar os formatos de arquivos e outras características essenciais de conteúdo. Construir relacionamentos com os criadores de conteúdo para encorajar a escolha de arquivos sustentáveis (formatos abertos).	Monitorar a obsolescência e a mudança de tecnologia às quais o conteúdo depende.	Executar migrações, normalizações, emulações e atividades similares que garantam que o conteúdo pode ser acessado.

Fonte: (LEVELS OF PRESERVATION REVISIONS WORKING GROUP, 2019, p. 1).

3.4.3 ISO 16363

A criação da norma ISO 14721, *Open Archival Information System* (OAIS), definiu um conjunto de requisitos, uma arquitetura de componentes e pacotes de informações a estarem presentes em um arquivo que tenha a intenção de realizar preservação confiável de dados e documentos a longo prazo. Em meio ao surgimento de soluções de preservação que se intitulavam aderentes ao padrão OAIS, houve a necessidade de se criar uma metodologia padronizada que pudesse verificar a real aplicação de requisitos do padrão nos repositórios existentes, podendo assim, declarar que um repositório é de fato confiável (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012).

A norma ISO 16363, conhecida como ACTDR (*Audit and Certification of Trustworthy Digital Repositories*), foi originada da recomendação técnica TRAC (*Trustworthy Repositories Audit & Certification: Criteria and Checklist*) desenvolvida pelo Comitê Consultivo para Sistemas de Dados Especiais (*Consultative Committee for Space Data Systems – CCSDS*) (CCSDS, 2011) que foi o produto da união de esforços entre a *Research Libraries Group* (RLG) e a *National Archives and Records Administration* (NARA), ambos órgãos norte-americanos. Em 2017, como parte do processo de revisão sistemática, a norma foi revista pelo respectivo comitê técnico e a versão de 2012 permaneceu em vigência.

Na Espanha, a norma encontra sua correspondente traduzida UNE-ISO 16363:2017 (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017). No Brasil não há uma tradução disponível da norma para a realização de certificações de repositórios, porém é possível utilizar o documento “Diretrizes para a Implementação de Repositórios Arquivísticos Digitais Confiáveis – RCD-Arq” (CONSELHO NACIONAL DE ARQUIVOS, 2015) para a verificação da presença de critérios que tornem confiável o respectivo repositório.

Para conceituar o que significa um repositório digital de confiança, a norma usa a definição criada pela RLG e OCLC “Um repositório digital confiável é aquele cuja missão é prover acesso confiável a longo tempo dos recursos digitais gerenciados para a sua Comunidade-Alvo, agora e no futuro” (RESEARCH LIBRARIES GROUP, 2002, p. i, tradução nossa). A partir dessa definição, é possível notar que a missão principal do repositório digital sempre deve ser voltada à sua Comunidade-Alvo (*Designated Community*), identificando as suas especificidades e demandas para que os processos do repositório as atendam e forneçam acesso confiável dos seus ativos digitais durante o tempo em que for necessário.

Para que a preservação a longo prazo aconteça, o repositório deve possuir processos de monitoramento constante de sua estrutura organizacional e técnica a fim de garantir que possui

equipe, financiamento e equipamentos adequados que o resguarde das possíveis ameaças e desafios que coloquem em risco o seu propósito. Ao dizer-se que um repositório é confiável, a partir da análise dos critérios da norma ISO 16363, pode-se afirmar que o mesmo possui planejamento para a execução da preservação digital, prevendo riscos e ameaças, ao mesmo tempo em que se adapta às mudanças de tecnologia e de sua Comunidade-Alvo. O conjunto de 109 critérios analisados pela norma são divididos em três grandes áreas: a) Infraestrutura organizacional; b) Gerenciamento do objeto digital, e; c) Gestão de riscos de infraestrutura e de segurança.

A área de infraestrutura organizacional verifica se o repositório possui condições que tornem viável a sua existência, analisando primeiramente se existe uma missão definida que inclua o seu compromisso com a preservação digital a longo prazo, assim como a definição de sua Comunidade-Alvo. A viabilidade financeira do repositório também é avaliada, analisando se existe um planejamento que permita o funcionamento do repositório pelo tempo que for necessário e a presença de um plano de continuidade que garanta a preservação dos documentos caso o repositório seja desativado. Estão presentes nessa área, ainda, critérios que avaliam a existência de uma equipe de pessoal dimensionada adequadamente, assim como a sua qualificação e ciclos contínuos de capacitação que permitam o acompanhamento da evolução tecnológica e de novos requisitos surgidos.

A segunda área, Gerenciamento do objeto digital, analisa principalmente, o tratamento do repositório quanto aos pacotes de informação utilizados nos modelos OAIS. Na ingestão de documentos, é avaliado se o repositório declara as especificações de documentos, formatos de arquivos e metadados aceitos, além da presença de processos que verificam e validam os objetos digitais recebidos no SIP (*Submission Information Package*).

Após o processamento da ingestão, o repositório gera o AIP (*Archival Information Package*), que contém o objeto digital recebido e armazena todas as informações necessárias para realizar a preservação. A norma contém critérios para analisar o tratamento desses pacotes, verificando se o repositório conta com procedimentos periódicos que examinam a sua integridade, se os formatos de arquivo preservados são compreendidos pela sua Comunidade-Alvo e se há a necessidade de se fazer alguma conversão ou migração de formatos de arquivos, por exemplo. Existem critérios ainda que verificam se todos os processamentos realizados sobre o AIP estão documentados e disponíveis para análise quando requisitada.

A política de acesso implementada pelo repositório é avaliada em critérios da norma que verificam se os documentos apropriados são tornados acessíveis, compreensíveis e íntegros à sua Comunidade-Alvo. A efetividade da difusão dos documentos sob custódia do repositório é realizada

através do DIP (*Dissemination Information Package*) que contém os documentos e metadados a serem enviados para as devidas plataformas.

A terceira área de requisitos, Gestão de riscos de infraestrutura e de segurança, avalia a presença de hardwares, softwares e processos adequados que garantam a preservação e acesso a longo prazo de forma segura à Comunidade-Alvo. A norma possui, por exemplo, requisitos que verificam se o repositório utiliza hardware adequado à preservação, assim como o seu monitoramento constante para evitar falhas ocorridas devido ao desgaste dos equipamentos, a presença de políticas de backup bem definidas e executadas corretamente, assim como a existência de equipamentos de rede bem dimensionados que permitam o acesso aos documentos preservados.

Do ponto de vista da segurança, existem critérios que avaliam se o repositório possui mecanismos que evitam o acesso indevido tanto aos servidores quanto às instalações que abrigam os equipamentos. Também é avaliada a presença de um plano de preparação e prevenção de desastres que permita ao repositório evitar ao máximo a parada de seu funcionamento e possa se recuperar sem comprometer a preservação de seus documentos.

Nas orientações para a verificação dos critérios de segurança, há uma grande relação com a norma ISO 17799⁵⁸ desenvolvida para avaliar os sistemas de gestão da informação e segurança de dados. Essa norma trata principalmente sobre a segurança dos dados, requisito primordial para a preservação de documentos a longo prazo, observando critérios bastante específicos para isso, no entanto, não contempla os demais aspectos presentes na ISO 16363 indispensáveis para a preservação de documentos digitais (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012, p. 22).

Outra norma que possui relação com a ISO 16363 é a ISO 27000⁵⁹ que trata sobre a análise de riscos para estabelecer um sistema de informações seguro. Caso o repositório digital siga os códigos de conduta estabelecidos na norma ISO 27000, pode utilizar os mesmos instrumentos para comprovar a metodologia empregada para a gestão de riscos, facilitando a verificação dos critérios que tratam desse assunto (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012, p. 74). O tratamento de riscos do repositório ainda pode utilizar a ferramenta DRAMBORA (*Digital Repository Audit Method Based on Risk Assessment*)⁶⁰, conforme sugestão da própria norma, com o intuito de se fazer uma varredura em toda a organização para identificação dos riscos aos quais está suscetível (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012, p. 75).

⁵⁸ (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2005)

⁵⁹ (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2018)

⁶⁰ (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007)

Um outro aspecto sobre segurança é ressaltado na ISO 16363, relativo à qualificação dos auditores responsáveis por sua aplicação. A norma torna claro que o repositório que tenha a intenção de obter a certificação deve contratar auditores qualificados e de procedência comprovada, uma vez que as evidências utilizadas para comprovar os critérios podem conter informações sensíveis e até confidenciais e, caso haja pessoas má intencionadas envolvidas no processo, elas podem ser vazadas, causando grandes prejuízos ao repositório (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012, p. 77). Ainda, pode haver a exposição de falhas de segurança detectadas no repositório que podem ser utilizadas em eventuais ataques.

O mesmo grupo de especialistas responsável pela criação da ISO 16363 criou o órgão PTAB (*Primary Trustworthy Digital Repository Authorisation Body*), primeira entidade acreditada habilitada a realizar a auditoria de repositórios que desejam obter a certificação. O órgão oferece cursos e treinamentos às equipes dos repositórios, contando inclusive com os próprios autores da norma, para o entendimento dos requisitos necessários para o cumprimento de cada critério, sendo aconselhado como o primeiro passo do processo de certificação (PRIMARY TRUSTWORTHY DIGITAL REPOSITORY AUTHORISATION BODY, 2022). A entidade implementa as normas ISO 17021⁶¹ e ISO 16919⁶² que estabelece regras e procedimentos para entidades que realizam auditoria e certificação.

O processo de auditoria para certificação realizado pelo PTAB possui diversos custos envolvidos. Uma pré-avaliação do repositório juntamente com uma assessoria de orientação sobre os documentos necessários para o início da auditoria atualmente custa £5.000. A auditoria, realizada em determinadas etapas remotamente e outras presencialmente, possui um custo mínimo de £13.500, adicionando-se ainda os valores para os deslocamentos dos membros assim como o custo das suas assessorias.

Dentre os repositórios existentes que possuem a certificação ISO 16363, pode-se citar o Centro Nacional de Artes Indira Gandhi (*Indira Gandhi National Centre for the Arts*), localizado na Índia, que recebeu sua certificação em novembro de 2017 com validade até julho de 2020 e o Escritório de Publicidade do Governo dos Estados Unidos (*United States Government Publishing Office*), que obteve sua certificação em dezembro de 2018 com validade até dezembro de 2021. O grau de exigência dos critérios da norma ISO 16363 requer documentos detalhados sobre todos os aspectos que ela cobre, “[...] requerendo um nível de organização e disposição de pessoal normalmente só alcançados por repositórios de instituições com uma economia privilegiada” (DE

⁶¹ (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2016)

⁶² (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2014)

GIUSTI, 2021, p. 57, tradução nossa), podendo ser um dos motivos pelos quais não existam mais repositórios certificados mundialmente.

3.4.4 ISO 17068

Dados os desafios de se produzir e manter documentos digitais autênticos e passíveis de serem usados como fontes de prova, há a possibilidade de as empresas e instituições contratarem repositórios de terceiros que fornecem o serviço de custódia. O serviço contratado torna-se responsável por custodiar os documentos recebidos mantendo uma estrutura que garanta o cumprimento dos requisitos de confiança necessários. Assim, a instituição terceiriza a sua responsabilidade de possuir equipes, equipamentos e estrutura necessários para a manutenção de repositórios com critérios de confiança contratando um serviço adequado a esse fim

A norma ISO 17068:2017 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012), tendo na Espanha sua equivalente UNE-ISO 17068:2020 (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2020) e, no Brasil, a ABNT NBR ISO 17068:2021 (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2021), estabelece a definição de um *Trusted Third Party Repositório* (TTPR – Repositório de Terceiros Confiável – RTC) como sendo um “serviço de retenção qualificado de terceiros que assegura que os documentos de arquivo digitais, disponibilizados por um cliente, permaneçam e sejam declarados confiáveis e autênticos” (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2021, p. 3). Nesse caso, o termo “terceiro” se refere a um organismo ou empresa independente da instituição que produz e deseja custodiar seus documentos ao mesmo tempo em que esse organismo também é independente dos clientes e interessados nos documentos da instituição.

Caso não haja regras, normas e padrões estabelecidos a serem cumpridos por esses repositórios de terceiros que garantam a confiabilidade dos documentos sob sua posse, a utilização desses documentos pode ser tornar bastante limitada, uma vez que a confiabilidade pode ser colocada em questão e tanto a sociedade quanto o próprio governo não poderão utilizá-los como fonte de prova. Formalmente, o RTC deve garantir “[...] autenticidade, confiabilidade, integridade e usabilidade dos documentos de arquivo digitais, durante o período do serviço contratado” (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2021, p. 4).

Assim como os repositórios próprios das instituições, os RTC devem receber os documentos e respectivos metadados dos seus clientes de forma a manter todas as evidências do período em que estava sob responsabilidade do sistema de gestão. Além disso, o RTC deve registrar todas as operações dos processos executados desde o momento em que recebeu as informações do

seu cliente, verificações periódicas de integridade e disponibilização dos documentos, por exemplo. Tais registros podem ser fornecidos às partes interessadas em verificar a autenticidade dos documentos custodiados.

Para que a confiabilidade de um RTC seja alcançada, ele deve possuir as características de estabilidade, competência e neutralidade (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2021, p. 8). A estabilidade se refere à capacidade de o RTC manter suas operações mesmo em caso de emergências ou desastres, tendo realizado um mapeamento de riscos que afetam a sua operação e um plano de recuperação a ser posto em prática sempre que necessário.

A competência se refere às qualificações do RTC e sua equipe para lidar com o desafio da preservação digital mantendo a confiabilidade dos documentos. Para atingir a competência necessária, o RTC deve possuir especialistas em preservação e gestão de documentos, estabelecendo processos bem definidos que englobem a aquisição, arquivamento, entrega, migração e destinação dos documentos sob sua posse.

A característica de neutralidade se refere à qualidade de o RTC não possuir favorecimento indevido a quaisquer partes interessadas, sejam elas o próprio cliente ou instituições políticas, possuindo uma gestão que suporte eventuais pressões externas. Caso haja indícios de que houve uma quebra na neutralidade do RTC, a confiabilidade dos documentos torna-se questionável, visto que pode ter havido algum processo tendencioso que possa ter beneficiado alguma das partes envolvidas.

Toda a relação entre o cliente o RTC é firmada através de um contrato que prevê os serviços fornecidos pelo repositório, prazos de retenção do material armazenado, forma de comunicação e eventuais particularidades do cliente. O contrato deve prever os serviços a serem fornecidos, as responsabilidades de cada uma das partes prevendo inclusive o comportamento a ser adotado em casos de alegação de falsificação, adulteração, vazamento e perdas de dados, por exemplo.

Junto ao contrato, deve existir um Acordo de Nível de Serviço (*Service Level Agreement* – SLA) que estabeleça o nível dos serviços ofertados e fatores de qualidade mensuráveis. Caso seja detectado o não atingimento de determinados níveis, o SLA deve prever a responsabilização aplicável ao RTC, podendo ser inclusive de natureza financeira através de multas. Em contrapartida, o SLA também pode prever bônus do cliente para o RTC sempre que determinados parâmetros forem superados positivamente (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2021, p. 9). Entre outros requisitos, o SLA deve prever também a taxa de disponibilidade dos serviços do RTC e o tempo médio de recuperação de falhas, visando que os serviços sejam minimamente interrompidos e diminuindo as consequências causadas pela indisponibilidade.

A contratação de terceiros para a execução dos serviços de repositório possui características peculiares a serem observadas pelas gestões das instituições de modo a embasá-los na tomada de decisão. Do lado do RTC, devem ser apresentadas todas as garantias que tornem a preservação a longo prazo contratada viável, demonstrando-se desde a estrutura e equipes disponíveis, interfaces entre RTC e cliente, até a verificação de sua saúde financeira para que a preservação seja realizada plenamente.

Do lado do cliente, principalmente quando se trata de instituições públicas, devem haver cuidados em relação à contratação para que não haja problemas com os órgãos de controle, levando a possíveis interrupções dos serviços e efetivação de uma nova contratação. Naturalmente, as questões financeiras também devem ser levadas em consideração pelo cliente, vislumbrando-se um horizonte futuro muitas vezes de difícil previsibilidade, principalmente tratando-se de documentos permanentes.

O processo de decisão entre a implementação de um repositório próprio, interno, ou a contratação de um RTC pelas instituições deve analisar os prós e contras de cada uma das alternativas, no entanto em ambos os casos, é possível contar com as normas internacionais vigentes para embasar tal processo. No caso de um RTC, utiliza-se a ISO 17068, enquanto que no caso de um repositório próprio, pode-se utilizar ISO 16363 para a verificação de conformidade para a preservação a longo prazo.

3.4.5 Nestor Seal

O grupo de trabalho “*Trusted Repository Certification*” da organização NESTOR (*Network of Expertise in long-term Storage and accessibility of digital resources*) criou um checklist de critérios que avalia a confiança de um repositório que armazena dados e documentos a longo prazo. O conjunto de critérios compõe o “Catálogo de critérios para repositórios digitais confiáveis” que teve a sua primeira versão publicada em 2006 e a sua segunda versão disponibilizada em 2009 com foco em “[...] identificar os critérios que permitem avaliar a confiabilidade de um repositório digital, tanto em nível organizacional quanto a nível técnico” (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 1, tradução nossa). A certificação fornecida pela organização Nestor usa como base a norma alemã DIN 31644:2012-04 (DEUTSCHES INSTITUT FÜR NORMUNG, 2012).

A crescente produção de dados e documentos digitais levou a produtores e consumidores de documentos científicos, históricos e patrimônio documental em geral à preocupação sobre o grau de confiança dos repositórios que armazenam tais dados. Tanto a

obsolescência tecnológica quanto à perda de capacidade de reter os dados ao longo do tempo apresentam ameaças aos ativos digitais das instituições, levando à necessidade de verificação de processos que impeçam a perda dos documentos e seu respectivo acesso.

O catálogo de critérios define um repositório digital como sendo “[...] uma organização (consistindo de pessoas e sistemas tecnológicos) que assumiram a responsabilidade pela preservação e acessibilidade a longo prazo dos objetos digitais, assim como a sua interpretabilidade, pelo propósito de eles serem usados por uma Comunidade-Alvo específica” (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 4, tradução nossa). É possível notar que a definição enfatiza a união de pessoas e sistemas de TI, deixando implícita necessidade de haver processos bem definidos para a interação entre ambos aconteça plenamente.

Complementarmente, a definição expressa o fato de que não basta conservar os dados e dar acesso a eles sem que haja uma garantia de que a Comunidade-Alvo do repositório conseguirá interpretá-los. Essa garantia é alcançada através de um constante monitoramento das mudanças e necessidades da Comunidade-Alvo do repositório, além das evoluções tecnológicas de hardwares e softwares adaptando-se para receber novos tipos de documentos assim como realizando migrações de formatos sempre que necessário.

O processo, ou a estrada, para a implementação de um repositório confiável pode ser realizado através de quatro etapas (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 6): a) Concepção; b) Planejamento e especificação; c) Realização e implementação, e; d) Avaliação. O documento deixa claro que esse não é um conjunto estático de fases a serem conquistadas, mas sim, fazem parte de um ciclo contínuo executado regularmente a fim de implementar melhorias no repositório em questão.

A elaboração do catálogo de critérios tem como princípios básicos a abstração e a conformidade com a terminologia utilizada no padrão OAIS⁶³. A abstração foi utilizada com o intuito de que os critérios pudessem ser aplicados nas mais diversas realidades existentes dos repositórios que se preocupam com a preservação de longo prazo. Os termos e nomes dos pacotes de informação do padrão OAIS também são usados na elaboração dos critérios com o intuito de facilitar a compreensão das equipes dos repositórios interessadas em obter a certificação NESTOR.

O processo de aplicação dos critérios no repositório segue princípios de documentação, transparência, adequação e mensurabilidade. O atendimento dos critérios é comprovado através de documentos bem especificados e, de preferência, disponíveis ao público para que seja possível a sua verificação e acompanhamento. A interpretação de cada critério deve seguir a realidade do repositório candidato à certificação, compreendendo as suas especificidades e características

⁶³ (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012)

singulares. A medição de atendimento de determinados critérios nem sempre pode ser realizada objetivamente, fazendo com que o examinador analise os indicadores fornecidos caso a caso, nem sempre podendo ser utilizado um número fixo para determinar a sua conformidade ou não.

O catálogo possui 42 critérios dividido em três grandes áreas: estrutura organizacional, gerenciamento de objetos e infraestrutura e segurança. Cada um dos critérios possui uma ficha contendo o nome do critério, uma explicação geral, exemplos que ajudam o examinador na verificação de atendimento e a literatura utilizada como referência para a elaboração do critério.

A primeira área dos critérios avalia se o repositório possui estrutura organizacional para realizar a preservação, observando primeiramente se há o claro comprometimento com a preservação através da especificação da sua missão e delimitação da sua Comunidade-Alvo. Também são analisados se a legislação aplicável é atendida, conforme o país ou região a qual o repositório está inserido.

Os critérios da segunda área analisam o tratamento dispensado aos objetos digitais em si que representa os pacotes de informação do padrão OAIS: SIP, AIP e DIP. Os três pacotes representam o fluxo das informações dentro do repositório digital sendo responsáveis, respectivamente pelo recebimento dos documentos no repositório, a administração e tratamento arquivístico de conservação dos documentos e a disseminação dos documentos para as plataformas de acesso da instituição. A principal preocupação dos critérios é sobre a garantia de autenticidade dos dados, verificando-se se existe uma padronização de informações estabelecida no recebimento de documentos, seja ela feita manualmente ou informatizada através de integrações com outros sistemas. É avaliada a documentação, através de metadados, dos procedimentos realizados internamente pelo repositório para a garantia da preservação a longo prazo, sendo requisito fundamental para a garantia de autenticidade dos objetos digitais.

A terceira área analisa a infraestrutura técnica dos equipamentos utilizados, analisando se o repositório protege adequadamente seus objetos de acessos indevidos assim como a presença de procedimentos que impedem a perda e corrupção de dados. A última área dos requisitos possui apenas três critérios a serem respondidos, que são apresentados em alto grau de abstração, sem o detalhamento sobre quais os processos devem estar presentes e como eles são executados, dessa forma a comprovação de atendimento desses critérios se torna uma tarefa um tanto subjetiva e dependente da interpretação dos avaliadores.

O processo de aplicação à certificação requer que o repositório realize o pagamento da taxa de €500 e indique duas pessoas que servirão como mediadores da comunicação com o Nestor, que por sua vez destacará um ou mais avaliadores responsáveis. A seguir, o repositório realiza a autoavaliação interagindo com os avaliadores sempre que houver dúvidas sobre os critérios. O

Nestor estima que o processo de análise dos critérios e consequente obtenção do selo não se estenda por mais que quatro meses desde o envio do formulário inicial até o parecer final (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 3).

Cada critério deve ser analisado quanto a sua aplicabilidade e, caso seja entendido que ele não se aplica à realidade do repositório, devem ser feitas as devidas justificativas e apresentação de documentos que embasem a decisão. A comprovação de atendimento dos critérios aplicáveis deve ser feita através de documentos preferencialmente públicos, no entanto, o Nestor garante o sigilo e preservação dos documentos fornecidos que não devem ser publicizados. Os respectivos documentos são aceitos nos idiomas alemão e inglês.

A resposta para cada critério segue a escala apresentada no Quadro 9. Para obtenção do selo Nestor, todos os critérios das áreas organizacional e gerenciamento de objetos deve ser aplicável e obter a pontuação máxima (10 pontos). O restante dos critérios, pertencentes à área de infraestrutura técnica e segurança, devem possuir média de 7 pontos no mínimo.

Quadro 9 – Tabela de pontuação Nestor Seal.

Resposta	Pontuação	Significado
Em aberto	0	Não existem planos de atendimento ao critério ou documentos que comprovem o atendimento.
Planejado	3	Existe um planejamento para o atendimento do critério levando em conta a realidade específica do repositório.
Planejado em detalhes	6	Existe um planejamento detalhado para o atendimento do critério reunindo as informações e aprovações internas necessárias. A aplicação do planejamento já foi iniciada.
Implementado	10	O planejamento foi implementado no repositório no âmbito organizacional ou técnico.

Fonte: (NESTOR CERTIFICATION WORKING GROUP, 2013, p. 4, tradução nossa).

Após a análise das respostas, caso o entendimento dos avaliadores seja positivo, o repositório recebe um selo com o ano de sua obtenção que pode ser adicionado no site do repositório reunindo preferencialmente, a documentação pública utilizada no processo. O Nestor não especifica um período fixo para a realização de uma nova rodada de avaliações, tendo o selo um caráter permanente, no entanto, a sua “[...] relevância provavelmente diminuirá após alguns anos, a menos que uma revisão seja feita posteriormente” (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 8, tradução nossa).

Atualmente, existem quatro repositórios com o selo Nestor, sendo metade deles obtida em 2016 e a outra metade em 2017 (NETWORK OF EXPERTISE IN LONG-TERM STORAGE OF DIGITAL RESOURCES, 2022). Ao todo, três repositórios são alemães e o restante trata-se do

repositório DANS (*Data Archiving and Networked Services*), repositório holandês do qual fazem parte os pesquisadores que elaboraram a certificação CoreTrust Seal.

3.4.6 CoreTrust Seal

Inicialmente conhecida como *Data Seal of Approval* (DSA), a certificação foi desenvolvida pelo grupo de pesquisadores DANS (*Data Archiving and Networked Services*) como forma de atender a demanda de organizações holandesas de verificar se os dados armazenados por determinados repositórios teriam a capacidade de serem encontrados, compreendidos e usados no futuro (L'HOURES, KLEEMOLA e LEEUW, 2019, p. 3). Em 2009, os criadores da certificação tomaram a decisão de formar um comitê internacional (*DSA Board*) para a sua administração, contando com membros de diversas organizações europeias.

Os critérios da certificação foram pensados para atender três papéis relacionados à pesquisa: produtores de dados, repositórios de dados e consumidores de dados. Dessa forma, os seus princípios iniciais foram (DATA SEAL OF APPROVAL, 2010, p. 1): a) dar garantia aos pesquisadores de que os seus dados estariam protegidos, armazenados de maneira segura e disponíveis para uso no futuro; b) garantir aos financiadores que os resultados das pesquisas estariam disponíveis no futuro para acesso e reuso; c) permitir aos pesquisadores uma forma de avaliar os repositórios nos quais os seus dados estariam armazenados, e; d) permitir aos repositórios uma maneira eficiente de armazenar e distribuir os dados de pesquisas sob sua guarda.

Inicialmente voltada aos repositórios de dados de pesquisa das ciências sociais e humanas, a certificação obteve a atenção de outros domínios da ciência, que fez com que o comitê administrador percebesse que os critérios de avaliação poderiam ser usados para instituições de memória, museus e arquivos (L'HOURES, KLEEMOLA e LEEUW, 2019, p. 5). Em 2010, houve uma revisão dos critérios de avaliação de modo a torná-los mais abrangentes a outros domínios e menos específicos a repositórios de pesquisa. Outra revisão foi realizada em 2013, resultando na segunda versão do DSA (DATA SEAL OF APPROVAL, 2013).

Sabendo-se da importância das certificações de nível básico e temendo a criação de várias outras certificações, o que tornaria os processos distintos e desiguais entre si, houve uma aproximação entre o DSA e o *International Council for Science's World Data System* (ICS-WDS) no sentido de alinhar os objetivos e criar uma única certificação com as mesmas características. Embora a origem das duas organizações fosse bastante distinta, a DSA vindo das Ciências Sociais e Humanas e a ICS-WDS vindo das Ciências do Espaço e da Terra, a parceria permitiu o conhecimento de diferentes domínios e o consequente atendimento de uma comunidade mais ampla

(L'HOURES, KLEEMOLA e LEEUW, 2019). A integração das duas entidades resultou na criação da fundação CoreTrustSeal, em 2018, como uma “organização internacional não governamental e sem fins lucrativos, baseada na comunidade, que promove infraestruturas de dados sustentáveis e confiáveis” (CORETRUSTSEAL BOARD, 2022, p. tradução nossa) composta por uma diretoria de 12 membros eleitos periodicamente. A última versão de requisitos da certificação foi publicada em 2019 estando válida para aplicação no período entre 2020 e 2022 (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019), sendo que a certificação obtida pelo repositório é válida por três anos.

Quadro 10 – Perguntas de contextualização do repositório do critério CoreTrust Seal R0.

Pergunta	Detalhamento
R0.1) Tipo de repositório (<i>Repository type</i>)	Deve-se selecionar um valor da lista fornecida na documentação, sendo possível ampliá-la, caso nenhum das opções seja a mais apropriada: repositório de domínio específico ou baseado em algum assunto; repositório institucional; sistema nacional de repositório, incluindo governamental; repositório de publicações; biblioteca; museu; arquivo, e; repositório de projeto de pesquisa.
R0.2) Descrição do repositório (<i>Brief description of repository</i>)	Fornece uma visão geral do repositório e seu funcionamento, tipos de dados e/ou documentos aceitos. O organograma da instituição pode ser informado adicionalmente, fornecendo a estrutura geral utilizada pelo repositório.
R0.3) Comunidade-Alvo (<i>Designated community</i>)	definição da comunidade atendida pelo repositório. É para a sua Comunidade-Alvo que o repositório deve se voltar para compreender suas demandas e planejar suas ações, estabelecendo, entre outros requisitos, os formatos de arquivo aceitos, suas estratégias de preservação e arquivos de disponibilização.
R0.4) Nível de curadoria (<i>Level of curation</i>)	Especifica se o repositório realiza a difusão dos documentos assim como os recebe ou se há algum tipo de intervenção como modificação do formato de arquivo ou a complementação de metadados. Em alguns casos, os repositórios podem optar por complementar os metadados originais a fim de facilitar a compreensão e uso dos documentos pela Comunidade-Alvo.
R0.5) Parceiros internos e externos (<i>Insource/Outsource partners</i>)	Relação dos parceiros do repositório, sejam eles a nível institucional, interno à organização, ou externos através da prestação de serviços, por exemplo. É importante relacionar os parceiros e seus eventuais contratos firmados informando os acordos de níveis de serviço aceitos (<i>Service Level Agreement</i>) assim como certificações conquistadas por tais parceiros.
R0.6) Resumo das principais mudanças desde a última aplicação da certificação (<i>Summary of significant changes since last application</i>)	Deve-se informar as mudanças ocorridas desde que o repositório realizou a última certificação, se for o caso. É importante ressaltar mudanças da Comunidade-Alvo, entidades financiadoras e eventuais mudanças drásticas na estrutura do repositório.
R0.7) Outras informações relevantes (<i>Other relevant information</i>)	Permite à equipe do repositório complementar as informações fornecidas que não se encaixam nos demais requisitos, como, por exemplo, alcance geográfico do uso do repositório, tipos de uso do seu material, etc.

Fonte: (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019, p. 3)

A certificação contém 16 requisitos a serem analisados pela equipe responsável pela autoavaliação do repositório com todas as respostas sendo fornecidas no idioma inglês e contendo

documentos e/ou *hyperlinks* para que o corpo avaliador possa comprovar a exatidão as informações fornecidas.

O primeiro requisito, chamado de “R0” prevê a contextualização do repositório, solicitando os dados apresentados no Quadro 10.

Pode-se notar que a primeira pergunta, R0.1 solicita o tipo do repositório que está sendo avaliado, em seguida, na pergunta R0.2, deve-se fornecer uma descrição do repositório que guiará os avaliadores no processo da auditoria.

A pergunta R0.3 solicita a definição da Comunidade-Alvo, sobre a qual é possível avaliar se determinadas ações do repositório estão alinhadas. A pergunta R0.4 busca identificar se o repositório realiza algum tipo de curadoria, principalmente para a difusão, no sentido de gerar derivadas de acesso, complementar metadados e outras ações que sejam cabíveis. A pergunta R0.5 visa identificar os parceiros do repositório com as devidas responsabilidades, assim, permite que a auditoria possua escopos de tratamento de pendências internas e outras a serem repassadas a esses parceiros. A pergunta R0.6 se utiliza no caso de reaplicação da auditoria, informando as principais modificações ocorridas no repositório desde a última auditoria realizada. A última pergunta, R0.7 permite que o repositório complemente informações relevantes do repositório que possam ser úteis para o processo de auditoria.

Os demais requisitos da certificação visam o mapeamento de aspectos específicos do repositório, verificando o atendimento de cada um de acordo com as seguintes as respostas apresentadas no

Quadro 11. As respostas possuem uma escala semântica que busca identificar em que grau os requisitos estão de fato aplicados, se estão mais para o campo da teoria e planejamento ou se já estão efetivamente aplicados e em funcionamento.

Quadro 11 – Respostas aos critérios das perguntas da certificação CoreTrust Seal

Resposta	Significado
Não se aplica	Nesses casos, o repositório deve justificar detalhadamente o motivo pelo qual o requisito não se aplica, adicionando evidências que comprovem a resposta.
O requisito ainda não foi considerado	Significando que o repositório desconhece o requisito ou ainda não iniciou o seu planejamento.
O requisito está presente somente na teoria	Significa que o repositório tem conhecimento do requisito, já formulou a sua aplicação, mas essa ainda não ocorreu na prática.
O requisito está em fase de implementação	Significa que há um planejamento de atendimento do requisito e as ações para o seu alcance estão sendo realizadas, porém não houve o seu atendimento por completo.
O requisito foi totalmente atendido	O repositório implementou ações que levaram ao atendimento do requisito por completo.

Fonte: (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019, p. 2)

Cada requisito possui o nível mínimo aceitável da resposta para que a certificação seja concedida ao repositório. Na grande maioria, os requisitos devem possuir no mínimo o nível 3, no qual já devem haver iniciativas em implementação para atendimento do requisito. O Quadro 12 exibe os requisitos da certificação com as suas devidas áreas.

Quadro 12 – Requisitos CoreTrust Seal

Área	Requisito
Infraestrutura organizacional	R1. Missão escopo. O repositório tem uma missão explícita de prover acesso e preservar os dados sob sua guarda?
	R2. Licenças. O repositório mantém todas as licenças que cobrem o acesso e uso dos dados, assim como monitora o cumprimento dessas licenças?
	R3. Continuidade de acesso. O repositório tem um plano de continuidade que garanta o acesso e a preservação dos dados sob sua guarda?
	R4. Confidencialidade. O repositório garante, até os seus limites, que os dados são criados, curados, acessados e usados de acordo com as normas éticas e disciplinares em vigência?
	R5. Infraestrutura organizacional. O repositório tem financiamento adequado e número suficiente de membros qualificados na equipe, gerenciados de acordo com um sistema transparente de governança para efetivamente alcançar a sua missão?
	R6. Orientação de especialistas. O repositório adota mecanismos para garantir orientação contínua de especialistas (interna, externa, incluindo orientação do meio científico, se for o caso)?
Gerenciamento de objetos digitais	R7. Integridade e autenticidade dos dados. O repositório garante a integridade e autenticidade dos dados.
	R8. Avaliação. O repositório aceita dados e metadados baseados em critérios definidos que garantem a relevância e compreensibilidade para os usuários dos dados?
	R9. Documentação de procedimentos de armazenamento. O repositório aplica processos e procedimentos documentados no gerenciamento do armazenamento arquivístico dos dados?
	R10. Plano de preservação. O repositório assume a responsabilidade pela preservação a longo tempo e gerencia essa função de uma maneira planejada e documentada?
	R11. Qualidade dos dados. O repositório possui conhecimentos adequados para garantir a qualidade dos

	dados técnicos e metadados e garante que informações suficientes estão disponíveis para que os usuários finais possam fazer análises baseadas na qualidade?
	R12. Fluxos de trabalho. O arquivamento é realizado de acordo com fluxos de trabalhos definidos desde a ingestão até a disseminação?
	R13. Descoberta de identificação de dados. O repositório permite aos usuários descobrir dados e os referenciar através de uma maneira persistente através de uma citação apropriada?
	R14. Reuso dos dados. O repositório permite o reuso dos dados ao longo do tempo, garantindo que os metadados apropriados estão disponíveis para apoiar a compreensão e uso dos dados?
Tecnologia	R15. Infraestrutura técnica. O repositório funciona com sistemas operacionais e demais softwares básicos bem suportados e usa tecnologias de hardware e software apropriadas para os serviços que provê para usa Comunidade-Alvo?
	R16. Segurança. A infraestrutura técnica do repositório provê proteção física e dos seus dados, produtos, serviços e usuários?

Fonte: (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019)

É possível notar que a primeira área dos requisitos, infraestrutura organizacional, avalia os aspectos institucionais do repositório, verificando se o mesmo possui a vocação para a preservação de dados e documentos através de uma estrutura, equipe e financiamento adequados, utilizando procedimentos bem documentados. É fundamental que o repositório possua uma administração organizada, transparente e baseada em padrões de governança para que o seu funcionamento seja adequado e esteja menos suscetível às adversidades podem ocorrer durante a preservação de dados a longo tempo.

É avaliado ainda o compromisso do repositório quanto às licenças dos materiais depositados, analisando se os direitos autorais, por exemplo, são cumpridos corretamente, sem a alteração do material original e alertando o usuário sobre as restrições impostas por cada licença. Outro aspecto analisado diz respeito ao cumprimento do direito à privacidade dos dados, objeto de legislações como LGPD (Lei Geral de Proteção de Dados) (BRASIL, 2018) e GPDR (*General Data Protection Regulation*) (EUROPEAN COMMISSION, 2018), podendo o repositório gerar derivadas de acesso dos documentos ocultando as partes que contém informações sensíveis.

A segunda área, gerenciamento de objetos digitais, foca nos cuidados e procedimentos que o repositório realiza para que os dados armazenados não se corrompam e sejam utilizáveis pelos usuários da sua Comunidade-Alvo ao longo do tempo. Os requisitos dessa área verificam se o repositório possui procedimentos de backup, por exemplo, assim como a existência de um plano de continuidade que garanta o funcionamento do repositório em caso de catástrofes ou descontinuidade.

A terceira área, tecnologia, analisa se o repositório possui um parque tecnológico adequado ao desafio da preservação de dados. O repositório deve contar com servidores, equipamentos de rede, equipamentos de backup e mecanismos de segurança que garantam que o repositório poderá se recuperar de falhas ocorridas intencionalmente ou não. O planejamento da

infraestrutura de Tecnologia da Informação deve prever a substituição dos dispositivos de armazenamento periodicamente, conforme os prazos recomendados pelos próprios fabricantes de modo a evitar que o seu mal funcionamento prejudique a integridade dos dados ou a sua disponibilidade de acesso. Além desses fatores, deve-se analisar a segurança dos servidores contra acessos não autorizados, sendo eles feitos virtualmente através de ataques remotos ou até mesmo ataques realizados presencialmente por pessoas maliciosas que tenham acesso aos equipamentos do *data center*. Tanto a segurança lógica quanto a segurança das instalações são avaliadas nos requisitos da terceira área da certificação.

Atualmente, existem 120 repositórios certificados CoreTrust Seal espalhados mundialmente, sendo a maioria (61%) em países do continente europeu (CORETRUSTSEAL BOARD, 2022). A Espanha e o Brasil não possuem nenhum repositório certificado, enquanto quase 25% dos repositórios certificados (29) se encontram nos Estados Unidos. O processo de aplicação envolve a realização de um cadastro no sistema, a utilização de uma ferramenta própria para a certificação e o pagamento de uma taxa de €1.000 que dá direito a até cinco revisões pelo comitê avaliador das informações do repositório. Sempre que necessário, o comitê pode ainda solicitar novas documentações que auxiliem a compreensão e verificação de atendimento dos requisitos da certificação.

É possível notar que os requisitos da certificação CoreTrust Seal possuem aspectos mínimos que um repositório deve possuir para que se possa dizer que ele se preocupa e tem condições de realizar a preservação dos dados sob sua posse. A certificação está no nível *core*/básico e todos os critérios são abordados com mais profundidade nas certificações de nível mais avançado, de tal modo que, a sua obtenção representa o marco inicial da jornada do repositório rumo ao seu aperfeiçoamento.

3.4.7 DRAMBORA

Com um enfoque diferente das certificações formais, existe o instrumento DRAMBORA (*Digital Repository Audit Method Based On Risk Assessment* – Método de Auditoria de Repositório Digital Baseado na Avaliação de Riscos) (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007) criado pelo grupo de trabalho formado pelo *Digital Curation Centre* (DCC) e o consórcio de instituições acadêmicas e culturais da Europa, *DigitalPreservationEurope* (DPE). A ferramenta (*toolkit*) foi desenvolvida para permitir uma autoanálise do repositório acerca dos riscos aos quais pode estar submetido, analisando desde a estrutura organizacional da instituição até a verificação das tecnologias empregadas para garantir a

preservação a longo prazo. Em conjunto com a documentação publicada do DRAMBORA, tida como ferramenta *offline*, uma ferramenta *online*⁶⁴ foi disponibilizada para que os auditores possam preencher o catálogo de riscos de suas instituições.

A utilização da ferramenta online facilita a criação dos relatórios de auditoria da instituição e contribui com o DCC e DPE na construção de uma melhor compreensão dos riscos envolvidos na curadoria digital a nível global e permite melhorias no DRAMBORA. É facultado às instituições a publicação dos resultados das auditorias, de maneira anonimizada, nos relatórios emitidos pelo DCC e DPE.

Os critérios de avaliação do DRAMBORA foram elaborados entre os anos 2006 e 2007 utilizando repositórios piloto de diversos domínios como bibliotecas nacionais, centros de armazenamento de dados científicos e arquivos de dados culturais e patrimoniais para o desenvolvimento, aprimoramento e validação da metodologia. Ainda em 2007 ocorreu o lançamento da primeira versão da ferramenta que visa fornecer meios para que os auditores identifiquem os riscos envolvidos nos repositórios digitais analisando as suas probabilidades de ocorrerem e o impacto gerado à instituição quando eles ocorrem.

O objetivo da metodologia é “[...] facilitar a auditoria interna, fornecendo aos administradores de repositório um meio de avaliar suas capacidades, identificar seus pontos fracos e reconhecer seus pontos fortes” (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 11, tradução nossa). O DRAMBORA não se autodenomina como prescritivo, no sentido de ser uma ferramenta estática e pronta para ser aplicada nas auditorias, mas sim, sugere áreas e critérios a serem avaliados, bem como fornece um catálogo de riscos, no Apêndice 3 (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 134), a ser usado como exemplo e ponto de partida para as instituições.

No contexto do DRAMBORA, a curadoria digital é “caracterizada como um processo de transformar incertezas controláveis e incontroláveis em uma estrutura de riscos gerenciáveis” (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 11, tradução nossa) que se molda às características do próprio repositório digital. O levantamento dos riscos do repositório permite estabelecer a implementação de estratégias que visam evitar o acontecimento de determinados riscos, mitigar a ocorrência de eventos relacionados aos riscos ou criar planos de reação aos riscos que por ventura sejam inevitáveis e estejam além do alcance do próprio repositório. A construção do catálogo de riscos do repositório traz como consequência reflexões sobre a extensão dos esforços que estão sendo feitos para que o repositório permaneça em funcionamento desempenhando suas funções alinhadas à missão à qual se propõe.

⁶⁴ <http://www.repositoryaudit.eu/>

Do ponto de vista dos financiadores e apoiadores do repositório, o DRAMBORA torna visível o planejamento das medidas para o enfrentamento de cada risco assim como auxilia na visualização das áreas que precisam maior alocação de verbas. Em certos casos, pode ser que seja necessária a aquisição de equipamentos de TI para o armazenamento de dados, por exemplo, a fim de se realizar uma replicação dos dados e aumentar o número de cópias de segurança para que não se corra o risco de perda de dados. Em outros casos, pode-se identificar que um determinado processo é executado por somente uma pessoa, necessitando que uma nova contratação seja feita, evitando o risco de dependência nessa única pessoa.

Enquanto a definição de repositório digital pode variar de uma fonte para outra e entre distintas certificações, o DRAMBORA estabeleceu um conjunto de dez critérios os quais um repositório digital deve cumprir, independente da sua natureza, tamanho e modelo de negócios (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 16, tradução nossa):

- 1) Comprometimento na manutenção da continuidade dos objetos digitais para a sua comunidade identificada;
- 2) Demonstrar aptidão organizacional (incluindo finanças, pessoal, estrutura, processos) para cumprir seu compromisso;
- 3) Adquire e mantém os direitos contratuais e legais necessários e cumpre com suas responsabilidades;
- 4) Tem uma política eficaz e eficiente;
- 5) Adquire e ingere objetos digitais com base em critérios declarados que correspondem aos seus compromissos e capacidades;
- 6) Mantém/garante a integridade, autenticidade e usabilidade dos objetos digitais que mantém ao longo do tempo;
- 7) Cria e mantém metadados necessários sobre as ações realizadas em objetos digitais durante a preservação, assim como os contextos relevantes sobre a produção, acesso e uso antes da preservação;
- 8) Cumpre com os requisitos necessários de divulgação;
- 9) Possui programa estratégico para planejamento e execução da preservação;
- 10) Possui infraestrutura técnica adequada para manutenção contínua e segurança de objetos digitais.

O DRAMBORA ainda faz questão de diferenciar as organizações que têm como principal missão a preservação de documentos, tendo atividades e processos voltados a esse fim, dos arquivos que fazem parte de organizações maiores. A maneira com que os riscos são avaliados e

tratados se diferencia entre tais organizações, enquanto em uma organização independente dedicada à preservação deve-se analisar todos os possíveis riscos, desde a gestão organizacional, financeira e tecnológica, uma organização pertencente à outra maior pode delegar o tratamento dos riscos à outras instâncias.

Um destaque feito pelos autores é em relação ao que é frequentemente entendido como risco. Normalmente são tidos como riscos as ameaças e vulnerabilidades às quais a organização está sujeita, no entanto, pode-se adicionar a esse conjunto a exposição às incertezas que podem comprometer a realização das atividades e o cumprimento da missão do arquivo (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 18). Nos casos de arquivos governamentais, por exemplo, a manutenção das suas atividades pode depender dos interesses do governo em questão, tornando-se um risco a ser avaliado quanto à sua probabilidade de acontecer durante a auditoria realizada.

De toda forma, sempre haverá riscos envolvidos nas atividades das organizações, uma vez que “o risco é inevitável e toda organização precisa tomar medidas para gerenciar o risco de uma maneira que possa ajustá-lo a um nível tolerável. A quantidade de risco que é considerada tolerável e justificável é o ‘apetite de risco’” (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 20, tradução nossa). Quando se trata de preservação digital, os principais riscos envolvidos se referem a preservar os objetos digitais autênticos, utilizáveis e acessíveis ao longo do tempo, necessitando de um gerenciamento de riscos que se adapte a esse ambiente em constante evolução e mudanças, enfrentando e mitigando a inevitável obsolescência tecnológica. Para Frank (2022, p. 46) “[...] há uma ampla aceitação da noção de que a preservação da informação digital e as práticas de gestão de risco estão relacionadas”.

Os mecanismos utilizados para combater e prevenir os riscos encontrados durante a auditoria podem ser comprovados mediante a apresentação de documentos aos auditores como, por exemplo, a política de preservação e acesso, planos de contingência e continuidade, planejamento de capacitação das equipes, relatórios contábeis, entre outros. No entanto, além da análise dos documentos, o auditor deve se atentar a fatores que muitas vezes não estão materializados nos documentos, assim como deve ter a sensibilidade de colher depoimentos de pessoas que não estejam envolvidas diretamente nos processos do repositório, mas que podem fornecer informações valiosas sobre aspectos não notados. Como exemplo, a equipe responsável pela limpeza das instalações físicas pode fornecer detalhes sobre a forma como as pessoas acessam as dependências do repositório, permitindo que se extraiam detalhes que impactam em determinados riscos (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007).

A metodologia do DRAMBORA utiliza a escala de probabilidade e a escala de impacto para traduzir os riscos em uma forma objetiva e quantificável. O Quadro 13 exibe a pontuação da probabilidade de um determinado risco ocorrer.

Quadro 13 – Pontuação DRAMBORA de probabilidade de ocorrência de um risco.

Pontuação da probabilidade	Interpretação
1	Probabilidade mínima. Ocorre uma vez a cada 100 anos ou mais.
2	Probabilidade muito baixa. Ocorre uma vez a cada 10 anos.
3	Probabilidade baixa. Ocorre uma vez a cada 5 anos.
4	Probabilidade média. Ocorre uma vez por ano.
5	Probabilidade alta. Ocorre uma vez por mês.
6	Probabilidade muito alta. Ocorre mais que uma vez por mês.

Fonte: (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 35, tradução nossa)

Para cada risco identificado no repositório, deve-se analisar quais as chances reais de que ele aconteça, atribuindo uma pontuação que inicia em 1, representando a menor probabilidade até a pontuação 5 que representa a maior probabilidade de o risco ocorrer.

O Quadro 14 exibe a escala que mede o potencial impacto no repositório na ocorrência de um determinado risco. O impacto de cada risco é medido através da pontuação que inicia em zero, que significa que o risco não oferece ameaça ao repositório, até a pontuação 6 que sinaliza que a ocorrência do risco causa um forte impacto no repositório e causa a interrupção permanente no seu funcionamento.

Quadro 14 – Pontuação DRAMBORA de impacto dos riscos.

Pontuação de impacto do risco	Interpretação
0	Impacto zero. Resulta em nenhuma perda da autenticidade e interpretabilidade ⁶⁵ dos objetos digitais.
1	Impacto negligenciável. Resulta em uma perda isolada, mas totalmente recuperável da autenticidade e interpretabilidade dos objetos digitais.
2	Impacto superficial. Resulta uma perda mais abrangente, mas totalmente recuperável da autenticidade e interpretabilidade dos objetos digitais.
3	Impacto médio. Resulta em perda total, mas totalmente recuperável da autenticidade e interpretabilidade dos objetos digitais.
4	Impacto alto. Resulta em uma perda isolada, incluindo perdas irre recuperáveis, da autenticidade e interpretabilidade dos objetos digitais.
5	Impacto considerável. Resulta em uma perda abrangente, incluindo perdas irre recuperáveis ou perdas que são recuperáveis somente em um terço da autenticidade e interpretabilidade dos objetos digitais.
6	Impacto cataclísmico. Resulta na perda total e irre recuperável da autenticidade e interpretabilidade dos objetos digitais.

⁶⁵ Os autores utilizam o termo para se referir à interpretabilidade dos objetos digitais nos sentidos técnico, contextual, sintático e semântico (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 35).

Fonte: (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 35)

A gravidade de determinado risco é o resultado do produto dos valores identificados para a sua probabilidade e impacto. Assim, os riscos com maior gravidade têm um resultado maior e devem ser prioridade para o repositório, necessitando a elaboração dos planos de mitigação e contingência, utilizados para evitar ou postergar a sua ocorrência e controlar os impactos quando da sua ocorrência, respectivamente.

O processo de auditoria do DRAMBORA estabelece seis estágios a serem seguidos, cada um com as suas particularidades e resultados finais (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 43):

1. Estágio 1: identificar o contexto organizacional. Busca identificar o papel do repositório analisando a sua missão e permitindo uma delimitação do seu escopo de atuação para a sua comunidade identificada;
2. Estágio 2: documentar o quadro de políticas e regulamentações. Busca identificar os documentos, normas e leis às quais o repositório está sujeito e deve estar em acordo;
3. Estágio 3: identificar atividades, ativos e seus respectivos proprietários. Tem como objetivo desenvolver um modelo conceitual das atividades e processos do repositório utilizados para o cumprimento da sua missão, as equipes responsáveis pelas suas execuções e os seus ativos. Nesse âmbito, os ativos do repositório podem ser representados por servidores de aplicação (hardware) ou softwares utilizados para realização das atividades, por exemplo;
4. Estágio 4: identificar os riscos. A partir dos resultados dos estágios anteriores, que identificaram a missão, marcos regulatórios, processos e ativos, observar quais riscos o repositório está sujeito à ocorrência;
5. Estágio 5: avaliação dos riscos. Realiza a análise minuciosa de cada risco identificado no estágio anterior, relacionando-os e informando-se as respectivas probabilidades e impactos, que resulta na gravidade de cada um.
6. Estágio 6: gerenciamento dos riscos. São informadas as decisões a serem tomadas pelo repositório sobre cada risco avaliado no estágio anterior. É nesse estágio que o tratamento dos riscos é efetivamente realizado, decidindo-se as ações necessárias para evitar determinados riscos, bem como quais riscos o repositório está disposto a correr. A maneira como os riscos são gerenciados no repositório, de certa forma, é limitada à sua estrutura organizacional, saúde

financeira, estrutura tecnológica e equipe disponível, uma vez que cada um desses fatores pode ser a chave na prevenção, mitigação ou contenção dos riscos.

A gestão de risco, em geral, é tida como uma atividade que nunca cessa, mas se retroalimenta gerando análises que servem como refinamento e melhoria das estratégias de prevenção de riscos além de identificar novos riscos. Adicionalmente, novas variáveis podem ser adicionadas ao contexto de atuação do repositório como, por exemplo, mudanças na sua comunidade, que podem necessitar a preservação de novos formatos de arquivo ou mudanças organizacionais que impactem na forma como a gestão é realizada.

Adicionalmente, os documentos que evidenciam as táticas adotadas para tratar os riscos precisam possuir ações práticas a serem acionadas de forma eficaz na ocorrência dos eventos previstos. Enquanto a documentação pode servir como prova para os auditores, internamente, em geral, as equipes dos repositórios se mostram céticas quanto a aplicabilidade dos planos, entre eles os planos de continuidade a serem acionados no caso de haver a impossibilidade financeira de continuar as atividades do repositório, por exemplo.

De acordo com Frank (2022), os planos de continuidade são traçados elaborados minuciosamente prevendo os procedimentos a serem executados na migração de custódia do arquivo descontinuado para o novo arquivo, porém, as equipes dos repositórios se mostram descrentes quanto a real viabilidade desses planos devido à dificuldade de simulá-los e testá-los. Outro fator que incrementa o ceticismo é que se o repositório deve ser descontinuado por problemas financeiros, deve-se passar o seu acervo para o outro repositório previsto no plano, porém há dúvidas sobre como a estrutura e orçamento desse outro repositório vai comportar essa carga de trabalho assumida (FRANK, 2022, p. 53). Como reflexo desse processo, um repositório pode obter certificações e ser aprovado em auditorias apresentando todos os documentos exigidos, no entanto, a aplicabilidade dos planos traçados talvez nunca possa ser comprovada.

Mesmo uma organização que não possua a cultura de análise de riscos em sua gestão pode se beneficiar com as auditorias realizadas com o DRAMBORA, uma vez que cria um autoconhecimento sobre os seus pontos fortes e mudanças que precisam ser implementadas para a sua melhoria. Os resultados das auditorias realizadas ao longo do tempo fornecem um arcabouço teórico fundamental que permite a visualização da evolução no tratamento dos riscos, propiciando um planejamento das ações a serem tomadas e favorecendo uma melhor gestão do repositório como um todo.

3.4.8 Guia de Avaliação de Repositórios Institucionais de Investigação REBIUN

O *Guía de Evaluación de Repositorios Institucionales de Investigación* teve sua primeira versão publicada em dezembro de 2010 com objetivo de fornecer uma ferramenta de auditoria para os repositórios que compõem a REBIUN (*Red de Bibliotecas Universitarias Españolas*) assim como fornecer um ponto de partida para novos repositórios. As diretrizes do guia buscam a melhoria de qualidade dos repositórios como um todo, procurando aumentar as suas visibilidades e fomentar a interoperabilidade entre repositórios (BARRUECO CRUZ, CABALLOS VILLAR, *et al.*, 2010).

O trabalho desenvolvido em colaboração entre a REBIUN e FECYT (*Fundación Española para la Ciencia y Tecnología*) faz parte do projeto RECOLECTA (*Recolector de Ciencia Abierta*) que busca criar uma infraestrutura nacional dados abertos através da cooperação e integração dos repositórios de pesquisa. Como um dos pilares do projeto RECOLECTA é a integração dos repositórios, esses devem disponibilizar serviços que permitam a coleta de metadados padronizados dos materiais preservados. Assim, o RECOLECTA “agrega todos os repositórios de acesso aberto numa única plataforma que garante a sua interoperabilidade e permite o acesso a toda a produção científica nacional em acesso aberto” (AZORÍN, BERNAL, *et al.*, 2021, p. 7, tradução nossa).

Dentre as referências para a elaboração do guia estão o modelo OAIS (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012), o *TDR: Attributes and Responsibilities* (RESEARCH LIBRARIES GROUP, 2002) e os 10 princípios para a preservação digital (CENTER FOR RESEARCH LIBRARIES, 2007), *NDSA Levels of Digital Preservation* (NATIONAL DIGITAL STEWARDSHIP ALLIANCE, 2022), além da utilização do protocolo OAI-PMH. Um fato a se destacar é que desde a sua primeira versão, o guia já possuía apelo bastante forte na interoperabilidade, acesso e reuso, precedendo os princípios FAIR que seriam publicados 2 anos após a segunda versão do guia publicada em 2014.

O Quadro 15 exhibe as áreas de distribuição dos critérios e a evolução das quantidades de critérios por área e por versão do guia. A área 8, que trata de serviços e funcionalidades de valor agregado só foi incluída na versão de 2017 do guia, por esse motivo, nos anos anteriores não são computados critérios para a área.

Quadro 15 – Áreas e evolução de critérios do Guia

Área	2010	2014	2017	2021
1. Visibilidadade	7	8	7	6

2. Políticas	5	6	6	8
3. Aspectos legais	4	4	4	4
4. Metadados descritivos e de publicação	4	19	22	32
5. Interoperabilidade de metadados e acesso aos conteúdos	7	10	11	17
6. Logs e estatísticas	2	4	5	8
7. Segurança, autenticidade e integridade dos dados	3	2	4	5
8. Serviços e funcionalidades de valor agregado	-	-	5	7
Total	32	53	64	87

Fonte: a partir de (BARRUECO, 2022, p. 8).

De uma forma geral, a quantidade de critérios se manteve uniforme entre as edições, no entanto, é possível notar um maior destaque aos metadados, com o incremento de critérios ao longo das versões nas áreas 4 e 5. Esse fato corrobora a importância dada pelo guia na verificação de presença de metadados que promovam a localização dos objetos e a interoperabilidade entre as plataformas dos repositórios.

Em 2015, o Grupo de Trabalho de Repositórios da REBIUN extraiu um conjunto de 25 critérios do guia, considerados essenciais e indispensáveis, para avaliar o estado geral dos repositórios componentes da rede. Inicialmente, esses critérios foram aplicados como piloto em alguns repositórios para validação tendo sido iniciada a avaliação de outros repositórios em 2016. O Quadro 16 exhibe os critérios escolhidos para a avaliação dos repositórios.

Quadro 16 – Áreas e critérios de avaliação dos repositórios REBIUN

1. Visibilidade
1.1 Menção do repositório na página principal da instituição
1.2 Presença em diretórios nacionais e internacionais Ex: ROAR, OpenDOAR, BuscaRepositorios, Hispana e OAI Data Providers
1.3 Presença em coletores nacionais e internacionais Ex: Google Scholar, OpenAire, Recolecta e BASE
1.4 Existência de um nome normalizado em todos eles Utilizar sempre o mesmo nome do repositório nos diretórios e coletores.
1.5 Existência de uma URL amigável
1.7 Ao menos 75% dos recursos textuais de pesquisa que o repositório disponibiliza se encontram em acesso aberto
1.8 A instituição aderiu a alguma das declarações de acesso aberto (Budapeste, Berlim ou Bethesda)
2. Políticas
2.1 Existe uma declaração sobre a missão e objetivos do repositório
2.2 Documento de acesso público sobre arquivamento no repositório, estabelecendo, pelo menos, os seguintes pontos: quem pode depositar, o que pode ser depositado, e em que depositado e em que formatos
2.3 Documento de acesso público sobre a preservação dos conteúdos
2.4 Documento de acesso público sobre o reuso de metadados
2.5 Se oferece visivelmente contato e assessoria Deve existir uma forma de contato visível para que os usuários possam ser assessorados, via formulário, e-mail ou telefone, por exemplo.
2.6 Política institucional de acesso aberto
4. Metadados
4.2 Todos os registros contêm o campo título (dc:title)
4.4 Todos os registros contêm o campo tipo de publicação (dc:type)
4.7 Todos os registros contêm o campo de data de publicação (dc:date)

4.11 Todos os registros contêm o campo autor (dc:Creator)
5. Interoperabilidade
5.2 Os recursos de pesquisa são identificados por meio de um ou mais conjuntos
5.4 Os registros excluídos são sinalizados Quando um registro é excluído da base de dados do repositório, deve-se sinalizar os coletores para que estes removam as referências para o registro eliminado.
5.5 O tempo de vida do token de continuação é de no mínimo 24h. O protocolo OAI-PMH prevê a utilização de tokens para os casos em que cliente necessite continuar a tarefa e extração dos registros posteriormente.
5.6 O e-mail do administrador do repositório está disponível na <i>tag AdminEmail</i> na resposta a um comando "Identificar".
5.8 A entrega de registros por meio do protocolo OAI-PMH é progressiva por lotes
5.9 Os tamanhos de lote para entrega de registros estão dentro do intervalo de 100 a 500 registros.
6. Logs e estatísticas
6.2 O repositório disponibiliza um serviço estatístico sobre a utilização dos documentos armazenados.
7.2 O repositório usa identificadores persistentes para seu conteúdo (DOIs, Handles, URNs, etc.).

Fonte: (RED DE BIBLIOTECAS UNIVERSITARIAS ESPAÑOLAS, 2015, p. 2, tradução nossa)

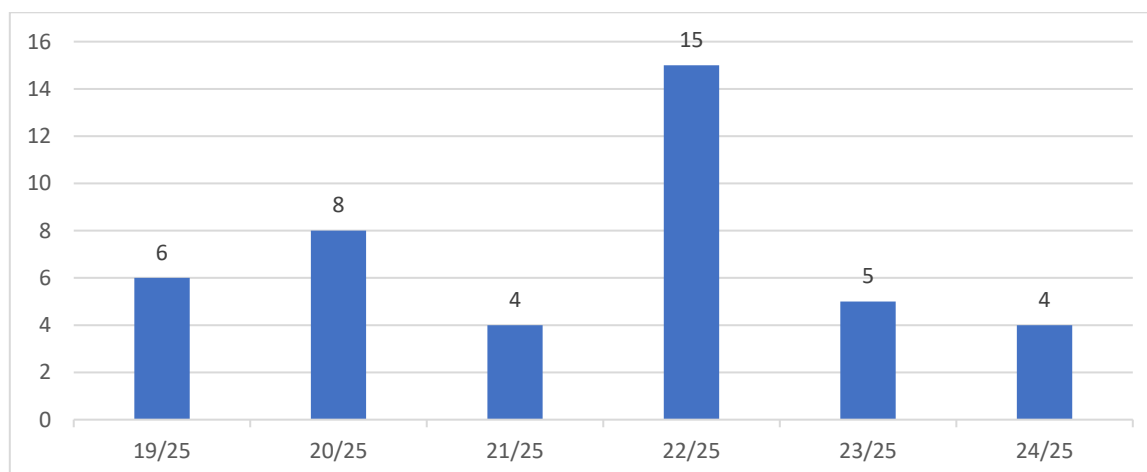
O rol de critérios de avaliação se manteve o mesmo desde a versão do ano 2014 do guia, facilitando assim, a realização de estudos de evolução do estado da preservação de cada repositório.

O guia contém ainda a área de Aspectos Legais que avalia se o repositório cumpre ou infringe as exigências legais aplicáveis, como os direitos intelectuais dos autores. Os critérios de avaliação não incluíram essa área, pois ela pode variar de acordo com as leis e normativas que regem cada repositório, não permitindo a comparação entre eles.

Mesmo que o guia possua a natureza de auditoria, cabe a cada repositório realizar a sua autoavaliação na busca de verificar o atendimento dos critérios, que se dividem em áreas. No entanto, conforme o interesse, o repositório pode submeter a documentação ao Grupo de Trabalho de Avaliação de Repositórios da REBIUN que faz a análise de cumprimento dos critérios e posteriormente publica em sua página web os resultados no formato 20/25 cumpridos (RED DE BIBLIOTECAS UNIVERSITARIAS ESPAÑOLAS, 2022).

A Figura 13 exhibe as quantidades de critérios de avaliação atendidos pelos repositórios REBIUN, sendo a última avaliação disponibilizada no ano de 2021. É possível notar que a menor quantidade de critérios atendidos é 19, enquanto a maior quantidade é 24, verificando-se que nenhum dos 42 repositórios avaliados cumpre com a totalidade de critérios. Os repositórios que mais cumpriram critérios, 24, representam 9,52% do total de repositórios avaliados, no entanto, é possível analisar que em 24 repositórios, 57,14%, são cumpridos ao menos 22 critérios de avaliação. O repositório documental GREDOS da *Universidad de Salamanca* foi avaliado em 2018, tendo cumprido 22 critérios do total de 25 (RED DE BIBLIOTECAS UNIVERSITARIAS ESPAÑOLAS, 2018).

Figura 13 – Totais de critérios atendidos pelos repositórios REBIUN.



Fonte: elaboração própria a partir de (RED DE BIBLIOTECAS UNIVERSITARIAS ESPAÑOLAS, 2022).

Analisando-se os critérios não cumpridos, no caso de 41 repositórios, quase a totalidade, o critério “2.3 Documento de acesso público sobre a preservação dos conteúdos” não é atendido. Esse critério verifica se o repositório disponibiliza um documento sobre o seu compromisso com a disponibilização permanente dos seus conteúdos, se comprometendo a executar ações de preservação que garantam esse acesso. O segundo critério menos cumprido, é o “1.1 Menção do repositório na página principal da instituição”, no qual 33 repositórios, 78,57%, declaram não possuir alguma menção ou link do repositório na página principal da instituição. O atendimento a esse critério pode ser considerado de fácil resolução e pode impactar positivamente na visibilidade do repositório frente a sua comunidade-alvo e entre a REBIUN (RED DE BIBLIOTECAS UNIVERSITARIAS ESPAÑOLAS, 2022).

Tanto os critérios de avaliação quanto a totalidade de critérios do Guia buscam cobrir uma grande extensão dos aspectos que envolvem a preservação digital nos repositórios, no entanto, nota-se um enfoque dedicado ao tratamento de metadados e disponibilização de serviços que se integram com a ferramenta RECOLECTA. Dessa forma, a verificação de cumprimento de critérios se torna bastante específica para os repositórios da REBIUN, porém a utilização de seus critérios em outras formas de avaliação e comunidade de repositórios ainda se torna válida, possibilitando inclusive a construção de benchmarks. No presente estudo, utilizaram-se principalmente os critérios ligados à visibilidade do repositório, uma vez que os demais critérios estão cobertos por outras normas ou são demasiadamente específicos para a REBIUN.

Parte II
Estudo Aplicado

Avaliação de Repositórios Digitais da Rede Cariniana

“Absolutamente nada é tão importante quanto saber em quem confiar”

Radovan Kavicky

4 METODOLOGIA

A metodologia científica se apresenta, de forma geral, como um guia de passos a serem seguidos para que se alcancem os objetivos definidos no estudo. Da mesma forma que um mapa guia um viajante em direção ao seu destino, a metodologia delimita o escopo e define as técnicas para que a realização do estudo culmine na realização dos objetivos pretendidos. Pode ser tratada também como um conjunto de regras e etapas que buscam garantir a validade, a precisão e a confiabilidade dos resultados obtidos em investigações científicas. Essa abordagem visa fornecer uma estrutura lógica para a resolução de problemas, a criação de teorias e a obtenção de conhecimentos que possam ser considerados verdadeiros e baseados em evidências sólidas.

Partindo-se do problema proposto e das hipóteses delineadas, aplicam-se os métodos adequados para a comprovação ou refutação dessas hipóteses. Assim, o capítulo atual apresenta primeiramente a abordagem metodológica utilizada e, em seguida, a divisão do estudo em fases a sua realização.

A primeira fase consistiu na seleção da população de repositórios objeto do estudo. A segunda fase envolveu a seleção de critérios de normas, certificações, auditorias e modelos de maturidade relacionados à preservação digital. A terceira fase tratou da elaboração de critérios utilizados na avaliação dos repositórios participantes. A quarta fase envolveu o envio do formulário de avaliação aos repositórios e a coleta dos respectivos dados. A quinta fase se refere à análise dos dados coletados, que permitem a constatação ou refutação das hipóteses levantadas pelo estudo.

4.1 ABORDAGEM METODOLÓGICA

Como ponto de partida para a definição dos métodos, utilizou-se os problemas levantados pelo próprio estudo, apresentados na seção 0.2, porém reinseridos aqui: **Como afirmar que a preservação digital dos repositórios brasileiros é confiável e de qualidade? Quais mecanismos são utilizados para avaliar os repositórios? Quais áreas funcionais de um repositório devem ser avaliadas para determinar a sua confiabilidade?**

Assim, entendeu-se que o estudo possui natureza exploratória e descritiva. A natureza exploratória é realizada quando há pouco conhecimento prévio sobre o tema de interesse ou quando é necessário investigar um assunto pouco explorado. Seu objetivo principal é proporcionar uma

compreensão mais aprofundada e mais clara do problema, identificando questões relevantes, levantando hipóteses preliminares e sugerindo possíveis abordagens de pesquisa. Complementa-se com a definição de Gil sobre a pesquisa exploratória (2008, p. 27):

Pesquisas exploratórias são desenvolvidas com o objetivo de proporcionar visão geral, de tipo aproximativo, acerca de determinado fato. Este tipo de pesquisa é realizado especialmente quando o tema escolhido é pouco explorado e torna-se difícil sobre ele formular hipóteses precisas e operacionalizáveis.

A natureza descritiva é realizada quando o objetivo é descrever características, comportamentos, atitudes ou fenômenos específicos de forma sistemática e detalhada. Esse tipo de pesquisa procura apresentar uma imagem fiel da situação ou dos eventos estudados, sem interferência ou manipulação significativa das variáveis em questão. Gil define as pesquisas descritivas como (2008, p. 28):

As pesquisas deste tipo têm como objetivo primordial a descrição das características de determinada população ou fenômeno ou o estabelecimento de relações entre variáveis. São inúmeros os estudos que podem ser classificados sob este título e uma de suas características mais significativas está na utilização de técnicas padronizadas de coleta de dados.

Para coleta e análise dos dados, utilizou-se o método qualitativo que busca compreender e interpretar fenômenos explorando significados, percepções e contextos sociais, culturais e individuais, objetivando responder a perguntas do tipo “por quê” e “como”. Em meio a diferentes interpretações sobre o real significado sobre o termo “pesquisa qualitativa” em diferentes áreas, Tesch resume (1990, p. 4, tradução nossa): “O único acordo que encontraríamos entre os pesquisadores qualitativos é que a análise é o processo de dar sentido aos dados narrativos.”

No método qualitativo, a amostragem utilizada é direcionada e não aleatória, buscando-se compreender a realidade desse conjunto, utilizando técnicas de coleta de dados qualitativos como entrevistas e questionários e focando na obtenção de respostas detalhadas e ricas em informações, permitindo uma compreensão profunda do fenômeno estudado. Godoy sintetiza o uso e aplicação das pesquisas qualitativas (1995, p. 21):

[...] um fenômeno pode ser melhor compreendido no contexto em que ocorre e do qual é parte, devendo ser analisado numa perspectiva integrada. Para tanto, o pesquisador vai a campo buscando “captar” o fenômeno em estudo a partir da perspectiva das pessoas nele envolvidas, considerando todos os pontos de vista relevantes. Vários tipos de dados são coletados e analisados para que se entenda a dinâmica do fenômeno.

A partir da análise dos dados, pode-se ainda utilizar o método indutivo, partindo-se do particular rumo a uma generalização. A partir da amostra dos respondentes do estudo pode-se tirar conclusões que contribuem para induzir a situação geral dos repositórios da própria Rede Cariniana.

Para a realização do presente estudo, optou-se pela elaboração de questionário de perguntas dividido em grandes áreas com questões, em sua maioria, fechadas ou semifechadas, de modo a não exaurir as alternativas apresentadas. De natureza autoaplicada, os próprios respondentes analisaram as perguntas e forneceram as respostas a partir dos seus julgamentos acerca das informações solicitadas. Para Gil, o uso de questionários permite a obtenção de informações por parte dos respondentes e suas perguntas devem se alinhar aos objetivos propostos pelo trabalho (2008, p. 121):

Pode-se definir questionário como a técnica de investigação composta por um conjunto de questões que são submetidas a pessoas com o propósito de obter informações sobre conhecimentos, crenças, sentimentos, valores, interesses, expectativas, aspirações, temores, comportamento presente ou passado etc. [...] Construir um questionário consiste basicamente em traduzir objetivos da pesquisa em questões específicas. As respostas a essas questões é que irão proporcionar os dados requeridos para descrever as características da população pesquisada ou testar as hipóteses que foram construídas durante o planejamento da pesquisa.

Com o objetivo de esclarecer o propósito do questionário, foi formulada uma apresentação sobre os objetivos do estudo, a não obrigatoriedade de respondê-lo, o seu caráter anônimo, assim como um contato para esclarecimento de dúvidas e orientações quando fosse o caso. O envio do questionário por mensagem possibilitou aos respondentes participarem do estudo fornecendo as respostas em momentos que julgassem mais oportunos, em oposição à entrevista, que necessita um encontro síncrono para aplicação de questionários e coleta das respostas. A entrevista requer ainda, a necessidade de alocação de tempo para o encontro se realizar, além da necessidade de se reunir, ao mesmo tempo, todas as pessoas com os conhecimentos técnicos requeridos, tornando-se pouco praticável. Por outro lado, pode-se citar como vantagem da realização de entrevistas um grau maior de participação e colaboração, “[...] posto que é mais fácil deixar de responder a um questionário do que negar-se a ser entrevistado” (GIL, 2008, p. 110).

A escolha pelo questionário também possibilita um distanciamento do pesquisador com os respondentes, evitando, assim a influência, enviesamento nas respostas e possíveis constrangimentos. Uma vez que o envio do questionário foi realizado por mensagens, ganha-se praticidade e alcance dos repositórios, uma vez que a população de repositórios pesquisada cobre todas as regiões geográficas do Brasil, impossibilitando as aplicações *in loco* dos questionários.

O questionário possui perguntas objetivas, quando se verifica a existência de determinados instrumentos, por exemplo, assim como perguntas de caráter mais subjetivo, quando é solicitada a adequação da estrutura organizacional do repositório para a realização da preservação digital, por exemplo. A ausência de uma escala objetiva nesses casos, que permita caracterizar com

precisão a adequação, faz com que essas respostas se baseiem na percepção dos respondentes, sendo afetada pelos seus graus de conhecimento do funcionamento e estrutura dos repositórios.

Dentre os desafios da elaboração de questionários para levantamento de dados está o dimensionamento adequado de perguntas e respostas, procurando não tornar o seu preenchimento cansativo, ao mesmo tempo em que deve possuir perguntas e respostas que permitam a extração das informações necessárias à realização do estudo. Dessa forma, o questionário aplicado foi subdividido em seções de perguntas com instruções de auxílio para o preenchimento, assim como justificativas a respeito do motivo de realizar tais perguntas, visando sensibilizar o respondente a respeito da importância em responde-las.

Conforme Sampaio, a elaboração e construção do questionário se realiza em etapas, que se inicia no âmbito das ideias, passa pela criação de rascunhos com conjuntos de respostas e termina com a sua validação através da realização de testes de preenchimento (2022, p. 43):

A elaboração do questionário inicia após a definição do tema, população-alvo e objetivos. A seguir, realiza-se um rascunho considerando o tipo e formato das perguntas, a forma de aplicação e as possibilidades de respostas. Uma vez que o rascunho tenha sido finalizado, é realizado um pré-teste para revisar e validar o instrumento.

Como etapa de pré-teste, a primeira versão do questionário foi aplicada em duas sessões distintas com profissionais arquivistas, que relataram as suas dúvidas e sugestões de modificações de enunciados de perguntas e respostas. Foi identificado que algumas respostas fechadas não permitiam uma fácil compreensão, uma vez que foi assumido um mesmo rol de respostas para várias perguntas distintas, buscando-se uma uniformização do questionário. Assim, optou-se por adaptar as respostas de acordo com cada enunciado das perguntas, facilitando a compreensão e diminuindo as possibilidades de más interpretações. Outro aspecto identificado foi a necessidade de dividir perguntas que solicitavam mais de uma situação ao mesmo tempo, o que poderia ocasionar desvios ou generalizações indesejadas na interpretação das respostas.

Após as primeiras adaptações do questionário, uma nova sessão de teste de aplicação foi realizada, quando se identificaram novas sugestões pontuais e também foi possível realizar a medição de tempo de resposta para a finalização do questionário. O questionário de teste foi completado em 10 minutos, porém foi decidido sugerir o prazo de 15 minutos para o preenchimento do questionário final pelos participantes, considerando-se possíveis consultas e discussões entre as equipes dos repositórios.

4.2 FASES DE DESENVOLVIMENTO DO TRABALHO

Nessa seção são apresentadas as fases executadas para a realização do presente estudo. A primeira fase se caracteriza pela escolha do rol de repositórios para a realização, ou seja, a população sobre a qual o estudo se aplica. A segunda fase se caracteriza pela escolha e seleção de fontes que permitiram a criação do formulário de avaliação dos repositórios, através de normas, auditorias, modelos de maturidade e guias de avaliação. A terceira fase apresenta a elaboração do formulário de avaliação dos repositórios utilizado no presente estudo, incluindo as seções do formulário, perguntas, respostas e justificativas detalhadas. A quarta fase trata dos métodos de envio do formulário de avaliação aos repositórios e a coleta das respectivas respostas, apresentando os métodos utilizados. A quinta fase se refere à análise dos dados recebidos dos formulários para avaliação dos resultados e identificação das situações das confiabilidades dos repositórios.

4.2.1 Fase 1: Seleção da População Objeto de Avaliação Exploratória: Os Repositórios da Rede Cariniana

Para o estudo aplicado, foram escolhidos os repositórios integrantes da Rede Brasileira de Serviços de Preservação Digital, conhecida por Rede Cariniana, criada e coordenada pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT). A Rede Cariniana tem o objetivo de criar uma rede de serviços de preservação digital que garantam a preservação de documentos tendo como critérios que os seus parceiros sejam “[...] instituições de ensino superior, ou instituições que possuam acervos digitais de acesso livre, assim como, de usuários e colaboradores individuais envolvidos com projetos de preservação digital” (INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA, 2014).

Dentre as justificativas para escolha da Rede Cariniana, está a sua trajetória e maturidade, possuindo atualmente 11 anos de existência e parcerias com instituições nacionais e internacionais como a *Stanford University* e a Sociedade Lusófona de Goa, por exemplo. Além desses fatos, destaca-se a presença de uma instituição coordenadora que gerencia o ingresso e o desligamento das instituições e demais parceiros, facilitando o estabelecimento de contato com os responsáveis pelos repositórios. O IBICT, como coordenador, mantém os cadastros de nomes e e-mails dos coordenadores de cada repositório, o que possibilita o envio de comunicações de convite para a participação do presente estudo.

O grande número de repositórios brasileiros registrados na plataforma OpenDOAR, atualmente 171 ao todo, apresenta um grande desafio para a realização de um estudo aplicado de

verificação de critérios de confiabilidade, assim, por esse prisma, justifica-se a escolha da Rede Cariniana para a realização do presente estudo (JOINT INFORMATION SYSTEMS COMMITTEE, 2023). Outro fator limitante é a inexistência de uma forma de contato dos repositórios na plataforma, o que impossibilitaria o envio das mensagens de convite para a participação do estudo.

O grande número de IES existentes no Brasil, 2.457 ao todo e 118 federais, também se mostra um fator impeditivo para a realização de um estudo mais amplo, uma vez que o gerenciamento da comunicação, coleta e análise dos dados recebidos podem agregar demasiadas complexidades para o estudo (INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2022). Assim, se favorece a escolha de um rol de repositórios mais delimitado como a Rede Cariniana.

O compromisso determinado pela Rede Cariniana é o de “[...] dimensionar as dificuldades e propor novos desafios capazes de assegurar o conteúdo preservado e permitir que seu acesso a longo prazo esteja fundamentado em normas e políticas internacionais” (INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA, 2014). Logo, as análises provenientes do presente estudo, originadas a partir de critérios de confiabilidade de normas internacionais, se alinham ao compromisso da Rede Cariniana, complementando as justificativas para a sua escolha.

Para se tornar parceiro da Rede Cariniana, deve-se escolher um dos quatro tipos de parcerias possíveis: parcerias integrais, parcerias institucionais, colaborador individual e instituição usuária. Para serem parceiros integrais, as instituições devem “ser públicas ou privadas, compartilhar infraestrutura de armazenamento digital e possuir documentos digitais em formatos adequados para a preservação digital” (INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA, 2014). Esse formato de parceria prevê um envolvimento constante das instituições, necessitando que elas participem ativamente da rede através de reuniões e no auxílio da sua administração, identificando mudanças tecnológicas das plataformas utilizadas e colaborando na construção de novos serviços da rede.

A formalização da parceria nesse caso ocorre através do estabelecimento de um Acordo de Cooperação Técnica em conjunto com um Plano de Trabalho. Ambos documentos definem as responsabilidades de cada parte do acordo assim como um cronograma de trabalho para que os objetivos da parceria sejam alcançados.

As parcerias institucionais se destinam às instituições que possuem um projeto ligado à preservação digital, não necessitando que a instituição tenha implementada serviços de preservação digital integralmente. Os parceiros devem participar da comunidade da rede a fim de contribuir nos assuntos ligados à preservação digital.

A categoria de colaborador individual se destina a pessoas que participam de pesquisas e projetos ligados à preservação digital que queiram aproveitar e contribuir com as discussões da rede. A categoria de instituição usuária é para instituições que queiram utilizar a infraestrutura da rede, tendo instalado a Plataforma OJS⁶⁶ ou o repositório DSpace⁶⁷.

O Quadro 17 apresenta as 31 instituições parceiras da Rede Cariniana com os respectivos tipos de parceria, Unidade Federativa, região geográfica e natureza da instituição. O IBICT mesmo atuando como coordenador da Rede Cariniana, foi incluído no estudo. As instituições Arquivo Nacional, Fundação Oswaldo Cruz (Fiocruz) e o Jardim Botânico de Brasília, embora não constem como parceiras da Rede Cariniana, possuem Acordo de Cooperação Técnico diretamente com o IBICT e foram incluídas no estudo igualmente. Adicionalmente, foram convidadas a participar do estudo as instituições que já demonstraram interesse em participar da Rede, no entanto, a parceria ainda não havia sido formalizada através dos documentos exigidos, constando no quadro como “em processamento”. Na modalidade de Colaborador Individual, não foi possível encontrar o número total de pesquisadores que a integram.

Pode-se notar a diversidade de parcerias existentes, naturezas de funcionamento das instituições e uma ampla cobertura geográfica, cobrindo todas as regiões brasileiras. Essas diversidades possibilitam que a Rede agregue diferentes pontos de vista, realidades de funcionamento e casos de uso favorecendo a identificação de desafios da preservação e realização de estudos sobre a temática como um todo. A seguir são apresentadas análises detalhadas quanto aos percentuais de distribuição geográfica e natureza das instituições, assim como são apresentados dados sobre os tamanhos das instituições que abrigam os repositórios.

⁶⁶ *Open Journal Systems: software* de gerenciamento e publicação de revistas eletrônicas, distribuído no Brasil pelo IBICT e anteriormente divulgado como SEER (Serviço Eletrônico de Editoração de Revistas) (INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA, 2021).

⁶⁷ *Software* para a criação de repositórios digitais abertos para organizações acadêmicas, comerciais ou sem fins lucrativos (DURASPACE, 2023).

Quadro 17 – Instituições integrantes da Rede Cariniana.

Nome da instituição	Tipo de parceria	Unidade Federativa	Região	Natureza
Arquivo Nacional	Acordo de Cooperação com o IBICT	Rio de Janeiro (RJ)	Sudeste	Pública federal
Centro Universitário Geraldo di Biase (UGB)	Em processamento	Rio de Janeiro (RJ)	Sudeste	Privada
Centro Universitário Internacional Uninter (Uninter)	Parceria Institucional	Paraná (PR)	Sul	Privada
Fundação Biblioteca Nacional (FBN)	Parceria Institucional	Rio de Janeiro (RJ)	Sudeste	Pública federal
Fundação Getúlio Vargas (FGV)	Parceria Institucional	Rio de Janeiro (RJ)	Sudeste	Privada
Fundação Oswaldo Cruz (Fiocruz)	Acordo de Cooperação com o IBICT	Rio de Janeiro (RJ)	Sudeste	Pública federal
Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)	Coordenação da Rede	Distrito Federal (DF)	Centro-Oeste	Pública federal
Instituto Música Brasilis	Parceria Institucional	Rio de Janeiro (RJ)	Sudeste	Privada
Jardim Botânico de Brasília (JBB)	Acordo de Cooperação com o IBICT	Distrito Federal (DF)	Centro-Oeste	Pública estadual
Museu Paraense Emílio Goeldi (MPEG)	Em processamento	Pará (PA)	Norte	Pública federal
<i>Scientific Electronic Library Online Brasil (SciELO)</i>	Parceria Institucional	São Paulo (SP)	Sudeste	Privada
Universidade de Brasília (UnB)	Parceria Integral	Distrito Federal (DF)	Centro-Oeste	Pública federal
Universidade de São Paulo (USP)	Parceria Institucional	São Paulo (SP)	Sudeste	Pública estadual
Universidade Estadual de Campinas (Unicamp)	Parceria Integral	São Paulo (SP)	Sudeste	Pública estadual
Universidade Estadual Paulista (Unesp)	Parceria Integral	São Paulo (SP)	Sudeste	Pública estadual
Universidade Federal da Bahia (UFBA)	Parceria Integral	Bahia (BA)	Nordeste	Pública federal
Universidade Federal da Paraíba (UFPB)	Parceria Integral	Paraíba (PB)	Nordeste	Pública federal
Universidade Federal de Goiás (UFG)	Parceria Integral	Goiás (GO)	Centro-Oeste	Pública federal
Universidade Federal de Juiz de Fora (UFJF)	Em processamento	Minas Gerais (MG)	Sudeste	Pública federal
Universidade Federal de Ouro Preto (UFOP)	Em processamento	Minas Gerais (MG)	Sudeste	Pública federal
Universidade Federal de Pernambuco (UFPE)	Em processamento	Pernambuco (PE)	Nordeste	Pública federal
Universidade Federal de Santa Catarina (UFSC)	Parceria Integral	Santa Catarina (SC)	Sul	Pública federal
Universidade Federal de Santa Maria (UFSM)	Parceria Integral	Rio Grande do Sul (RS)	Sul	Pública federal
Universidade Federal de São Paulo (Unifesp)	Em processamento	São Paulo (SP)	Sudeste	Pública federal
Universidade Federal de Uberlândia (UFU)	Em processamento	Minas Gerais (MG)	Sudeste	Pública federal
Universidade Federal do Ceará (UFC)	Parceria Institucional	Ceará (CE)	Nordeste	Pública federal
Universidade Federal do Pará (UFPA)	Em processamento	Pará (PA)	Norte	Pública federal
Universidade Federal do Rio de Janeiro (UNIRIO)	Em processamento	Rio de Janeiro (RJ)	Sudeste	Pública federal
Universidade Federal do Rio Grande do Norte (UFRN)	Parceria Integral	Rio Grande do Norte (RN)	Nordeste	Pública federal
Universidade Federal do Rio Grande do Sul (UFRGS)	Parceria Integral	Rio Grande do Sul (RS)	Sul	Pública federal
Universidade Tecnológica Federal do Paraná (UTFPR)	Em processamento	Paraná (PR)	Sul	Pública federal

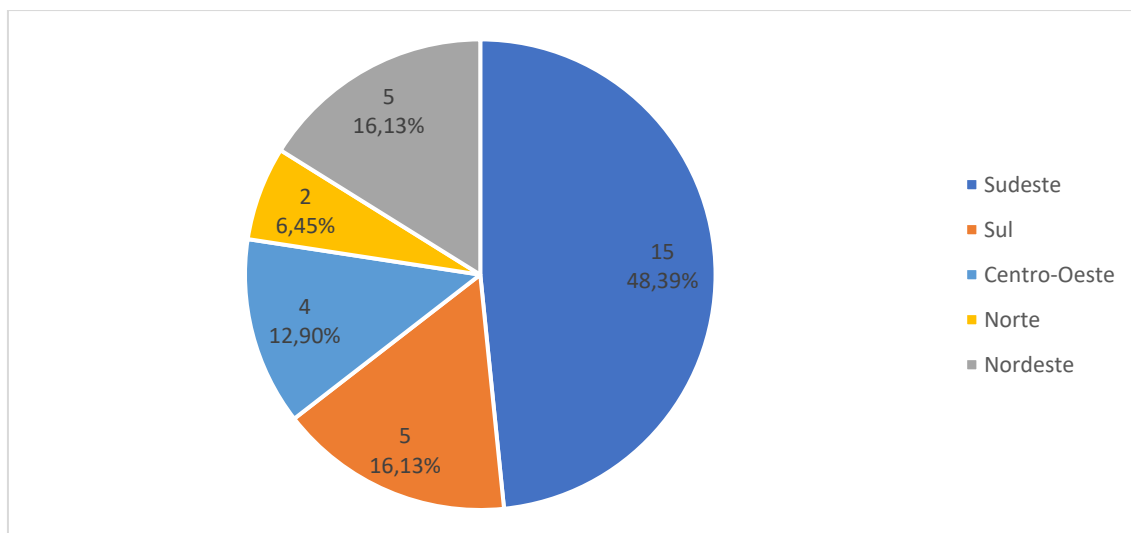
Fonte: (INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA, 2014)

Ao todo, a Rede Cariniana é integrada por um total de 31 instituições, contando com vários tipos de parcerias. Na modalidade de Colaborador Individual, não foi possível encontrar o número total de pesquisadores que a integram.

4.2.1.1 Distribuição Geográfica das Instituições por Regiões

A Figura 14 exibe as instituições parceiras com distribuição por região geográfica do Brasil. É possível notar a predominância de instituições da região Sudeste, contando com 15 instituições que representam uma participação de 48,39%, com destaque para o estado do Rio de Janeiro que possui 7 instituições parceiras da rede.

Figura 14 – Distribuição das instituições parceiras da Rede Cariniana pelas regiões brasileiras.



Fonte: elaboração própria.

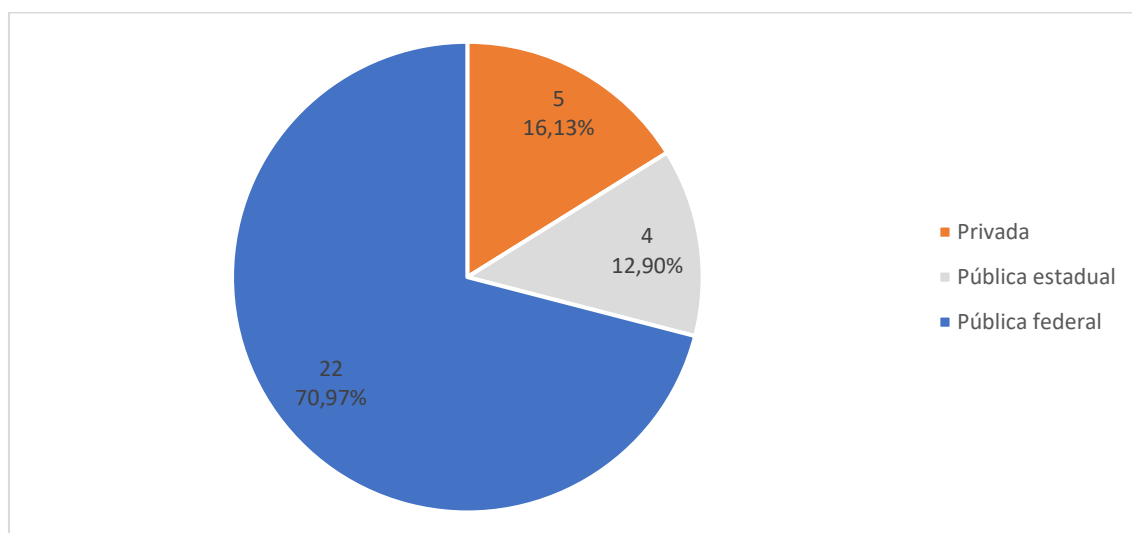
A região Nordeste registra um percentual de 16,13% de participação das instituições parceiras, assim como a região Sul, contando cada uma com 5 instituições. Um fato a ser observado é a presença dos três estados que compõem a região Sul (Rio Grande do Sul, Santa Catarina e Paraná) possuírem instituições parceiras na rede.

Na região Centro-Oeste nota-se a presença de 4 instituições que representam 12,90% de participação, dentre essas, 3 instituições estão localizadas no Distrito Federal. A região Norte é a que possui menor representatividade na Rede Cariniana, contando com 2 instituições, o que representa 6,45% do total de instituições parceiras.

4.2.1.2 Natureza das Instituições

A Figura 15 demonstra a natureza da administração das instituições parceiras da Rede Cariniana. É possível notar um grande predomínio de instituições públicas federais, somando um total de 22 instituições, o que representa 71% de todas as instituições. As instituições de administração estadual apresentam 13% de participação na rede, através de 4 instituições parceiras. Ressalta-se que, como forma de simplificação, optou-se por somar a instituição Jardim Botânico de Brasília às instituições estaduais, mesmo sendo administrada pelo governo do Distrito Federal devido à forma análoga de administração com os demais estados brasileiros. Por fim, as instituições particulares possuem um total de 5 parceiras, que representam 16% de todas as instituições parceiras da Rede Cariniana.

Figura 15 – Natureza das instituições parceiras da Rede Cariniana.



Fonte: elaboração própria.

4.2.1.3 População Geral dos Repositórios de Instituições de Ensino Superior

Como forma de se quantificar os tamanhos das comunidades dos repositórios, apresenta-se na Tabela 1 uma compilação de dados quantitativos das Instituições de Ensino Superior (IES), tanto públicas quanto privadas, parceiras da Rede Cariniana, uma vez que elas apresentam 64,74% de participação na rede como um todo, contando com 21 instituições. Como fonte de dados para obtenção das informações, utilizou-se a Sinopse Estatística da Educação Superior realizada no ano de 2020 (INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2022). As colunas do apresentam, respectivamente, o somatório de cursos de graduação na modalidade presencial e à distância oferecidos pela instituição, a quantidade de alunos matriculados em cursos de graduação, a quantidade total de docentes contratados, a quantidade de

funcionários técnico-administrativos que dão suporte ao funcionamento das instituições e o público total da instituição. Para o cálculo do público total da instituição, utilizou-se o somatório dos alunos de graduação, docentes e funcionários técnico-administrativos. Importante notar que o total de alunos disponibilizado se refere aos cursos de graduação apenas, não incluindo os discentes de cursos de ensino médio e técnicos, muitas vezes oferecidos pelas instituições, além de não incluir os discentes de cursos de pós-graduação, como especializações, mestrados e doutorados.

Através do número do público total é possível se ter uma dimensão da quantidade de pessoas atendidas pelas instituições, que consomem os seus documentos arquivísticos, materiais de pesquisa e diversos outros tipos de materiais físicos e digitais. Muitos desses materiais integram as bases dos respectivos repositórios institucionais, como as monografias e artigos de pesquisa, por exemplo, enquanto outros materiais podem se referir a documentos arquivísticos produzidos e geridos por seus sistemas de gestão documental.

Cabe ressaltar que cada IES possui uma área de abrangência referente às regiões em que estão instaladas, de onde se assume que, além dos seus públicos internos, as comunidades ao seu redor também são impactadas por suas operações, consumindo serviços e materiais digitais de um modo geral. A quantificação desse impacto geral das IES agregando seu público interno e externo não foram quantificadas no presente estudo.

Tabela 1 - Dados das Instituições de Ensino Superior (IES) parceiras da Rede Cariniana.

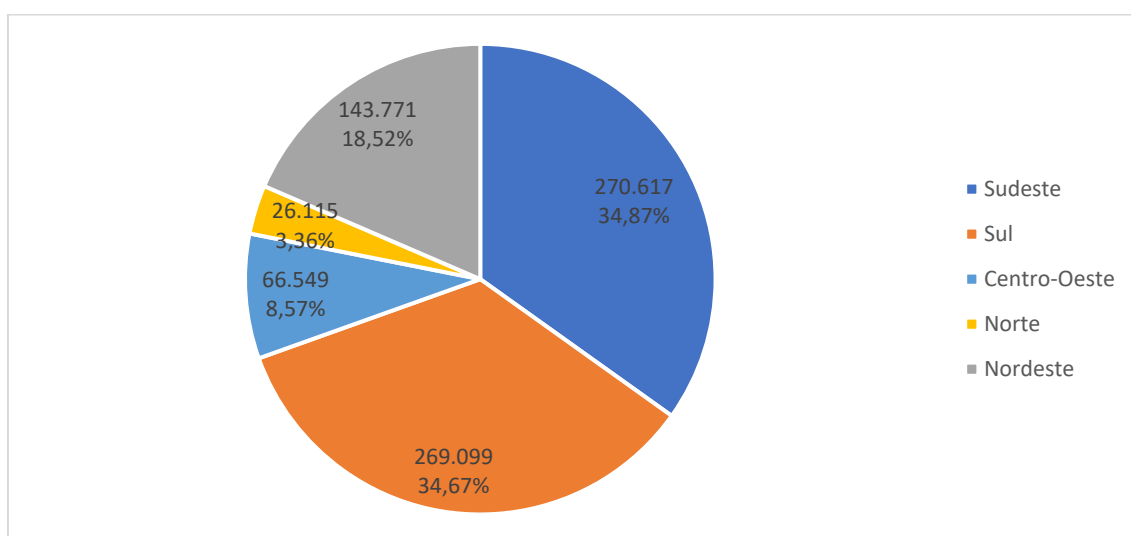
Nome da instituição	Cursos de graduação	Alunos de graduação	Docentes	Funcionários Técnico-administrativos	Público total
Centro Universitário Internacional Uninter (Uninter)	47	129.710	456	875	131.041
Universidade de Brasília (UnB)	132	37.918	2.958	3.204	44.080
Universidade de São Paulo (USP)	113	60.668	5.954	9.333	75.955
Universidade Estadual de Campinas (Unicamp)	64	18.577	1.858	3.512	23.947
Universidade Estadual Paulista (Unesp)	94	33.822	2.925	2.904	39.651
Universidade Federal da Bahia (UFBA)	124	4.511	2.979	3.103	10.593
Universidade Federal da Paraíba (UFPB)	122	23.878	2.798	2.608	29.284
Universidade Federal de Goiás (UFG)	104	18.420	2.277	1.772	22.469
Universidade Federal de Juiz de Fora (UFJF)	90	22.305	1.710	1.251	25.266
Universidade Federal de Ouro Preto (UFOP)	55	12.169	910	719	13.798
Universidade Federal de Pernambuco (UFPE)	108	30.376	2.761	3.840	36.977
Universidade Federal de Santa Catarina (UFSC)	119	31.903	2.596	2.057	36.556
Universidade Federal de Santa Maria (UFSM)	126	22.722	2.350	2.677	27.749
Universidade Federal de São Paulo (Unifesp)	52	12.071	1.778	1.807	15.656
Universidade Federal de Uberlândia (UFU)	97	20.520	2.010	1.852	24.382
Universidade Federal do Ceará (UFC)	125	28.733	2.386	3.370	34.489
Universidade Federal do Pará (UFPA)	151	20.775	2.670	2.670	26.115
Universidade Federal do Rio de Janeiro (UNIRIO)	176	45.581	4.949	1.432	51.962
Universidade Federal do Rio Grande do Norte (UFRN)	115	27.228	2.569	2.631	32.428
Universidade Federal do Rio Grande do Sul (UFRGS)	97	33.810	2.851	2.611	39.272
Universidade Tecnológica Federal do Paraná (UTFPR)	120	30.630	2.703	1.148	34.481
Totais	2.231	666.327	54.448	55.376	776.151

Fonte: elaboração própria a partir de (INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2022).

Dentre as instituições parceiras, chama a atenção o Centro Universitário Internacional Uninter por possuir 129.710 alunos de graduação, representando 19,47% do total de 666.327 alunos de todas as instituições. A Uninter possui sua sede localizada na cidade de Curitiba (PR) e atua principalmente com educação a distância, ofertando mais de 400 cursos e possuindo 800 polos de apoio presencial aos seus alunos (CENTRO UINVERSITÁRIO INTERNACIONAL UNINTER, 2022).

A Figura 16 apresenta o público geral computado para as IES parceiras da Rede Cariniana distribuído pelas regiões brasileiras. A partir dos dados, nota-se que as regiões Sul e Sudeste possuem praticamente o mesmo quantitativo de público entre si e representam o equivalente a 69,54% do público total de todas as instituições. A região Norte é a que possui menor participação de instituições parceiras, contando com 2 instituições, o seu quantitativo de público também é o menor dentre as regiões, registrando 3,36% do total do público geral de todas as IES.

Figura 16 – Público geral das IES parceiras da Rede Cariniana distribuído pelas regiões brasileiras.



Fonte: elaboração própria a partir de (INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2022).

4.2.2 Fase 2: Seleção de Fontes e Identificação de Critérios de Avaliação para a Certificação de Repositórios Confiáveis

O objetivo dessa fase foi selecionar as fontes que deveriam ser tomadas como documentos de referência para a identificação de diretrizes e orientações para a certificação de repositórios digitais. Em geral, deveriam ser documentos normativos ou elaborados por instituições de referência na área. A partir da análise das fontes

escolhidas, foram estabelecidos os critérios de avaliação, que foram agrupados em áreas de acordo com sua afinidade. Assim, essa fase foi dividida em dois estágios, que são apresentados a seguir.

4.2.2.1 *Seleção de Fontes Normativas*

Nessa seção são detalhadas as fontes selecionadas para a elaboração dos critérios de avaliação dos repositórios participantes do estudo. Buscou-se utilizar em primeiro lugar, normas internacionais bem estabelecidas, regulamentadas pela ISO, em seguida, foram incluídos selos de confiança vastamente utilizados por repositórios mundiais dos quais as suas equipes, inclusive, interagiram com a elaboração de normas ISO. Complementarmente, foram analisados instrumentos de auditoria e modelo maturidade que, quando aplicados, permitem aos repositórios o alcance de um autoconhecimento que favorece a mudança de comportamento e adaptações de estruturas. Em seguida, buscou-se por instrumentos de escopos nacionais tanto do Brasil quanto da Espanha. Assim, foram selecionadas as seguintes fontes normativas:

- ISO 16363:2012. “Sistemas de transferência de informação e dados especiais. Auditoria e Certificação de Repositórios Digitais Confiáveis”. Trata-se de uma norma que proporciona um vasto conjunto de critérios de avaliação dos repositórios, avaliando-os em três macro áreas: infraestrutura da organização, gestão dos objetos digitais e gestão de riscos de segurança e confiança (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012). A seção 3.4.3 apresenta a norma em detalhes;
- Nestor Seal. Trata-se de um catálogo de 42 critérios dividido em três grandes áreas: estrutura organizacional, gerenciamento de objetos e infraestrutura e segurança. De acordo com o nível de conformidade com os critérios, o repositório obtém o selo de confiança e pode exibi-lo em seu site. A equipe elaboradora do catálogo possui forte envolvimento com a criação da norma alemã DIN 31644:2012-04⁶⁸ que estabelece critérios para repositórios digitais confiáveis. (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009). O Nestor Seal é apresentado em mais detalhes na seção 3.4.5;
- CoreTrust Seal. Apresenta-se como uma alternativa inicial na busca de certificações, contendo 16 requisitos e permitindo que o repositório realize uma auditoria interna

⁶⁸ (DEUTSCHES INSTITUT FÜR NORMUNG, 2012)

e posteriormente envie os resultados e comprovantes para apreciação e conferência do corpo técnico responsável pelo selo (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019). O CoreTrust Seal, seus requisitos e história são apresentados na seção 3.4.6;

- DRAMBORA. Apresenta-se como um método de auditoria que analisa os riscos os quais o repositório está exposto, pontuando-os de acordo com o grau de impacto e probabilidade de ocorrência, além de analisar as ações planejadas para evitá-los ou mitigá-los. Os repositórios devem elaborar um catálogo próprio de riscos, no entanto, é fornecido um catálogo de exemplo que pode ser usado como base para os trabalhos (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007). O DRAMBORA é apresentado em detalhes na seção 3.4.7;
- Matriz de níveis de preservação NDSA. Trata-se de uma matriz que analisa o grau de evolução das soluções e estratégias utilizadas pelo repositório em 5 áreas funcionais. Pode ser usado como uma espécie de fotografia do estado atual do repositório, assim como para a medição de sua evolução (NATIONAL DIGITAL STEWARDSHIP ALLIANCE, 2022). A matriz NDSA é apresentada em detalhes na seção 3.4.2.
- Em termos nacionais, do Brasil, utilizou-se a Resolução 43 do CONARQ que estabelece requisitos para a implementação de um Repositório Arquivístico Digital Confiável (RDC-Arq) (CONSELHO NACIONAL DE ARQUIVOS, 2015), que por sua vez se baseia na ISO 14721:2012⁶⁹. A Resolução é apresentada em mais detalhes na seção 2.5;
- Da Espanha, utilizou-se o Guia de Avaliação de Repositórios REBIUN que analisa principalmente a forma como os repositórios científicos interoperam entre si e disseminam os seus materiais digitais (AZORÍN, BARRUECO, *et al.*, 2018). O Guia é apresentado em detalhes na seção 3.4.8;

A seguir são apresentadas as correlações entre as fontes utilizadas para a elaboração dos critérios de avaliação do presente estudo.

⁶⁹ (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012)

4.2.2.2 *Identificação, Seleção e Classificação de Critérios*

A partir das fontes selecionadas, indicadas na seção anterior, se iniciou um trabalho de análise com dois objetivos concretos: 1) Identificar e selecionar critérios similares de avaliação de repositórios no conjunto de fontes selecionadas, e; 2) Agrupar os critérios selecionados em áreas afins.

De forma a embasar a elaboração das perguntas do questionário aplicado nas instituições, primeiramente, buscou-se identificar critérios com finalidades coincidentes ou similares entre as normas e instrumentos de auditoria e certificação estudadas. Essa análise permitiu a formulação de perguntas de avaliação que buscassem cobrir vários critérios e aspectos dos instrumentos ao mesmo tempo.

Cabe esclarecer que cada instrumento possui particularidades e uma ótica específica para análise dos seus requisitos, assim, uma instituição ao atender o aspecto analisado em uma pergunta do presente estudo, não está necessariamente cumprindo todos os requisitos desses instrumentos, porém é possível se conhecer a quais critérios se relaciona.

Os critérios foram agrupados em grandes áreas, apresentando os critérios dos instrumentos usados como referência para a sua elaboração, assim como as quantidades de critérios equivalentes em cada um desses instrumentos.

Ressalta-se que alguns critérios dos instrumentos poderiam pertencem a mais de uma área ao mesmo tempo, devido a sua abrangência, porém foram alocados em somente uma das áreas para evitar recontagem de critérios. O monitoramento e verificação de formatos de arquivos suportados pelo repositório, critério avaliado no presente estudo, perpassa diversas áreas dos processos do repositório, como a ingestão e o acesso, assim, os critérios relacionados a esse tema estão cobertos pelas Políticas de Preservação e de Acesso.

No caso do DRAMBORA, embora se saiba que a avaliação de riscos deve ser personalizada de acordo com a realidade de cada repositório, se utilizou o catálogo de riscos de exemplos como base para a elaboração dos critérios do presente estudo. Outra situação a ser destacada é o uso de critérios, tanto da terceira versão (AZORÍN MILLARUELO, CÍVICO MARTÍN, *et al.*, 2015), quanto da quarta e última versão do Guia de avaliação de repositórios institucionais FECYT-REBIUN. Nos critérios em que essa situação ocorre, é sinalizada a versão utilizada como base para o critério. Os critérios dos respectivos níveis das áreas funcionais da Matriz de Preservação NDSA

não foram contabilizados para os totais de critérios cobertos, uma vez que eles evoluem em grau de sofisticação e atendimento dos requisitos, podendo um mesmo repositório estar em mais de um nível ao mesmo tempo.

Assim sendo, foram identificados 13 critérios de avaliação que foram agrupados em 7 grandes áreas funcionais, apresentados no Quadro 18.

Quadro 18 – Seleção de critérios de avaliação e suas áreas funcionais.

Critérios	Áreas
1. Tipo/natureza do repositório 2. Definição ou delimitação da Comunidade-Alvo 3. Definição de missão do repositório	1. Contexto Geral
4. Existência de Política ou Plano Estratégico de Preservação 5. Existência de Plano de Continuidade	2. Estratégia e responsabilidade
6. Infraestrutura organizacional e orçamento adequados 7. Equipe adequada (quantidade e capacitação) à preservação digital	3. Infraestrutura organizacional
8. Infraestrutura física e técnica adequadas à preservação digital	4. Infraestrutura física e técnica
9. Adoção de medidas de segurança contra ataques e acessos indevidos 10. Adoção de medidas que garantam a integridade, autenticidade e recuperação dos dados em caso de perdas	5. Segurança, integridade e autenticidade dos dados
11. Existência e padrões de metadados adotados	6. Metadados
12. Existência de Plano ou Política de Acesso 13. Facilidade de acesso ao repositório e promoção de seu uso	7. Acesso e visibilidade

Fonte: elaboração própria.

A seguir, são analisados os critérios agrupados por cada uma das 7 grandes áreas funcionais, apresentando-se os instrumentos atendidos e as respectivas quantidades de critérios atendidas.

A primeira área analisada, “1) Contextualização geral”, é apresentada no Quadro 19 e agrupa critérios que têm como objetivo uma caracterização geral dos repositórios, sendo que no critério “1.1) Tipo/natureza do repositório” busca-se conhecer qual a tipologia e/ou tipos de materiais custodiados pelo repositório. Esse critério engloba 1 critério das certificações CoreTrust Seal e Nestor Seal, cada uma. Ao todo, o critério cobre 2 critérios das certificações citadas. O critério “1.2) Definição ou delimitação da Comunidade-Alvo” busca analisar se o repositório possui definida ou de certa forma delimitada sobre o público que utiliza os seus serviços. É voltada à essa comunidade que o funcionamento do repositório e estabelecimento de suas políticas se direciona, assim, ao não se conhecer a Comunidade-Alvo, o repositório corre o risco de não saber a quem servir e quais requisitos cumprir.

Quadro 19 – Análise de critérios e similaridades da área “1) Contexto Geral”.

Área:
1) Contexto geral
Critério:
1.1) Tipo/natureza do repositório (institucional, de domínio/temático, biblioteca, museu, arquivo, dados de pesquisa, etc)
Instrumentos atendidos:
1) CoreTrust Seal: requisito “R0. Informações básicas e contexto”, item “2 Tipo de repositório” (2022, p. 6, tradução nossa); 2) Nestor Seal: critério “1.1 O repositório digital desenvolveu critérios para a seleção de seus objetos digitais” (2009, p. 10, tradução nossa);
Quantidade de critérios:
CoreTrust Seal: 1 ; Nestor Seal: 1 . Total: 2 critérios.
Critério:
1.2) Definição ou delimitação da Comunidade-Alvo
Instrumentos atendidos:
1) CoreTrust Seal: requisito “R0. Informações básicas e contexto”, item “4 Comunidade-Alvo” (2022, p. 6, tradução nossa); 2) Nestor Seal: critério “1.3 O repositório digital definiu sua(s) Comunidade(s)-Alvo” (2009, p. 11, tradução nossa); 3) ISO 16363:2012: critério “3.3.1 O repositório deve ter sua Comunidade-Alvo e base(s) de conhecimento associada(s) definida(s) e deve ter essas definições adequadamente acessíveis” (2017, p. 27, tradução nossa); 4) DRAMBORA (2007, p. 140, 141, 143, 144, tradução nossa): a. R06 Os requisitos da comunidade mudam substancialmente; b. R07 Requisitos da comunidade mal compreendidos ou mal comunicados; c. R09 <i>Feedback</i> da comunidade não recebido; d. R10 O <i>feedback</i> da comunidade não foi levado em consideração. 5) Guia de avaliação de repositórios FECYT-REBIUN: critério “2.5.- Política de conteúdo do repositório” (2021, p. 13, tradução nossa).
Quantidade de critérios:
CoreTrust Seal: 1 ; Nestor Seal: 1 ; ISO 16363:2012: 1 ; DRAMBORA: 4 ; Guia de avaliação de repositórios FECYT-REBIUN: 1 . Total: 8 critérios.
Critério:
1.3) Definição de missão do repositório
Instrumentos atendidos:
1) CoreTrust Seal: requisito “R01. O repositório tem uma missão explícita de fornecer acesso e preservar objetos digitais.” (2022, p. 8, tradução nossa); 2) Nestor Seal (2009, p. 10, 11, tradução nossa): a. 1. O repositório digital seus objetivos definidos; b. 1.2 O repositório digital assume a responsabilidade pela preservação a longo prazo das informações representadas pelos objetos digitais. 3) ISO 16363:2012: critério “3.1.1 O repositório deve ter uma declaração de missão que reflita o compromisso com as informações digitais para preservação, retenção a longo prazo, gerenciamento e acesso.” (2017, p. 23, tradução nossa); 4) DRAMBORA (2007, p. 138, 153, 154, tradução nossa): a. R04 Objetivos de negócio não atingidos; b. R19 Incapacidade de avaliar o sucesso do repositório; c. R20 Falsa percepção da extensão do sucesso do repositório. 5) Guia de avaliação de repositórios FECYT-REBIUN: critério “2.3.- Existe uma declaração sobre a missão e objetivos do repositório” (2021, p. 13, tradução nossa)
Quantidade de critérios:
CoreTrust Seal: 1 ; Nestor Seal: 2 ; ISO 16363:2012: 1 ; DRAMBORA: 3 ; Guia de avaliação de repositórios FECYT-REBIUN: 1 . Total: 8 critérios.

Fonte: elaboração própria.

De um modo geral, as auditorias confrontam o cumprimento dos critérios em relação ao esperado pelas Comunidades-Alvo dos repositórios. Ao todo, o critério

possui uma cobertura de 8 critérios dos instrumentos CoreTrust Seal, Nestor Seal, ISO 16363:2012, DRAMBORA e Guia de avaliação de repositórios FECYT-REBIUN.

O critério “1.3) Definição de missão do repositório” analisa se o repositório declara a sua missão, pela qual orienta as suas ações e processos. É através da declaração de missão e dos termos utilizados que o repositório torna explícito o seu compromisso em realizar a preservação digital confiável a longo prazo, formalizando esse compromisso. Ao todo, o critério cobre 8 critérios dos instrumentos CoreTrust Seal, Nestor Seal, ISO 16363:2012, DRAMBORA e Guia de avaliação de repositórios FECYT-REBIUN.

A área “2) Estratégia e responsabilidade”, apresentada no Quadro 20, se refere aos critérios que analisam a presença de documentos estratégicos para o repositório e que definem as suas responsabilidades e limites de atuação.

O critério “2.1) Existência de Política de Preservação ou Plano Estratégico de Preservação” busca verificar se o repositório possui esses documentos que estabelecem a forma como a preservação é realizada, quais os critérios mínimos a serem seguidos, permissões de uso e estabelecem as responsabilidades por cada órgão ou setor da instituição envolvidos na preservação digital. Ao todo, o critério possui uma cobertura de 18 critérios dos instrumentos CoreTrust Seal, Nestor Seal, ISO 16363:2012, DRAMBORA e Guia de avaliação de repositórios FECYT-REBIUN.

O critério “2.2) Existência de Plano de Continuidade” tem o objetivo de verificar se os repositórios possuem planos que garantam a preservação digital confiável em situações em que a operacionalização do próprio repositório é comprometida. Ao todo, o critério possui uma cobertura de 5 critérios dos instrumentos CoreTrust Seal, Nestor Seal, ISO 16363:2012 e DRAMBORA.

Quadro 20 – Análise de critérios e similaridades da área “2) Estratégia e responsabilidade”.

Área: 2) Estratégia e responsabilidade
Critério: 2.1) Existência de Política de Preservação ou Plano Estratégico de Preservação
Instrumentos atendidos: 1) CoreTrust Seal: requisito “R09. O repositório assume a responsabilidade pela preservação de longo prazo e gerencia essa função de forma planejada e documentada.” (2022, p. 14, tradução nossa); 2) Nestor Seal: critério “8. O repositório digital possui um plano estratégico para suas medidas técnicas de preservação (plano de preservação)” (2009, p. 11, tradução nossa); 3) ISO 16363:2012 (2007, p. 138, 153, 154, tradução nossa): a. 3.1.2 O repositório deve ter um Plano Estratégico de Preservação que defina o foco que o repositório manterá no suporte de longo prazo de sua missão; b. 3.1.3 O repositório deve ter uma Política de coleção/fundo ou outro documento que especifique o tipo de informação que ele reterá, conservará, gerenciará e fornecerá acesso; c. 3.3.2 O repositório deve ter Políticas de Preservação para garantir o cumprimento de seu Plano Estratégico de Preservação; d. 3.3.2.1 O repositório deve ter mecanismos para revisão, atualização e desenvolvimento contínuos de suas Políticas de Preservação à medida que o repositório cresce e à medida que a tecnologia e as práticas da comunidade evoluem; e. 4.3.1 O repositório deve ter estratégias de preservação documentadas relevantes para seu acervo; f. 4.3.2.1 O repositório deve ter dispositivos para monitorar e notificar quando as Informações de Representação são inadequadas para a Comunidade-Alvo compreender os conjuntos de dados; g. 4.3.3 O repositório deve possuir mecanismos para alterar seus planos de preservação em decorrência de suas atividades de monitoramento; h. 4.3.4 O repositório deve fornecer evidências da eficácia de suas atividades de preservação. 4) DRAMBORA (2007, p. 145, 146, 147, 148, 199, 201, tradução nossa): a. R11 O negócio falha em preservar as características essenciais da informação digital; b. R12 As políticas e procedimentos de negócios são desconhecidos; c. R13 Políticas e procedimentos de negócios são ineficientes; d. R14 As políticas e procedimentos de negócios são inconsistentes ou contraditórios; e. R65 Os planos de preservação não podem ser implementados; f. R67 Incapacidade de validar a eficácia da preservação. 5) Guia de avaliação de repositórios FECYT-REBIUN (2021, p. 14, 15, tradução nossa): a. 2.6.- Política de preservação digital dos conteúdos do repositório; b. 3.3.- Existe documentação disponível que ajuda o autor a decidir se deve ou não depositar o seu documento.
Quantidade de critérios: CoreTrust Seal: 1 ; Nestor Seal: 1 ; ISO 16363:2012: 8 ; DRAMBORA: 6 ; Guia de avaliação de repositórios FECYT-REBIUN: 2 . Total: 18 critérios.
Critério: 2.2) Existência de Plano de Continuidade
Instrumentos atendidos: 1) CoreTrust Seal: requisito “R03. O Repositório tem um plano para garantir o acesso contínuo e a preservação de seus dados e metadados” (2022, p. 10, tradução nossa); 2) Nestor Seal: critério “4.6 A continuidade das tarefas de preservação é assegurada mesmo além da existência do repositório digital” (2009, p. 17, tradução nossa); 3) ISO 16363:2012 (2017, p. 24, tradução nossa): a. 3.1.2.1 O repositório deve ter um plano de continuidade, planos de contingência e/ou contratos de garantia adequados no caso de o repositório deixar de funcionar ou a instituição que o financia ou dirige alterar substancialmente o seu âmbito; b. 3.1.2.2 O repositório deve monitorar seu ambiente organizacional para determinar quando executar seu plano de sucessão, planos de contingência e/ou contratos de garantia. 4) DRAMBORA: risco “R08 Cessação forçada das operações do repositório” (2007, p. 142, tradução nossa).
Quantidade de critérios: CoreTrust Seal: 1 ; Nestor Seal: 1 ; ISO 16363:2012: 2 ; DRAMBORA: 1 . Total: 5 critérios.

Fonte: elaboração própria.

A área “3) Infraestrutura organizacional”, apresentada no Quadro 21, avalia os aspectos ligados à forma como o repositório se organiza para a distribuição de responsabilidades e execução das atividades de preservação.

Quadro 21 – Análise de critérios e similaridades da área “3) Infraestrutura organizacional”.

Área:
3) Infraestrutura organizacional
Critério:
3.1) Existência de infraestrutura organizacional e orçamento adequados à preservação digital
Instrumentos atendidos:
1) CoreTrust Seal: requisito “R05. O repositório tem financiamento adequado e número suficiente de funcionários gerenciados por meio de um sistema claro de governança para realizar a missão com eficácia” (2022, p. 11, tradução nossa); 2) Nestor Seal (2009, p. 15, 16, tradução nossa): a. 4. A forma organizacional é apropriada para o repositório digital; b. 4.1 É assegurado financiamento adequado ao repositório digital; c. 4.3 Existem estruturas organizacionais apropriadas para o repositório digital. 3) ISO 16363:2012 (2017, p. 25, 26, 66, 68, tradução nossa): a. 3.2.1 O repositório deve ter identificado e estabelecido as responsabilidades necessárias a serem cumpridas e deve ter designado pessoal com habilidades e experiência adequadas para cumprir essas funções; b. 3.2.1.1 O repositório deve ter identificado e estabelecido as funções que precisa realizar; c. 5.1.1.1.4 O repositório deve ter procedimentos, compromisso e financiamento para substituir o <i>hardware</i> quando a avaliação indicar que isso é necessário; d. 5.1.1.1.8 O repositório deve ter procedimentos, compromisso e financiamento para substituir o <i>software</i> quando a avaliação indicar que isso é necessário. 4) DRAMBORA (2007, p. 137, 159, 160, 162, 163, tradução nossa): a. R03 A atividade é negligenciada ou os recursos alocados são insuficientes; b. R25 Finanças insuficientes para atender aos compromissos do repositório; c. R26 Alocação incorreta de recursos financeiros; d. R28 Deficiências financeiras ou restrições de renda; e. R29 Redução orçamentária.
Quantidade de critérios:
CoreTrust Seal: 1; Nestor Seal: 3; ISO 16363:2012: 4; DRAMBORA: 5. Total: 13 critérios.
Critério:
3.2) Existência de equipe adequada (quantidade e capacitação) à preservação digital
Instrumentos atendidos:
1) CoreTrust Seal (2022, p. 11, 12, tradução nossa): critério “R06. O repositório adota mecanismos para garantir conhecimento especializado, orientação e <i>feedback</i> contínuos, sejam eles internos ou externos”; 2) Nestor Seal: critério “4.2 Existe disponível um número suficiente de funcionários devidamente qualificados” (2009, p. 15, tradução nossa); 3) ISO 16363:2012 (2017, p. 26, tradução nossa): a. 3.2.1.2 O repositório deve ter o número adequado de funcionários para dar suporte a todas as funções e serviços; b. 3.2.1.3 O repositório deve ter um programa de desenvolvimento profissional ativo que ofereça oportunidades de desenvolvimento para a equipe com habilidades e conhecimentos. 4) DRAMBORA (2007, p. 155, 156, 157, 158, tradução nossa): a. R21 Perda de membro(s) chave da equipe; b. R22 A equipe sofre deterioração de suas habilidades; c. R23 As habilidades da equipe se tornam obsoletas; d. R24 Incapacidade de avaliar a eficácia ou a adequação da equipe.
Quantidade de critérios:
CoreTrust Seal: 1; Nestor Seal: 1; ISO 16363:2012: 2; DRAMBORA: 4. Total: 8 critérios.

Fonte: elaboração própria.

O critério “3.1) Existência da infraestrutura organizacional e orçamento adequados à preservação digital” avalia se o repositório possui divisões e/ou setores

adequados em tamanho e responsabilidades para que a preservação digital aconteça plenamente. O critério ainda avalia se os recursos financeiros empregados no repositório são adequados para a manutenção de suas atividades, investimentos em melhorias de evolução e capacitações para as equipes. Ao todo, o critério possui uma cobertura de 13 critérios dos instrumentos CoreTrust Seal, Nestor Seal, ISO 16363:2012 e DRAMBORA.

O critério “3.2) Existência de equipe adequada (quantidade e capacitação) à preservação digital” visa identificar se existem pessoas suficientes no repositório para que a preservação ocorra de maneira adequada, sem que haja sobrecarga de responsabilidades e dependência de em pessoas-chave da instituição. O critério ainda avalia os aspectos de capacitação dessas equipes, buscando identificar se os conhecimentos existentes e em uso são suficientes para a realização da preservação digital confiável a longo prazo e se existem previsões e planejamentos de capacitações para aquisição ou atualização de conhecimentos. Ao todo, o critério possui uma cobertura de 8 critérios dos instrumentos CoreTrust Seal, Nestor Seal, ISO 16363:2012 e DRAMBORA.

A área “4) Existência de infraestrutura física e técnica adequadas à preservação digital”, apresentada no Quadro 22, possui como único critério “4.1) Existência de infraestrutura física e técnica adequadas à preservação”. Esse critério avalia a estrutura física em si do repositório em relação à presença de salas e prédios que permitam que as suas equipes realizem suas atividades adequadamente, além da adequação dos espaços físicos para a instalação dos equipamentos utilizados. Outro aspecto analisado se refere à adequação dos próprios equipamentos, se referindo aos servidores de aplicação, armazenamento e demais equipamentos de rede necessários para a preservação digital. Ao todo, o critério possui uma cobertura de 18 critérios dos instrumentos CoreTrust Seal, Nestor Seal, ISO 16363:2012 e DRAMBORA.

Quadro 22 – Análise de critérios e similaridades da área “4) Infraestrutura física e técnica”.

Área:
4) Infraestrutura física e técnica
Critério:
4.1) Existência de infraestrutura física e técnica adequadas à preservação digital
Instrumentos atendidos:
1) CoreTrust Seal: requisito “R15. O repositório é gerenciado em sistemas operacionais bem suportados e outros <i>softwares</i> e <i>hardwares</i> de infraestrutura básica adequados aos serviços que presta à sua Comunidade-Alvo” (2022, p. 18, tradução nossa); 2) Nestor Seal: critério “13 A infraestrutura de TI é adequada” (2009, p. 36, tradução nossa); 3) ISO 16363:2012 (2017, p. 64, 65, 66, 67, 68, tradução nossa): a. 5.1.1.1.1 O repositório deve possuir tecnologias de <i>hardware</i> adequadas aos serviços oferecidos às Comunidades-Alvo; b. 5.1.1.1.2 O repositório deve ter procedimentos para monitorar e receber notificações quando mudanças na tecnologia de <i>hardware</i> forem necessárias; c. 5.1.1.1.3 O repositório deve ter procedimentos para avaliar quando é necessário fazer alterações no <i>hardware</i> existente; d. 5.1.1.1.4 O repositório deve ter procedimentos, compromisso e financiamento para substituir o <i>hardware</i> quando a avaliação indicar que isso é necessário; e. 5.1.1.1.5 O repositório deve possuir tecnologias de <i>software</i> adequadas aos serviços que oferece às suas Comunidades-Alvo; f. 5.1.1.1.6 O repositório deve ter procedimentos para monitorar e receber notificações quando forem necessárias alterações no <i>software</i>; g. 5.1.1.1.7 O repositório deve ter procedimentos para avaliar quando são necessárias alterações no <i>software</i> existente; h. 5.1.1.1.8 O repositório deve ter procedimentos, compromisso e financiamento para substituir o <i>software</i> quando a avaliação indicar que isso é necessário. 4) DRAMBORA (2007, p. 164, 165, 166, 167, 168, 177, 181, 208, tradução nossa): a. R30 Falha ou incompatibilidade de <i>hardware</i>; b. R31 Falha ou incompatibilidade de <i>software</i>; c. R32 <i>Hardware</i> ou <i>software</i> incapaz de suportar os objetivos emergentes do repositório; d. R33 Obsolescência de <i>hardware</i> ou <i>software</i>; e. R34 Degradação ou obsolescência da mídia; f. R43 indisponibilidade de serviços básicos; g. R47 Incapacidade de avaliar a eficácia da infraestrutura técnica e segurança; h. R74 Indisponibilidade de serviços de entrega de informações.
Quantidade de critérios:
CoreTrust Seal: 1 ; Nestor Seal: 1 ; ISO 16363:2012: 8 ; DRAMBORA: 8 . Total: 18 critérios.

Fonte: elaboração própria.

A área “5) Segurança, integridade e autenticidade dos dados”, apresentada no Quadro 23 possui critérios para a verificação dos aspectos relacionados à garantia de segurança dos repositórios, além da manutenção de integridade e autenticidade dos materiais custodiados. O critério “5.1) Adoção de medidas de segurança contra ataques e acessos indevidos” analisa se o repositório implementa mecanismos contra invasões, tanto físicas quanto virtuais, que impeçam que os dados sejam corrompidos ou vazados, comprometendo a preservação realizada. Ao todo, o critério possui uma cobertura de 16 critérios dos instrumentos CoreTrust Seal, Nestor Seal, ISO 16363:2012, DRAMBORA e Guia de avaliação de repositórios FECYT-REBIUN.

O critério “5.2 Adoção de medidas que garantam a integridade, autenticidade e recuperação dos dados em caso de perdas” procura identificar se o

repositório adota padrões de verificação de integridade como fixidez e algoritmos de CRC, por exemplo, sobre os materiais custodiados, além de verificar a existência de registros de operações de preservação que afetam os materiais preservados. Esses registros, juntamente com os metadados dos materiais recebidos na ingestão apoiam a autenticidade dos materiais, assim, pode-se notar uma relação entre os critérios dessa área com a área “6) Metadados”. Outro aspecto verificado é a existência de planos de backup que garantam que os dados armazenados copiados estejam seguros e possam ser utilizados para recuperação em casos de perdas de informações por quaisquer motivos. Ao todo, o critério possui uma cobertura de 35 critérios dos instrumentos CoreTrust Seal, Nestor Seal, ISO 16363:2012, DRAMBORA e Guia de avaliação de repositórios FECYT-REBIUN.

Quadro 23 – Análise de critérios e similaridades da área “5) Segurança, integridade e autenticidade dos dados”.

Área:
5) Segurança, integridade e autenticidade dos dados
Critério:
5.1) Adoção de medidas de segurança contra ataques e acessos indevidos
Instrumentos atendidos:
1) CoreTrust Seal: requisito “R16. O repositório protege a instalação e seus dados, metadados, produtos, serviços e usuários” (2022, p. 19, tradução nossa); 2) Nestor Seal (2009, p. 36, 37, tradução nossa): a. 13.2 A infraestrutura de TI implementa os requisitos de segurança do sistema de segurança de TI; b. 14 A infraestrutura protege o repositório digital e seus objetos digitais. 3) ISO 16363:2012 (2017, p. 74, 75, 76, tradução nossa): a. 5.2.1 O repositório deve manter uma análise sistemática de segurança dos fatores de risco associados aos dados, sistemas, pessoal e instalações; b. 5.2.2 O repositório deve ter controles adequadamente implementados visando cada um dos riscos de segurança definidos; c. 5.2.3 A equipe do repositório deve ter funções, responsabilidades e autorizações definidas relacionadas a mudanças de implementação dentro do sistema; d. A norma ainda possui o Anexo A, de caráter informativo e de apoio, para tratar da segurança do repositório (2017, p. 77); 4) DRAMBORA (2007, p. 169, 170, 171, 172, 175, 181, 187, 209, 210, tradução nossa): a. R35 Exploração de vulnerabilidade de segurança; b. R36 Comprometimento de segurança não identificado, vulnerabilidade ou degradação de informações; c. R37 Intrusão física do espaço de armazenamento de <i>hardware</i>; d. R38 Intrusão remota ou local via <i>software</i>; e. R41 Sabotagem deliberada do sistema; f. R47 Incapacidade de avaliar a eficácia da infraestrutura técnica e segurança; g. R53 Perda de disponibilidade de informação e/ou serviço; h. R75 O subsistema de autenticação falha; i. R76 O subsistema de autorização falha. 5) Guia de avaliação de repositórios FECYT-REBIUN: critério “7.5 - Perfis e permissões de gerenciamento de conteúdo” (2021, p. 28, tradução nossa); 6) Matriz NDSA: níveis 1, 2, 3, 4 da área funcional “Controle” (2019, p. 1, tradução nossa).
Quantidade de critérios:
CoreTrust Seal: 1; Nestor Seal: 2; ISO 16363:2012: 3; DRAMBORA: 9; Guia de avaliação de repositórios FECYT-REBIUN: 1. Total: 16 critérios.
Critério:
5.2) Adoção de medidas que garantam a integridade, autenticidade e recuperação dos dados em caso de perdas.
Instrumentos atendidos:

- 1) CoreTrust Seal (2022, p. 17, tradução nossa):
 - a. R07. O repositório garante a autenticidade dos objetos digitais e fornece informações de proveniência;
 - b. R14. O repositório aplica processos documentados para garantir o armazenamento e a integridade dos dados e metadados.
- 2) Nestor Seal (2009, p. 19, 21, 33, 37, tradução nossa):
 - a. 6 O repositório digital garante a integridade dos objetos digitais durante todas as etapas de processamento;
 - i. O critério possui 3 subcritérios que analisam a integridade dos processos de Ingestão, Arquivamento e Disseminação e seus respectivos pacotes: SIP, AIP, DIP.
 - b. 7 O repositório digital garante a autenticidade dos objetos digitais durante todas as etapas do processamento;
 - i. O critério possui 3 subcritérios que analisam a autenticidade dos processos de Ingestão, Arquivamento e Disseminação e seus respectivos pacotes: SIP, AIP, DIP.
 - c. 12.4 O repositório digital registra metadados adequados para documentar todas as alterações feitas pelo repositório digital nos objetos digitais;
 - d. 14 A infraestrutura protege o repositório digital e seus objetos digitais;
- 3) ISO 16363:2012 (2017, p. 68, 69, 70, 71, 73, 76, tradução nossa):
 - a. 4.1.5 O repositório deve ter um processo de ingestão que verifique cada SIP quanto à integridade e precisão;
 - b. 5.1.1.2 O repositório deve possuir *hardware* e *software* adequados para fazer cópias de segurança, preservar o conteúdo do repositório e monitorar suas funções;
 - c. 5.1.1.3 O repositório deve possuir mecanismos eficazes para detectar a perda e corrupção de bits;
 - d. 5.1.1.3.1 O repositório deve registrar e informar aos seus administradores todas as ocorrências decorrentes da perda ou corrupção de dados, devendo ser tomadas medidas para reparar/substituir os dados perdidos ou corrompidos;
 - e. 5.1.1.4 O repositório deve possuir um processo de registro e reação à disponibilidade de atualizações de segurança, com base na avaliação risco-benefício;
 - f. 5.1.1.5 O repositório deve ter processos definidos para alterações na mídia de armazenamento e/ou hardware (por exemplo, migração ou atualização);
 - g. 5.1.2 O repositório deve gerenciar o número e a localização das cópias de todos os objetos digitais;
 - h. 5.2.4 O repositório deve ter um(s) plano(s) adequado(s) de prevenção e preparação para desastres, incluindo pelo menos um *backup* externo de todas as informações mantidas associadas a uma cópia externa do(s) plano(s) de prevenção contra desastres.
- 4) DRAMBORA (2007, p. 173, 174, 176, 180, 186, 188, 189, 190, 192, 193, 194, 195, 200, 203, tradução nossa):
 - a. R39 Fenômeno ambiental local destrutivo ou perturbador;
 - b. R40 Interrupção acidental do sistema;
 - c. R42 Destruição ou indisponibilidade do site do repositório;
 - d. R46 Destruição da documentação primária;
 - e. R52 Perda do sigilo das informações;
 - f. R54 Perda de autenticidade das informações;
 - g. R55 Perda de integridade das informações;
 - h. R56 Não identificação de alterações de informações;
 - i. R58 Perda de confiabilidade da informação;
 - j. R59 Perda de proveniência da informação;
 - k. R60 Perda ou inadequação dos *backups*;
 - l. R61 Inconsistência entre cópias redundantes;

- m. R66 As estratégias de preservação resultam em perda de informações;
 - n. R69 A integridade referencial dos metadados da informação é comprometida.
- 5) Guia de avaliação de repositórios FECYT-REBIUN (2021, p. 27, tradução nossa):
- a. 7.1.- Cópias de segurança;
 - b. 7.2.- Existem pelo menos três cópias dos registros (metadados e arquivos) e, pelo menos, uma delas está localizada em uma localização geográfica diferente;
 - c. 7.3.- *Checksum*;
- 6) Matriz NDSA: níveis 1, 2, 3, 4 das áreas funcionais “Armazenamento” e “Integridade” (2019, p. 1, tradução nossa).

Quantidade de critérios:

CoreTrust Seal: **2**; Nestor Seal: **8**; ISO 16363:2012: **8**; DRAMBORA: **14**; Guia de avaliação de repositórios FECYT-REBIUN: **3**. Total: **35** critérios.

Fonte: elaboração própria.

Quadro 24 – Análise de critérios e similaridades da área “6) Metadados”.

Área:	
6) Metadados	
Critério:	
6.1) Existência e padrões de metadados adotados.	
Instrumentos atendidos:	
1)	CoreTrust Seal: requisito “R08. O repositório aceita dados e metadados com base em critérios definidos para garantir relevância e compreensão para os usuários” (2022, p. 13, tradução nossa);
2)	Nestor Seal (2009, p. 32, 33, 34, tradução nossa): a. 12.2 O repositório digital registra metadados adequados para descrição formal e de conteúdo e identificação dos objetos digitais; b. 12.3 O repositório digital registra metadados adequados para descrição estrutural dos objetos digitais; c. 12.4 O repositório digital registra metadados adequados para documentar todas as alterações feitas pelo repositório digital nos objetos digitais; d. 12.5 O repositório digital adquire metadados adequados para descrição técnica dos objetos digitais; e. 12.6 O repositório digital adquire metadados adequados para registrar os direitos e correspondentes condições de uso.
3)	ISO 16363:2012 (2017, p. 35, 38, 40, 41, 42, 51, 54, 56, 58, 69, tradução nossa): a. 3.5.2 O repositório deve rastrear e gerenciar os direitos de propriedade intelectual e as restrições de uso do conteúdo do repositório conforme exigido pelo acordo, contrato ou licença do repositório; b. 4.1.2 O repositório deve especificar claramente as informações que precisam ser associadas às Informações de Conteúdo específicas no momento de seu depósito; c. 4.1.5 O repositório deve ter um processo de entrada que verifique cada SIP quanto à integridade e precisão; d. 4.1.6 O repositório deve obter controle suficiente sobre os Objetos Digitais para preservá-los; e. 4.1.8 O repositório deve possuir documentos atualizados das ações dos processos de administração pertinentes à aquisição de conteúdo; f. 4.2.10 O repositório deve possuir documentos atualizados das ações e processos de gestão relevantes para a criação do AIP; g. 4.3.4 O repositório deve fornecer evidências da eficácia de suas atividades de preservação; h. 4.4.2 O repositório deve possuir registros contemporâneos das ações e processos de gestão pertinentes ao armazenamento e preservação dos AIP; i. 4.5.2 O repositório deve capturar ou criar a informação descritiva mínima e assegurar que está associada ao AIP; j. 4.5.3 O repositório deve manter uma ligação bidirecional entre cada AIP e suas informações descritivas; k. 5.1.1.3 O repositório deve possuir mecanismos eficazes para detectar a perda e corrupção de bits.
4)	DRAMBORA (2007, p. 169, 170, 171, 172, 175, 181, 187, 209, 210, tradução nossa): a. R48 Invalidadez estrutural ou malformação de pacotes recebidas; b. R49 Incompletude dos pacotes enviados; c. R63 Incapacidade de validar a eficácia do processo de ingestão; d. R68 Não rastreabilidade de um pacote recebido, arquivado ou difundido; e. R69 A integridade referencial dos metadados da informação é comprometida; f. R70 Histórico documentado de alterações incompleto ou incorreto; g. R71 Perda de acesso de objetos de informação.
5)	Guia de avaliação de repositórios FECYT-REBIUN (2021, p. 21, 24, 28, tradução nossa): a. 4.31. É utilizado algum esquema de metadados técnicos e/ou de preservação; b. 5.12. Integração com outros sistemas de informação da instituição; c. 8.3. Visualizar/exportar metadados em diferentes esquemas e formatos.
6)	Matriz NDSA: níveis 1, 2, 3, 4 da área funcional “Metadados” (2019, p. 1, tradução nossa).
Quantidade de critérios:	
CoreTrust Seal: 1 ; Nestor Seal: 5 ; ISO 16363:2012: 11 ; DRAMBORA: 7 ; Guia de avaliação de repositórios FECYT-REBIUN: 3 . Total: 27 critérios.	

Fonte: elaboração própria.

A área “6) Metadados”, apresentada no Quadro 24, avalia como o repositório se relaciona com os metadados das submissões de materiais e os registrados durante a realização da preservação propriamente dita.

“Absolutamente nada é tão importante quanto saber em quem confiar”

Radovan Kavicky

Quadro 25 – Análise de critérios e similaridades da área “7) Acesso e visibilidade”.

Área: 7) Acesso e visibilidade
Critério: 7.1) Existência de Plano ou Política de Acesso
Instrumentos atendidos: 1) CoreTrust Seal (2022, p. 14, 16, 17, tradução nossa): a. R10. O repositório atinge qualidade técnica e conformidade com padrões e garante que existem informações disponíveis suficientes para que o usuário final faça avaliações relacionadas à qualidade; b. R12. O repositório permite que os usuários localizem os objetos digitais e se refiram a eles de forma persistente por meio de citação adequada; c. R13. O repositório permite a reutilização dos objetos digitais ao longo do tempo, garantindo que estejam disponíveis informações apropriadas para apoiar a compreensão e o uso. 2) Nestor Seal (2009, p. 12, tradução nossa): a. 2 O repositório digital concede à(s) comunidade(s)-alvo acesso adequado à informação representada pelos objetos digitais; b. 2.1 O repositório digital garante que sua(s) comunidade(s)-alvo possa(m) acessar os objetos digitais; c. 2.2 O repositório digital garante que sua(s) comunidade(s)-alvo possa(m) interpretar os objetos digitais. 3) ISO 16363:2012 (2017, p. 60, 61, tradução nossa): a. 4.2.5.2 O repositório deve ter ferramentas e métodos para determinar quais Informações de Representação são necessárias para tornar cada Objeto de Dados compreensível para cada Comunidade-Alvo; b. 4.2.7.1 O repositório deve ter um processo documentado para verificar a inteligibilidade das Informações de Conteúdo dos AIP em sua criação para sua Comunidade-Alvo; c. 4.2.7.2 O repositório deve executar o processo de verificação para cada classe de Informação de Conteúdo dos AIP; d. 4.2.7.3 O repositório deve trazer as informações de conteúdo AIP até o nível exigido de inteligibilidade se a verificação de inteligibilidade falhar; e. 4.6.1 O repositório deve cumprir suas Políticas de Acesso; f. 4.6.1.1 O repositório deve registrar e analisar todas as falhas e anomalias na gestão de acessos; g. 4.6.2 O repositório deve seguir políticas e procedimentos que permitam a consulta de objetos digitais rastreáveis aos originais, com evidências que comprovem sua autenticidade; h. 4.6.2.1 O repositório deve registrar e atuar sobre problemas relacionados a erros nos dados ou nas respostas dos usuários. 4) DRAMBORA (2007, p. 206, 207, 211, tradução nossa): a. R72 Ambigüidade da definição de interpretabilidade; b. R73 Deficiências na compreensão semântica ou técnica da informação; c. R77 Incapacidade de validar a eficácia do mecanismo de difusão. 5) Guia de avaliação de repositórios FECYT-REBIUN: critérios “8.3.- Visualizar/exportar metadados em diferentes esquemas e formatos” (2021, p. 28, tradução nossa)
Quantidade de critérios: CoreTrust Seal: 3; Nestor Seal: 3; ISO 16363:2012: 8; DRAMBORA: 3; Guia de avaliação de repositórios FECYT-REBIUN: 1. Total: 18 critérios.
Critério: 7.2) Facilidade de acesso ao repositório e promoção de seu uso
Instrumentos atendidos: 1) Guia de avaliação de repositórios FECYT-REBIUN (2021, p. 12, 14, 28, tradução nossa): a. 1.5.- Existência de iniciativas para promover a visibilidade do repositório dentro da própria instituição; b. 2.8.- Existem dados de contato e uma oferta visível de assessoria; c. 8.1.- Possibilidade de vinculação com redes sociais; d. O critério “1.1.- Menção do repositório na página principal da instituição” constava na versão anterior do guia de avaliação e foi removido na versão atual (2015, p. 11, tradução nossa).
Quantidade de critérios: Guia de avaliação de repositórios FECYT-REBIUN: 4. Total: 4 critérios.

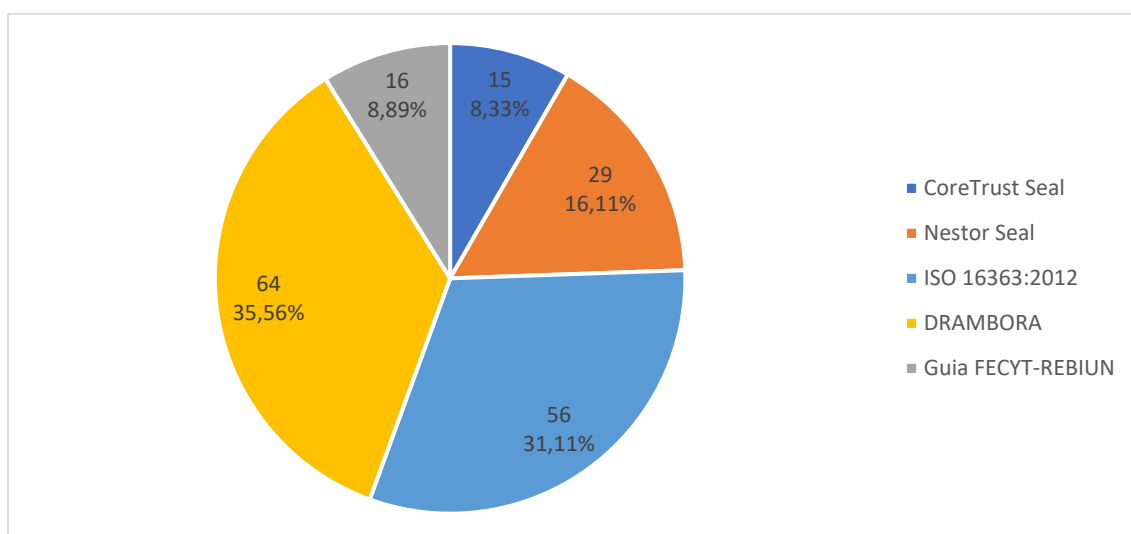
Fonte: elaboração própria.

A área “7) Acesso e visibilidade”, apresentada no Quadro 25, possui critérios que avaliam com os usuários se relacionam com os repositórios para acesso aos

materiais preservados. O critério “7.1) Existência de Plano ou Política de Acesso” avalia se o repositório possui definidas e documentadas as regras que regem o acesso aos seus materiais, especificando quais requisitos para a realização desse acesso, custos, além dos formatos de arquivos e metadados disponibilizados. A ausência dessas definições pode levar ao repositório um não atendimento das expectativas e necessidades da sua Comunidade-Alvo. Ao todo, o critério possui uma cobertura de 18 critérios dos instrumentos CoreTrust Seal, Nestor Seal, ISO 16363:2012, DRAMBORA e Guia de avaliação de repositórios FECYT-REBIUN.

O critério “7.2) Facilidade de acesso ao repositório e promoção do seu uso” verifica se o usuário consegue facilmente encontrar o repositório para acessar os seus materiais a partir dos sites das suas instituições e a existência de ações que promovam o uso dos repositórios em si. Uma vez que os materiais são preservados para serem acessíveis, a divulgação de acesso contribui para esse fim. Ao todo, o critério possui uma cobertura de 4 critérios do Guia de avaliação de repositórios FECYT-REBIUN.

Figura 17 – Cobertura de critérios



Fonte: elaboração própria.

Analisando-se a cobertura de critérios dos instrumentos usados como referência, atingiu-se um total de 180 critérios cobertos. A Figura 17 exibe as percentagens de cobertura de critérios entre os instrumentos. Ao todo, foram atingidos 15 critérios CoreTrust Seal, representando 8,33% do total, 29 critérios Nestor Seal, representando 16,11% do total, 56 critérios ISO 16363:2012, representando 31,11%, 64 critérios DRAMBORA, representando 35,56% e 16 critérios do Guia de avaliação de repositórios FECYT-REBIUN, representando 8,89%.

Em relação ao atendimento de critérios de cada instrumento, atingiu-se a cobertura de 15 dos 17 critérios CoreTrust Seal, representando 88,23% dos critérios, sendo o instrumento com maior cobertura total. O segundo instrumento com mais critérios cobertos é o DRAMBORA, com 64 critérios, dos 78 sugeridos, representando 82,05%. Em relação ao catálogo de critérios Nestor Seal, atingiu-se uma cobertura de 29 dos 42 critérios, o que representa 69,05% de cobertura. A norma ISO 16363:2012 foi coberta por 56 dos 109 critérios especificados, atingindo-se 51,38% de cobertura. Por fim, o Guia de avaliação de repositórios FECYT-REBIUN possuiu cobertura de 16 dos seus 53 critérios, representando 30,19% de cobertura. Esse fato se justifica pelo foco do guia em critérios que avaliam as compatibilidades dos repositórios com os requisitos do projeto RECOLECTA, voltado ao acesso aberto de publicações.

A obtenção de uma certificação formal de confiança de repositórios, em geral, não é um processo simples, pois requer um alto nível de maturidade e conhecimento dos próprios processos internos, equipe capacitada e dedicada, além dos custos financeiros envolvidos. No entanto, o fato de o repositório possuir uma certificação de confiança pode refletir em vantagens à sua Comunidade-Alvo assim como para os seus financiadores, pois estabelecem uma relação baseada no cumprimento de aspectos que garantem o seu funcionamento a longo prazo.

Sabe-se que o processo de aplicação das certificações de confiança disponíveis atualmente requer grandes esforços para serem realizadas, por isso, o estudo atual teve por objetivo a elaboração de critérios baseados nas diferentes certificações de confiança de repositórios. A verificação de atendimento dos critérios realizada no presente estudo nos repositórios selecionados não pretende se caracterizar como uma certificação por parte do autor, sendo as respostas obtidas utilizadas para se conhecer a situação atual dos repositórios quanto à confiança.

A elaboração dos critérios visou analisar primeiramente o compromisso do repositório em relação à preservação digital de documentos a longo prazo, sendo posteriormente analisado o tratamento dispensado aos objetos digitais mantidos pelas instituições. Foram desenvolvidos critérios que cobrem aspectos comuns às certificações CoreTrust Seal⁷⁰, ISO 16363⁷¹ e Nestor Seal⁷², devido a elas serem as atuais utilizadas para se percorrer os três níveis de certificação existentes: *core*,

⁷⁰ (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019)

⁷¹ (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012)

⁷² (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009)

extended e *formal*. Ainda que não representem certificações em si, foram adicionadas questões referentes relacionadas às ferramentas *NDSA Levels of Preservation*⁷³, *DRAMBORA*⁷⁴ e do *Guía para Evaluación de Repositorios Institucionales de Investigación*⁷⁵.

4.2.3 Fase 3: Elaboração do Formulário de Coleta de Dados

Com o intuito de coletar informações a respeito do estado da preservação digital dos repositórios brasileiros, elaborou-se um questionário de perguntas dividido em 11 seções que solicitam desde aspectos relacionados ao perfil do respondente até aspectos relacionados à visibilidade do repositório. Como ferramenta para a aplicação do questionário e coleta das respostas, utilizou-se a plataforma *Google Forms*⁷⁶ para a configuração das seções, perguntas e respostas.

A apresentação do questionário, exibida no Quadro 26, introduz o propósito da pesquisa, o autor e seus orientadores, além de esclarecer que a participação é voluntária e não condiciona ao repositório a obtenção de nenhum tipo de certificação. Posteriormente, é explicado o tratamento dos dados obtidos perante à privacidade e recolhimento de dados pessoais visando não ferir a LGPD e instruindo que todas as informações fornecidas fossem de conhecimento público e que não houvesse o fornecimento de dados pessoais. Ao respondente, é facultado o fornecimento de um e-mail institucional para o caso de desejar receber os resultados alcançados na pesquisa. A seguir, são apresentadas instruções para que as respostas não sejam perdidas enquanto o preenchimento do formulário não é finalizado. De forma a fornecer uma estimativa de tempo necessário para responder o questionário, é apresentado o tempo médio de resposta e o prazo final ao qual são aceitas respostas.

Após a apresentação, fez-se a pergunta sobre o nome e sigla da instituição que está respondendo o questionário. A próxima pergunta buscou identificar o interesse em responder o questionário da pesquisa, contendo três possíveis respostas:

- a) Compreendi o propósito da pesquisa e estou disposto a fornecer as informações requisitadas;
- b) Desejo indicar outra pessoa de minha instituição para responder, e;

⁷³ (NATIONAL DIGITAL STEWARDSHIP ALLIANCE, 2022)

⁷⁴ (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007)

⁷⁵ (AZORÍN, BERNAL, *et al.*, 2021)

⁷⁶ <https://docs.google.com/forms>

c) Não tenho interesse em responder o questionário.

Sendo selecionada a primeira resposta, as demais seções do questionário são exibidas. Caso a segunda resposta fosse selecionada, o formulário apresenta campos para informar o nome e endereço de e-mail de outra pessoa mais adequada a responder as questões. A última resposta interrompe o preenchimento do questionário e apresenta a mensagem de agradecimento pela participação na pesquisa.

Quadro 26 – Apresentação do questionário de avaliação.

Olá!

Me chamo Marcos Vinícius Bittencourt de Souza, sou doutorando pela Universidad de Salamanca (Espanha), no Programa de Formação na Sociedade do Conhecimento - Informação e Sociedade, sob orientação da Prof^a. Dr^a. María Manuela Moro Cabero e Prof. Dr. José Luis Bonal Zazo. Além disso, sou Analista de Tecnologia da Informação na Universidade Federal de Santa Maria, trabalhando com projetos relacionados à preservação digital de documentos.

A tese que desenvolvo avalia o atendimento de critérios relacionados à confiança de repositórios institucionais, baseando-se em padrões de avaliação e certificação tais como ISO 1636, CoreTrust Seal, Nestor Seal, NDSA Levels e DRAMBORA.

Solicitamos a sua importante colaboração respondendo aos questionamentos a seguir, que possui questões abertas e fechadas. A participação é voluntária e a conclusão do questionário não condiciona o fornecimento de nenhuma certificação, sendo os dados coletados utilizados para traçar um panorama em relação a critérios de confiança dos repositórios selecionados.

Em acordo com a LGPD, nenhum dado pessoal, sensível ou confidencial é solicitado. Todas as informações prestadas devem ser de conhecimento público.

Caso esteja interessado em receber os resultados da pesquisa, por favor, forneça o seu e-mail institucional para podermos contatá-lo posteriormente.

Para que as respostas preenchidas sejam salvas automaticamente, por favor, faça o login com uma conta do Google. Assim, os dados não serão perdidos e o progresso do questionário pode ser retomado em outro instante. Essa ferramenta auxilia o preenchimento nos casos em que seja necessário contatar outras pessoas da sua instituição para obter determinadas informações.

O tempo médio para resposta do questionário é de 15 minutos e o prazo para o seu preenchimento é até 10/05/2022.

Em caso de dúvidas, por favor, entre em contato comigo através do e-mail: marcos.cpd@ufsm.br

Fonte: elaboração própria.

O tipo de resposta das perguntas das seções do questionário, em sua maioria, é do tipo semiaberta, apresentando um rol de opções de respostas mutuamente exclusivas e uma resposta na qual o respondente pode complementar a sua resposta com suas próprias palavras, caso nenhuma das opções apresentadas seja apropriada. Em todas as perguntas, foi adicionada a resposta “Prefiro não responder” que possibilita ao respondente não fornecer a informação solicitada seja por falta de conhecimento, falta de embasamento ou nos casos em que a informação possa afetar a imagem da instituição,

por exemplo. Assim, objetivou-se facultar ao respondente o não fornecimento de informações sensíveis ou restritas de acordo com a sua interpretação.

O Quadro 27 exibe um resumo do questionário apresentando as áreas e quantidades de perguntas em cada uma. As próximas seções apresentam detalhadamente cada área com as suas respectivas perguntas e respostas. O questionário completo com as áreas, perguntas, respostas e as justificativas teóricas para cada pergunta é apresentado no Anexo A.

Quadro 27 – Resumo das seções do questionário.

Área do questionário	Quantidade de perguntas
0. Informações sobre o respondente	2
1. Contextualização do repositório	6
2. Estratégia e responsabilidade	3
3. Certificações e normas	3
4. Infraestrutura organizacional	5
5. Infraestrutura técnica e organizacional	4
6. Segurança	4
7. Integridade dos dados	4
8. Metadados	2
9. Acesso e consulta	4
10. Visibilidade	4

Fonte: elaboração própria.

4.2.3.1 Área 0: Informações Sobre o Respondente

O prosseguimento no questionário apresenta a área denominada “Área 0) Informações sobre o respondente”, apresentada no Quadro 28.

Quadro 28 – Perguntas da área 0) Informações sobre o respondente.

Pergunta	Tipo	Tipo de resposta
0.1. Qual o seu cargo ou função desempenhada no órgão que gerencia o repositório?	Semiaberta	Múltiplas opções com opção discursiva.
0.2. Qual a sua formação acadêmica?	Semiaberta	Múltiplas opções com opção discursiva.

Fonte: elaboração própria.

A pergunta 0.1 busca identificar o perfil dos respondentes, permitindo-se se conhecer as suas responsabilidades perante o repositório institucional, proporcionando maior ou menor grau de confiabilidade das respostas. Pessoas envolvidas diretamente

com o repositório têm maior conhecimento sobre a sua estrutura, processos e características como um todo.

A pergunta 0.2 busca identificar a formação dos respondentes, permitindo identificar as suas bases de conhecimento, além de servir como identificação das formações profissionais e acadêmicas existentes nos repositórios.

4.2.3.2 *Área 1: Contextualização do Repositório*

A área “1) Contextualização do repositório” é apresentada no Quadro 29 e contém perguntas que auxiliam a compreender o comprometimento do repositório com a preservação a longo prazo assim como a natureza do material depositado. A pergunta 1.1 visa identificar se o repositório possui missão, visão e valores declarados, declarações essas que modelam a sua forma de funcionamento e norteiam a maneira como os seus processos são planejados e executados. Caso a resposta dessa pergunta seja positiva, é solicitado ao respondente que insira as informações ou forneça uma URL onde seja possível acessá-las. É fundamental que no texto das declarações requisitadas exista o comprometimento com a preservação digital e acesso a longo prazo, demonstrando a vocação e razão de existência do repositório.

A ISO 16363:2012 (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 23, tradução nossa) estabelece que “O repositório deve ter uma declaração da missão que reflita um compromisso com a informação digital para a sua conservação, retenção a longo prazo, gestão e acesso”, através do critério 3.1.1. O CoreTrust Seal, através do requisito R1, analisa se o repositório possui uma missão explícita que contemple a preservação e acesso dos dados sob sua custódia (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019, p. 10).

O Nestor Seal analisa, através dos critérios 1 e 1.2 se o repositório possui uma clara concepção dos seus objetivos, firmando um compromisso e assumindo a sua responsabilidade com a preservação a longo prazo dos objetos digitais custodiados (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 10-11).

A pergunta 1.2 analisa se a Comunidade-Alvo do repositório foi identificada e delimitada, buscando-se identificar os requisitos necessários para atender a sua demanda. Somente com a Comunidade-Alvo identificada é possível estabelecer o escopo de atuação do repositório, buscando compreender os usos dos documentos,

formatos de arquivos a serem aceitos e difundidos e o estabelecimento das rotinas de preservação adequadas. Há casos em que a Comunidade-Alvo do repositório pode ser a população inteira de um país ou região, fazendo com que as ações de difusão, por exemplo, tenham que abranger uma vasta gama de necessidades e seja necessário optar por formatos de arquivo largamente utilizados.

Em contrapartida, “[...] outros repositórios podem ter várias "comunidades específicas" diferentes com necessidades altamente especializadas, cada uma exigindo diferentes funcionalidades ou suporte do repositório.” (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 18, tradução nossa). Repositórios de dados de pesquisa, por exemplo, podem atender a uma gama específica de usuários que depositam dados em formatos de arquivo especiais utilizados em softwares específicos e, com isso, necessitam estratégias de preservação e difusão voltadas às necessidades dessa Comunidade-Alvo. Adicionalmente, um repositório de documentos históricos pode armazenar os representantes digitais de documentos físicos, disponibilizando imagens digitalizadas em alta resolução.

A norma ISO 16363:2012 analisa se o repositório possui a sua Comunidade-Alvo delimitada através do critério 3.3.1 (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 27) e é a partir dessa informação que as demais verificações poderão se basear para saber se atendem às necessidades da Comunidade-Alvo em questão. O Nestor Seal analisa a existência da definição da Comunidade-Alvo no critério 1.3 (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 11) e enfatiza que caso haja mudanças nas Comunidade-Alvo, o repositório deve reagir adaptando-se, o que leva à uma constante necessidade de monitoramento dos requisitos de sua Comunidade-Alvo.

O CoreTrust Seal em sua primeira seção de análise, que trata das informações do contexto do repositório, solicita ao repositório no requisito R0 uma breve descrição da Comunidade-Alvo, porém que contenha informações suficientes para que os revisores possam avaliar a adequação das medidas de curadoria e preservação descritas na aplicação da certificação (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019, p. 6). O DRAMBORA exemplifica os riscos R06 e R07 que tratam, respectivamente, da mudança substancial da Comunidade-Alvo e da falta de compreensão dos requisitos da Comunidade-Alvo repositório, analisando quais estratégias são adotadas para mitigar tais riscos (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 140-141).

A pergunta 1.3 questiona se existe alguma particularidade identificada na comunidade designada do repositório buscando complementar a pergunta 2.2.1 que solicita a definição da comunidade designada. Assim, mesmo que não seja fornecida uma resposta à pergunta 2.2, o fornecimento de uma resposta à pergunta 2.3 sinaliza que há uma certa delimitação e identificação de necessidades da comunidade designada do repositório. A pergunta 1.4 procura identificar a natureza dos materiais depositados no repositório, uma vez que as necessidades de ingestão e difusão da sua Comunidade-Alvo estão intimamente relacionadas aos tipos de documentos custodiados. O fornecimento dessa informação complementa a informação fornecida na pergunta 2.2 do questionário, pois mesmo que não haja uma delimitação formal da Comunidade-Alvo do repositório, pode-se conhecer o tipo de material tratado.

Quadro 29 – Perguntas da área 1) Contextualização do repositório.

Pergunta	Tipo	Tipo de resposta
1.1 O repositório institucional possui missão, visão e valores elaborados?	Semiaberta	Sim ou Não com opção discursiva
1.1.1 Em caso positivo, qual a missão, visão e valores do repositório institucional?	Aberta, condicionada à anterior.	Discursiva
1.2 O repositório institucional possui a sua Comunidade-Alvo delimitada?	Semiaberta	Sim ou Não com opção discursiva
1.2.1 Em caso positivo qual a Comunidade-Alvo do repositório institucional?	Aberta, condicionada à anterior	Discursiva
1.3 Particularidades da Comunidade-Alvo. A Comunidade-Alvo do repositório possui especificidades em suas demandas de uso dos documentos? Caso julgue que não haja particularidades, deixe a resposta em branco.	Aberta	Discursiva
1.4 Qual a natureza do material depositado no repositório?	Fechada	Múltiplas opções com opção discursiva.

Fonte: elaboração própria.

4.2.3.3 Área 2: Estratégia e Responsabilidade

A área 2, nomeada “Estratégia e responsabilidade”, apresentada no Quadro 30, tem como objetivo avaliar as estratégias do repositório a respeito da preservação a longo prazo, contendo perguntas sobre a presença de instrumentos que regulam a sua atividade.

A pergunta 2.1 solicita ao respondente que informe se o repositório possui uma Política de Preservação definida que contenha os requisitos, delimitações e tipos de materiais a serem depositados. Caso a resposta seja negativa, a pergunta 2.1.1 analisa se

o repositório possui definido um Plano Estratégico de Preservação que contém de uma forma ampla, os objetivos do repositório quanto à preservação digital de seu material, assim como as metas sob as quais o repositório está sujeito e deve trabalhar para alcançá-las. Enquanto o Plano Estratégico de Preservação é um documento mais genérico alinhado à missão, visão e valores do repositório, a Política de Preservação deve ser elaborada para conter ações que se alinhem ao plano. A existência desses instrumentos é avaliada no critério 3.1.2 da norma ISO 16363 (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 23).

A pergunta 2.2 verifica se há um Plano de Continuidade, usado nos casos em que, por qualquer razão, o funcionamento do repositório é interrompido ou descontinuado. Dentre os motivos que podem levar à descontinuidade estão a falta de financiamento, mudança de gestão ou a ocorrência de desastres naturais, por exemplo.

O Plano de Continuidade deve ter mapeados os riscos que podem levar à interrupção do funcionamento do repositório, procurando soluções que evitem, mitiguem ou reajam às situações identificadas. Ao depositarem seus materiais, os usuários do repositório creem que a preservação digital a longo prazo será efetivamente realizada, independente das condições que possam vir a ocorrer num futuro próximo ou distante, fazendo com que o Plano de Continuidade possua papel fundamental na percepção de confiança do repositório e no cumprimento de sua missão.

O Plano de Continuidade deve conter os processos de recuperação a serem executados quando da ocorrência de um desastre natural, procurando minimizar os impactos aos usuários e à própria preservação. Nos casos de haver problemas de sustentabilidade financeira do repositório, deve ser prevista a migração do repositório para outra instituição que tenha condições de assumir o material custodiado para continuar a preservação a longo prazo, sem que haja falhas que comprometam a confiabilidade e autenticidade do material.

A norma ISO 16363 possui o requisito 3.1.2.1 (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 24) que avalia a presença do Plano de Continuidade no repositório digital. O CoreTrust Seal possui o requisito R3 (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019, p. 12) que verifica a existência do Plano de Continuidade nos mesmos moldes que a ISO 16363, no entanto, os aspectos técnicos de execução desse plano são complementados pelo requisito R15, que analisa as estratégias previstas para a recuperação de desastres (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019, p. 24).

O Nestor Seal avalia se “a continuação das tarefas de preservação é garantida mesmo além da existência do repositório digital” (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 17, tradução nossa) no critério 4.6, verificando se existem planos de contingência a serem acionados nos casos de descontinuidade do repositório. O DRAMBORA fornece como exemplo o risco R08 denominado “Cessação forçada das operações do repositório” e prevê como estratégia de mitigação o estabelecimento de planos de contingência e sucessão (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 142).

A matriz de preservação NDSA não prevê a existência de um plano de continuidade formalizado, no entanto, a área funcional de armazenamento foca principalmente na existência de cópias de segurança armazenadas em localizações geográficas distintas e com mais de uma tecnologia de suporte. À medida em que o repositório progride nos níveis da matriz, as soluções vão se tornando mais sofisticadas e tendem a prever uma grande gama de riscos que podem acometer o conteúdo digital armazenado (LEVELS OF PRESERVATION REVISIONS WORKING GROUP, 2019, p. 1).

Quadro 30 – Perguntas da área 2) Estratégia e responsabilidade.

Pergunta	Tipo	Tipo de resposta
2.1 Existe uma Política de Preservação com os requisitos mínimos dos documentos recebidos, tratamento de preservação e acesso?	Fechada	Escala de diferencial semântico com opção discursiva.
2.1.1 Caso não exista uma Política de Preservação, existe um Plano Estratégico de Preservação?	Fechada, condicionada à anterior.	Escala de diferencial semântico com opção discursiva.
2.2 Existe um Plano de Continuidade que garante que os documentos estarão preservados e acessíveis a longo prazo, caso ocorra algum desastre ou descontinuidade do funcionamento?	Fechada	Escala de diferencial semântico com opção discursiva.

Fonte: elaboração própria.

O conjunto de possíveis respostas das perguntas da área 2 busca identificar a adoção efetiva dos instrumentos declarados pelo repositório, e são apresentadas no Quadro 31.

Quadro 31 – Respostas das perguntas da Área 2) Estratégia e responsabilidade

Resposta	Semântica
Existe em teoria	O instrumento da pergunta foi elaborado de acordo, mas as suas ações não foram colocadas em prática
Existe na prática, mas ainda não foi totalmente implantado	As ações necessárias para cumprir o objetivo de cada instrumento ainda não foram implantadas totalmente na prática. Para o presente estudo, não foi buscado identificar os motivos pelos quais as ações não foram completamente implantadas
Existe na prática e está totalmente implantado	O repositório aplica na prática todas as ações necessárias para alcançar o objetivo previsto no instrumento elaborado
Não existe	O repositório ainda não dispendeu esforços para a elaboração do instrumento ou que ainda está em fase de elaboração e não consta como finalizado
Prefiro não responder	O responsável pelo preenchimento do formulário prefere não manifestar o estado atual do seu repositório, podendo representar que o respondente não possui o conhecimento necessário para responder à pergunta formulada ou que não há interesse em expor a situação atual do repositório

Fonte: elaboração própria.

4.2.3.4 Área 3: Certificações e Normas

A área 3, nomeada “Certificações e Normas”, apresentada no Quadro 32, visa identificar a relação do repositório com normas e certificações existentes. A pergunta 3.1 solicita que se responda quais certificações, junto aos seus anos de obtenção, o repositório possui. Optou-se por uma pergunta do tipo aberta para que o respondente tenha a possibilidade de preencher qualquer tipo de certificação que o repositório possua, incluindo certificações de qualidade como a ISO 9000⁷⁷ ou as normas de segurança da informação da série ISO 27000⁷⁸, por exemplo.

A pergunta 3.2 verifica quais certificações o repositório tem a intenção de obter, possibilitando que o respondente sinalize que o repositório tem a intenção de obter uma certificação, mas ainda não definiu qual será, ou ainda informando que o repositório não tem intenções de obter nenhuma certificação.

A norma ISO 16363:2012, no critério 3.3.6, avalia se o repositório se compromete com um programa regular de autoavaliação e certificação externa, buscando incentivar a obtenção de outras certificações assim como a utilização de métodos de avaliação variados (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 30). Esse comprometimento visa garantir o grau de confiança adquirido pelo repositório, uma vez que outros instrumentos de avaliação

⁷⁷ (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2015)

⁷⁸ (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2018)

podem cobrir outros aspectos diferentes da ISO 16363, ou ainda cobrir os mesmos aspectos, porém sob uma ótica diferente e complementar.

O Nestor Seal avalia a presença de uma adequada gestão de qualidade, no sentido de favorecer o alcance dos objetivos do repositório, através do critério 5 (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 17). É sugerida a obtenção da certificação ISO 9000⁷⁹ para gerenciar a qualidade dos processos existentes em todas as partes do repositório, no entanto, a norma não é obrigatória para o atendimento do critério, uma vez que o repositório pode apresentar outras documentações que demonstram a sua gestão de qualidade.

A pergunta 3.3, a última da área, trata da adoção de normas no repositório, sendo do tipo semiaberta com opções referentes à normas comuns no meio arquivístico. Dentre as alternativas apresentadas, estão as normas de descrição arquivística: ISAD(G)⁸⁰ e NOBRADE⁸¹. Além das normas, são apresentadas duas alternativas de padrões estabelecidos para a descrição arquivística: Dublin Core⁸² e PREMIS⁸³. Uma das alternativas de resposta se refere a adoção do padrão ISO 14721 (OAIS)⁸⁴ pela ferramenta de repositório utilizada. A última opção apresentada, METS⁸⁵, trata do empacotamento de documentos e metadados para comunicação e intercâmbio com outras ferramentas.

De acordo com o Conarq (2015, p. 11), “Um repositório digital deve estar em conformidade com as normas e padrões estabelecidos, de forma a possibilitar níveis de interoperabilidade com outros repositórios digitais e sistemas informatizados que tratam de documentos arquivísticos”. Daí a importância de o repositório adotar normas e padrões em sua estrutura e processos como um todo.

A norma ISO 16363:2012 trata em si da adequação e adoção do repositório à norma ISO 14721:2012 (OAIS) em relação à estrutura do repositório em si. Logo, ao responder que o repositório adota essa norma, pressupõe-se que ele, no mínimo, adota uma ferramenta alinhada ao padrão OAIS, utilizando a sua estrutura de pacotes, por exemplo. A adoção de padrões bem estabelecidos no repositório promove a uniformização dos dados e possibilita a integração com outras ferramentas e

⁷⁹ (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2015)

⁸⁰ (INTERNATIONAL COUNCIL ON ARCHIVES, 2000).

⁸¹ (CONSELHO NACIONAL DE ARQUIVOS, 2006).

⁸² (DUBLIN CORE METADATA INITIATIVE, 2020)

⁸³ (PREMIS EDITORIAL COMMITTEE, 2015).

⁸⁴ (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012)

⁸⁵ (METS BOARD, 2010)

repositórios externos, realizadas através de interfaces, protocolos de informação e estrutura de pacotes bem definidos.

Quadro 32 – Perguntas da área 3) Certificações e normas.

Pergunta	Tipo	Tipo de resposta
3.1 O repositório possui certificações alcançadas?	Aberta	Discursiva
3.2 Quais certificações o repositório tem a intenção de alcançar?	Semiaberta	Múltiplas opções, com opção discursiva.
3.3 O repositório adota padrões e normas estabelecidos?	Semiaberta	Múltiplas opções, com opção discursiva.

Fonte: elaboração própria.

4.2.3.5 Área 4: Infraestrutura Organizacional

A área 4, nomeada “Infraestrutura organizacional”, apresentada no Quadro 33, avalia a estrutura do repositório em relação ao seu propósito de preservação a longo prazo através de uma análise de sua estrutura organizacional e de equipe. A maneira com que o repositório se organiza internamente e, muitas vezes, está inserido dentro de uma instituição maior, permite o estabelecimento de escopos de atuação, delimitação das responsabilidades e a modelagem de processos bem definidos. Essa estrutura estabelecida deve ser viavelmente executável e estar voltada e alinhada aos objetivos do repositório na execução da preservação a longo prazo.

A pergunta 4.1 pede que o respondente informe se no seu repositório institucional há uma estrutura organizacional adequada à preservação digital e o seu rol de respostas tem por objetivo identificar se a estrutura foi somente planejada, executada em partes ou está totalmente implantada e em funcionamento. Em certos casos, pode acontecer de o repositório realizar o seu planejamento de estrutura, porém a execução pode ter sido prejudicada ou até impedida devido a fatores como indisponibilidade financeira, por exemplo. Em outros casos, a estrutura pode ter sido planejada, implantada e adequada à preservação em uma época diferente da atual e já não reflete a demanda e volume de trabalho atual.

A norma ISO 16363:2012, nos critérios 3.2.1 e 3.2.1.1 (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 25), assim como o critério Nestor Seal 4.3 (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 16) analisam a presença de estruturas organizacionais com as respectivas definições escopo e responsabilidades.

O DRAMBORA (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 135 e 139) cita como exemplos os riscos R01 (*Management Failure* – Falha de gerenciamento) e R04 (*Business objectives not met* – Objetivos de negócio não alcançados) relacionados a problemas de gestão do repositório e que podem ser interpretados como falhas de infraestrutura organizacional.

O CoreTrust Seal possui o requisito R5 que trata da infraestrutura organizacional englobando a definição de processos, responsabilidades, governança, dimensionamento e qualificação da equipe e financiamento do repositório (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019, p. 14).

Um aspecto de suma importância na operacionalização do repositório se refere ao orçamento destinado ao seu funcionamento, tema abordado na pergunta 4.2. O orçamento deve cobrir os custos de funcionamento do repositório assim como a previsão de futuras manutenções para evolução da sua estrutura, contendo as origens de cada verba a ser utilizada.

A ISO 16363:2012 possui a seção 3.4 com três critérios que avaliam a sustentabilidade financeira do repositório. O critério 3.4.1 especifica que o repositório deve ter publicados documentos de plano de negócios a curto e longo prazos que demonstrem “[...] a viabilidade financeira do repositório durante o período de tempo o qual se comprometeu em proporcionar acesso ao seu conteúdo à sua Comunidade-Alvo” (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 31). O critério 3.4.2 trata da utilização de práticas transparentes e que estejam de acordo com as normas contábeis aplicáveis ao repositório, no entanto, não basta a gestão do repositório se auto denominar transparente, mas sim estar disposta à eventuais auditorias externas, por exemplo (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 32). O critério 3.4.3 trata do compromisso que o repositório deve ter com uma contínua análise financeira a fim de estar monitorando a sua situação e reagindo o mais cedo possível às ameaças que possam surgir (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 32).

O conjunto de critérios da norma ISO 16363:2012 sobre os aspectos financeiros permite à Comunidade-Alvo a constatação de que existem meios de manter o repositório em funcionamento e que a preservação digital é viável, sendo aspecto fundamental para o estabelecimento da confiança do repositório.

O CoreTrust Seal, no requisito R5, analisa se o repositório possui evidências que comprovem que o repositório possui fundos suficientes para o seu funcionamento que incluam os gastos de pessoal e recursos de TI, possuindo um planejamento ideal entre três e cinco anos à frente (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019, p. 14).

O Nestor Seal possui o critério 4.1 que verifica se o repositório possui recursos financeiros adequados ao funcionamento dos serviços aos quais se propõe demonstrando através de documentos que englobem o curto e o longo prazo (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019, p. 15).

O DRAMBORA fornece como exemplo, a identificação de cinco riscos ligados às finanças do repositório, todos eles com alto grau de severidade (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 159-163):

- R25 *Finances insufficient to meet repository commitments* (Finanças insuficientes para alcançar os compromissos firmados pelo repositório): quando não há recursos financeiros que permitam a execução das atividades do repositório;
- R26 *Misallocation of finances* (Má alocação de finanças): quando não há um planejamento adequado levando a investimentos realizados acima do que o orçamento do repositório pode comportar;
- R27 *Liability for non-adherence to financial law or regulations* (Suscetibilidade a não cumprimento de leis ou regulações financeiras): quando o repositório está suscetível ao não cumprimento das regulações a que está sujeito, podendo ocorrer por falhas de natureza contábil ao não serem pagos os impostos devidos, por exemplo;
- R28 *Financial shortfalls or income restrictions* (Deficiências financeira ou restrições de renda): quando ocorrem circunstâncias que não foram devidamente planejadas e requerem mais orçamento do que o repositório teria disponível;
- R29 *Budgetary reduction* (Redução orçamentária): quando há algum evento que acaba forçando a diminuição dos fundos destinados ao repositório. Em casos de repositórios de instituições públicas a redução pode ser devido a uma recessão que ocasiona uma diminuição de arrecadação de impostos ou devido a uma realocação de recursos para outros setores.

Na pergunta 4.3 do questionário, o respondente deve informar se o quantitativo pessoal do repositório é adequado à preservação digital somente em teoria,

na prática ou definitivamente não é adequada. Muitas vezes o projeto do repositório pode prever um quantitativo de pessoal com responsabilidades e atribuições bem definidas e com a divisão igualitária de tarefas, porém a execução do projeto pode ser prejudicada devido à indisponibilidade de pessoas, falta de pessoas interessadas a trabalhar com o tema ou devido à falta de priorização da organização superior em alocar pessoas no repositório.

A efetivação da preservação digital a longo prazo, de forma confiável, necessita de equipes bem dimensionadas que permitam a execução dos processos arquivísticos por completo. A norma ISO 16363, no critério 3.2.1.2 (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 26) e o critério Nestor Seal 4.2 (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 15) analisam se o repositório possui a quantidade adequada de pessoas para manter o funcionamento do repositório digital. O CoreTrust Seal analisa a questão relacionada às equipes do repositório no requisito R5 (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019, p. 14).

O DRAMBORA cita como exemplo o risco R21, relacionada à perda de pessoas chave da equipe do repositório, que podem afetar o seu funcionamento (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 155). Dentre as possíveis causas atribuídas à perda de pessoas, pode estar a saída para atuar em outra empresa ou àquelas ligadas a causas naturais que impossibilitem a continuação da pessoa no trabalho que desempenhava. Uma das formas de mitigar esse risco é o dimensionamento correto das equipes com o compartilhamento de funções, de modo a evitar a dependência em pessoas específicas, por exemplo.

A pergunta 4.4 analisa se a capacitação das equipes do repositório é adequada na prática, em teoria ou não é adequada. Muitas vezes durante a formação das equipes, pode-se selecionar pessoas com o perfil e conhecimentos necessários para a realização dos trabalhos ou pode-se realizar capacitações que permitam o nivelamento dos membros, de forma a haver uma distribuição mais uniforme de conhecimentos dentro das equipes. Porém, na prática, certas ações de capacitação podem não ter sido executadas, podendo ocasionar lacunas de conhecimento que podem impactar na execução dos processos ligados à preservação digital a longo prazo.

A pergunta 4.5 verifica a existência de um plano contínuo de capacitação da equipe que busque a atualização da equipe. À medida em que as tecnologias avançam, podem ser necessárias mudanças na maneira com que as ações de preservação são

executadas, fazendo com que o repositório analise tais mudanças e proporcione capacitações que se alinhem aos novos desafios encontrados.

A presença de um programa de desenvolvimento profissional é analisada na norma ISO 16363:2012 através do critério 3.2.1.3, justificando que “a tecnologia e as práticas gerais para a preservação digital seguirão mudando, da mesma forma que os requisitos da sua Comunidade-Alvo, por isso o repositório tem que assegurar que o conjunto de habilidades do seu pessoal evolua” (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 26, tradução nossa).

O Nestor Seal, no critério 4.2 analisa tanto o quantitativo pessoal da equipe quanto a sua qualificação, analisando se o repositório possui esquemas de capacitação que garantam os conhecimentos adequados à sua equipe (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 15-16). Uma das formas citadas como treinamento é a participação dos membros da equipe em eventos, conferências e congressos relevantes, trazendo com isso, visibilidade para as ações de capacitação desempenhadas pelo repositório, o que ajuda a melhorar a percepção de confiança da sua Comunidade-Alvo. O incentivo à pesquisa na equipe do repositório favorece o acompanhamento das descobertas assim que elas são publicadas, assim como permite à equipe contribuir com o seu conhecimento para a evolução da comunidade arquivística mundial.

A certificação CoreTrust Seal trata a viabilidade financeira, quantitativo e capacitação do pessoal do repositório, no requisito R5, que inclusive analisa a governança do repositório (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019, p. 14). A aplicação desse requisito avalia se a equipe possui acesso contínuo a treinamentos e desenvolvimento profissional.

O DRAMBORA exemplifica três riscos relacionados à equipe do repositório (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 156-158): a) R22: Deterioração das habilidades da equipe; b) R23: Habilidades da equipe se tornam obsoletas, e; c) R24: Incapacidade de avaliar a eficácia e ou adequação da equipe. A forma aconselhada de mitigar esses riscos é através do fornecimento e promoção de capacitações que permitam à equipe estar atualizada com as últimas mudanças tecnológicas ocorridas relacionadas à preservação digital.

Quadro 33 – Perguntas da área 4) Infraestrutura organizacional.

Pergunta	Tipo	Tipo de resposta
4.1 Existe uma estrutura organizacional adequada ao funcionamento do repositório para a preservação a longo prazo?	Semiaberta	Escala de diferencial semântico com opção discursiva
4.2 Existe orçamento adequado ao funcionamento do repositório para a preservação a longo prazo?	Semiaberta	Escala de diferencial semântico com opção discursiva
4.3 Em relação à quantidade de pessoas, a equipe responsável pelo repositório é adequada?	Semiaberta	Escala de diferencial semântico com opção discursiva
4.4 Em relação às capacitações e conhecimentos, a equipe responsável pelo repositório é adequada?	Semiaberta	Escala de diferencial semântico com opção discursiva
4.5 Existe um plano de capacitação permanente da equipe do repositório?	Semiaberta	Escala de diferencial semântico com opção discursiva

Fonte: elaboração própria.

4.2.3.6 Área 5: Infraestrutura Técnica e Tecnológica

A área 5 do formulário, nomeada “Infraestrutura técnica e tecnológica”, apresentada no Quadro 34, avalia o tratamento do repositório dispensado à sua estrutura física e de equipamentos voltados à preservação de longo prazo. Deve-se avaliar a estrutura física do repositório verificando se ela possui condições de comportar a sua equipe, assim como o conjunto de equipamentos necessários ao seu funcionamento.

De uma forma geral, as normas não possuem critérios específicos sobre a estrutura física do repositório, tratando-a como um aspecto incluído na análise da estrutura organizacional. A verificação do espaço físico é tratada de uma forma voltada à verificação de segurança, analisando-se as possibilidades de acessos indevidos. As perguntas elaboradas no presente estudo relativas aos critérios de segurança do repositório são tratadas na área 6.

No presente estudo, escolheu-se elaborar uma pergunta específica para esse item buscando-se identificar a percepção dos responsáveis quanto a adequação da sua estrutura física para a realização da preservação e, conseqüentemente, se a deficiência de um espaço físico adequado pode acarretar conseqüências para a realização da preservação a longo prazo plena. Assim, a pergunta 5.1 busca identificar se o espaço físico do repositório é adequado somente em teoria, na prática ou não é adequada à preservação digital.

A pergunta 5.2 foca na estrutura técnica do repositório, analisando se ela é adequada em teoria apenas ou se na prática, os equipamentos utilizados são suficientes

para realizar a preservação digital. A efetivação da preservação digital a longo prazo inevitavelmente se dá através da utilização de equipamentos apropriados na forma de servidores de aplicação, armazenamento e estrutura de rede para que os documentos possam ser recebidos corretamente assim como a sua difusão ocorra conforme a definição da política de preservação e acesso institucional.

A norma ISO 16363 verifica a presença de tecnologias de *hardware* adequadas à preservação através do critério 5.1.1.1.1 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012, p. 64), enquanto o critério 5.1.1.1.5 analisa as tecnologias de *software* (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012, p. 66). O monitoramento constante da comunidade designada do repositório pode identificar uma nova demanda de requisitos que impactem na mudança ou atualização dos hardwares do repositório, tais como a utilização de formatos de arquivos que exijam mais espaço de armazenamento ou a difusão de documentos através de streaming, por exemplo, que pode requerer um aumento da largura da banda de rede utilizada anteriormente.

O Nestor Seal analisa se a infraestrutura de TI, como um todo, implementa os requisitos mínimos necessários à preservação digital, através do critério 13.1 (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 36), levando em consideração a execução dos processos OAIS de ingestão, arquivamento, armazenamento e acesso.

O CoreTrust Seal analisa, através do requisito R15, a presença de tecnologias de *hardware* e *software* adequados à preservação dos documentos do repositório (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019, p. 24). Como forma de medir o cumprimento desse requisito, sugere-se verificar se a disponibilidade, largura de banda de rede e demais aspectos de conectividade suprem as necessidades da comunidade designada do repositório. Outra análise feita é quanto à existência de um plano de desenvolvimento da infraestrutura de TI que comporte a manutenção e a evolução do repositório de acordo com o crescimento de sua demanda.

O DRAMBORA apresenta os exemplos de riscos R30, R31, R32, R33 e R34 que tratam sobre problemas técnicos que podem ocorrer e que podem ser evitados através da execução ativa de monitoramento (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 164-168).

A pergunta 5.3 analisa a existência de um plano periódico de atualização dos hardwares e softwares utilizados no repositório. O respondente pode informar que o plano existe somente em teoria, está parcialmente implementado ou está implementado e em execução na prática. É essencial que haja um planejamento de monitoramento das tecnologias empregadas buscando-se evitar que a obsolescência tecnológica de *hardware* e *software* prejudique a preservação, assim como a implantação de novos meios tecnológicos que visem adaptar o repositório às novas demandas da sua comunidade designada.

A norma ISO 16363, nos critérios 5.1.1.1.2 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012, p. 64) e 5.1.1.1.6 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012, p. 67), respectivamente, analisa se o repositório possui procedimentos implantados para monitorar a necessidade de atualização dos *hardwares* e *softwares* utilizados. O monitoramento e atualização de *hardware* se faz necessário para que os níveis de serviço estipulados sejam cumpridos, evitando-se assim que haja interrupção dos serviços por indisponibilidade de acesso ou por falhas de armazenamento, por exemplo.

Deve-se implementar o monitoramento dos equipamentos de armazenamento buscando-se manter os seus *firmwares* atualizados, que melhoram o seu desempenho e corrigem falhas, além de monitorar os respectivos tempos de vida útil. Equipamentos de armazenamento, como HDs por exemplo, possuem um tempo de vida útil estipulado pelos seus fabricantes, período no qual a probabilidade de erros acontecer é menor do que quando esse prazo é ultrapassado. Assim, o planejamento de substituição dos equipamentos se faz necessário para a realização da preservação digital.

Da mesma forma, deve existir um planejamento de atualização dos *softwares* utilizados no repositório, seja para a implantação de versões mais atualizadas que contêm correções de erros ou para versões que disponibilizam novas funcionalidades à comunidade designada do repositório. Em particular, as atualizações de *software* possuem um aspecto ligado à segurança, implantando-se correções para falhas identificadas que podem fornecer acesso indevido ao material preservado, possibilitando o vazamento de informações ou o seu corrompimento.

O Nestor Seal, no critério 4.5 (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 16, tradução nossa), verifica se “O repositório digital reage a mudanças substanciais”, buscando identificar se existem processos implantados para o monitoramento e consequente reação às mudanças

técnicas, organizacionais ou baseadas na comunidade designada. Quanto aos aspectos tecnológicos, busca-se analisar se a prevenção do repositório à obsolescência tecnológica, assim como o monitoramento de utilização de novos formatos de arquivo e a adequação do armazenamento ao crescimento de demanda do repositório digital.

A pergunta 5.4 solicita a identificação do *software* utilizado no repositório, apresentando algumas das principais ferramentas utilizadas: DSpace⁸⁶, Archivematica⁸⁷ ou RODA⁸⁸. As opções apresentadas se referem a *softwares* livres (*open source*) e possuem comunidade engajada no desenvolvimento de novas funcionalidades assim como para a correção de problemas relatados.

O objetivo da pergunta foi verificar se o repositório em questão utiliza alguma ferramenta para o gerenciamento do seu acervo digital, fornecendo opções bem difundidas na comunidade arquivística. Além desse fato, ao selecionar alguma das ferramentas, é possível conhecer as características do software utilizado, tais como tecnologias, linguagem de programação, comunidade de desenvolvimento, licenças de *software*, etc. Há ainda a possibilidade de o respondente sinalizar que o repositório utiliza outro software, nenhuma ferramenta é utilizada ou se prefere não fornecer resposta.

O desenvolvimento do *software* DSpace é gerenciado pela organização sem fins lucrativos DuraSpace que tem como missão “[...] fornecer liderança e inovação para tecnologias abertas que promovem acesso durável e persistente a dados digitais” (DURASPACE, 2023, p. tradução nossa). A organização possui um conjunto de membros que tratam desde o estabelecimento de parcerias com empresas para a realização de desenvolvimento de suas ferramentas, até a assessoria técnica para iniciativas de desenvolvimento que visem integração entre as ferramentas suportadas.

A ferramenta Archivematica se intitula como sendo “[...] um aplicativo de código aberto baseado na Web e em padrões que permite à sua instituição preservar o acesso de longo prazo a conteúdo digital confiável, autêntico e seguro” (ARTEFACTUAL SYSTEMS INC, 2022). A ferramenta é composta por diversas outras ferramentas de código livre com o intuito de prover funcionalidades desde a ingestão até a difusão, estando em conformidade com o modelo de referência OAIS⁸⁹.

⁸⁶ <https://dspace.lyrasis.org/>

⁸⁷ <https://www.archivematica.org/>

⁸⁸ <https://www.roda-community.org/#welcome>

⁸⁹ (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012)

A ferramenta RODA (Repositório de Objetos Digitais Autênticos) surgiu como iniciativa da Direção Geral do Livro, dos Arquivos e das Bibliotecas (DGLAB), órgão do Arquivo Nacional de Portugal, que estabeleceu parceria técnica com a Universidade do Minho para o desenvolvimento de um repositório digital. A solução desenvolvida se baseou em tecnologias de software livre e objetivou atender aos padrões utilizados na preservação digital: OAIS, METS, EAD e PREMIS (RAMALHO, FERREIRA, *et al.*, 2014).

Quadro 34 – Perguntas da área 5) Infraestrutura técnica e tecnológica.

Pergunta	Tipo	Tipo de resposta
5.1 O repositório possui estrutura física de prédios e/ou salas adequada à preservação e garantia de preservação a longo prazo?	Semiaberta	Escala de diferencial semântico com opção discursiva
5.2 O repositório possui estrutura de equipamentos (servidores de aplicação, armazenamento, rede, etc.) adequada à preservação e garantia de preservação a longo prazo?	Semiaberta	Escala de diferencial semântico com opção discursiva
5.3 Existe um plano de atualização periódico de hardware e software?	Semiaberta	Escala de diferencial semântico com opção discursiva
5.4 Qual plataforma (<i>software</i>) utilizada pelo repositório?	Semiaberta	Múltiplas opções com opção discursiva

Fonte: elaboração própria.

4.2.3.7 Área 6: Segurança

A área 6 do formulário, nomeada “Segurança”, apresentada no Quadro 35, busca identificar como é feita a gestão de riscos relativos a possíveis problemas de segurança que possam comprometer a preservação digital do repositório. A análise de riscos é uma importante ferramenta que permite a identificação e avaliação de ameaças permitindo o planejamento de medidas para evitá-las, postergá-las ou minimizar os seus efeitos quando acontecerem.

Assim, a pergunta 6.1 solicita que o respondente informe se os riscos que possivelmente afetem a preservação digital são identificados e avaliados. Complementarmente, a pergunta 6.2 solicita que se informe quais ferramentas o repositório utiliza para essa avaliação de riscos, caso ela ocorra no repositório.

A ISO 16363:2012 possui a seção de critérios 5.2 (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 74) voltada exclusivamente para a gestão de riscos de segurança, abordando os riscos envolvidos desde os sistemas de TI integrantes do repositório até os acessos físicos e definição de

responsabilidades da sua equipe. A própria norma, inclusive cita a possibilidade de se utilizar o DRAMBORA para a avaliação geral dos riscos de forma a auxiliar no cumprimento dos requisitos de segurança. Os critérios de segurança da norma ainda fazem uma interligação com a série de normas ISO 27000 que trata de Sistemas de Gestão de Segurança da Informação (*Information Security Management System*)⁹⁰, no sentido de aplicar os seus instrumentos de verificação e controle de ameaças de segurança (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 75).

As perguntas 6.3 e 6.4 buscam analisar como quais medidas o repositório aplica para se prevenir de acessos ou invasões aos seus equipamentos, assim como em relação a sua estrutura física no que se refere ao acesso às instalações onde se localizam os seus equipamentos.

O DRAMBORA apresenta os riscos R35, R36, R37 e R38 (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 169-172) como exemplos que podem ser utilizados para a análise dos riscos que tratam a exploração de vulnerabilidades do repositório, exposição de dados protegidos e a invasão dos equipamentos ou das instalações físicas.

A matriz NDSA de níveis de preservação trata a segurança na área funcional Controle, iniciando no Nível 1 com a identificação dos agentes que interagem com os documentos e sistemas de TI, até chegar no Nível 4 com a revisão periódica dos registros (*logs*) das ações executadas pelos agentes (LEVELS OF PRESERVATION REVISIONS WORKING GROUP, 2019, p. 1)

A Nestor Seal possui o critério 13.2 (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 36) que verifica se a infraestrutura de TI cumpre com os requisitos mínimos de segurança dos documentos, analisando se a integridade, autenticidade, confidencialidade e disponibilidade dos objetos é de fato garantida. O critério 14 (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 37), analisa se a infraestrutura de TI do repositório protege os documentos custodiados, abordando aspectos que vão desde a prevenção de desastres e invasões até a verificação de vírus e programas maliciosos que possam corromper ou expor os dados armazenados.

⁹⁰ (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2018)

Quadro 35 – Perguntas da área 6) Segurança.

Pergunta	Tipo	Tipo de resposta
6.1 Os riscos que afetam a preservação dos objetos digitais são identificados e avaliados?	Semiaberta	Escala de diferencial semântico com opção discursiva
6.2 Caso os riscos tenham sido identificados, quais as ferramentas de avaliação de riscos utilizadas?	Aberta	Discursiva
6.3 A estrutura de TI (<i>software</i> e <i>hardware</i>) do repositório adota medidas de segurança contra ataques?	Semiaberta	Escala de diferencial semântico com opção discursiva
6.4 A estrutura física do repositório adota medidas de segurança para evitar acesso indevido?	Semiaberta	Múltiplas opções com opção discursiva

Fonte: elaboração própria.

4.2.3.8 Área 7: Integridade dos Dados

A área 7, nomeada “Integridade dos dados”, apresentada no Quadro 36, avalia o tratamento dos dados pelo repositório e suas garantias quanto à integridade dos mesmos. Primeiramente, o repositório deve estabelecer o monitoramento dos formatos de arquivos recebidos garantindo que eles estejam de acordo com o seu plano de preservação.

Todo o tratamento e operacionalização dos documentos digitais sob custódia do repositório é realizado a partir do estabelecimento dos formatos suportados, permitindo assim que haja tratamentos especiais em determinados casos. A conversão de um documento recebido em formato proprietário, por exemplo, pode ser realizada para um formato livre que seja mais adequado à preservação a longo prazo. Outra estratégia que pode ser implementada é a migração para uma versão mais recente de um formato obsoleto que permita uma melhor disponibilização do documento à Comunidade-Alvo do repositório. Mesmo nas operações que manipulam os dados originais submetidos, deve haver o compromisso de que o documento original se mantenha original e íntegro, tornando-se elemento essencial para a comprovação da sua autenticidade. Nesse sentido, a pergunta 7.1 visa verificar se o repositório monitora e controla os formatos de arquivos dos materiais submetidos para preservação.

A ISO 16363:2012 possui o critério 4.2.5.1 que determina que o repositório deve possuir ferramentas de monitoramento para a identificação dos formatos de arquivos recebidos, buscando torná-los sempre acessíveis e interpretáveis à Comunidade-Alvo conforme a sua demanda necessita, “esses aspectos são necessários para assegurar que os objetos digitais do repositório sejam compreensíveis a sua Comunidade-Alvo” (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y

CERTIFICACIÓN, 2017, p. 47, tradução nossa). A norma evidencia a sua preocupação de que o conteúdo digital custodiado e difundido pelo repositório sempre atenda às necessidades da sua Comunidade-Alvo tornando-o utilizável.

A preocupação com a integridade física dos objetos digitais custodiados é um dos principais fatores a serem observados pelos repositórios pois, uma vez que haja o corrompimento dos dados armazenados, o acervo pode se tornar inutilizável e o seu principal propósito deixa de existir. A norma ISO 16363:2012 possui o critério 5.1.1.3 e seu subcritério 5.1.1.3.1 que tratam sobre os mecanismos empregados pelo repositório para evitar a perda de dados pelo corrompimento, o estabelecimento de processos para a recuperação dos dados corrompidos e a geração de relatórios disponíveis aos administradores do repositório sobre a extensão dos problemas detectados (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 69). As preocupações sobre os corrompimentos de dados dos documentos devem abranger também os seus respectivos metadados, visto que estes possuem muitas vezes o registro do ciclo de vida da sua fase corrente e intermediária, além das operações realizadas pelo próprio repositório enquanto sob sua guarda.

O Nestor Seal possui o critério 6 que verifica se o repositório digital garante a integridade dos seus objetos digitais em todos os estágios de processamento dos documentos, levando em consideração os instrumentos utilizados para a verificação dos dados assim como as medidas tomadas em relação à infraestrutura de TI que pode sofrer com roubos, por exemplo (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 19). Os subcritérios 6.1, 6.2 e 6.3 tratam da integridade durante a ingestão, arquivamento e difusão, respectivamente (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 20), analisando a presença de processos que evitam e detectam o corrompimento dos dados.

De forma análoga, o critério Nestor Seal 7 e seus subcritérios 7.1, 7.2 e 7.3 (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 21-23) analisam a capacidade de o repositório digital garantir a autenticidade dos objetos armazenados, que só pode existir apoiada na integridade dos mesmos.

O CoreTrust Seal possui o critério R7 para verificar se o repositório garante a integridade e autenticidade dos dados armazenados (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019, p. 16). A verificação da integridade dos dados é feita através de uma análise dos mecanismos e ferramentas

utilizados na verificação de possíveis corrompimentos nos processos arquivísticos de ingestão, arquivamento e difusão de documentos.

Entre os possíveis mecanismos para a verificação de corrompimento, está a utilização de códigos de verificação de fixidade (*checksums*) que são calculados no momento do recebimento dos arquivos e sinalizam que houve a alteração de bits quando há um novo cálculo que fornece um resultado diferente do resultado prévio. Nos casos em que há a necessidade de modificar o arquivo original fornecido, um novo cálculo é realizado e o resultado é armazenado para futuras verificações de integridade.

Tais medidas implantadas servem para “[...] garantir que as mudanças não intencionais ou não autorizadas podem ser detectadas e as versões corretas dos dados e metadados possa ser recuperada” (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019, p. 16, tradução nossa). A verificação de mecanismos e processos que garantam o não corrompimento dos dados preservados é o foco da pergunta 7.2.

A matriz NDSA possui a área funcional **Integridade** voltada a analisar os aspectos implementados pelo repositório no tratamento da integridade dos seus documentos (LEVELS OF PRESERVATION REVISIONS WORKING GROUP, 2019, p. 1). No **Nível 1**, é analisado se o repositório faz as verificações de fixidade no conteúdo recebido comparando os valores informados ou gerando essas informações quando elas não estão presentes na ingestão dos documentos. É analisada a utilização de ferramentas de verificação de vírus no conteúdo e a remoção para quarentena para posterior tratamento dos documentos infectados ou suspeitos de possuírem códigos-fonte mal intencionados.

À medida em que se avança nos níveis dessa área funcional, é analisada a presença de mecanismos adicionais que permitam a detecção de corrompimentos e a recuperação dos dados usando-se rotinas periódicas automáticas para essa finalidade. Vários outros aspectos relacionados à integridade são analisados como a utilização de backups distribuídos que diminui a probabilidade de perda de dados por desastres naturais e a realização de auditorias esporádicas para a identificação de possíveis corrompimentos.

A área 7 do questionário ainda aborda, na pergunta 7.3, a existência de um plano de *backup* nas instituições que salvguarde os dados e evite, diminua ou postergue os riscos mais comuns identificados na instituição em questão. Complementarmente, a pergunta 7.4 busca identificar o uso de diferentes estratégias de backup pelo repositório,

buscando diminuir riscos de perda de dados e alinhando-se principalmente à matriz de preservação NDSA.

Quadro 36 – Perguntas da área 7) Integridade dos dados.

Pergunta	Tipo	Tipo de resposta
7.1 O repositório implementa o monitoramento e tratamento dos formatos de arquivos recebidos?	Semiaberta	Escala de diferencial semântico com opção discursiva
7.2 O repositório adota medidas para prevenir o corrompimento dos dados?	Semiaberta	Escala de diferencial semântico com opção discursiva
7.3 O repositório possui um plano de <i>backup</i> ?	Semiaberta	Escala de diferencial semântico com opção discursiva
7.4 Caso o repositório possua um plano de <i>backup</i> , quais estratégias são empregadas?	Semiaberta	Múltiplas opções com opção discursiva

Fonte: elaboração própria.

4.2.3.9 Área 8: Metadados

É através do conjunto de metadados recebido junto com os documentos para guarda do repositório que se conhece as suas histórias e etapas do ciclo de vida cada documento. Assim, é de fundamental importância que haja o estabelecimento sobre a forma como os metadados serão aceitos, seus formatos e obrigatoriedades buscando preservá-los juntos aos respectivos documentos. As perguntas da área 8, exibidas no Quadro 37, buscam identificar se o repositório pesquisado estabelece, seja na política de preservação e acesso ou através de algum outro instrumento, a forma como os metadados são tratados.

A pergunta 8.1 verifica se os metadados recolhidos estão de acordo com algum padrão pré-estabelecido e reconhecido no meio arquivístico. A escolha de padrões bem aceitos e utilizados por outros repositórios tende a promover a interoperabilidade entre diferentes sistemas, além de facilitar o reuso dos dados custodiados, tornando o repositório aderente aos princípios FAIR⁹¹.

O critério ISO 16363:2012 4.1.2 (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 38) analisa se o repositório especifica quais informações devem estar presentes junto ao objeto digital no momento do seu depósito, podendo-se interpretar que deve haver, de alguma maneira, a formalização desse conjunto de metadados deve acompanhar os documentos durante a

⁹¹ Findable, Accessible, Interoperable and Reusable (WILKINSON, DUMONTIER, *et al.*, 2016).

ingestão. Já o critério ISO 16363:2012 4.5.1 (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 57), da seção de gestão da informação, verifica se o repositório especifica entre os requisitos mínimos para recebimento de documentos, informações que permitam à sua Comunidade-Alvo o acesso e localização do material custodiado, novamente, se alinhando aos princípios FAIR.

Similarmente, o CoreTrust Seal no requisito R8, analisa se os dados e metadados recebidos pelo repositório são especificados da partir de critérios bem definidos e que sejam relevantes à sua Comunidade-Alvo (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019, p. 12). Adicionalmente, o requisito CoreTrust Seal R11 trata da verificação da qualidade dos dados, verificando se o repositório implanta dispositivos, sejam eles automáticos ou manuais, de verificação da completude e semântica dos metadados e quais os procedimentos adotados quando esses dados não estão na qualidade requerida (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019, p. 15). Para o CoreTrust Seal, (2019, p. 15, tradução nossa) “a avaliação de qualidade torna-se mais relevante quando a Comunidade Designada é multidisciplinar, onde os usuários podem não ter experiência pessoal para fazer uma avaliação da qualidade apenas a partir dos dados”⁹².

A Nestor Seal apresenta na seção de gestão de qualidade, o critério 5.2 nomeado “O repositório digital documenta todos os seus elementos com base em um processo definido” (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 18, tradução nossa). Esse critério analisa se o repositório estabelece e cumpre com as suas regras, que englobam desde os seus processos até os conjuntos de metadados e a respectiva qualidade dos metadados recebidos.

A quarta versão do Guia de Avaliação de Repositórios Institucionais analisa no critério 2.7 se o repositório adota uma política de reuso de metadados (AZORÍN, BERNAL, *et al.*, 2021, p. 14), enquanto a seção 4 conta com 32 critérios (AZORÍN, BERNAL, *et al.*, 2021, p. 15-21) e a seção 5 conta com 17 critérios (AZORÍN, BERNAL, *et al.*, 2021, p. 22-25) que tratam, respectivamente, sobre o conjunto de metadados presentes nos documentos sobre a interoperabilidade e acesso desses

⁹² O termo Comunidade Designada é a tradução literal do termo *Designated Community*. Na norma ABNT NBR 14721, o termo foi traduzido para Comunidade-Alvo (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2021, p. 19). A norma espanhola AENOR ISO 14721 traduziu o termo para *Comunidad Específica* (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2015, p. 21).

metadados. O critério 8.3 analisa ainda se o repositório permite a exportação dos metadados de seus documentos em formatos conhecidos e utilizados na comunidade arquivística (AZORÍN, BERNAL, *et al.*, 2021, p. 28).

A pergunta 8.2 tem o intuito de identificar se o repositório recolhe metadados dos seus documentos que permitam se conhecer as respectivas proveniências. A presença de informações, nesse caso de metadados, que testemunhem a proveniência de um documento é essencial para a verificação da autenticidade dos documentos, permitindo-se conhecer os detalhes do contexto de produção desse documento, bem como os processos pelos quais passou durante o seu ciclo de vida.

O propósito da preservação digital a longo prazo só pode ser alcançado quando os documentos custodiados são autênticos e possuem formas e mecanismos de comprovar as suas autenticidades, caso contrário, corre-se o risco de o repositório armazenar documentos que não servem como fontes de prova. Ainda, a autenticidade desses documentos pode ser desqualificada caso não haja elementos que evidenciem o contexto em que foram produzidos, necessitando-se conhecer as funções e cargos ocupados pelos agentes envolvidos além de haver mecanismos de verificação de integridade dos arquivos. Complementarmente, deve-se reunir metadados que especificam os aspectos legais envolvidos como os direitos autorais e licenças de uso aplicáveis aos documentos. Essas informações serão propagadas para as plataformas de difusão do repositório, explicitando como cada documento pode ser usado por sua Comunidade-Alvo.

O critério ISO 16363:2012 4.1.4 (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 40) trata, respectivamente, da presença de mecanismos que permitam ao repositório identificar o(s) produtor(es) dos documentos recebidos. O critério 4.2.1.2 se volta à verificar se o Pacote de Arquivamento de Informações (PAI) (*Archival Information Package – AIP*) possui as informações necessárias para realizar plenamente a preservação a longo prazo dos materiais, possuindo mecanismos que permitam a extração de dados relativos à procedência, direitos de acesso, contexto, referência e fixidez (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 43).

O Nestor Seal possui os critérios 6.1, 6.2, 6.3 (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 20-21) para avaliar se o repositório dispõe de informações que garantem a integridade dos objetos digitais nos processos de ingestão, arquivamento e difusão, respectivamente. Já os critérios

Nestor Seal 7.1, 7.2 e 7.3 (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 21-23) avaliam se o repositório garante a autenticidade dos objetos digitais durante os processos de ingestão, arquivamento e difusão, respectivamente. Tanto a garantia da integridade como da autenticidade dos objetos digitais se inicia com o recebimento de metadados adequados que apoiem os procedimentos de verificação do repositório e vão sendo complementados à medida em que esses procedimentos são executados e registrados.

A Matriz de Níveis de Preservação NDSA apresenta uma dimensão específica para metadados, analisando desde o seu recebimento, passando pela análise da presença de metadados que permitam determinar a proveniência dos documentos, passando pela adaptação de procedimentos para a captura e preenchimento de metadados importantes que estejam em falta, terminando na implantação dos conjuntos de metadados estabelecidos pela instituição e verificação dos procedimentos de preservação dos metadados em si (NATIONAL DIGITAL STEWARDSHIP ALLIANCE, 2022, p. 1). Tanto a pergunta 8.1 quanto a pergunta 8.2 se alinham à dimensão de metadados da Matriz de Níveis de Preservação NDSA.

O DRAMBORA possui listada uma seção completa de exemplos de riscos que se referem ao gerenciamento de metadados. O risco R69, por exemplo, trata sobre o comprometimento dos metadados que suportam a integridade dos objetos de informação, possuindo grau doze de severidade, podendo ser causado pelo corrompimento dos metadados ou a perda da relação entre os metadados e os objetos de informação aos quais eles se referem (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 203). O risco DRAMBORA R71 trata sobre a ausência de metadados dos objetos de informação que permitam a sua localização pela Comunidade-Alvo do repositório (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 205). Esse risco muitas vezes pode ter origem na falta de metadados enviados para o repositório ou, ainda, na presença de metadados que não refletem as necessidades da Comunidade-Alvo, fazendo com que os objetos de informação sejam difíceis ou impossíveis de serem localizados. A ocorrência desse risco pode ainda ferir o princípio **Localizável** (*Findable*) dos princípios FAIR.

Quadro 37 – Perguntas da área 8) Metadados.

Pergunta	Tipo	Tipo de resposta
8.1 O repositório estabelece um padrão de metadados reconhecido a ser adotado pelos documentos recebidos?	Semiaberta	Escala de diferencial semântico com opção discursiva
8.2 São recolhidos/registrados metadados que permitam identificar o contexto, referência, direitos e autenticidade dos documentos?	Semiaberta	Escala de diferencial semântico com opção discursiva

Fonte: elaboração própria.

4.2.3.10 Área 9: Acesso e Consulta

A área 9, nomeada “Acesso e consulta” busca analisar a definição de critérios do repositório para o acesso e consulta de seus documentos e é apresentada no Quadro 38. A pergunta 9.1 analisa se o repositório possui uma política estabelecida de acesso aos objetos digitais custodiados e, em caso positivo, solicita ao respondente que sinalize se essa política só existe na teoria ou é aplicada na prática arquivística por completo ou parcialmente.

O critério ISO 16363:2012 4.6.1 se encarrega de verificar se o repositório cumpre com a política de acesso estabelecida, sendo o instrumento responsável por definir quais documentos e quais informações devem ser disponibilizadas à Comunidade-Alvo e sob quais condições (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 60). O repositório pode determinar que certos materiais não podem ser disponibilizados devido a restrições de direitos de acesso e propriedade ou ainda por conterem dados pessoais que possam infringir direitos pessoas, observando a LGPD (BRASIL, 2018) e GPDR (EUROPEAN COMMISSION, 2018), por exemplo.

A política emanada pode decidir ainda que todos os acessos realizados no acervo devem ser realizados somente após a identificação do usuário através de login e senha. Assim, tem-se o controle de quem acessou determinados documentos, permitindo a realização de auditorias quando necessário. Da mesma forma, é estabelecido se devem existir APIs⁹³ de interação entre outros *softwares* com o repositório e em quais condições essas interações estão autorizadas. Adicionalmente, a política de acesso deve se preocupar com os acessos aos documentos originais em meio físico, especificando quais procedimentos os membros da Comunidade-Alvo devem seguir para ter acesso a esse material e sob quais condições esse acesso pode ocorrer.

⁹³ *Application Programming Interface* (Interface de Programação de Aplicações).

O CoreTrust Seal, no requisito R13, verifica se o repositório permite aos usuários localizarem os dados custodiados através de mecanismos de busca adequados e se o material disponibilizado possui identificadores persistentes que permitem a sua identificação e referência permanente e única através da internet (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019, p. 17). Além disso, o requisito verifica se existe a possibilidade de realizar pesquisas nos metadados armazenados, preferencialmente em padrões internacionais, e se existe algum mecanismo implantado que permita a extração desses metadados de forma automatizada.

O Nestor Seal apresenta a seção 2 de critérios relacionados aos acessos dos objetos digitais do repositório. O critério 2.1 avalia se o repositório permite que a sua Comunidade-Alvo acesse os objetos digitais, enquanto o critério 2.2 avalia se os mesmos objetos digitais disponibilizados podem ser interpretáveis pela Comunidade-Alvo, voltando a atenção à necessidade de se manter os objetos utilizáveis, atingindo-se o propósito da preservação a longo prazo (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 12).

O DRAMBORA possui a seção 7.3.8 destinada a exemplos de riscos relacionados ao acesso e disseminação de informações do repositório. Dentre os riscos exemplificados, o risco R75 trata de falhas ocorridas no subsistema de autenticação de usuários, impossibilitando o acesso aos documentos do repositório (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 209). O risco R76 trata da ocorrência de falhas no subsistema de autorização de acesso, fazendo com que os usuários não consigam acessar os documentos adequadamente, conforme a política de acesso estabelece, ou ainda permitindo que documentos restritos estejam inapropriadamente disponíveis (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 210). O risco R78 se refere à perda de desempenho ou a quebra de nível do serviço estabelecido de acesso, impossibilitando ou dificultando o acesso e consulta aos documentos do repositório digital (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 212).

O nível de critérios estabelecidos para a implementação da política de acesso deve balancear os interesses da Comunidade-Alvo com o nível de detalhes a serem capturados ou exigidos nos acessos. Por exemplo, ao tornar o acesso aos documentos completamente público, sem a exigência de identificação do usuário, pode-

se favorecer a difusão dos documentos e, talvez a sua popularização, no entanto, abre-se mão de se conhecer as ações realizadas pelos usuários durante os seus acessos.

A pergunta 9.2 analisa se as necessidades da Comunidade-Alvo são monitoradas a fim de se identificar possíveis mudanças a serem efetuadas na política de acesso ou dos mecanismos implementados para o acesso e consulta dos documentos. O rol de opções de respostas permite que se forneça a informação sobre a real aplicação desse monitoramento, que pode não existir, existir em teoria, estar aplicado parcialmente ou estar totalmente aplicado na prática do repositório.

Entre as mudanças que podem ser identificadas no monitoramento, estão a inclusão de formatos de arquivos ofertados ou a mudança para outros formatos mais adequados, disponibilização de outros suportes além dos já existentes, preferência dos usuários por pacotes de softwares diferentes dos oferecidos, etc (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2015, p. 51).

O repositório deve se comprometer com o monitoramento dos requisitos de acesso e consulta da sua Comunidade-Alvo a fim de se adaptar e promover uma melhor usabilidade. A política de acesso do repositório pode estabelecer a forma e a periodicidades de identificação desses requisitos, optando por realizar pesquisas, enquetes ou entrevistas, por exemplo. Paralelamente, a própria equipe do repositório deve desempenhar função semelhante, buscando identificar mudanças de tecnologias que favoreçam melhores formas de acesso e consulta aos documentos custodiados, atenta ao seu entorno tecnológico.

Sempre que possível o repositório deve escolher formatos de arquivos abertos, não proprietários, evitando que haja dependência em soluções e softwares proprietários que podem descontinuar o suporte aos seus programas de interpretação e leitura. Dentre as boas práticas recomendadas está o uso de formatos sustentáveis que possuem uma maior probabilidade de permanecerem acessíveis ao longo do tempo (NATIONAL DIGITAL STEWARDSHIP ALLIANCE, 2022, p. 8).

No entanto, nos casos de repositórios digitais de domínio específico, por exemplo, pode ser necessário que a política de acesso especifique o recebimento e a difusão de arquivos em formatos proprietários não sustentáveis, com o intuito de se adaptar às necessidades da sua Comunidade-Alvo. Consequentemente, o repositório deverá adaptar as suas estratégias de preservação de forma a manter esses documentos íntegros e interpretáveis pelo tempo de preservação necessário.

A pergunta 9.3 solicita ao respondente a informação sobre quais os formatos de arquivos mais comumente disponibilizados à sua Comunidade-Alvo, informando o grau de frequência de cada um dos formatos sugeridos. Como opções de resposta, foram oferecidos alguns formatos utilizados para a difusão de documentos, como PDF, imagens, áudio, vídeo, texto e planilhas de cálculo. Para a escala de frequência foram fornecidas as opções: a) Nunca; b) Raramente; c) Ocasionalmente; d) Frequentemente, e; e) Muito frequentemente. Para cada um dos formatos de arquivo assinalados, o respondente deveria selecionar qual a opção de frequência mais apropriada de acordo com a realidade do repositório pesquisado.

A pergunta 9.4 complementa a sua anterior, buscando identificar se o repositório disponibiliza outros formatos de arquivo, além dos apresentados na pergunta 9.3. O propósito das perguntas 9.3 e 9.4 é identificar, de forma ampla, o tipo de material custodiado e difundido pelo repositório, analisando-se os formatos de arquivo mais usados.

Quadro 38 – Perguntas da área 9) Acesso e consulta.

Pergunta	Tipo	Tipo de resposta
9.1 Existe uma política de acesso aos objetos digitais?	Semiaberta	Escala de diferencial semântico com opção discursiva.
9.2 Existe uma política de monitoramento das necessidades da sua Comunidade-Alvo a fim de disponibilizar os documentos da forma correta e nos formatos de arquivo mais adequados?	Semiaberta	Escala de diferencial semântico com opção discursiva.
9.3 Quais os formatos de arquivo mais habituais que o repositório costuma disponibilizar aos usuários?	Semiaberta	Múltiplas opções e escala de diferencial semântico.
9.3.1 Existem outros formatos habituais disponibilizados aos usuários?	Aberta	Discursiva.

Fonte: elaboração própria.

4.2.3.11 Área 10: Visibilidade

A última área do questionário, nomeada “Visibilidade”, é apresentada no Quadro 39 e apresenta questões que buscam analisar como ocorre a interação entre o repositório e a sua Comunidade-Alvo. Os usuários do repositório devem localizá-lo e acessá-lo facilmente para que consigam realizar as suas consultas e acessos aos materiais custodiados, promovendo a difusão. A visibilidade do repositório traz como consequência positiva a promoção da transparência ativa dos documentos da instituição, contribuindo para o exercício da cidadania da sua Comunidade-Alvo.

Embora o critério ISO 16363:2012 4.5.1 (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 57) não trate especificamente da visibilidade do repositório, ao se voltar mais aos metadados e aos mecanismos que permitem à comunidade localizar o seu material de interesse, pode-se inferir que primeiramente a Comunidade-Alvo deve localizar o próprio repositório para então realizar as suas buscas através das funcionalidades voltadas a esse fim. Portanto, a presença ou a ausência de ações que difundam o próprio repositório impacta diretamente, positivamente ou negativamente, na difusão dos seus documentos.

A pergunta 10.1 verifica se a página principal da instituição à qual o repositório pertence possui algum meio de direcionar o usuário para o repositório em si. Ao acessar a página principal do sítio da instituição, é recomendado que haja alguma forma de o usuário navegar até o repositório dessa instituição, facilitando que o usuário tome conhecimento dos detalhes de seu funcionamento, sua política de preservação e demais instrumentos que o regulam, por exemplo, além de direcioná-lo para as funcionalidades de pesquisa existentes. Essa pergunta sobre a visibilidade do repositório se alinha ao critério Recolecta 1.1 referente a uma modificação da segunda versão do Guia de Avaliação de Repositórios Institucionais (AZORÍN MILLARUELO, CÍVICO MARTÍN, *et al.*, 2015, p. 11), no entanto, na última versão do Guia (AZORÍN, BERNAL, *et al.*, 2021), esse critério não foi incluído.

A pergunta 10.2 verifica se existem estratégias de promoção do repositório institucional que fazem com que a sua Comunidade-Alvo saiba da sua existência. Através das respostas fornecidas, é possível saber se essas estratégias existem em teoria, são aplicadas totalmente, parcialmente ou se inexistem na instituição. A pergunta 10.3 permite ao respondente informar, através de resposta aberta, quais são as iniciativas de promoção implementadas pelo repositório pesquisado.

O propósito do repositório digital, nos moldes de um Arquivo OAIS, é “[...] preservar informação para acesso e uso por uma Comunidade-Alvo” (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2021, p. 18), logo, se os seus usuários desconhecem a sua existência e forma de funcionamento, o acesso aos documentos torna-se limitado e, muitas vezes, pode ser subutilizado. Assim, é fundamental que haja iniciativas que promovam o repositório, difundindo as suas condições de uso, tipos de materiais disponíveis e formatos suportados, por exemplo, buscando aproximar os usuários dos documentos custodiados da instituição.

A pergunta 10.4 analisa a página do repositório disponibiliza algum meio de contato entre o público e sua equipe. Dentre as opções, o respondente pode assinalar se existe um endereço de e-mail formulário para contato, comunicação por redes sociais, telefone ou nenhuma forma de comunicação é disponibilizada. Torna-se importante o repositório estabelecer um canal de comunicação com a sua Comunidade-Alvo, permitindo que os usuários tirem suas dúvidas, solicitem esclarecimentos, sugiram o acréscimo de materiais, e até mesmo sugiram melhorias no seu funcionamento.

Da mesma forma como o repositório disponibiliza os seus documentos digitalmente, sob a própria demanda da sua comunidade, os meios de contato também devem acompanhar essas mudanças, permitindo que as pessoas contatem o repositório facilmente. Assim, os meios de contato digitais favorecem a comunicação e torna o repositório mais acessível e amigável à sua Comunidade-Alvo ao permitir troca de mensagens. Há que se ressaltar que essas qualidades de visibilidade do repositório podem não se aplicar a arquivos obscuros (*dark archives*), uma vez que a natureza desses repositórios é permanecer fora de acesso ao público. Dentre as razões para um repositório permanecer obscuro, existe a falta de mecanismos de processamento e análise do material para classificá-lo e tornar acessível somente o permitido, assim, opta-se por tornar o repositório inacessível, custodiando documentos secretos ou com informações sensíveis geralmente.

Quadro 39 – Perguntas da seção 10) Visibilidade.

Pergunta	Tipo	Tipo de resposta
10.1 É possível localizar o acesso ao repositório a partir da página principal da instituição?	Semiaberta	Sim ou não com opção discursiva.
10.2 Existem iniciativas para a promoção do repositório institucional?	Semiaberta	Escala de diferencial semântico com opção discursiva.
10.3 Caso haja iniciativas de promoção do repositório, quais estratégias e instrumentos utilizados?	Aberta	Discursiva.
10.4 Na página web do repositório, são utilizadas quais as formas de contato com os responsáveis pelo repositório?	Aberta	Múltiplas opções com opção discursiva.

Fonte: elaboração própria.

4.2.4 Fase 4: Coleta de Dados

Para a coleta das respostas dos formulários recebidos, utilizou-se o *Google Forms*⁹⁴, configurando-se as seções, perguntas e opções de respostas. Essa ferramenta

⁹⁴ <https://docs.google.com/forms>

permite aos respondentes o preenchimento de forma online, remotamente, sem a necessidade de envio e recebimento de arquivos, ao mesmo tempo em que se integra a outra ferramenta de planilha de cálculos que armazena as respostas em tempo real e permite a análise dos dados. Com o intuito de se identificar falhas que pudessem comprometer a coleta dos dados, foi realizada uma testagem guiada antes da aplicação efetiva com os repositórios selecionados, o que permitiu identificar o tempo médio para o preenchimento completo do formulário.

Como forma de contato com os 31 repositórios da Rede Cariniana, primeiramente foi escolhida mensagem por e-mail para os contatos dos responsáveis contendo a apresentação do estudo e o *hiperlink* para acesso ao formulário de respostas. Essa forma de contato, em geral, é ágil e permite a troca de mensagens para a resolução de eventuais dúvidas, no entanto, é necessário conhecer previamente o endereço de e-mail correto dos responsáveis pelos repositórios. Como desvantagens, não é possível saber se houve o recebimento da mensagem corretamente e se alguém de fato a leu.

O prazo inicial para responder ao questionário foi estipulado em 10 dias, iniciando em abril de 2022, com monitoramento diário dos recebimentos das respostas. Nesse primeiro período, foi registrada a participação de 6 repositórios, sendo 1 de natureza privada. Buscando contar com uma maior participação das instituições, foi prevista uma ampliação do prazo para respostas, compreendendo o período entre abril e maio de 2022, tendo sido enviadas novas mensagens informando sobre o novo prazo de participação. No caso dos repositórios de natureza privada ou estadual, novas mensagens de e-mail foram enviadas, no caso dos repositórios de instituições públicas federais, essa segunda comunicação foi realizada através de manifestações de ouvidoria, solicitando as suas adesões ao estudo.

A manifestação representa um tipo de comunicação formal com os órgãos públicos para Pedido de Acesso à Informação, através da Plataforma Integrada de Ouvidoria e Acesso à Informação (Fala.BR) do governo federal (CONTROLADORIA-GERAL DA UNIÃO, 2022). A Lei 13.460 de 26 de junho de 2017 trata sobre os direitos dos usuários de serviços públicos e, em seu artigo 5º inciso VI, determina que os órgãos e agentes públicos devem cumprir com os prazos estabelecidos, que no caso das manifestações de ouvidoria são de 30 dias prorrogáveis sob justificativa por igual período uma única vez, conforme artigo 16 (BRASIL, 2017). No decorrer desses prazos, as ouvidorias dos órgãos podem solicitar complementações e esclarecimentos de

informações, quando for o caso. O não cumprimento dos prazos estipulados favorece a realização de uma avaliação negativa da instituição que pode sofrer sanções futuramente.

Esse tipo de comunicação possui vantagens como o controle das comunicações realizadas, identificação dos agentes e cumprimento de prazos, ao contrário das mensagens de e-mail que possuem caráter mais informal. No entanto, se impõe burocracia e um certo grau de formalidade à troca de mensagens, uma vez que elas são recebidas pelos ouvidores das instituições e podem envolver a tramitação para diversos setores internos da instituição até alcançar o repositório em si. Ainda, devido ao monitoramento do governo federal sobre as manifestações, esse tipo de comunicação pode representar imposição e obrigatoriedade na participação da atual pesquisa, qualidades não desejáveis à realização do atual estudo que preza pela participação livre e espontânea.

Ao todo, foram criadas 19 manifestações de Pedido de Acesso à Informação destinadas às instituições públicas responsáveis ou controladoras dos repositórios. No caso de 4 instituições, a pedido dos seus responsáveis, foram necessárias criações de manifestações adicionais para maior esclarecimento dos propósitos da pesquisa e o fornecimento de um arquivo PDF com o conteúdo do formulário para apreciação, totalizando-se 23 manifestações.

Dentre os 5 repositórios de natureza privada pertencentes à Rede Cariniana, 1 participou da pesquisa e respondeu ao questionário. Dos 4 repositórios de esfera estadual ou distrital, 1 repositório respondeu ao questionário. Dentre os demais 22 repositórios de instituições federais, registrou-se a participação de 17 repositórios, totalizando-se 19 repositórios participantes do estudo. Uma análise mais detalhada sobre a participação dos repositórios é apresentada na seção 5.1 do presente trabalho.

4.2.5 Fase 5: Elaboração da Ferramenta de Pontuação

Com o objetivo de ser possível quantificar os desempenhos dos repositórios participantes frente às áreas e perguntas realizadas, foi elaborada uma ferramenta de pontuação, facilitando a identificação dos *scores* alcançados pelos repositórios em relação às diferentes dimensões avaliadas.

Utilizando-se uma planilha de cálculos, inseriu-se todas as áreas e perguntas do questionário de avaliação, onde se atribuiu pesos iguais para cada uma das 10 áreas de avaliação. Dessa forma, se favorece uma avaliação uniforme entre os critérios, não

favorecendo determinadas áreas em detrimento de outras, o que poderia beneficiar ou prejudicar o desempenho de certos repositórios.

Uma vez que determinadas áreas possuem quantidades de perguntas diferentes entre si, optou-se por atribuir os pesos de cada pergunta como sendo o peso da área dividido pela quantidade de perguntas da área. Assim, a pontuação dos repositórios em cada área não é afetada positivamente ou negativamente pela quantidade de perguntas dessa área.

O Quadro 40 apresenta um exemplo da planilha de cálculos utilizada com algumas particularidades utilizadas para certas perguntas, contendo as pontuações máximas das perguntas e a escala de valores das suas respostas. No caso da pergunta “1.1 O repositório institucional possui missão, visão e valores elaborados”, o seu valor total máximo é de 3,33 pontos, uma vez que a área “1. Contextualização do repositório” possui peso 10 e 3 perguntas ao todo. Nesse caso, para as respostas “Sim” e “Não”, se atribuiu, respectivamente 100% (3,33 pontos) e 0,00% (0 ponto).

Quadro 40 - Exemplo de pontuação de perguntas das áreas de avaliação.

Pergunta	Pontuação máxima
1.1 O repositório institucional possui missão, visão e valores elaborados?	3,33
Respostas	Valor
Sim	3,33 (100%)
Não	0,00
Prefiro não responder	0,00
Pergunta	Pontuação máxima
1.4 Qual a natureza do material depositado no repositório?	3,33
Respostas	Valor
1 Categoria de material	0,67 (20%)
2 Categorias de materiais	1,33 (40%)
3 Categorias de materiais	2,00 (60%)
4 Categorias de materiais	2,67 (80%)
5 ou mais categorias de materiais	3,33 (100%)
Prefiro não responder	0,00
Pergunta	Pontuação máxima
2.1 Existe uma Política de Preservação com os requisitos mínimos dos documentos recebidos, tratamento de preservação e acesso?	3,33
Respostas	Valor
Existe em teoria	0,83 (25%)
Existe na prática, mas ainda não foi totalmente implantado	1,67 (50%)
Existe na prática e está totalmente implantado	3,33 (100%)
Não existe	0,00
Prefiro não responder	0,00

Fonte: elaboração própria.

A pergunta “1.4 Qual a natureza do material depositado no repositório”, com pontuação máxima de 3,33 pontos, buscou atribuir pontos de acordo com a quantidade de categorias de materiais tratadas pelo repositório, informações extraídas a partir das respostas fornecidas pelos respondentes. A pontuação das respostas inicia com

0,67 ponto (20% do valor da pergunta) no caso de o repositório possuir apenas uma categoria de material e evolui até o valor máximo da pergunta, equivalente a 100% do valor máximo da pergunta.

No caso de perguntas que apresentam resposta em uma espécie de escala, como a “2.1 Existe uma Política de Preservação com os requisitos mínimos dos documentos recebidos, tratamento de preservação e acesso”, que possui valor máximo 3,33 pontos, os valores das respostas são 0,83, 1,67 e 3,33, respectivamente, 25%, 50% e 100% da pontuação da pergunta. Em todas as perguntas, adotou-se a pontuação de 0,00 (zero) ponto para a resposta “Prefiro não responder”.

No caso da pergunta “6.2 Caso os riscos sejam identificados, quantas ferramentas de avaliação de risco são utilizadas?”, da área “6. Segurança”, optou-se por não a incluir nos critérios de pontuação, uma vez que a quantidade informada pelos respondentes não leva em consideração a abrangência e escopo das ferramentas utilizadas. Métodos como o DRAMBORA (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007), por exemplo, abrangem aspectos de todo o repositório, que vão desde a infraestrutura organizacional até a avaliação de ameaças, enquanto outras ferramentas podem ser específicas de uma determinada área operacional do repositório como, por exemplo, os ativos de rede.

O instrumento completo de avaliação quantitativa para a atribuição da pontuação dos repositórios é apresentado no Anexo B.

4.2.6 Fase 6: Análise de Dados

Como a plataforma escolhida para elaboração do questionário foi a *Google Forms*, à medida em que as respostas vão sendo recebidas, a plataforma as armazena em uma planilha de cálculos e, ao final do prazo de recebimento do questionário, utilizou-se essa planilha como fonte de dados para as análises realizadas. A planilha de cálculos foi importada para o software Microsoft Excel para a realização de análises e agrupamentos detalhados, assim como para a geração de gráficos.

Nos casos das perguntas em que foi sugerido ao respondente informar o endereço (URL) onde seria possível encontrar as informações como, por exemplo, as políticas de preservação, foi realizado o acesso para extração das devidas informações.

Não foram registradas respostas com endereços inválidos ou inacessíveis, favorecendo a coleta de dados.

A análise dos resultados utilizou a pontuação dos repositórios em relação a diferentes dimensões, buscando compreender e justificar os comportamentos dos repositórios frente às pontuações obtidas. Primeiramente, foram analisados os dados sobre os respondentes dos questionários na seção 5.1. Após, foram feitas análises globais dos resultados, estabelecendo-se o *ranking* global dos repositórios na seção 5.2 e na seção 5.3 os *rankings* por área. Na seção 5.4 foram analisados os desempenhos dos repositórios área a área, identificando fatores internos aos repositórios que influenciam nos seus desempenhos. Em seguida, na seção 5.5 foram feitas análises sobre a influência de fatores externos aos repositórios que influenciam nos seus desempenhos.

Uma vez que um grande número de repositórios está subordinado a universidades, a seção 5.5.5 apresenta análises focadas e direcionadas a características e dados encontrados para as Instituições de Ensino Superior (IES). Por fim, a seção 5.6 apresenta algumas conclusões e análises gerais sobre os desempenhos dos repositórios.

Como forma de medir a variabilidade das pontuações dos repositórios, escolheu-se utilizar o Desvio Médio (DM), que se refere à Média Aritmética (MA) das diferenças das pontuações em relação às suas MA. Esse cálculo permite constatar o quanto as pontuações dos repositórios se afastam em relação às médias de pontuação. Em alguns casos, se utilizou também o Desvio Padrão (DP) que indica o quanto os valores individuais de um conjunto de dados se afastam, em média, da média dos valores. Interpretando-se o Desvio Padrão, quanto maior o seu valor, maior é a dispersão dos dados em relação à média, indicando uma maior variabilidade. Por outro lado, um Desvio Padrão menor indica que os valores estão mais próximos da média, sugerindo menor variabilidade (ALTMAN e BLAND, 2005). O DP ainda indica um intervalo em relação à Média Aritmética onde a maioria dos valores da amostra se encontra.

Outro cálculo utilizado foi o Coeficiente de Variação (CV) expresso em porcentagem, calculado como o produto da multiplicação por 100 da razão entre o Desvio Padrão pela Média Aritmética geral ($CV = (DP/MA) \cdot 100$), que demonstra o grau de variabilidade da amostra de valores apresentados e serve como medida de comparação da variabilidade entre os repositórios (BEDEIAN e MOSSHOLDER, 2000). Uma vez que o CV se expressa em porcentagem, ele permite comparar a variação de

conjuntos diferentes de dados, como as pontuações dos repositórios em diferentes áreas de avaliação.

O Quadro 41 exemplifica a realização dos cálculos a partir de pontuações fictícias dos repositórios em duas áreas diferentes. No caso da área A, a MA calculada das pontuações foi 5,44 pontos, o DM encontrado foi 1,48 pontos, o que indica que as pontuações se afastam da MA em média 1,48 pontos, enquanto o DP calculado foi 1,61 pontos, indicando que a maioria das pontuações se encontra no intervalo entre $5,44 \pm 1,61$ pontos, ou seja, entre 2,83 pontos (MA-DP) e 7,05 pontos (MA+DP). Em comparação com a área B, calculou-se um DM de 1,67 pontos, indicando que as pontuações dessa área são ligeiramente menos uniformes que a área A, fato corroborado pelo maior DP que abrange um intervalo maior de pontuação, na faixa entre $2,92 \pm 1,97$ pontos.

Ao se comparar o CV, percebe-se uma nítida diferença entre os valores das duas áreas, permitindo-se concluir que as pontuações da área A são mais uniformes que as da área B, uma vez que os valores dos CV foram, respectivamente, 29,50% e 67,51%. Ao analisar as pontuações reais dos repositórios nas diferentes áreas avaliadas, o CV permite identificar áreas em que o tratamento e atendimento aos requisitos das áreas é mais uniforme entre os repositórios do que outras áreas, indicando um maior consenso.

Quadro 41 – Exemplo de cálculos de MA, DM, DP e CV.

Repositório	Pontuação área A	Pontuação área B
R01	4,67	6,67
R02	4,00	0,83
R03	4,00	1,67
R04	7,33	1,67
R05	8,00	2,50
R06	4,67	4,17
Média Aritmética (MA)	5,44	2,92
Desvio Médio (DM)	1,48	1,67
Desvio Padrão (DP)	1,61	1,97
Coefficiente de Variação (CV)	29,50%	67,51%

Fonte: elaboração própria.

5 RESULTADOS

O presente capítulo apresenta os resultados obtidos no presente estudo, analisando primeiramente o respondente na seção 5.1. Após, são realizadas as análises globais dos desempenhos dos repositórios na seção 5.2, seguida pelas análises por áreas na seção 5.3, pelas análises por regiões geográficas brasileiras na seção 5.5.1, pelas análises por tipo de parceria dos repositórios com a Rede Cariniana na seção 5.5.2, pelas análises por tipo de instituição na seção 5.5.3 e, por último, pelas análises por tempo de criação dos repositórios na seção 5.5.4. Logo após, na seção 5.5.5 são feitas análises direcionadas aos repositórios que são ligados a Instituições de Ensino Superior (IES), visto que representam a grande maioria dos repositórios participantes.

Buscando compreender o comportamento dos repositórios individualmente nas áreas de avaliação, a seção 5.4 apresenta os resultados por área e perguntas realizadas.

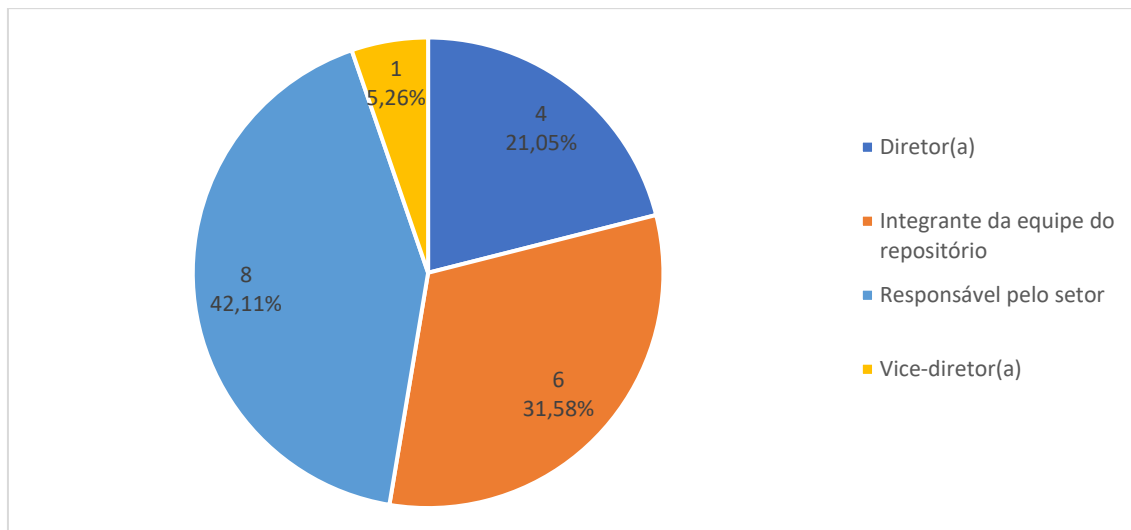
5.1 ANÁLISE DO RESPONDENTE

Nessa seção são analisadas as respostas da seção 1 do questionário que visa identificar o perfil da pessoa respondente. A Figura 18 analisa as respostas da pergunta “1.1 Qual o seu cargo ou função desempenhada no órgão que gerencia o repositório?”. Do total de respondentes, 8 pessoas declararam-se responsáveis pelo setor do repositório, 4 pessoas declararam-se diretoras e 1 pessoa se declarou vice-diretor(a) do repositório. Com isso, pode-se concluir que 68,42% das pessoas que participaram da pesquisa são responsáveis e/ou possuem certo grau de responsabilidade sobre o funcionamento dos repositórios das instituições. As 6 pessoas restantes, que representam 31,58%, são integrantes das equipes dos repositórios e, portanto, detêm conhecimento a respeito do funcionamento dos seus respectivos repositórios. Assim, pode-se assumir um alto grau de confiança das respostas fornecidas ao estudo.

O preenchimento das respostas de 1 questionário foi realizado por várias pessoas ligadas ao respectivo repositório, com o intuito de formar um consenso e

melhor compreensão a respeito das informações solicitadas. Nesse caso, para fins de análise, contabilizou-se essa resposta como sendo “Integrante da equipe do repositório”.

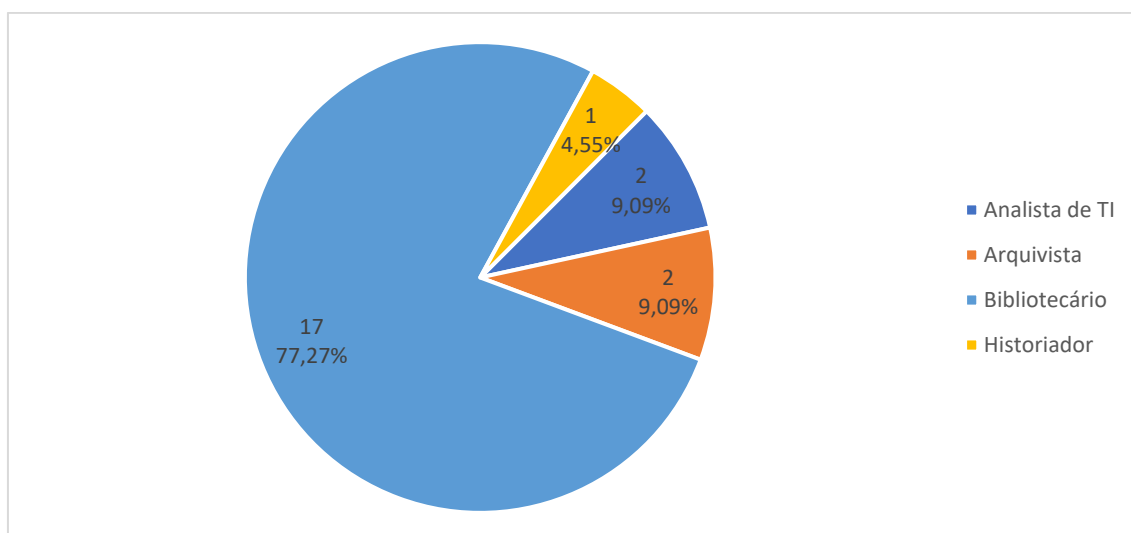
Figura 18 – Cargo ou função desempenhada pelo respondente no repositório



Fonte: elaboração própria.

A Figura 19 analisa as respostas à pergunta “1.2 Qual a sua formação acadêmica”, que solicitou ao respondente o preenchimento da resposta mais adequada à sua formação acadêmica.

Figura 19 – Formação acadêmica dos respondentes



Fonte: elaboração própria.

Como o preenchimento do formulário de um dos repositórios foi preenchido por uma equipe de vários profissionais, a resposta à essa questão apresentou mais de um valor nesse caso. Em um outro formulário, o respondente declarou possuir duas

formações acadêmicas: Bibliotecário e Arquivista. Assim, todas as respostas foram consideradas na análise, fato demonstrado pela quantidade de respostas, 22, ser maior que a quantidade de repositórios participantes da pesquisa, 19.

A maioria dos respondentes, 77,27%, declarou possuir a formação de Bibliotecário(a), totalizando 17 pessoas. A formação acadêmica Arquivista foi declarada por 2 pessoas respondentes, o que equivale a 9,09% das pessoas, mesmo total de pessoas que declarou possuir formação acadêmica de Analista de Tecnologia da Informação, 9,09%. Dentre os respondentes, 1 pessoa declarou possuir a formação acadêmica de Historiador(a), equivalente a 4,55% total.

Nota-se uma grande disparidade nas formações dos respondentes, uma vez que a grande maioria possui formação de Bibliotecário, enquanto que as formações de Arquivista e Historiador somam 13,64%, mesmos que estas também possuam como objeto de trabalho materiais a serem preservador e depositados em repositórios digitais. A predominância ainda pode ser justificada pelo fato de que a natureza dos materiais da maioria dos repositórios institucionais respondentes ser materiais de pesquisa como artigos, monografias, dissertações e teses, conforme se observa na pergunta 2.4, analisada seção 5.4.1 do presente estudo.

Enquanto a tecnologia é parte fundamental e indispensável na realização da preservação digital, nota-se a presença de apenas 2 respondentes, 9,09%, com formação na área de Tecnologia da Informação (TI), demonstrando um grande distanciamento das demais formações e permite constatar a carência de profissionais com esse tipo de formação atuando nos repositórios brasileiros, ao menos nas funções e cargos declarados na pergunta anterior, 1.1.

5.2 ANÁLISE GLOBAL DOS REPOSITÓRIOS

Nessa seção do trabalho são analisados os dados de desempenho dos repositórios participantes de um modo geral, de onde é possível se estabelecer um ranking das pontuações. Uma vez que foram consideradas 10 áreas de avaliação e cada área possui valor máximo de pontuação de 10 pontos, as pontuações gerais máximas e mínimas possíveis são 100 e 0 (zero) pontos, respectivamente. De forma análoga, para as pontuações das perguntas das áreas, escolheu-se dividir o valor da área, 10 pontos, pela sua quantidade de perguntas, conforme detalhado na seção 4.1.

O Quadro 42 apresenta os repositórios com as suas respectivas pontuações gerais, que são a soma das pontuações em cada área, e as Médias Aritméticas (MA) das pontuações, que é a pontuação geral dividida pelas 10 áreas de avaliação. Como Média Aritmética das pontuações gerais, se obteve 42,71 pontos e a Média Aritmética das Médias Aritméticas das pontuações foi de 4,27 pontos.

Quadro 42 – Repositórios, suas pontuações totais e Médias Aritméticas.

Repositório	Pontuação	MA	Repositório	Pontuação	MA
R01	46,03	4,60	R11	53,94	5,39
R02	21,01	2,10	R12	37,65	3,77
R03	47,94	4,79	R13	63,40	6,34
R04	27,43	2,74	R14	66,86	6,69
R05	64,65	6,47	R15	31,26	3,13
R06	36,10	3,61	R16	17,90	1,79
R07	47,24	4,72	R17	66,74	6,67
R08	51,11	5,11	R18	23,49	2,35
R09	49,28	4,93	R19	24,85	2,48
R10	34,63	3,46			

Fonte: elaboração própria.

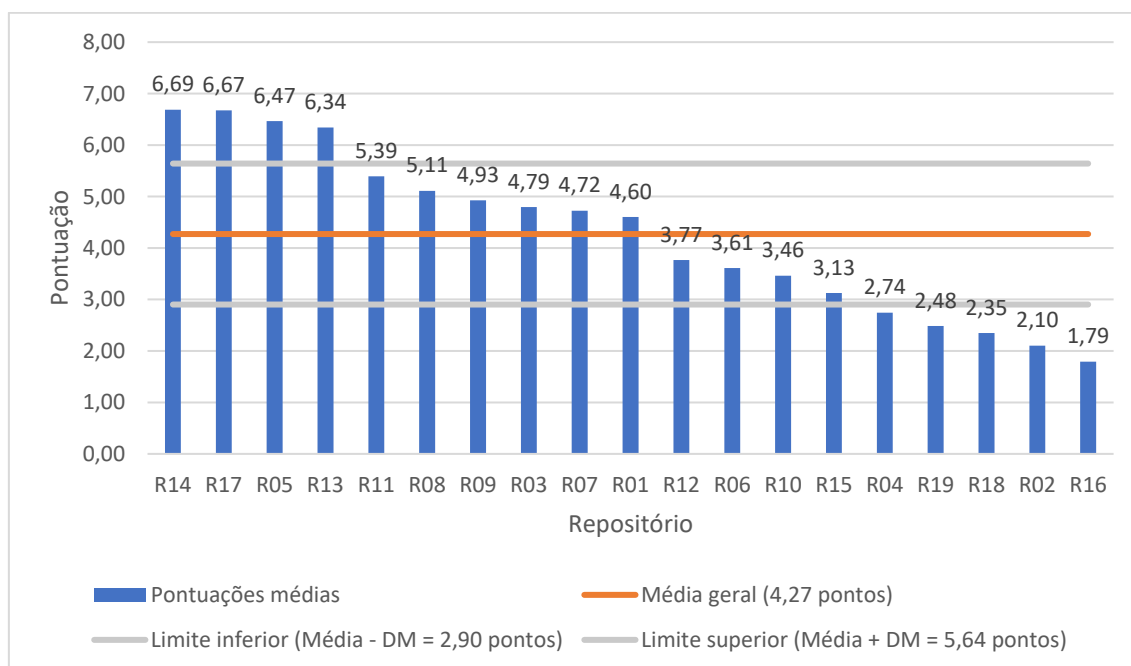
Com o objetivo de medir a variabilidade das pontuações apresentadas, utilizou-se o Desvio Médio (DM), que é a média das diferenças das pontuações em relação à média geral (4,27 pontos), o Desvio Padrão (DP) que indica o quanto os valores individuais de um conjunto de dados se afastam, em média, da média dos valores. Interpretando-se o Desvio Padrão, quanto maior o seu valor, maior é a dispersão dos dados em relação à média, indicando uma maior variabilidade. Por outro lado, um Desvio Padrão menor indica que os valores estão mais próximos da média, sugerindo menor variabilidade (ALTMAN e BLAND, 2005).

Assim, utilizando-se as Médias Aritméticas de pontuação dos repositórios, se obteve como Desvio Médio o valor de 1,37 pontos, significando que as pontuações médias dos repositórios variam, em média, 1,37 pontos da Média Aritmética geral. O Desvio Padrão da amostra foi de 1,56 pontos, significando que a maioria das pontuações médias se encontra no intervalo $4,27 \pm 1,56$ pontos, ou seja, entre 2,71 e 5,83 pontos. O Coeficiente de Variação (CV) encontrado foi de 36,60%, identificando-se uma grande variação das pontuações dos repositórios em relação à MA.

A Figura 20 apresenta o *ranking* geral dos repositórios obtido a partir das MA de suas pontuações em ordem decrescente, a MA geral de 4,27 pontos e a faixa de pontuação que considera a MA geral acrescentada ou subtraída do Desvio Médio de

1,37 pontos, entre 2,90 e 5,64 pontos. Em uma primeira análise, pode-se observar que 10 repositórios possuem pontuação média acima da média e 9 repositórios se posicionam com pontuação média abaixo da média de pontuação. Pode-se assumir, assim que a média das pontuações divide a amostra efetivamente em dois conjuntos de quantidades muito similares.

Figura 20 – *Ranking* geral dos repositórios pelas médias de pontuações.



Fonte: elaboração própria.

O Quadro 43 apresenta uma análise de desempenho das pontuações médias dos repositórios com as suas colocações, dividindo-os em 3 categorias de acordo com a faixa de pontuação calculada. Os repositórios que atingiram pontuação média acima de 5,64 pontos, foram considerados com Desempenho excelente, os repositórios com pontuação média entre 2,90 e 5,64 pontos foram considerados com Desempenho mediano e os repositórios com pontuação inferior a 2,90 pontos foram considerados com Desempenho inferior.

Nota-se que 4 repositórios, 21,05% dos repositórios participantes, obtiveram pontuações altas e foram classificados com Desempenho superior, ultrapassando os 5,64 pontos. Com Desempenho mediano, foram classificados 10 repositórios, 52,63% dos repositórios participantes, com pontuações acima de 2,90 pontos e abaixo de 5,64 pontos. A última porção de repositórios classificados com Desempenho inferior possui 5

repositórios, equivalente a 26,32% dos repositórios participantes, com pontuações inferiores a 2,90 pontos.

Quadro 43 – Classificação dos repositórios pelos seus desempenhos gerais.

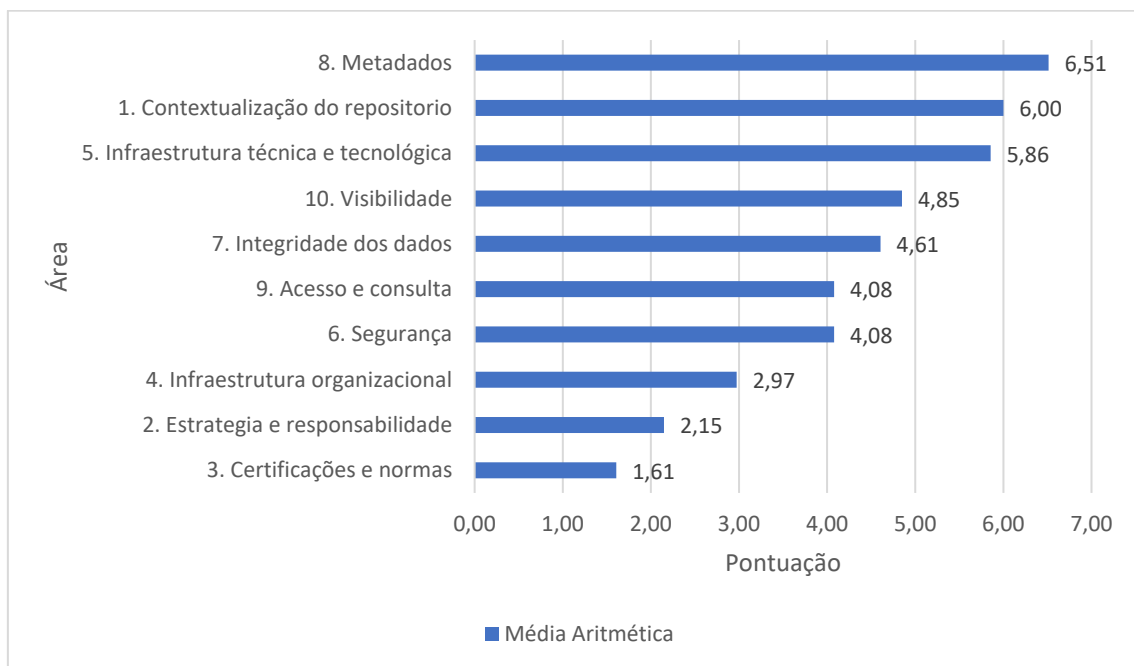
Desempenho excelente MA > 5,64 pontos	Desempenho mediano 2,90 pontos < MA < 5,64 pontos	Desempenho inferior MA < 2,90 pontos
1º) R14	5º) R11	15º) R04
2º) R17	6º) R08	16º) R19
3º) R05	7º) R09	17º) R18
4º) R13	8º) R03	18º) R02
	9º) R07	19º) R16
	10º) R01	
	11º) R12	
	12º) R06	
	13º) R10	
	14º) R15	
Total: 4	Total: 10	Total: 5

Fonte: elaboração própria.

Pode-se perceber um grande contraste ao se analisar a pontuação média do primeiro colocado, R14, com a do último colocado, R16, observando-se uma diferença de 4,90 pontos, equivalente a 273,74%. Esse fato demonstra uma alta variação das pontuações, significando uma grande disparidade de tratamento dos repositórios em relação aos aspectos avaliados no presente estudo.

5.3 RESULTADOS GERAIS POR ÁREAS

Nessa seção do trabalho são analisadas as pontuações dos repositórios em relação às áreas de avaliação. A Figura 21 exhibe o ranking das áreas de acordo as suas respectivas MA das pontuações dos repositórios. É possível notar que a área “8. Metadados” é a que possui maior pontuação, registrando 6,51 pontos, representando um alto desempenho dos repositórios nos critérios dessa área, que avalia o estabelecimento de padrões de metadados e a presença de metadados que permitam a identificação do contexto dos materiais preservados. Opostamente, a área “3. Certificações e normas” é a que apresenta menor MA, com 1,61 pontos, denotando um baixo grau de interesse ou conhecimento dos repositórios sobre as questões relacionadas ao alcance de certificações ou uso de normas que possam melhorar as suas confiabilidades ou padronizar os seus processos.

Figura 21 - *Ranking* das áreas pela Média Aritmética de pontuação.

Fonte: elaboração própria.

O Quadro 44 exibe a pontuação total individual de cada repositório participante em relação às áreas avaliadas, as Médias Aritméticas (MA) das pontuações dos repositórios, o Desvio Médio (DM) das pontuações dos repositórios em relação à MA de cada área e o Coeficiente de Variação (CV) das pontuações de cada área por repositório.

Ao se detalhar as pontuações dos repositórios, é possível identificar repositórios que não obtiveram pontuação em diferentes áreas, assim como o fato de apenas 4 áreas não registrarem pontuação zero para nenhum repositório: 1. Contextualização do repositório; 3. Certificações e normas; 5. Infraestrutura técnica e tecnológica, e; 10 Visibilidade. Em todas as demais áreas houve registros de pontuações zero, destacando-se a área “2. Estratégia e responsabilidade” como a área que mais obteve pontuações zero entre os repositórios, com 6 repositórios ao todo, equivalente a 31,58% dos repositórios participantes. Esse fato demonstra uma deficiência nos quesitos que tratam sobre o estabelecimento de instrumentos e políticas que regem os repositórios como a Política de Preservação e o Plano de Continuidade, por exemplo.

Quadro 44 – Pontuações dos repositórios por área.

Repositório/Área	R01	R02	R03	R04	R05	R06	R07	R08	R09	R10	R11	R12	R13	R14	R15	R16	R17	R18	R19
1. Contextualização do repositório	4,67	4,00	4,00	7,33	8,00	4,67	7,33	4,00	7,33	0,67	8,67	4,00	8,00	8,00	8,00	4,00	9,33	4,00	8,00
2. Estratégia e responsabilidade	6,67	0,83	1,67	1,67	2,50	4,17	0,00	3,33	5,83	2,50	0,00	0,00	5,00	1,67	0,00	0,00	3,33	1,67	0,00
3. Certificações e normas	2,78	0,56	1,11	0,56	1,11	0,56	1,11	1,11	1,11	1,67	1,11	2,78	4,44	1,11	0,56	2,78	4,44	1,11	0,56
4. Infraestrutura organizacional	3,50	2,00	4,50	0,00	3,50	3,00	4,00	3,00	4,00	1,00	7,50	0,50	4,00	6,50	0,00	2,00	6,50	1,00	0,00
5. Infraestrutura técnica e tecnológica	7,50	6,25	4,38	2,50	8,13	7,50	6,25	4,38	8,75	10,00	5,00	8,75	7,50	5,00	3,75	2,50	5,00	3,13	5,00
6. Segurança	4,17	0,00	1,67	0,00	6,67	3,33	6,67	1,67	0,00	6,67	6,67	7,50	8,33	8,33	3,33	0,00	5,00	0,83	6,67
7. Integridade dos dados	6,25	1,88	8,13	0,00	6,25	4,38	3,13	5,63	3,75	4,38	5,00	8,13	6,88	8,75	3,13	1,25	8,13	2,50	0,00
8. Metadados	7,50	5,00	10,00	5,00	10,00	0,00	7,50	10,00	7,50	3,75	10,00	2,50	10,00	10,00	7,50	2,50	10,00	2,50	2,50
9. Acesso e consulta	2,50	0,00	5,00	6,25	10,00	2,50	5,00	10,00	7,50	0,00	2,50	0,00	5,00	10,00	0,00	1,25	7,50	2,50	0,00
10. Visibilidade	0,50	0,50	7,50	4,13	8,50	6,00	6,25	8,00	3,50	4,00	7,50	3,50	4,25	7,50	5,00	1,63	7,50	4,25	2,13
Total	46,03	21,01	47,94	27,43	64,65	36,10	47,24	51,11	49,28	34,63	53,94	37,65	63,40	66,86	31,26	17,90	66,74	23,49	24,85
MA	4,60	2,10	4,79	2,74	6,47	3,61	4,72	5,11	4,93	3,46	5,39	3,77	6,34	6,69	3,13	1,79	6,67	2,35	2,48
DM	1,78	2,25	1,90	2,23	2,29	1,63	1,28	2,12	2,01	2,24	2,14	2,71	2,19	2,78	1,77	2,71	2,57	1,92	2,70
CV (%)	49,36	125,79	45,34	95,48	43,68	65,86	30,33	52,44	48,41	82,39	46,06	76,58	38,81	48,96	66,62	162,46	41,81	94,97	118,96

Fonte: elaboração própria.

Analisando-se o DM, nota-se que o repositório R19 apresenta o maior valor, 2,70 pontos, significando que as suas pontuações, em média, possuem um grande afastamento das médias das áreas. Por outro lado, o repositório R07 apresenta o menor DM, com 1,28 pontos, significando uma maior aproximação das suas pontuações em relação às médias das áreas.

Analisando-se o CV, é possível notar que o repositório R16 apresenta 162,46% de variação entre as pontuações das áreas em relação às suas respectivas médias, representando um alto grau de variabilidade que reflete um alto grau de discrepância de tratamento entre as questões avaliadas pelo presente estudo. Opostamente, o repositório R07 apresenta CV 30,33%, refletindo um tratamento mais uniforme entre as questões elaboradas no presente estudo.

O Quadro 45 sumariza os dados referentes às pontuações dos repositórios por área, exibindo as respectivas pontuações totais, a Média Aritmética, o Desvio Médio e o Coeficiente de Variação.

Quadro 45 – Áreas de avaliação e pontuações totais.

Área	Total	MA	DM	CV(%)
1. Contextualização do repositório	114,00	6,00	2,11	38,15
2. Estratégia e responsabilidade	40,83	2,15	1,70	95,24
3. Certificações e normas	30,56	1,61	0,97	75,10
4. Infraestrutura organizacional	56,50	2,97	1,82	74,69
5. Infraestrutura técnica e tecnológica	111,25	5,86	1,89	37,01
6. Segurança	77,50	4,08	2,72	73,78
7. Integridade dos dados	87,50	4,61	2,28	58,28
8. Metadados	123,75	6,51	2,98	50,81
9. Acesso e consulta	77,50	4,08	3,11	86,59
10. Visibilidade	92,13	4,85	2,12	50,94
Total	811,51	42,71	-	
Média	81,15	4,27	-	

Fonte: elaboração própria.

É possível notar que a área “9. Acesso e consulta” é a que apresenta maior DM, com média de 3,11 pontos de afastamento da MA da área, 4,08 pontos, enquanto a área “3. Certificações e normas” apresenta o menor DM, com média de 0,97 pontos de afastamento da MA da área, 1,61 pontos. Complementarmente, a área “3. Estratégia e responsabilidade” é que possui maior CV, com 95,24%, representando um alto grau de variabilidade das pontuações dos repositórios nessa área e se destacando como a área

que possui mais distorções de atendimento aos critérios avaliados entre os repositórios. O menor CV registrado foi o da área “5. Infraestrutura técnica e tecnológica”, com 37,01%, significando que os repositórios apresentam um tratamento mais uniforme nas questões avaliadas nessa área.

5.4 RESULTADOS ESPECÍFICOS POR ÁREAS

Nessa seção, são realizadas análises dos resultados dos repositórios em cada uma das 10 áreas pesquisadas, individualmente. Da mesma forma, dentro de cada área, as perguntas são analisadas individualmente, buscando compreender as características e comportamentos dos repositórios participantes.

O Quadro 46 exibe um resumo geral das pontuações e indicadores gerais das áreas, exibindo as Médias Aritméticas (MA), os Desvios Médios (DM), os Desvios Padrão (DP) e os Coeficientes de Variação (CV) em porcentagem e a Figura 22 exibe o gráfico relativo a esses dados, com o objetivo de facilitar as análises.

Quadro 46 – Pontuações e indicadores gerais das áreas avaliadas.

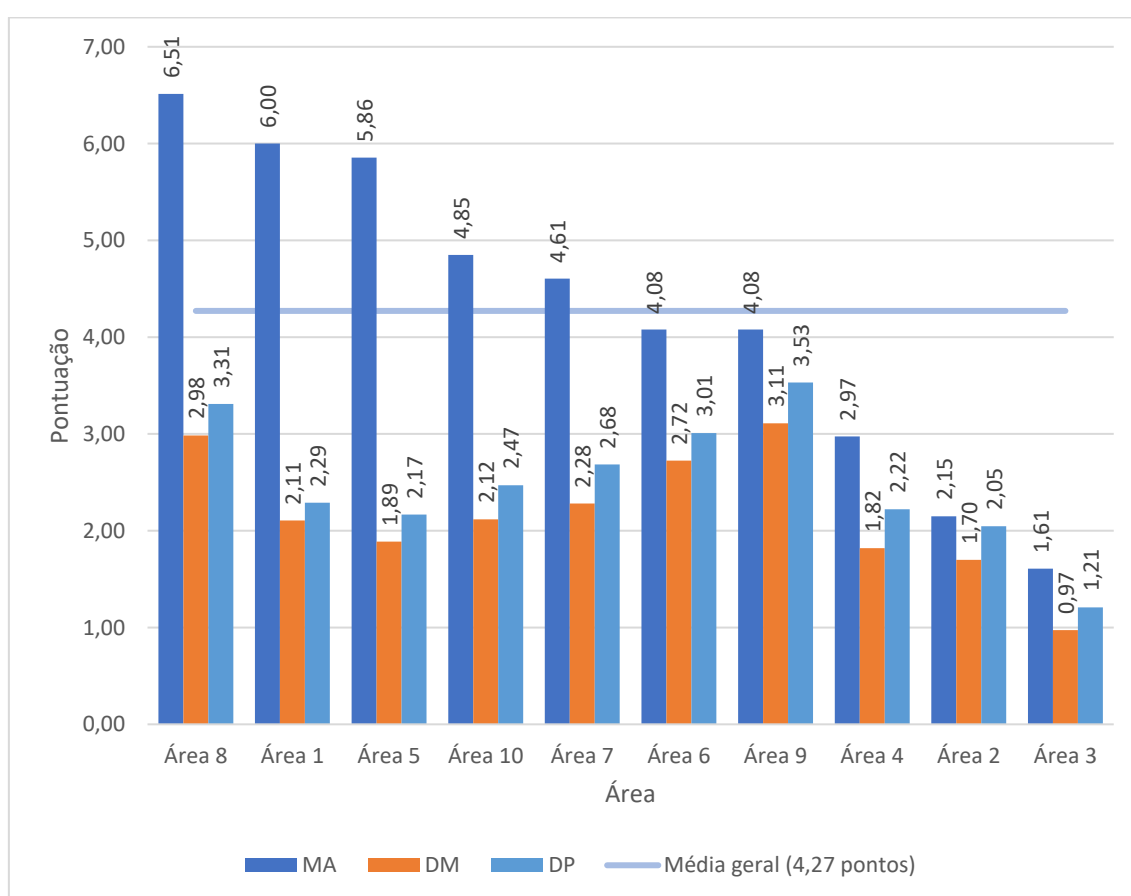
Área	MA	DM	DP	CV (%)
1. Contextualização do repositório	6,00	2,11	2,29	38,15%
2. Estratégia e responsabilidade	2,15	1,70	2,05	95,24%
3. Certificações e normas	1,61	0,97	1,21	75,10%
4. Infraestrutura organizacional	2,97	1,82	2,22	74,69%
5. Infraestrutura técnica e tecnológica	5,86	1,89	2,17	37,01%
6. Segurança	4,08	2,72	3,01	73,78%
7. Integridade dos dados	4,61	2,28	2,68	58,28%
8. Metadados	6,51	2,98	3,31	50,81%
9. Acesso e consulta	4,08	3,11	3,53	86,59%
10. Visibilidade	4,85	2,12	2,47	50,94%
Média geral	4,27			

Fonte: elaboração própria.

Destaca-se a área 8 (Metadados) com a maior MA geral, 6,51 pontos, caracterizando-se como um ponto forte entre os repositórios nas questões que tratam sobre o estabelecimento e uso de padrão de metadados e o recolhimento de metadados de contexto dos materiais digitais. Em seguida, a área 1 (Contextualização do repositório) apresenta a segunda maior MA, 6,00 pontos, indicando um bom desenvolvimento dos aspectos que tratam do estabelecimento das missões dos repositórios assim como a delimitação das suas Comunidades-Alvo.

A área 3 (Certificações e normas) apresentou a menor MA, 1,61 ponto, indicando uma necessidade que os repositórios devem trabalhar no sentido de adquirir e aprimorar conhecimentos sobre as certificações do universo da preservação digital, com a adoção de práticas de auditorias, por exemplo, assim como a adoção de normas formais nos processos arquivísticos dos repositórios. A área 2 (Estratégia e responsabilidade) apresenta a segunda menor MA, 2,15 pontos, indicando que os repositórios devem realizar esforços para o estabelecimento de planos que formalizam as suas políticas como um todo.

Figura 22 – Pontuações médias, Desvios Médios e Desvios Padrão das áreas avaliadas.



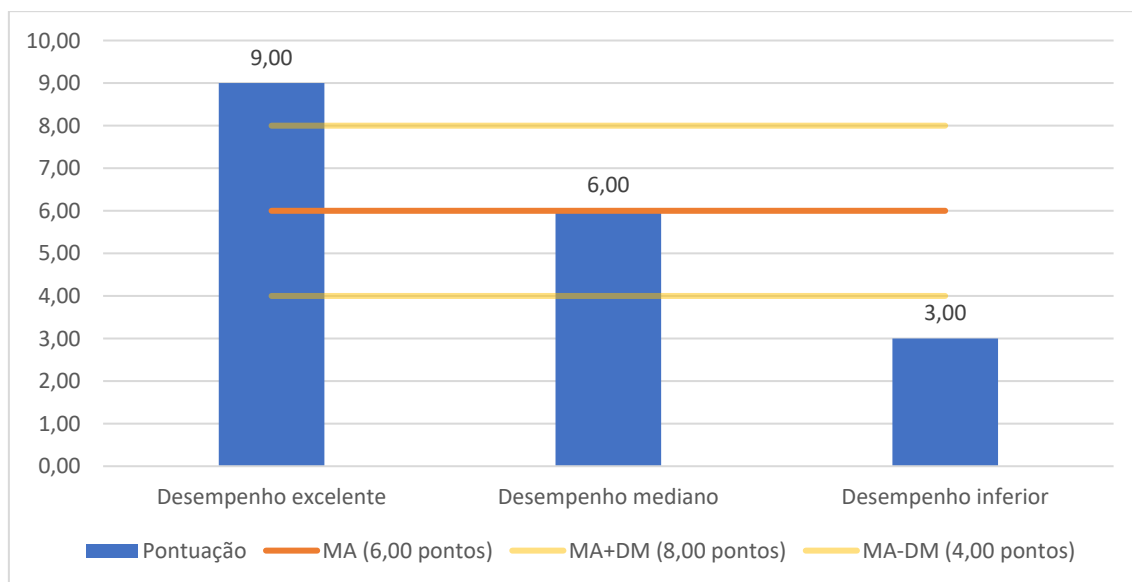
Fonte: elaboração própria.

Observando-se o Desvio Médio, é possível perceber que a área 9 (Acesso e consulta) possui o maior índice, 3,11 pontos, indicando um grande afastamento das pontuações dos repositórios em relação à MA da área, fato corroborado pelo maior DP, 3,53 pontos, registrado para a mesma área, o que indica uma alta variação das pontuações dos repositórios em relação à MA da área. O indicador CV registrou maior índice, 95,24%, na área 2, indicando que os repositórios possuem tratamentos muito

distintos entre si nessa área, que também registra a segunda menor MA das áreas. O menor CV foi registrado para a área 5, com 37,01% de variação entre as pontuações dos repositórios, que apresentou a terceira maior MA, o que permite concluir que os aspectos ligados às infraestruturas técnicas e tecnológicas dos repositórios apresentam um bom desempenho e pouca variabilidade entre si, se tornando um aspecto positivo.

Com o objetivo de criar uma classificação dos repositórios, em cada área foram criadas três faixas de classificação, de acordo com as pontuações obtidas em relação à Média Aritmética e o Desvio Médio. A Figura 23 exibe um exemplo de classificação dos repositórios com as faixas de classificação, com MA de 6,00 pontos e DM de 2,00 pontos. Os repositórios com pontuações maiores ou iguais a MA+DM (8,00 pontos) são classificados com “**Desempenho excelente**”, os repositórios com pontuações maiores ou iguais a MA-DM (4,00 pontos) e menores que MA+DM (8,00 pontos) são classificados com “**Desempenho mediano**” e os repositórios com pontuações menores que MA-DM (4,00 pontos) são classificados com “**Desempenho inferior**”. Em cada área, são adotadas essas faixas de classificação, porém observando a MA e DM individuais das áreas.

Figura 23 - Exemplo de classificação dos repositórios de acordo com as suas pontuações nas áreas.



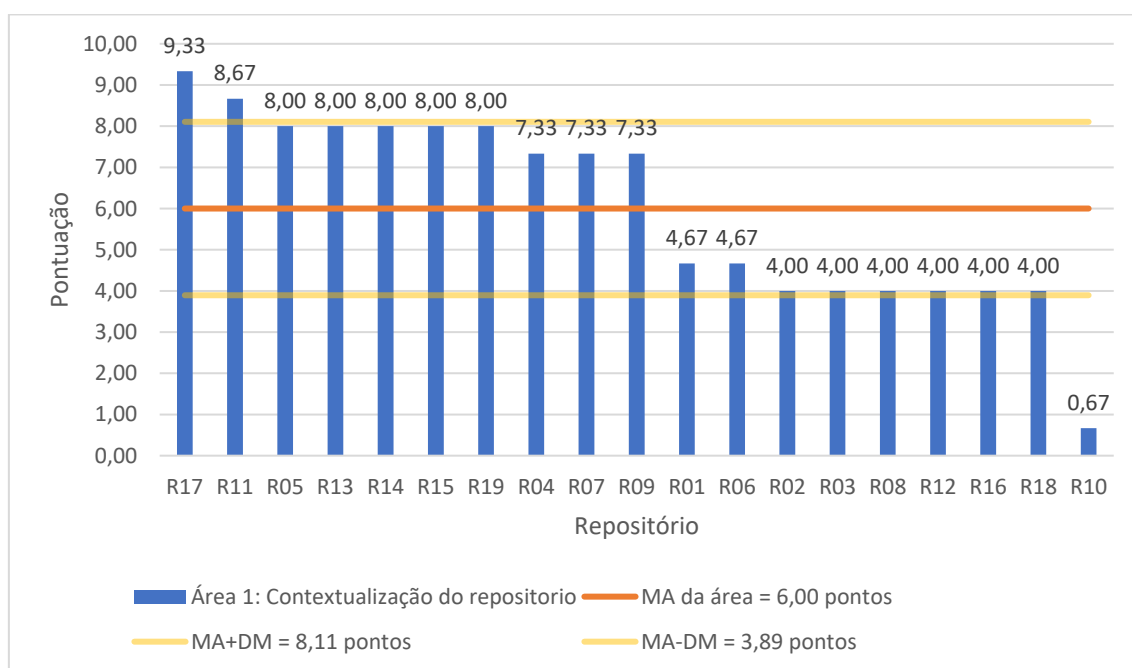
Fonte: elaboração própria.

5.4.1 Resultados da Área 1: Contextualização do Repositório

A seguir, são apresentados os resultados acerca da área 1: Contextualização do repositório, com base nas respostas fornecidas pelos respondentes. Essa

contextualização do repositório tem como objetivo compreender, de forma ampla, a sua natureza, através da declaração da sua missão, delimitação da sua comunidade-alvo e tipo de material preservado. A Figura 24 exibe as pontuações totais dos repositórios nas perguntas da área 1, que teve como Média Aritmética, 6,00 pontos, Desvio Médio de 2,11 pontos em relação à MA e Coeficiente de Variação de 38,15%, indicando uma grande variabilidade das pontuações. O repositório com maior pontuação, R17, obteve 9,33 pontos, enquanto o repositório com menor pontuação, R10, obteve 0,67 ponto, uma diferença de 8,66 pontos entre os dois extremos.

Figura 24 – Pontuações individuais dos repositórios na Área 1: Contextualização do repositório.



Fonte: elaboração própria.

O Quadro 47 exibe a classificação dos repositórios na área 1, de acordo com as suas faixas de pontuação. Pode-se notar que apenas 2 repositórios se classificaram com desempenho excelente, enquanto a grande maioria foi classificada com desempenho mediano. Apenas 1 repositório foi classificado com desempenho inferior, no entanto, a sua pontuação de 0,67 ponto, é 5,33 pontos distante da MA da área de 6,00 pontos, caracterizando a área como uma forte deficiência do repositório R10.

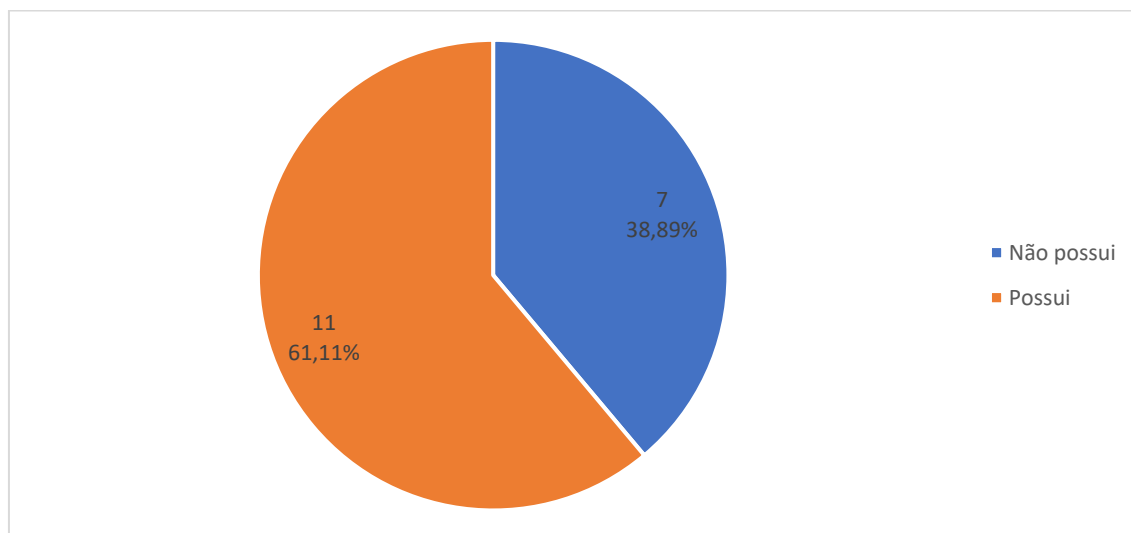
Quadro 47 – Classificação dos repositórios pelos seus desempenhos na Área 1: Contextualização do repositório

Desempenho excelente ($\geq 8,11$ pontos)	Desempenho mediano ($>3,89$ e $< 8,11$ pontos)	Desempenho inferior ($<3,89$ pontos)
1º) R17	3º) R05, R13, R14, R15, R19	7º) R10
2º) R11	4º) R04, R07, R09	
	5º) R01, R06	
	6º) R02, R03, R08, R12, R16, R18	
Total: 2	Total: 16	Total: 1

Fonte: elaboração própria.

A Figura 25 exibe o gráfico das respostas registradas para a pergunta “1.1 O repositório institucional possui missão, visão e valores elaborados?”. Dentre os 19 participantes, 1 repositório manifestou não responder à pergunta. Considerando os 18 repositórios restantes, 11 declararam possuir missão, visão e valores, representando 61,11%, enquanto em 7 repositórios, 38,89%, esses atributos não estão presentes. O estabelecimento da missão, visão e valores em um repositório demonstra importância ao se colocarem como uma bússola que guia a produção das suas políticas de preservação e acesso, além de demonstrar o seu comprometimento com a preservação digital confiável. A presença desse trio de informações na maioria dos repositórios respondentes se constata como um fator positivo nesse caso.

Figura 25 – Repositórios com missão, visão e valores declarados.



Fonte: elaboração própria.

De modo a ser fazer uma analogia, no guia para avaliação de repositórios REBIUN é possível encontrar a pergunta “2.1 - Existe uma declaração sobre a missão e objetivos do repositório?” (AZORÍN MILLARUELO, CÍVICO MARTÍN, *et al.*, 2015,

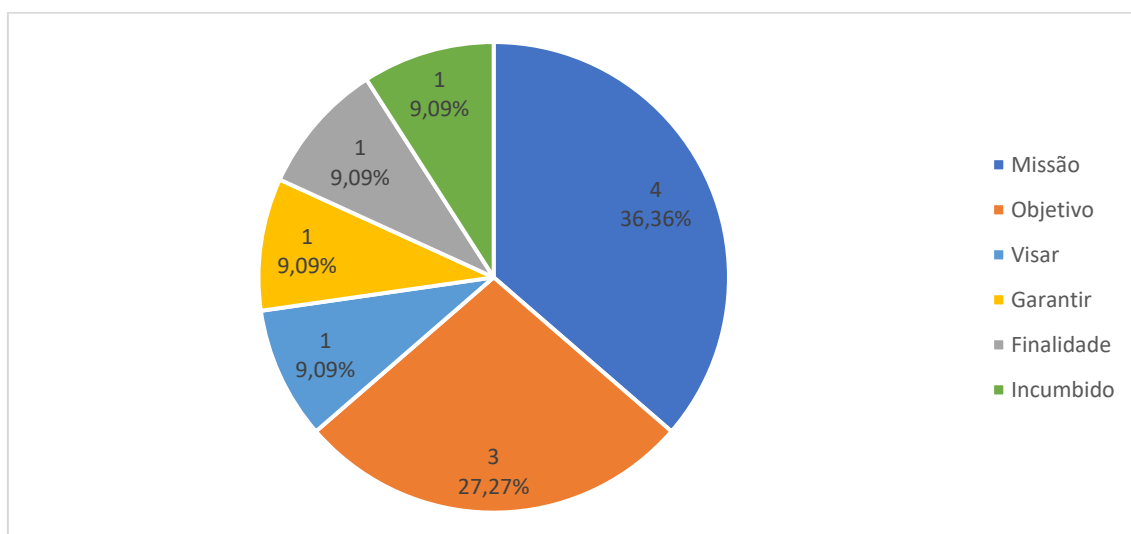
p. 13). Ainda que as perguntas não sejam idênticas, é possível a realização da comparação uma vez que ambas tratam da missão do repositório. No caso dos repositórios espanhóis, do total de 42 repositórios avaliados, somente 3 não possuem missão e objetivos elaborados, representando 7% do total. Nota-se que esse percentual é bastante diferente entre as duas realidades, identificando-se que em 39% dos repositórios brasileiros, não existem missão, visão ou valores declarados.

A pergunta 1.1 solicitou que o respondente informasse a missão, visão e valores do repositório, caso elas existam, sendo possível o preenchimento da URL onde essas informações podem ser extraídas. A análise das respostas identificou que nenhum dos repositórios possui declaradas as três informações na maneira formal como foi solicitada. Somente a missão ou outros termos com o mesmo significado foram encontrados nas respostas fornecidas.

A Figura 26 exibe a análise dos termos encontrados nas declarações dos repositórios utilizados para delimitar os seus respectivos propósitos. Do Total de 11 repositórios que manifestaram haver as respectivas missões declaradas, o termo “**Missão**” é encontrado em apenas 4 declarações, representando 36,36% das respostas, seguido pelo termo “**Objetivo**”, que pode ser encontrado em 3 declarações, o que representa 27,27% das respostas. Ainda que ambos termos não possuam exatamente o mesmo significado, pode-se assumir que eles denotam a finalidade dos repositórios, o que representaria que os 7 repositórios, 63,63% das respostas dessa pergunta e equivalente a 36,84% do total de 19 respondentes, possuem um norte para o qual guiam seus processos e atividades de preservação.

Os demais termos encontrados, em conjunto, totalizam 36,37% das respostas. Mesmo com a disparidade de termos usados, as respostas fornecidas pelos 11 repositórios contêm o termo “**preservação**” ou o verbo “**preservar**”, tornando clara a preocupação dos órgãos responsáveis pelos repositórios em deixar essa informação explícita.

Figura 26 – Utilização de termos para a definição da missão do repositório.



Fonte: elaboração própria.

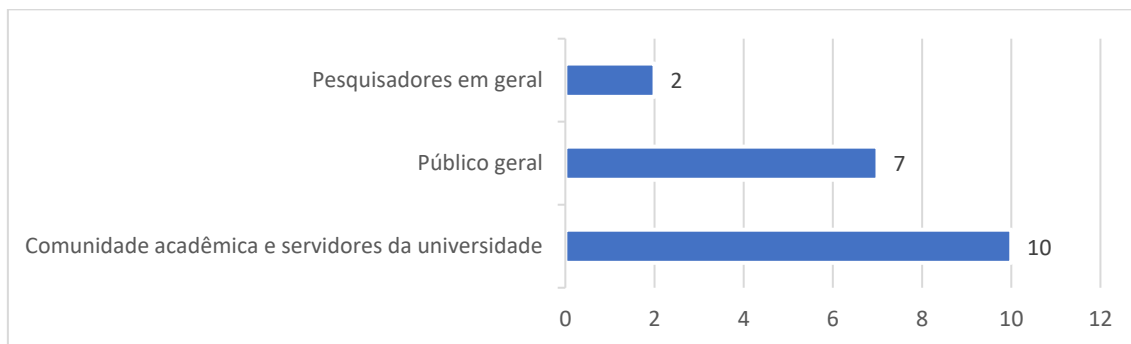
De um modo geral, as certificações e auditorias de confiança não exigem a declaração formal de missão com essa nomenclatura, porém deve ser possível encontrar um documento que manifeste os princípios estabelecidos que norteiam as suas ações e expressam o seu compromisso com a preservação a longo prazo. Ainda que nem todos os repositórios possuam missão, visão e valores explicitamente declarados, conclui-se que em 11 repositórios, equivalente a 58% do total de 19 repositórios, possuem alguma espécie de lema que inclui o ato de preservação.

Na pergunta 1.2, foi analisada a presença de uma definição da Comunidade-Alvo dos repositórios, através do enunciado “O repositório institucional possui a sua Comunidade-Alvo delimitada?”. Dentre as instituições respondentes, 79%, equivalente a 15 repositórios, declararam possuir uma delimitação das suas respectivas comunidades-alvo, enquanto 21%, equivalente a 4 repositórios declararam não haver essa definição sobre o perfil de público usuário do repositório com as suas possíveis especificidades. Foram registradas respostas de todos os repositórios participantes. O fato de a grande maioria possuir a sua Comunidade-Alvo delimitada é um fator positivo, uma vez que os repositórios podem conhecer e se adaptar às demandas do público ao qual servem.

A Figura 27 exhibe o quantitativo das respostas à pergunta “2.2.1 Em caso positivo qual a Comunidade-Alvo do repositório institucional?”, que solicitou que o respondente preenchesse de forma livre qual a Comunidade-Alvo do seu repositório, sendo registradas as respostas dos 15 repositórios que declararam haver essa

delimitação, conforme relatado na pergunta anterior 2.2. As respostas fornecidas foram processadas e agrupadas pelos termos de acordo as suas similaridades, sendo agrupadas em 3 itens, totalizando 19 respostas ao todo.

Figura 27 – Delimitação da comunidade-alvo dos repositórios.



Fonte: elaboração própria.

No caso de 10 repositórios, foi relatada a comunidade como sendo a “Comunidade acadêmica e servidores de universidade”, representando a grande maioria das respostas. Em 7 repositórios, foi relatado o “Público geral” como sendo a sua comunidade. No caso de 2 repositórios, foram relatados os “Pesquisadores em geral”, diferenciando-se da comunidade acadêmica por não se tratarem de universidades. Particularmente, 4 repositórios, relataram atender a “Comunidade acadêmica e servidores da universidade” e o “Público geral”, tendo sido os resultados computados em ambos os termos.

Como diversos repositórios pertencem a instituições acadêmicas, prevaleceu nas respostas a delimitação da própria comunidade acadêmica e de servidores como sendo a Comunidade-Alvo do repositório. O público geral foi citado em 7 repositórios, sinalizando que qualquer pessoa pode utilizar os serviços do repositório e sua plataforma disponibilizada.

Cabe notar que o entendimento sobre a definição de Comunidade-Alvo por parte dos respondentes se refere muito mais aos usuários que acessam os materiais preservados, os consumidores, do que aos usuários que depositam materiais, os produtores. Por mais que 7 instituições tenham relatado o público geral como parte da sua Comunidade-Alvo, é sabido que a ingestão de materiais, nesses casos, não está disponível ao público.

A pergunta 1.3 solicitou que os respondentes fornecessem as particularidades observadas na comunidade-alvo do seu repositório, buscando

compreender se são oferecidos serviços especializados a determinados usuários ou se são disponibilizados formatos de arquivo não convencionais para atender à demanda da sua Comunidade-Alvo. Nenhuma resposta foi registrada para essa pergunta, concluindo-se que as definições das Comunidades-Alvo, quando respondidas, são suficientes para se compreender o tipo de serviço prestado.

Essa análise traz à tona o fato de que, no caso das universidades, as demandas entre as diferentes instituições é uniforme, ou pelo menos, se acredita ser, uma vez que não foram relatadas particularidades. Esse fato pode indicar que uma iniciativa de um repositório de uma universidade como, por exemplo, a disponibilização de novas funcionalidades ou formatos de arquivos, poderia ser aplicada às outras de maneira muito aproximada, uma vez que a natureza das suas comunidades são as mesmas.

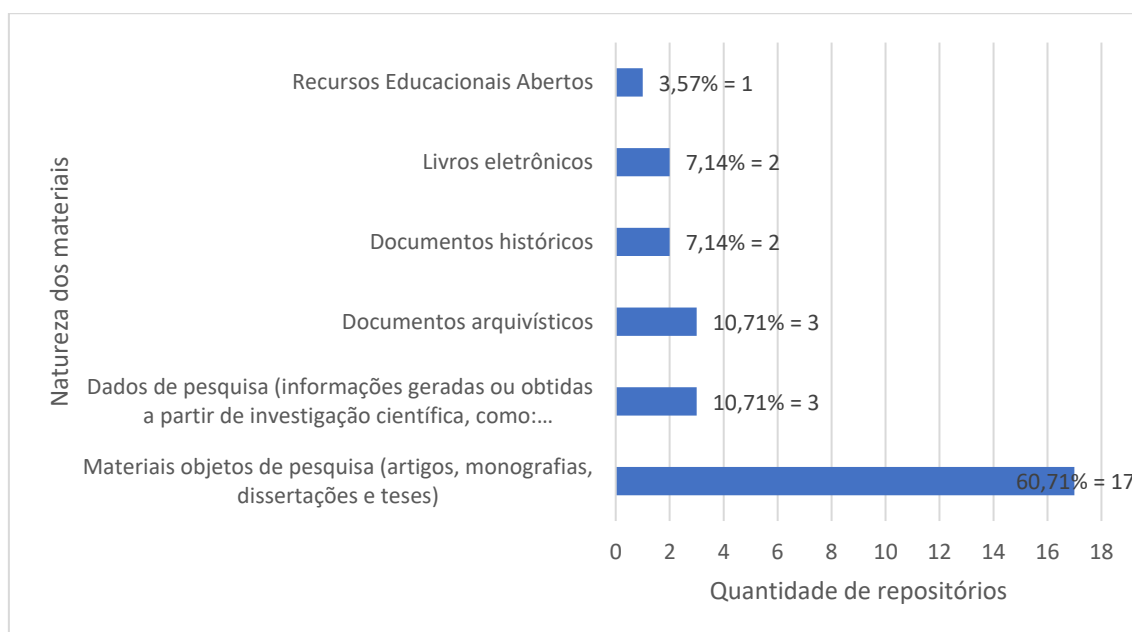
A pergunta “1.4 Qual a natureza do material depositado no repositório?” permite compreender as necessidades da comunidade-alvo, ao mesmo tempo em que complementa a pergunta anterior, que solicitou a definição da comunidade-alvo. Foi registrado o preenchimento de respostas para todos os 19 repositórios que participaram da pesquisa, inclusive os quatro repositórios que declararam não possuir uma comunidade-alvo definida assinalaram sobre a natureza do material depositado nos seus repositórios.

As respostas à pergunta 1.4, que permite ao respondente selecionar mais de uma resposta, de acordo com o que mais se adequasse à realidade do repositório, são exibidas na Figura 28. Ao todo, foram informadas 28 respostas, dentre essas, 17 repositórios, equivalente a 60,71%, declararam custodiar materiais objetos de pesquisa, tais como, artigos, monografias, dissertações e teses. A custódia de dados de pesquisa, que podem ser interpretados como informações-base sobre as quais as pesquisas são realizadas, foi declarada por 3 repositórios, equivalente a 10,71% dos repositórios, mesma quantidade de repositórios que declarou custodiar documentos arquivísticos. A custódia de documentos históricos foi sinalizada por 2 repositórios, equivalente a 7,14%, mesmo percentual de repositórios que declararam custodiar livros eletrônicos. A custódia de Recursos Educacionais Abertos (REA)⁹⁵ foi sinalizada por 1 repositório.

⁹⁵ Recursos Educacionais Abertos (*Open Educational Resources*) são “os materiais de ensino, aprendizagem e investigação em quaisquer suportes, digitais ou outros, que se situem no domínio público ou que tenham sido divulgados sob licença aberta que permite acesso, uso, adaptação e redistribuição gratuitos por terceiros, mediante nenhuma restrição ou poucas restrições” (UNESCO, 2012).

Mesmo que não haja uma definição formal estabelecida para a comunidade-alvo de certos repositórios, as respostas à essa pergunta permitem conhecer o público que interage com o repositório, assim como os materiais utilizados na ingestão e difusão de documentos. Um total de 20 respostas, equivalente a 71,43% sinaliza que os repositórios custodiam materiais e dados de pesquisa, corroborando a informação de que os repositórios, em sua maioria, possuem a comunidade acadêmica como sendo as suas comunidades-alvo.

Figura 28 – Natureza dos materiais custodiados pelos repositórios.



Fonte: elaboração própria.

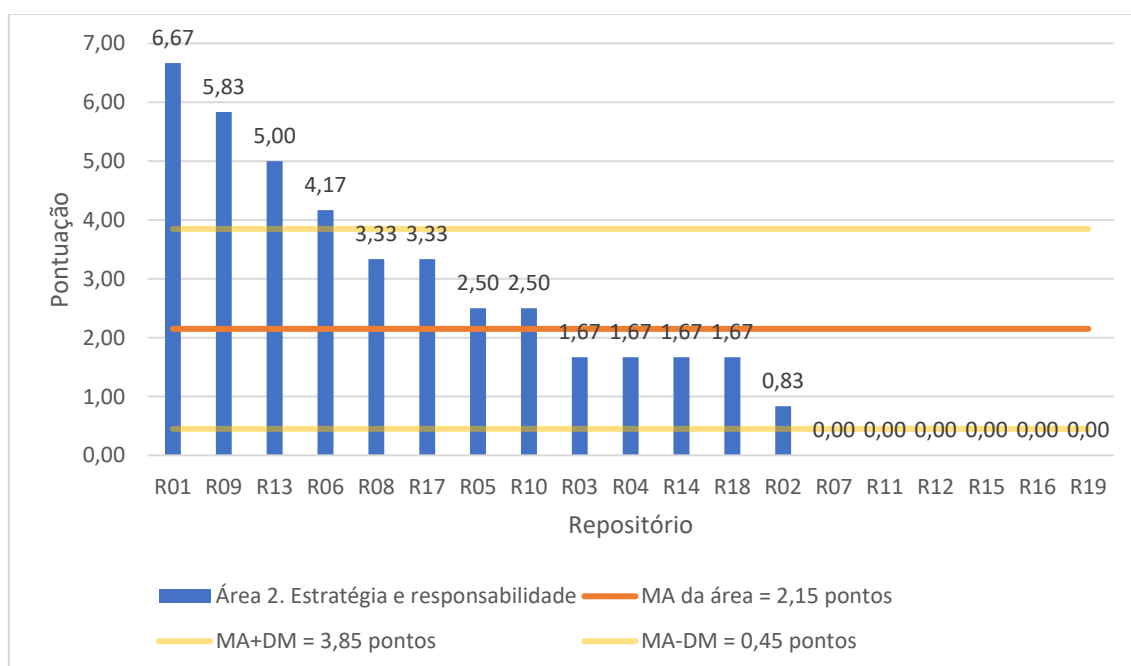
A partir das análises, pode-se perceber um desempenho relativamente bom dos repositórios, uma vez que 16, 84,21% apresentaram um desempenho mediano, tendo 2 repositórios apresentado desempenho excelente, significando que os repositórios possuem pontos fortes nessa área.

5.4.2 Resultados da Área 2: Estratégia e Responsabilidade

Nessa seção, são analisadas as repostas das perguntas da área 2: Estratégia e Responsabilidade. Os instrumentos elaborados e adotados pelos repositórios são essenciais na formalização dos seus respectivos comprometimentos, norteando as estratégias e ações a serem realizadas para que se alcance a plena preservação a longo prazo dos seus conteúdos custodiados.

A Figura 29 exhibe as pontuações totais dos repositórios nas perguntas da área 2, que teve como Média Aritmética, 2,15 pontos, Desvio Médio de 1,70 ponto em relação à MA. O Coeficiente de Variação calculado foi de 95,24%, o maior entre todas as áreas avaliadas, indicando uma altíssima variabilidade das pontuações entre os repositórios. O repositório com maior pontuação, R01, obteve 6,67 pontos, enquanto a menor pontuação, 0,00 ponto, foi registrada para os repositórios R07, R11, R12, R15, R16 e R19, uma diferença de 6,67 pontos entre os dois extremos. A área 2 foi a que mais registrou pontuações zero dentre todas as áreas avaliadas, caracterizando-a como uma forte deficiência nos repositórios e sinalizando recomendando-se uma priorização de tratamento e atendimento aos critérios avaliados nessa área, que trata do estabelecimento macro das responsabilidades dos repositórios.

Figura 29 – Pontuações individuais dos repositórios na Área 2: Estratégia e Responsabilidade.



Fonte: elaboração própria.

O Quadro 48 exhibe a classificação dos repositórios na área 2, de acordo com as suas faixas de pontuação. Pode-se notar que 4 repositórios se classificaram com desempenho excelente, enquanto 9 repositórios foram classificados com desempenho mediano. Um total de 6 repositórios foi classificado com desempenho inferior, sendo esta área de avaliação a que apresentou a segunda menor média geral entre todas as 10 áreas.

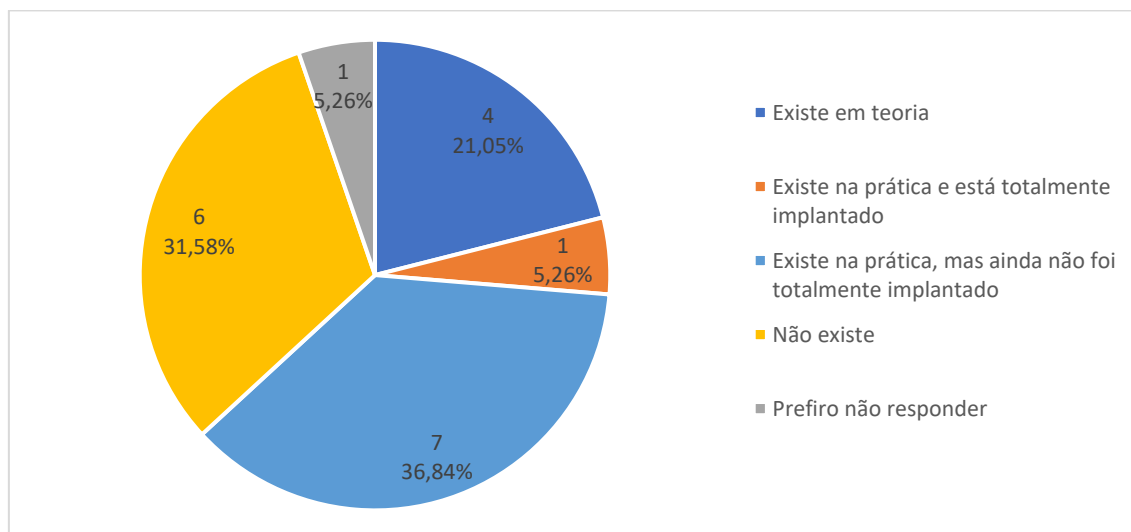
Quadro 48 – Classificação dos repositórios pelos seus desempenhos na Área 2: Estratégia e Responsabilidade.

Desempenho excelente ($\geq 3,85$ pontos)	Desempenho mediano ($\geq 0,45$ e $< 3,85$ pontos)	Desempenho inferior ($< 0,45$ ponto)
1º) R01	5º) R08 e R17	9º) R07, R11, R12, R15, R16 e R19
2º) R09	6º) R05 e R10	
3º) R13	7º) R03, R04, R14 e R18	
4º) R06	8º) R02	
Total: 4	Total: 9	Total: 6

Fonte: elaboração própria.

A Figura 30 exibe as repostas fornecidas para a pergunta “2.1) Existe uma Política de Preservação com os requisitos mínimos dos documentos recebidos, tratamento de preservação e acesso?”. Do total de 19 questionários, 1 preferiu não responder à questão, equivalente a 5,26%, enquanto 6 repositórios relataram não possuir política de preservação, equivalente a 31,58%.

Figura 30 – Existência de Política de Preservação nos repositórios.



Fonte: elaboração própria.

Apenas 1 repositório, 5,26%, relatou a existência e a aplicação total na prática da sua política de preservação, enquanto 4 repositórios, 21,05%, declararam existir a Política de Preservação somente em teoria, ou seja, a prática do dia-a-dia desses repositórios não se alinha aos processos, atividades e ações desenhados na Política de Preservação. Um total de 7 repositórios, 36,84%, possuem a Política de Preservação estabelecida, porém não a aplicam em sua totalidade.

A pergunta “2.1.1 Caso não exista uma Política de Preservação, existe um Plano Estratégico de Preservação?” complementa a sua anterior, no sentido de

compreender quais instrumentos são usados nos repositórios para delimitar as suas funções e definir as estratégias utilizadas para a preservação a longo prazo.

Analisando-se a duas perguntas em conjunto, 1 repositório que declarou não possuir uma Política de Preservação, declarou existir em teoria um Plano Estratégico de Preservação. Outro repositório que declarou a existência da Política de Preservação, porém com aplicação parcial na prática, declarou que o Plano Estratégico de Preservação se encontra em fase de elaboração. Assim, dos 6 repositórios que não possuem Política de Preservação, 5 também não possuem um Plano Estratégico de Preservação. De todos os repositórios pesquisados, 26% não possuem esses documentos que estabelecem os princípios e objetivos do repositório, critérios de ingestão, curadoria e difusão de materiais. Portanto, torna-se difícil identificar os seus comprometimentos com a preservação a longo prazo e todas as responsabilidades exigidas.

O guia de avaliação de repositórios REBIUN possui o critério “2.2 Documento de acesso público sobre o arquivo no repositório, onde se estabelecem ao menos os seguintes pontos: quem pode depositar, o que se pode depositar e em quais formatos” (AZORÍN MILLARUELO, CÍVICO MARTÍN, *et al.*, 2015, p. 13, tradução nossa), que analisa se o repositório especifica formalmente quem está autorizado a depositar materiais, quais os gêneros aceitos e quais formatos esperados, buscando transparecer à sua comunidade os seus critérios mínimos exigidos. Outro critério existente no guia REBIUN é o “2.3 Documento de acesso público sobre a preservação dos conteúdos” (AZORÍN MILLARUELO, CÍVICO MARTÍN, *et al.*, 2015, p. 13, tradução nossa), que expressa o compromisso do repositório com a preservação, especificando quais ações realiza para que ela de fato aconteça, citando, inclusive, as estratégias de preservação utilizadas.

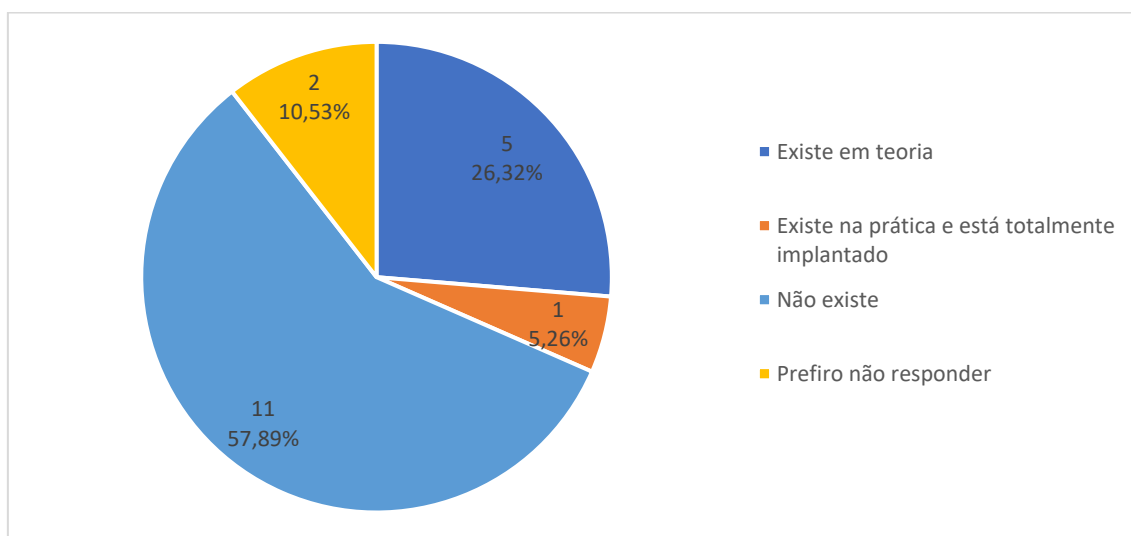
Na última avaliação realizada nos repositórios espanhóis, registrou-se que 86% dos repositórios cumpriam o critério 2.2 enquanto que no critério 2.3, somente 5% dos repositórios atendiam. Pode-se perceber uma grande discrepância entre os resultados desses critérios, permitindo identificar que nos repositórios espanhóis, existe a definição sobre como e quem depositar os materiais, porém não se tem conhecimento dos seus comprometimentos com a preservação em si.

Em outro estudo realizado por Azorín *et al.*, avaliou-se os repositórios REBIUN em relação à matriz de preservação NDSA, buscando compreender qual o grau de cumprimento dos repositórios em cada um dos níveis. Como resultados, identificou-se que somente 2 repositórios, dos 52 participantes, 3,8%, possuem planos

de preservação aprovados e em funcionamento, de onde concluem “as técnicas de preservação digital estão sendo aplicadas de forma incipiente e parcial na maioria dos repositórios” (2018, p. 6, tradução nossa). Pode-se notar um certo grau de coincidência entre os estudos, uma vez que os resultados do guia de avaliação de repositórios REBIUN identificou que a grande maioria não possui um documento público que demonstra como a preservação é realizada efetivamente.

No caso dos repositórios brasileiros, a política de preservação pode unir essas informações em único documento, porém, essa política existe e é aplicada na prática em apenas 1 repositório, representando 5,26% da amostra e em 7 repositórios, 36,84%, a política existe, mas não é aplicada totalmente na prática. Em ambos os casos, esse fato se mostra preocupante, pois a ausência dessas informações, materializadas em documentos formais ou informais, pode comprometer a confiança nos repositórios por parte das suas Comunidades-Alvo.

Figura 31 – Presença de Plano de Continuidade nos repositórios.



Fonte: elaboração própria.

A pergunta 2.2 analisou a existência de um Plano de Continuidade nos repositórios, tendo das suas respostas exibidas na Figura 31. Das respostas recebidas, 2 repositórios, 10,53%, preferiram não responder à pergunta, 5 repositórios, 26,32%, possuem o Plano de Continuidade somente em teoria, sem aplicação prática. Apenas 1 repositório, 5,26%, possui o Plano de Continuidade com aplicação total na prática. A maioria dos repositórios, 11, equivalente a 57,89%, declarou não possuir o Plano de Continuidade.

A ausência de um documento que preveja as ações a serem executadas na ocorrência de eventos que causem a interrupção temporária ou definitiva das atividades na maioria dos repositórios traz preocupações a respeito da capacidade da realização efetiva da preservação a longo prazo. Dentre as ameaças à continuidade dos repositórios podem-se citar falhas tecnológicas, ataques por invasões ou a cessação das atividades por mudança de interesses da administração superior, por exemplo, fazendo com que seja necessário o mapeamento dessas ameaças e a criação de estratégias que agreguem características de resiliência aos repositórios.

A Lei nº 8.159, de 8 de janeiro de 1991, que dispõe sobre a política nacional de arquivos públicos e privados, conhecida popularmente como a Lei dos Arquivos, em seu artigo 18º, estabelece (BRASIL, 1991):

Compete ao Arquivo Nacional a gestão e o recolhimento dos documentos produzidos e recebidos pelo Poder Executivo Federal, bem como preservar e facultar o acesso aos documentos sob sua guarda, e acompanhar e implementar a política nacional de arquivos.

Nesse âmbito, o artigo estabelece a competência do Arquivo Nacional em relação à gestão e recolhimento dos documentos da Poder Executivo Federal, compreendendo tanto os documentos físicos quanto os digitais desses órgãos. Ao que se conclui que cabe a esse órgão a responsabilidade de guarda permanente dos documentos produzidos na administração federal, sendo o destino final para o qual os órgãos subordinados devem enviar os seus documentos. Logo, em uma eventual cessão de atividades dos repositórios, o Arquivo Nacional seria o responsável por prosseguir com a preservação e acesso de todos os seus patrimônios documentais permanentes.

Assim, recai sobre os repositórios a elaboração de planos de continuidade que tenham como sucessor o Arquivo Nacional, seja através da sincronização periódica de seus acervos ou através da transferência dos acervos na ocorrência de eventos que afetem as suas continuidades. Essa abordagem traz ao Arquivo Nacional a responsabilidade por se manter operante e possuir ele mesmo um plano de continuidade de preservação, uma vez que o órgão deve preservar os seus próprios documentos digitais, assim como pode receber os acervos digitais dos demais órgãos da administração federal.

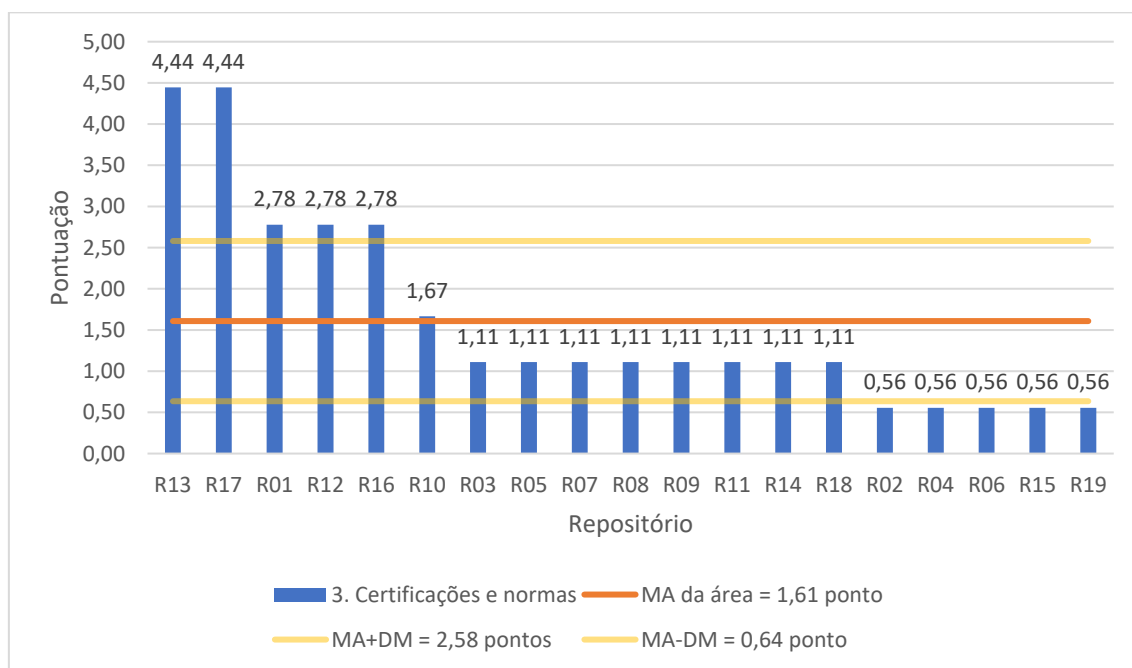
As análises dessa área 2 permitem identificar uma grande deficiência de instrumentos e políticas que definem os critérios, requisitos da preservação digital e

planos de contingência, necessitando ser uma área trabalhada com prioridade nos repositórios.

5.4.3 Resultados da Área 3: Certificações e Normas

Nessa seção, são analisados os resultados das perguntas da Área 3: Certificações e Normas do formulário, tratando das certificações de interesse do repositório além das normas por ele adotadas. A Figura 32 exibe as pontuações totais dos repositórios nas perguntas da área 3, que teve como Média Aritmética, 1,61 ponto, Desvio Médio de 0,97 ponto em relação à MA e Coeficiente de Variação de 75,10%, indicando uma altíssima variabilidade das pontuações. Os repositórios R13 e R17 obtiveram a maior pontuação, com 4,44 pontos, enquanto a menor pontuação, 0,56 ponto, foi registrada para os repositórios R02, R04, R06, R15 e R19, uma diferença de 3,88 pontos entre os dois extremos.

Figura 32 – Pontuações individuais dos repositórios na Área 3: Certificações e Normas.



Fonte: elaboração própria.

O Quadro 49 exibe a classificação dos repositórios na área 3, de acordo com as suas faixas de pontuação. Pode-se notar que 5 repositórios se classificaram com desempenho excelente, enquanto 9 repositórios foram classificados com desempenho

mediano. Um total de 5 repositórios foi classificado com desempenho inferior, sendo esta área de avaliação a que apresentou a menor média geral entre todas as 10 áreas.

Quadro 49 – Classificação dos repositórios pelos seus desempenhos na Área 3: Certificações e Normas.

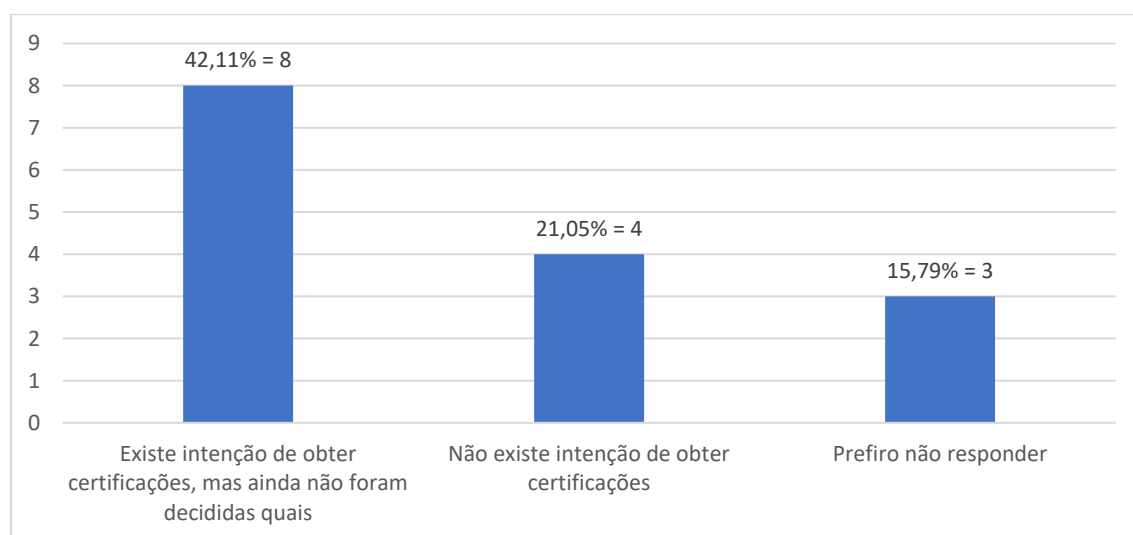
Desempenho excelente ($\geq 2,58$ pontos)	Desempenho mediano ($\geq 0,64$ e $< 2,58$ pontos)	Desempenho inferior ($< 0,64$ ponto)
1º) R13 e R17	3º) R10	5º) R02, R04, R06, R15 e R19
2º) R01, R12 e R16	4º) R03, R05, R07, R08, R09, R11, R14 e R18	
Total: 5	Total: 9	Total: 5

Fonte: elaboração própria.

A pergunta 3.1 solicitou ao respondente a informação a respeito da existência de certificações já obtidas pelos respectivos repositórios. Do total de 19 repositórios, 18 declararam não possuir nenhuma certificação, enquanto 1 repositório apontou a existência de um certificado obtido em 2012, porém não relacionado à verificação de critérios de confiança do repositório.

As repostas à pergunta “3.2 Quais certificações o repositório tem a intenção de alcançar?”, do tipo múltipla escolha, que permite mais de uma resposta, são analisadas na Figura 33.

Figura 33 – Interesse dos repositórios em obter certificações.



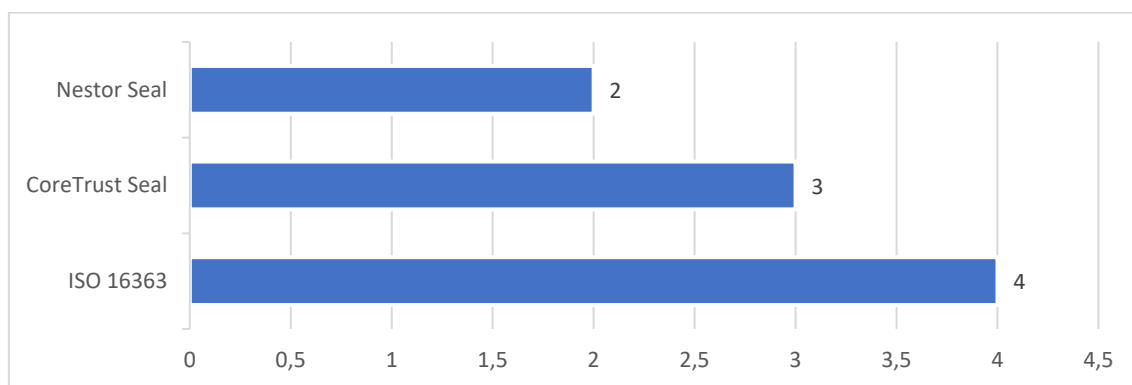
Fonte: elaboração própria.

Em 8 dos 19 repositórios participantes da pesquisa, o equivalente a 42,11%, existe a intenção de obtenção de uma certificação, porém não há a definição de quais seriam. Dentre as prováveis causas dessa indefinição podem estar o desconhecimento

sobre o universo de certificações ou ainda a falta de estudos internos direcionados a esse tema. Para 4 repositórios, o equivalente a 21,05%, não existe nenhuma intenção na obtenção de certificações, enquanto 3 repositórios preferiram não responder à pergunta.

Os 4 repositórios restantes, equivalente a 21,05% dos participantes, demonstraram interesse em obter as certificações apresentadas na Figura 34. Todos os 4 repositórios, equivalente a 45% dos tipos de certificações, manifestaram que têm interesse em obter a certificação ISO 16363, enquanto 3 repositórios, 33%, deseja obter a certificação CoreTrust Seal e 2 repositórios, 22%, demonstraram interesse na certificação NestorSeal.

Figura 34 – Tipos de certificações de interesse dos repositórios.



Fonte: elaboração própria.

Chama a atenção o fato de a quantidade de repositórios que possui definidas as certificações que deseja alcançar ser a mesma quantidade de repositórios que não deseja obter nenhuma certificação. Ao mesmo tempo, para 8 repositórios existe o interesse, porém não se sabe ao certo quais certificações serão escolhidas para obtenção.

Essas evidências podem revelar um alto grau de desconhecimento das certificações ou pouco interesse em adequação a normas que são usadas e validadas mundialmente. Essa situação pode vir a impactar os repositórios futuramente quando for requerida alguma certificação para a contratação de novos recursos ou pessoas para as suas equipes. Se esse for o caso, os repositórios terão que buscar conhecimentos, assessorias e promover mudanças, muitas vezes estruturais, sob a pressão de não conseguirem se adaptar às evoluções inevitáveis que a preservação digital a longo prazo impõe. Sob outra ótica, o baixo interesse nas certificações pode revelar que já existem lacunas que não permitem que os repositórios dediquem recursos e aloquem pessoas

para o estudo das certificações, uma vez que as demandas atuais já sobrecarregam as estruturas atualmente disponíveis.

A pergunta “3.3 O repositório adota padrões e normas estabelecidos?” teve como objetivo conhecer os padrões e normas adotados pelos repositórios. Dentre o rol de respostas, foram fornecidas opções de padrões de metadados e empacotamento de metadados, descrição arquivística e a adoção do repositório ao padrão ISO 14721 (OAIS)⁹⁶. A pergunta, do tipo múltipla escolha, permitiu aos respondentes a escolha de mais de uma opção, além de permitir a complementação, através de um campo de texto livre, de outras normas adotadas. Ao todo foram recebidas 28 respostas e os resultados são exibidos na Figura 35.

O padrão de metadados Dublin Core⁹⁷ foi o mais citado entre as respostas, ocupando um total de 94,74% sendo respondido por 18 repositórios. Pode-se notar que quase a totalidade dos repositórios utiliza o padrão Dublin Core para a descrição dos seus materiais, tornando-se uma característica positiva aos repositórios, uma vez que ele se tornou um padrão “*de facto*”, quando é adotado vastamente por diversas iniciativas ao redor do mundo. Posteriormente, o conjunto de termos Dublin Core foi tomado como base para a criação do padrão estabelecido ISO 15836:2017 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2017), o que mostra um alto grau de amadurecimento, estabilidade e adoção por diferentes soluções de descrição arquivística.

Cabe notar que o padrão Dublin Core possui somente elementos para o registro de metadados descritivos dos materiais digitais, não possuindo elementos destinados ao registro completo do contexto ou história do ciclo de vida dos materiais, possuindo para esse fim o elemento *provenance* definido como “Uma declaração de quaisquer alterações na propriedade e custódia do recurso desde sua criação que sejam significativas para sua autenticidade, integridade e interpretação” (DUBLIN CORE METADATA INITIATIVE, 2020).

Em seguida, com 15,79% dos repositórios, o padrão METS⁹⁸ foi citado por 3 repositórios. Com isso, pode-se inferir que a sua utilização é devida à necessidade de haver um certo grau de integração entre as ferramentas utilizadas no repositório, uma vez que esse padrão é principalmente destinado ao empacotamento de metadados.

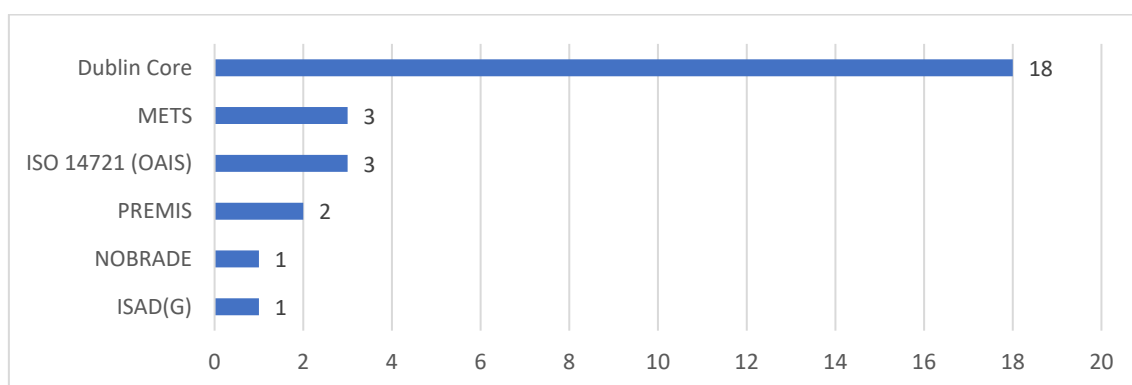
⁹⁶ (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012)

⁹⁷ (DUBLIN CORE METADATA INITIATIVE, 2020)

⁹⁸ (METS BOARD, 2010)

O padrão ISO 14721 (OAIS) foi citado por 3 repositórios, representando 15,79% dos repositórios da pesquisa. Uma vez que a norma cobre desde a estrutura organizacional do repositório, existência de políticas e planos de preservação, até a estrutura técnica/tecnológica e a que as respostas recebidas nas perguntas anteriores foram negativas em sua maioria sobre esses aspectos, conclui-se que a compreensão sobre a adoção do padrão OAIS se referiu às ferramentas de preservação utilizadas. A partir das respostas, nesses 3 repositórios, as respectivas ferramentas adotadas para a preservação de documentos seguem o padrão estabelecido, adotando os pacotes de informação para a ingestão, arquivamento e difusão, por exemplo, além das demais definições inerentes ao padrão. Ao mesmo tempo, pode-se observar que os outros 16 repositórios participantes da pesquisa, equivalente a 84,21%, não adotam esse padrão, o que se torna um fato preocupante, visto que esse é o principal padrão utilizado para o estabelecimento de Repositórios Arquivísticos Digitais Confiáveis (RDC-Arq).

Figura 35 – Adoção de certificações e normas pelos repositórios.



Fonte: elaboração própria.

O padrão PREMIS⁹⁹ ocupa 10,53% dos repositórios, sendo sinalizado por 2 repositórios. Esse *framework* permite que se registrem todos os metadados necessários para a realização da preservação digital a longo prazo, contendo elementos voltados ao detalhamento dos objetos, ambiente, direitos, eventos e agentes que fizeram parte do ciclo de vida dos documentos, além de permitir que sejam adicionadas as próprias ações realizadas para a preservação (PREMIS EDITORIAL COMMITTEE, 2015, p. 6). Assim, a presença desse padrão nos repositórios é desejável e até mesmo aconselhável, porém, a baixa adesão nos repositórios pesquisados é fato preocupante, pois não se tem

⁹⁹ (PREMIS EDITORIAL COMMITTEE, 2015)

registro dos mecanismos utilizados para armazenar os metadados dos documentos e registrar as ações de preservação desenvolvidas.

Os padrões de descrição arquivística NOBRADE¹⁰⁰ e ISAD(G)¹⁰¹ foram citados por apenas 1 repositório cada um, representando 5,26% dos repositórios cada um. Tais padrões são usados no sentido de se promover uma uniformização das descrições dos documentos custodiados, considerando-se importantes para que possa promover a interoperabilidade entre sistemas, por exemplo. Enquanto o primeiro padrão é de escopo nacional, brasileiro, o segundo padrão é adotado internacionalmente, portanto, o número de repositórios passíveis de sua adoção se sobressai aos nacionais, tornando a sua adoção de extrema importância.

Dentre todas as áreas avaliadas, a área 3 foi a que apresentou as menores pontuações, gerando a menor Média Aritmética, de 1,61 ponto. Assim, identifica-se um baixo grau de interesse dos repositórios nos assuntos relacionados às certificações, principalmente, sendo uma área que pode ser melhor priorizada nos repositórios, de modo a comprovarem os seus trabalhos através de auditorias, por exemplo, que poderão impactar as suas reputações frente às suas Comunidades-Alvo.

5.4.4 Resultados da Área 4: Infraestrutura Organizacional

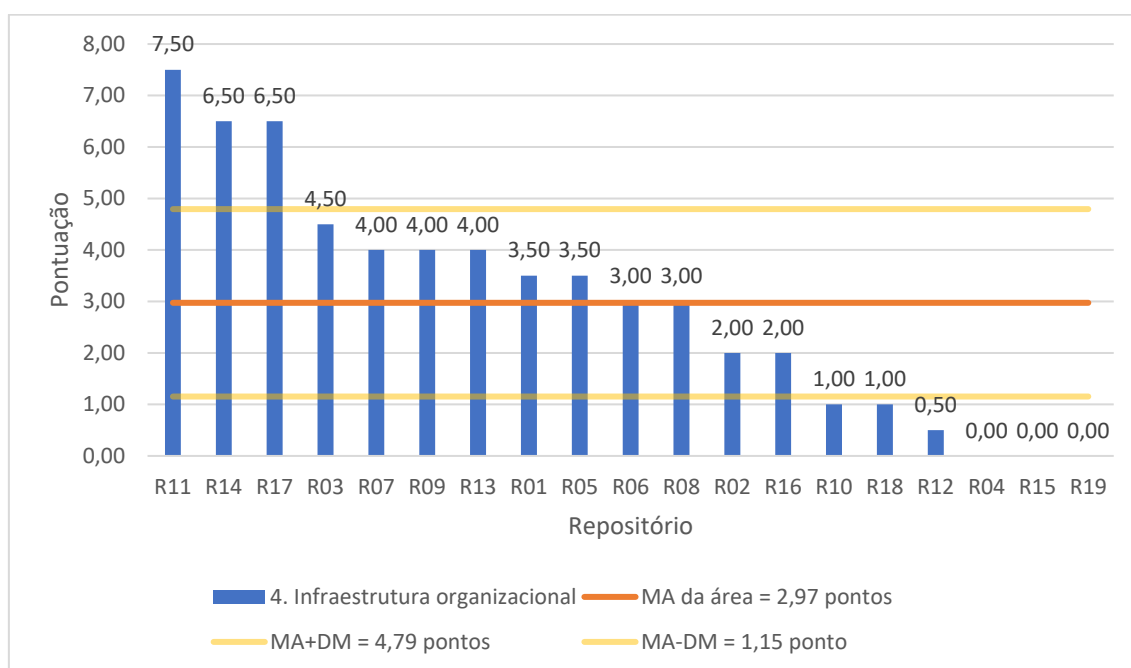
Nessa seção são analisadas as respostas referentes à Área 4: Infraestrutura Organizacional, que avaliou a adequação da infraestrutura organizacional dos repositórios aos propósitos da preservação digital confiável a longo prazo. Desde os aspectos relacionados à presença de divisões internas até o dimensionamento dos respectivos orçamentos foram avaliados nessa seção.

A Figura 36 exibe as pontuações totais dos repositórios nas perguntas da área 4, que teve como Média Aritmética, 2,97 pontos, Desvio Médio de 1,82 ponto em relação à MA e Coeficiente de Variação de 74,69%, indicando uma altíssima variabilidade das pontuações. O repositório R11 obteve a maior pontuação, com 7,50 pontos, enquanto a menor pontuação, 0,0 ponto, foi registrada para os repositórios R04, R15 e R19, uma diferença de 7,50 pontos entre os dois extremos.

¹⁰⁰ (CONSELHO NACIONAL DE ARQUIVOS, 2006)

¹⁰¹ (INTERNATIONAL COUNCIL ON ARCHIVES, 2000)

Figura 36 – Pontuações individuais dos repositórios na Área 4: Infraestrutura Organizacional.



Fonte: elaboração própria.

O Quadro 50 exibe a classificação dos repositórios na área 4, de acordo com as suas faixas de pontuação. Pode-se notar que 3 repositórios se classificaram com desempenho excelente, enquanto 10 repositórios foram classificados com desempenho mediano. Um total de 6 repositórios foi classificado com desempenho inferior.

Quadro 50 – Classificação dos repositórios pelos seus desempenhos na Área 4: Infraestrutura Organizacional.

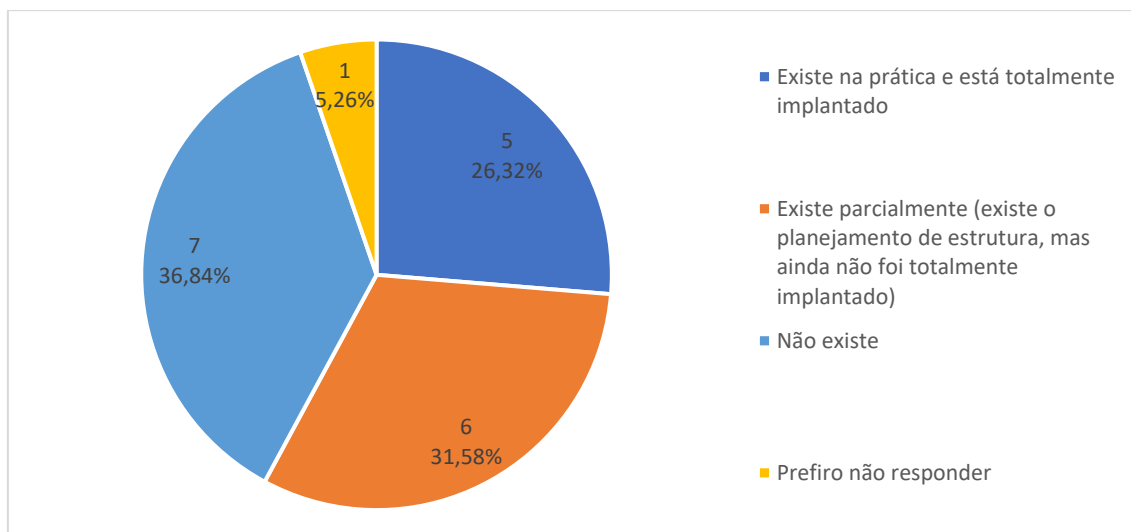
Desempenho excelente ($\geq 4,79$ pontos)	Desempenho mediano ($\geq 1,15$ e $< 4,79$ pontos)	Desempenho inferior ($< 1,15$ ponto)
1º) R11	3º) R03	8º) R10 e R18
2º) R14 e R17	4º) R07, R09 e R13	9º) R12
	5º) R01 e R05	10º) R04, R15 e R19
	6º) R06 e R08	
	7º) R02 e R16	
Total: 3	Total: 10	Total: 6

Fonte: elaboração própria.

A Figura 37 exibe a distribuição dos resultados referentes à pergunta “4.1 Existe uma estrutura organizacional adequada ao funcionamento do repositório para a preservação a longo prazo?”. Dentre os 19 repositórios participantes, 1 repositório preferiu não responder à questão. A não existência de uma estrutura organizacional

adequada à preservação foi registrada por 7 repositórios, o que equivale a 36,84% dos repositórios.

Figura 37 – Adequação da estrutura organizacional dos repositórios.



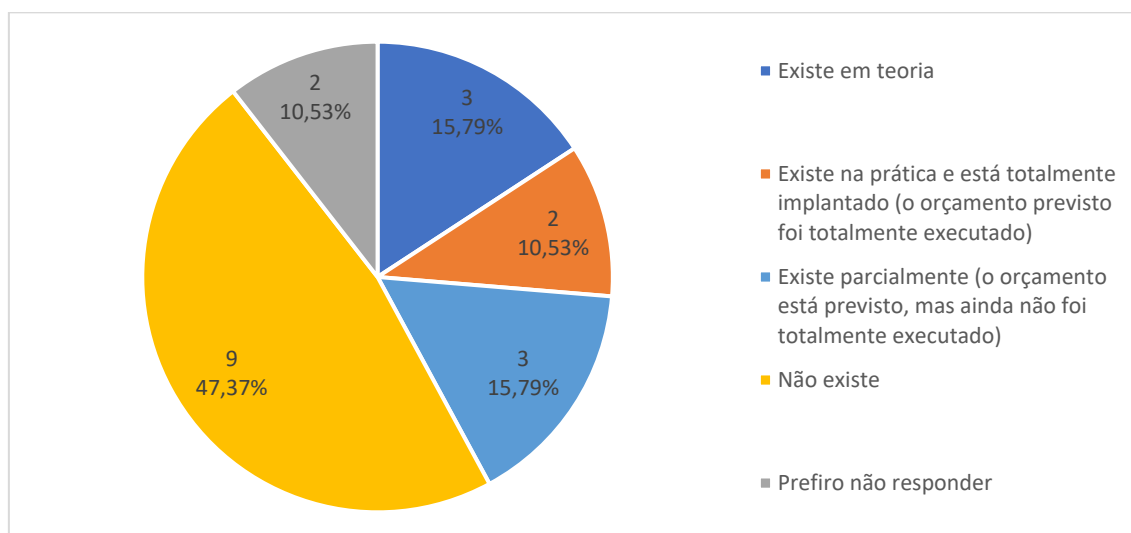
Fonte: elaboração própria.

Para 6 repositórios, 31,58% o planejamento para a implementação de uma estrutura adequada existe, porém não foi totalmente aplicado na prática. A análise desses casos não permite saber se a implementação da estrutura foi pausada temporariamente por algum motivo ou se ela ainda se encontra em execução. Somente em 5 repositórios, 26,32%, foi reportado haver uma estrutura organizacional adequada e aplicada na prática. A ausência de uma estrutura que permita a correta alocação de pessoas e divisão de responsabilidades pode, em certos casos, impactar no acúmulo de funções e atividades, fazendo com que haja a necessidade de se priorizar determinados processos em detrimento a outros podendo impactar negativamente na realização da preservação a longo prazo. Cabe ressaltar que a estrutura organizacional é um dos componentes da ISO 14721:2012 (OAIS) (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012), e a verificação de sua adequação é um dos critérios da ISO 16363:2012 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012).

As respostas da pergunta “4.2 Existe orçamento adequado ao funcionamento do repositório para a preservação a longo prazo?” são analisadas na Figura 38. Dos 19 repositórios participantes, 2 preferiram não responder à pergunta, o que equivale a 10,53% do total. Para 3 repositórios, o equivalente a 15,79%, o orçamento é adequado somente em teoria, quando existe o planejamento de recursos, mas eles não são de fato

investidos sustentar o funcionamento do repositório. O mesmo número de repositórios relatou existir o planejamento do orçamento para a adequação do funcionamento do repositório, porém ainda não houve a sua execução total.

Figura 38 – Adequação dos orçamentos dos repositórios à preservação de longo prazo.



Fonte: elaboração própria.

No caso de 2 repositórios, 10,53%, o orçamento adequado existe e foi aplicado em sua totalidade na adequação do funcionamento do repositório às necessidades da preservação a longo prazo. Do lado oposto, 9 repositórios, 47,37%, relataram a não existência de orçamento que deveria ser destinado ao repositório para a realização da preservação a longo prazo.

Ao todo, somente 5 repositórios, 26,32%, declararam a existência de um orçamento para o repositório, dentre os totalmente e parcialmente executados. Esse baixo número demonstra a situação frágil em que a custódia de documentos das instituições pode estar inserida, uma vez que a preservação a longo prazo requer constantes investimentos para a manutenção e evolução da sua estrutura como um todo. Como forma de ressaltar a importância desse quesito, a ISO 16363:2012, possui a seção “3.4 Sustentabilidade financeira” de critérios que verificam se existem meios de se comprovar que o repositório tem condições de funcionamento atualmente e com planejamento futuro (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 31).

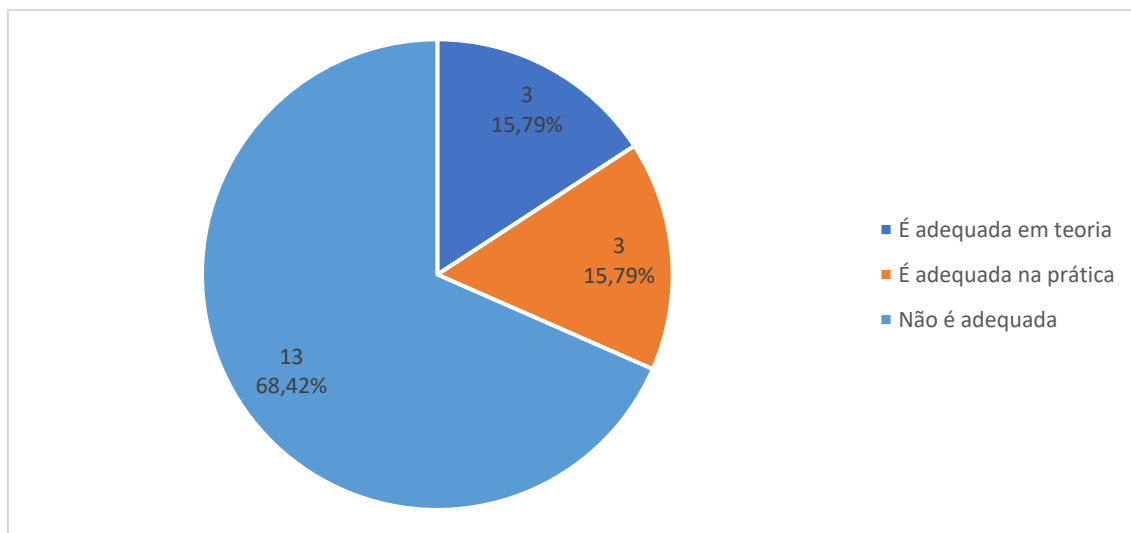
A verificação de orçamento também é tratada no Nestor Seal através do critério 4.1 (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 15, tradução nossa):

4.1 O financiamento adequado do repositório digital está assegurado: O Repositório Digital deve poder fundamentar a sua alegação de que os serviços propostos podem ser financiados, tanto a curto como a longo prazo. O financiamento do repositório digital deve ter uma base legalmente assegurada. No caso de repositórios digitais financiados pelo estado, o financiamento deve ser incluído nos documentos formais de planejamento (pelo menos a médio prazo). Um Repositório Digital privado deve ser capaz de garantir sua sustentabilidade financeira com base no uso cobrado de seus serviços e em um plano de negócios de longo prazo.

Ainda que possa parecer óbvio ou redundante, a operação dos repositórios envolve um rol de custos fixos e variáveis, que passam pelo fornecimento de energia elétrica, rede, aquisição e manutenção de equipamentos, aquisições de licenças de *software*, equipes e vai até as capacitações necessárias, por exemplo. Daí a necessidade de os gestores dos repositórios planejarem as formas de financiamento que permitam que a preservação digital seja plenamente realizada.

A situação sobre o quantitativo de pessoal dos repositórios foi solicitada na pergunta “4.3 Em relação à quantidade de pessoas, a equipe responsável pelo repositório é adequada?”, que possui os resultados demonstrados na Figura 39. Foram registradas respostas para todos os 19 repositórios participantes.

Figura 39 – Adequação da quantidade de pessoas nos repositórios.



Fonte: elaboração própria.

Em 3 repositórios, 15,79%, foi registrado que o quantitativo de pessoas envolvidas no funcionamento do repositório é adequado na prática, ou seja, existem pessoas o suficiente para a execução dos processos e atividades implementados em seus respectivos repositórios. O mesmo número de repositórios relata que o quantitativo de

peçoas somente é adequado em teoria, mas na prática não existe um número suficiente de peçoas para o funcionamento adequado do repositório.

Na grande maioria dos repositórios, 13 ao todo, o que representa 68,42%, existe uma carência de peçoas para a realização dos trabalhos dos seus repositórios. Essa carência pode acarretar sobrecarga física e psicológica nas equipes, uma vez que, provavelmente, há o acúmulo de papéis desenvolvidos pelas mesmas peçoas.

Devido à natureza dos repositórios participantes, que em sua maioria são geridos por órgãos dos governos, mesmo que de diferentes esferas, a contratação de peçoas somente pode ser realizada através de concursos públicos que requerem autorização explícita dos seus órgãos superiores. Ainda que haja a saída de servidores devido às aposentadorias, por exemplo, a reposição de vagas de trabalho só pode ocorrer mediante autorizações para novos concursos. Assim, os repositórios podem apresentar muitas vezes déficits em seus quadros pessoais, pois a reposição de servidores nem sempre ocorre e, quando ocorre, não é imediatamente após a saída das peçoas. Nos casos de repositórios de iniciativa privada, a dinâmica de contratação tende a seguir uma metodologia diferente, uma vez que a gestão pode contratar peçoas à medida em que identifica as devidas necessidades.

O Quadro 51 exibe uma análise da proporção de servidores Técnico Administrativos das IES, cujos repositórios participaram do estudo, em relação aos seus públicos totais, utilizando dados da Sinopse Estatística da Educação Superior de 2020 (INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2022). Nessas instituições, são os servidores que realizam os trabalhos administrativos necessários à operacionalização das suas atividades, atendendo outros servidores, docentes e alunos. Pode-se perceber uma grande disparidade das proporções entre as instituições, destacando-se os casos dos repositórios R07 e R15 que possuem, respectivamente, 29,29% e 3,33% de proporção de servidores em relação aos públicos totais das instituições.

Ainda que não seja possível estabelecer uma proporção ideal de servidores em relação aos públicos totais e que, dentro do universo de servidores das instituições não tenha sido possível identificar aqueles que desempenham atividades diretamente ligadas aos repositórios e a preservação, chama a atenção essas disparidades encontradas. Esses dados ajudam a corroborar o indicativo de que 68% dos repositórios relatou não haver quantidade de peçoas suficiente para compor as suas equipes.

Quadro 51 – Proporção de funcionários em relação ao público total das IES.

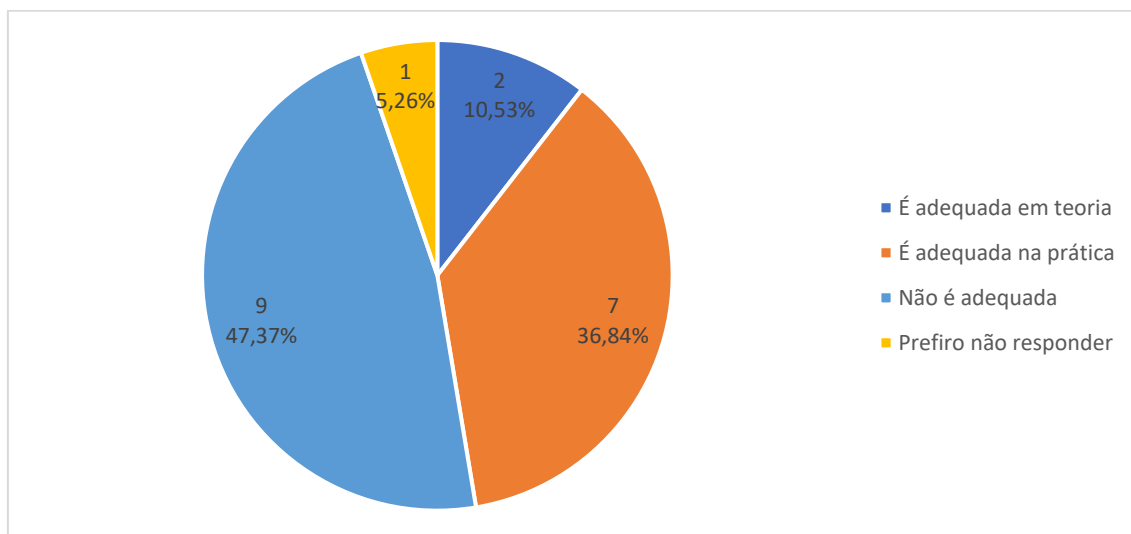
Repositório	Público total	Servidores	Proporção
R01	Não se aplica	-	-
R02	Não se aplica	-	-
R03	39.272	2.611	6,65%
R04	Não se aplica	-	-
R05	23.497	3.512	14,95%
R06	27.749	2.677	9,65%
R07	10.593	3.103	29,29%
R08	34.489	3.370	9,77%
R09	32.428	2.631	8,11%
R10	22.469	1.772	7,89%
R11	36.977	3.840	10,38%
R12	29.284	2.608	8,91%
R13	Não se aplica	-	-
R14	26.115	2.670	10,22%
R15	34.481	1.148	3,33%
R16	25.266	1.251	4,95%
R17	24.382	1.852	7,60%
R18	15.656	1.807	11,54%
R19	44.080	3.204	7,27%

Fonte: elaboração própria

A pergunta “4.4 Em relação às capacitações e conhecimentos, a equipe responsável pelo repositório é adequada?” tem os seus resultados analisados na Figura 40. Dentre os 19 repositórios participantes, 1 preferiu não responder à pergunta.

As respostas de 2 repositórios, 10,53%, relataram que as suas equipes possuem capacitações e conhecimentos adequados em teoria às demandas dos repositórios, significando que a aplicação desses conhecimentos não ocorre na prática do dia-a-dia do funcionamento dos repositórios. Em 7 repositórios, 36,84%, a equipe possui conhecimentos adequados e os aplicam na prática, favorecendo a maneira como esses repositórios operam. Em 47,37% dos repositórios, 9 ao todo, em certo grau as capacitações e conhecimentos das suas equipes não são adequadas, o que pode levar, em certos casos, à má execução de atividades no repositório e ao comprometimento da preservação a longo prazo dos materiais sob sua custódia.

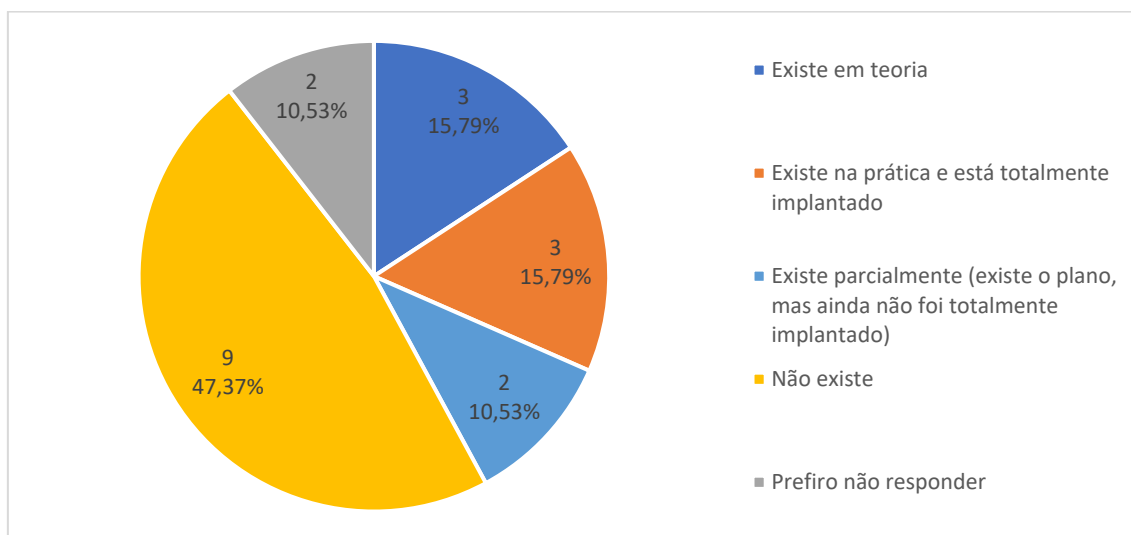
Figura 40 – Adequação de capacitações e conhecimentos das equipes dos repositórios



Fonte: elaboração própria.

A pergunta “4.5 Existe um plano de capacitação permanente da equipe do repositório?” tem o objetivo de complementar a análise da pergunta anterior e seus resultados são exibidos na Figura 41. Dentre todos os repositórios participantes, 2 preferiram não responder à pergunta.

Figura 41 – Existência de um plano de capacitação nos repositórios.



Fonte: elaboração própria.

Para 3 repositórios, equivalente a 15,79%, existe somente em teoria um plano de capacitação da equipe, ou seja, não existe uma aplicação prática desse plano. Em 3 repositórios, 15,79%, foi reportado haver o plano de capacitação e a sua aplicação ocorre na prática por completo. No caso de 2 repositórios, 10,53%, foi reportado que

existe o plano de capacitação, porém ele não está totalmente implantado na prática. Para os 9 repositórios restantes, 47,37%, os respondentes relataram não haver nenhum plano de capacitação das equipes.

É através de um plano de capacitação permanente que o repositório identifica as necessidades da sua equipe, buscando cursos, capacitações, palestras e eventos, por exemplo, que permitam que a equipe esteja em contato com novos conhecimentos ou para que o nível do conhecimento da equipe seja mais uniforme entre os seus integrantes. Em 12 repositórios, 63,16%, as suas equipes não contam, ou contam somente em teoria, com um plano que promova a busca de novos conhecimentos relacionados à preservação digital, o que pode fazer com que os próprios membros tenham que buscar cursos e eventos, por exemplo, muitas vezes custeados por eles próprios.

Por outro lado, a não existência de um plano de capacitação implantado pode vir a favorecer somente determinados membros das equipes, promovendo diferenças de tratamento das pessoas dentro das equipes. Dessa forma, a disseminação dos conhecimentos e o consequente nivelamento de capacitações não se dá uniformemente, estabelecendo ou aumentando lacunas na formação das pessoas.

A análise conjunta das perguntas 4.4 e 4.5 permite algumas análises adicionais. Considerando os 9 repositórios que sinalizaram não haver nenhum plano de capacitação das equipes, 4 responderam que a capacitação das equipes é adequada na prática, concluindo-se que os membros das equipes buscam conhecimentos e capacitações por conta própria, sem levar em consideração nenhum planejamento da instituição. Outros 4 repositórios declaram que a capacitação das equipes não é adequada às atividades dos repositórios e, por final, 1 repositório respondeu que a capacitação da equipe é adequada somente em teoria. Assim, do universo de 47,37% dos repositórios que não possuem plano de capacitação, apenas 44% possuem equipe com capacitações adequadas às atividades desenvolvidas.

No caso dos 3 repositórios que possuem plano de capacitação aplicado totalmente na prática, 1 relatou que a equipe possui capacitação adequada, enquanto os outros 2 repositórios relataram que a capacitação da equipe não é adequada às atividades do repositório. Dentre os fatores que podem levar a essa situação pode ser a elaboração de planos de capacitação que não estejam alinhados aos conhecimentos necessários para o desenvolvimento das atividades e processos dos repositórios.

No caso dos 5 repositórios que declararam que o plano de capacitação existe em teoria ou está parcialmente implantado, 1 repositório declarou que a capacitação da sua equipe é adequada em teoria, enquanto 2 repositórios declararam que a capacitação da equipe é adequada na prática e 2 repositórios relataram que a capacitação da equipe não é adequada. Esses casos podem revelar que a implementação total do plano de capacitação do repositório na prática pode reverter a situação da capacitação das equipes nos 3 repositórios, fazendo com que as pessoas possuam os conhecimentos necessários para a realização das atividades requeridas para a preservação a longo prazo nos repositórios.

Dentre os 2 repositórios que não responderam sobre a existência do plano de capacitação em suas instituições, 1 relatou que a sua equipe não possui capacitação adequada e o outro repositório também preferiu não responder sobre a adequação de capacitação da sua equipe.

A partir das análises da área 4, percebe-se uma grande disparidade entre as infraestruturas organizacionais dos repositórios, enquanto 3 repositórios apresentaram desempenhos excelentes, 3 repositórios não apresentaram nenhuma pontuação na área. Ainda, caso a execução total dos planejamentos já existentes de diversos critérios seja realizada, será observada uma melhoria nos desempenhos dos repositórios, recomendando-se essa ação. Cabe observar que as perguntas e respostas da área 4 não permitem identificar as razões pelas quais tais planejamentos não foram totalmente aplicados na prática, devendo cada repositório, a partir da sua realidade, traçar as estratégias mais adequadas para essa execução.

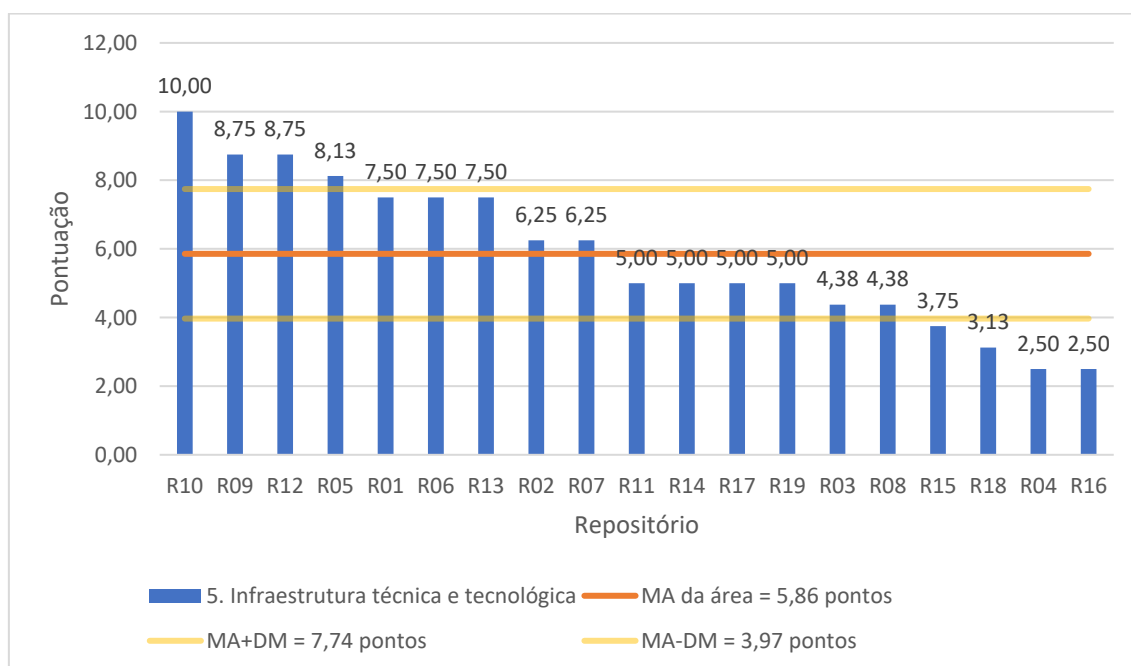
5.4.5 Resultados da Área 5: Infraestrutura Técnica e Tecnológica

Nessa seção são analisadas as respostas fornecidas no formulário relativas à Área 5: Infraestrutura Técnicas e Tecnológica, que trata sobre a estrutura e equipamentos utilizados pelos repositórios para o seu funcionamento. Foram analisadas a presença de uma estrutura física e de equipamentos adequadas, além da presença de processos de atualização dos *softwares* envolvidos.

A Figura 42 exibe as pontuações totais dos repositórios nas perguntas da área 5, que teve como Média Aritmética, 5,86 pontos, Desvio Médio de 1,89 ponto em relação à MA. O Coeficiente de Variação calculado foi de 37,01%, o menor entre todas as áreas, indicando uma variabilidade relativamente baixa das pontuações, indicando

que a infraestrutura técnica e tecnológica é tratada de maneira mais uniforme entre os repositórios participantes. O repositório R10 obteve a maior pontuação, com 10,00 pontos, enquanto a menor pontuação, 2,50 pontos, foi registrada para os repositórios R04, e R16, uma diferença de 7,50 pontos entre os dois extremos.

Figura 42 – Pontuações individuais dos repositórios na Área 5: Infraestrutura Técnica e Tecnológica.



Fonte: elaboração própria.

O Quadro 52 exibe a classificação dos repositórios na área 5, de acordo com as suas faixas de pontuação. Pode-se notar que 4 repositórios se classificaram com desempenho excelente, enquanto 11 repositórios foram classificados com desempenho mediano. Um total de 4 repositórios foi classificado com desempenho inferior.

Quadro 52 – Classificação dos repositórios pelos seus desempenhos Área 5: Infraestrutura Técnica e Tecnológica.

Desempenho excelente ($\geq 7,74$ pontos)	Desempenho mediano ($\geq 3,97$ e $< 7,74$ pontos)	Desempenho inferior ($< 3,97$ pontos)
1º) R10	4º) R01, R06 e R13	8º) R15
2º) R09 e R12	5º) R02 e R07	9º) R18
3º) R05	6º) R11, R14, R17 e R19	10º) R04 e R16
	7º) R03 e R08	
Total: 4	Total: 11	Total: 4

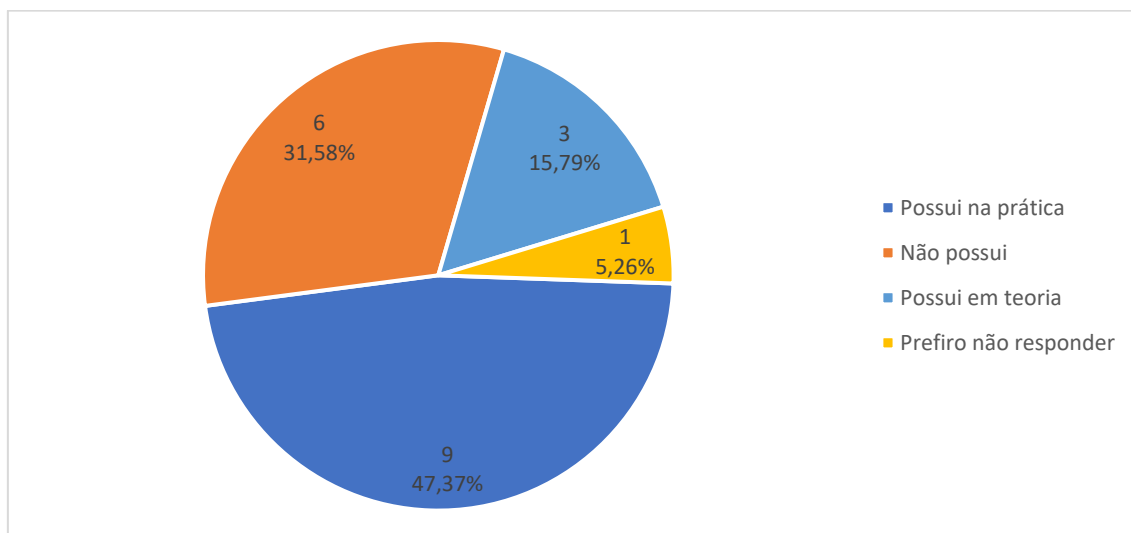
Fonte: elaboração própria.

As respostas da pergunta “5.1 O repositório possui estrutura física de prédios e/ou salas adequada à preservação e garantia de preservação a longo prazo?” são

apresentadas na Figura 43. Do total de 19 respostas recebidas, 1 repositório preferiu não responder, enquanto 9 repositórios, 47,37%, declarou possuir a estrutura física do repositório adequada à preservação a longo prazo.

No caso de 3 repositórios, 15,79%, a estrutura física é adequada somente em teoria, de onde se pode concluir que houve um planejamento para a construção ou adaptação da estrutura física do repositório para adequá-la às necessidades identificadas, porém, por motivos não identificados, a execução do planejamento não foi realizada. No caso de 6 repositórios, o equivalente a 31,58%, não existe uma estrutura física de prédios e salas adequada às necessidades da preservação a longo prazo.

Figura 43 – Adequação da estrutura física dos repositórios



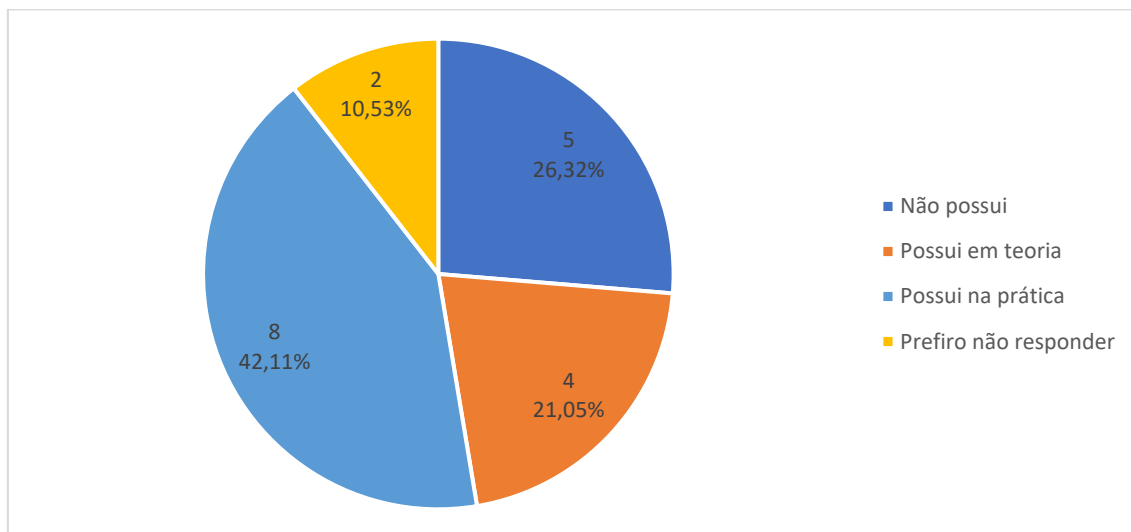
Fonte: elaboração própria.

A pergunta “5.2 O repositório possui estrutura de equipamentos (servidores de aplicação, armazenamento, rede, etc.) adequada à preservação e garantia de preservação a longo prazo?” tem as suas respostas analisadas na Figura 44. Dentre os repositórios participantes, 2 preferiram não responder à pergunta. No caso de 8 repositórios, 42,11%, a estrutura de equipamentos existe na prática e é adequada aos requisitos necessários à preservação a longo prazo, enquanto 4 repositórios, 21,05%, relataram que a estrutura de equipamento somente existe em teoria. Os representantes de 5 repositórios, 26,32%, declaram não possuir uma estrutura adequada de equipamentos.

Analisando-se as respostas fornecidas, caso o planejamento da estrutura de equipamentos dos 4 repositórios que relataram a sua existência somente em teoria seja executado, poderia se chegar a um total de 63,16% de repositórios adequados. Ainda

que essa porcentagem pudesse representar a maioria dos repositórios, a pesquisa realizada não possui dados que permitam inferir que essa execução dos planejamentos seja aplicada na prática para esses repositórios.

Figura 44 – Adequação da infraestrutura técnica e tecnológica dos repositórios.



Fonte: elaboração própria.

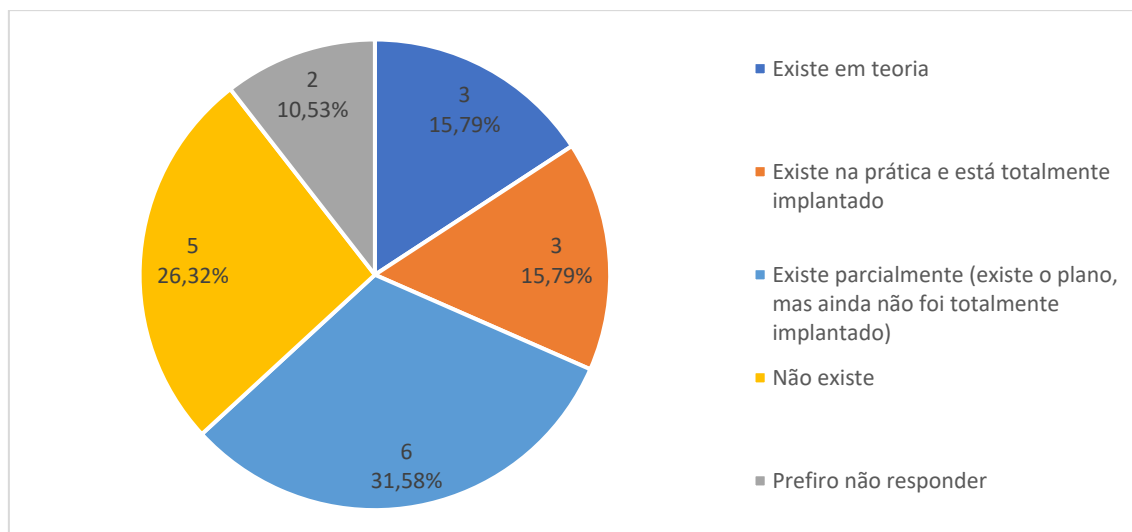
Devido à crescente produção de documentos natos digitais, muitos documentos que necessitam preservação não possuem mais os seus representantes físicos em papel, fazendo com que a sua existência se restrinja somente ao mundo digital. O total de repositórios que declarou haver estrutura de equipamentos adequada à preservação a longo prazo representa menos da metade do total dos repositórios participantes da pesquisa, tornando-se um fato preocupante, uma vez que o restante dos repositórios não pode assumir o compromisso de preservar os documentos sob a sua custódia sem possuir a infraestrutura que garanta que os documentos existirão no futuro pelo tempo que for necessário.

Ao se analisar as respostas das perguntas anteriores 5.1 e 5.2, sobre as estruturas físicas e tecnológicas dos repositórios, pode-se perceber um certo grau de similaridade e aproximação. Pode-se perceber que 9 e 8 repositórios, respectivamente, possuem estrutura física e tecnológica adequadas à preservação digital. Ao mesmo tempo, 3 e 4 repositórios declararam possuir em teoria as estruturas físicas e tecnológicas adequadas.

A pergunta “5.3 Existe um plano de atualização periódico de *hardware* e *software*?” analisa se o repositório busca monitorar os estados dos seus equipamentos e *softwares* utilizados, atualizando-os sempre que necessário. A análise das respostas é

apresentada na Figura 45, sendo possível observar que 2 repositórios, 10,53%, preferiram não fornecer uma resposta.

Figura 45 – Existência de um plano de atualização de *hardware* e *software* nos repositórios.



Fonte: elaboração própria.

Apenas 3 repositórios, o equivalente a 15,79%, declararam haver um plano de atualização dos seus equipamentos executado na prática, tendo o mesmo quantitativo, 3 repositórios, respondido que existe somente em teoria um plano, ou seja, ele não é aplicado nas práticas do repositório.

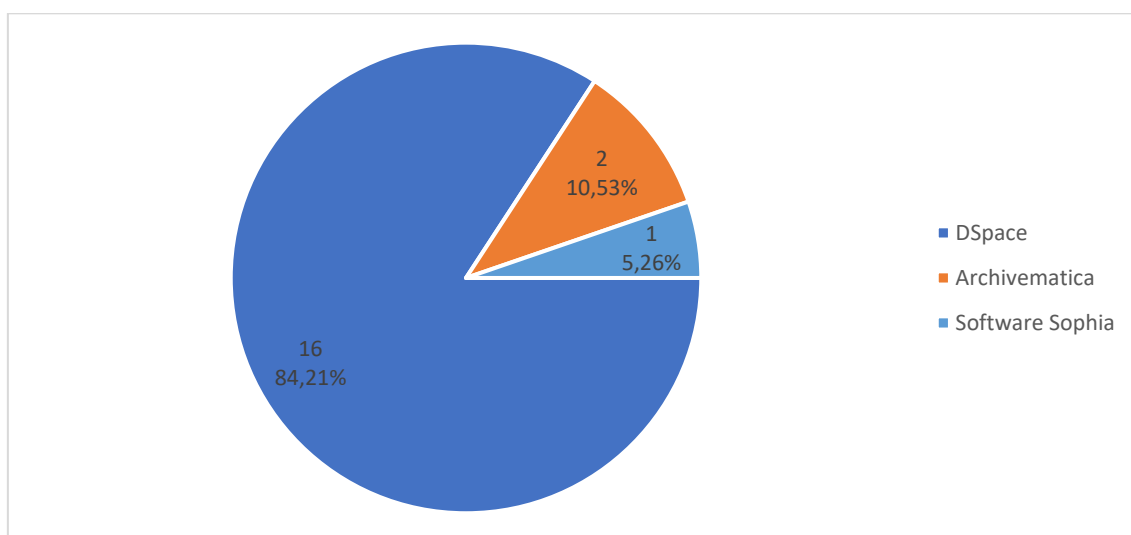
No caso de 6 repositórios, correspondente a 31,58% do total de repositórios, existe um plano de atualização, porém ele não se encontra totalmente aplicado na prática, faltando ações a serem tomadas para que a aplicação ocorra na prática. Em 5 repositórios, equivalente a 26,32%, não existe nenhum plano que contenha processos, atividades e ações voltadas ao monitoramento dos estados dos equipamentos e *softwares* que forneça parâmetros para realização de atualizações quando necessárias.

As respostas fornecidas permitem a criação de um cenário em que, caso a implantação do plano de atualização seja concluída nos 6 repositórios que sinalizaram a sua existência com execução parcial, seria alcançado um total de 9 repositórios com o plano de atualização existente e totalmente implantando, o que representaria um total de 47,37% do total dos repositórios. No entanto, esse cenário só pode ser alcançado se o único impeditivo da aplicação prática do plano de atualização for o instante de tempo no qual a pesquisa foi realizada, assumindo-se que a implantação do plano estava em andamento e com a sua finalização prevista para um momento posterior ao do fornecimento das respostas do formulário.

Ao mesmo tempo, pode-se estabelecer outro cenário, quando houvesse a aplicação prática do plano de atualização dos 3 repositórios que declararam a sua existência somente em teoria, assim, se alcançaria 12 repositórios, ou 63,16% dos repositórios participantes, alcançando-se a maioria dos repositórios. Nesse caso, porém, não são conhecidas as causas que levaram os 3 repositórios a não aplicar os seus planos, podendo haver restrições de pessoal, orçamentárias ou políticas, por exemplo.

A última pergunta da seção, “5.4 Qual plataforma (*software*) utilizada pelo repositório?” solicita ao respondente que informe a solução de software utilizada pelo repositório para a preservação dos seus documentos, buscando identificar as suas características, arquiteturas, comunidades, disponibilização de versões, e demais características técnicas. As respostas são exibidas na Figura 46, tendo todos os repositórios fornecido respostas.

Figura 46 – Plataformas (*softwares*) utilizados pelos repositórios.



Fonte: elaboração própria.

Um total de 16 repositórios, equivalente a 84,21%, declararam utilizar a plataforma DSpace para a realização da preservação dos seus documentos, representando a grande maioria das respostas. Em 2 repositórios, equivalente a 10,53%, o *software* utilizado para a preservação é o Archivematica, enquanto 1 repositório declarou utilizar o *software* Sophia.

A predominância de repositórios que utilizam o software DSpace vai de encontro ao que o estudo de Costa e Leite realizaram em 2017, quando identificaram que esse software é o mais usado nos repositórios institucionais da América Latina (2017, p. 161).

O *software* DSpace se caracteriza por ser uma solução web de código-livre (*open source*) vastamente utilizada em bibliotecas mundialmente, possuindo ferramentas que permitem a personalização da instalação para adaptá-la às necessidades de cada repositório ao mesmo tempo em que agrega funcionalidades de Sistemas de Gerenciamento de Conteúdo (*Content Management System* - CMS) e gerenciamento de documentos. De uma forma geral, o DSpace organiza os materiais através de coleções de itens, que podem conter vários níveis hierárquicos, possuindo mais a metáfora como de como as bibliotecas se organizam do que os arquivos.

A ferramenta é mantida pela organização não governamental e sem fins lucrativos DuraSpace que possui uma comunidade apoio ao DSpace e outros projetos de código-livre. O DSpace é definido como um *software* que visa atender às necessidades de arquivos digitais com foco no armazenamento a longo prazo, acesso e preservação de conteúdo digital, voltando o seu uso principalmente a organizações acadêmicas, sem fins lucrativos e comerciais (DURASPACE, 2023). Embora o DSpace seja uma ferramenta livre, a organização DuraSpace¹⁰² é a coordenadora do desenvolvimento e recebe contribuições de parceiros institucionais que colaboram com horas de contribuição técnica ou através de contribuições financeiras. Essas parcerias trabalham no desenvolvimento de funcionalidades novas e oferecem serviços que podem ser contratados pelos repositórios interessados.

A quantidade de respostas recebidas de repositórios que utilizam o *software* DSpace reflete a grande comunidade de usuários mundial, que possui registrados em sua base 3.705 repositórios. No Brasil, atualmente estão registrados 153 repositórios que utilizam o DSpace, dentre esses, 103 caracterizam a sua organização como sendo do tipo Acadêmica, o que representa um percentual de 67,32% sobre o total brasileiro (DURASPACE, 2022).

Uma das principais características do DSpace é agregar as funcionalidades de ingestão, arquivamento e difusão em uma mesma ferramenta, possuindo as três principais entidades da arquitetura OAIS: produtor, gerenciador e consumidor (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2021). Assim, é permitido que os produtores enviem os seus materiais pela interface web juntamente com um conjunto pré-estabelecido de metadados, que posteriormente segue o fluxo de aprovação

¹⁰² <https://duraspace.org/>

configurado pelos administradores do repositório e, conforme as decisões tomadas, já pode ser disponibilizado para consultas.

No entanto, a forma como os documentos são armazenados internamente pelo DSpace não utiliza a arquitetura de pacotes de informação, destacando-se principalmente, a ausência do AIP (*Archival Information Package* – Pacote de Arquivamento de Informações), que reúne os documentos recebidos, suas versões de preservação e acesso, assim como a trilha de ações de preservação realizadas. O DSpace armazena essas informações de uma forma dependente da ferramenta, utilizando principalmente um banco de dados relacional, no entanto, possui meios de exportar todas essas informações para AIPs que podem abastecer outras ferramentas (ROCHA, JUNIOR, *et al.*, 2021).

Essa característica pode trazer impactos negativos na ocorrência de eventuais problemas ou erros da ferramenta, uma vez que a geração dos pacotes para exportação pode não funcionar de maneira adequada, impossibilitando a migração para outros repositórios. Adicionalmente, a construção dos pacotes AIP para exportação pode acrescentar perigos à autenticidade dos documentos, uma vez que a ocorrência de falhas pode corromper os documentos ou gerar informações e metadados incorretos a respeito dos ciclos de vida dos documentos custodiados.

Em 2 repositórios foi relatado o uso da ferramenta Archivematica para a realização dos processos envolvidos na preservação a longo prazo. O Archivematica se autodenomina como “[...] um conjunto integrado de ferramentas de software de código aberto que permite que os usuários processem objetos digitais desde o ingresso até o acesso, em conformidade com o modelo funcional ISO-OAIS” (ARTEFACTUAL SYSTEMS INC, 2023), tendo o seu desenvolvimento gerenciado pela empresa Artefactual Systems Inc.

A empresa é responsável por coordenar o desenvolvimento da ferramenta, uma vez que ela é *open source* e é desenvolvida por uma comunidade de desenvolvedores espalhada mundialmente. Pessoas interessadas em corrigir erros ou adicionar funcionalidades podem se inscrever para a realização de tarefas existentes, acessar o código fonte e, posteriormente, submeter as suas alterações.

Caso alguma instituição possua uma demanda específica pode ainda patrocinar o desenvolvimento pela empresa Artefactual o que faz com que as suas funcionalidades adquiram prioridade de desenvolvimento, ao mesmo tempo em que após a conclusão desse desenvolvimento, essas novas funcionalidades passam a ser

incorporadas no código-fonte disponível a toda comunidade do *software*. Essa modalidade se torna útil para repositórios que possuem demandas específicas que impactam no seu funcionamento e não possuem equipes de desenvolvimento, porém preferem não esperar que a própria comunidade as desenvolva de acordo com as prioridades próprias.

No Brasil, a Resolução nº 43 (CONSELHO NACIONAL DE ARQUIVOS, 2015) regulamenta a definição de um RDC-Arq e sob a qual o Archivemática se caracteriza, segundo Gomes e Autran (2020, p. 111):

Dessa forma, podemos realçar que os aspectos citados na Resolução nº 43 estão afinados à incumbência do Archivemática. Este *software* é considerado RDC-Arq, pois dispõe dos elementos necessários para tal. Sua função principal é armazenar a documentação em formato digital, de acordo com os padrões de preservação exigidos.

A utilização do Archivemática tem mais destaque na comunidade internacional do que na comunidade brasileira, fato comprovado pelas respostas recebidas dos repositórios, apontando-se como um dos motivos para essa baixa adesão um certo grau de escassez de literatura sobre a adoção da ferramenta. Gomes e Autran (2020, p. 118) concluem que:

[...] o *software* se configura como eficiente estratégia de preservação digital, com vantagem da interoperabilidade para criar via de comunicação com outros tipos de dispositivos, já empregados por arquivos, bibliotecas, museus e instituições de ensino.

Internamente, o Archivemática funciona baseado em microserviços de natureza *open source* que atendem a diversas funcionalidades como verificação de vírus, identificação de tipo de arquivo, por exemplo. Essa abordagem “[...] usa ferramentas fracamente acopladas para fornecer serviços de curadoria digital granular e ortogonal construídos em torno do sistema de armazenamento de arquivos” (GARDEREN, 2010, p. 1, tradução nossa). Dentre as iniciativas de preservação, o *University of California Curation Center* (UC3) declara que “a sua nova abordagem para a infraestrutura de curadoria digital é baseada na ideia de delegar as funções necessárias a um conjunto de microserviços independentes, mas interoperáveis, que incorporam valores e estratégias de curadoria” (ABRAMS, KUNZE e LOY, 2010, p. 173, tradução nossa). Assim, a forma como a infraestrutura do Archivemática é construída se alinha às práticas mais atuais de preservação digital.

O aspecto positivo dessa abordagem é a sua modularização, permitindo que determinados serviços sejam substituídos ou ainda que se personalize fluxos de execução com serviços distintos de diferentes ferramentas. Por outro lado, essa modularização pode tornar difícil o gerenciamento da plataforma fazendo com que as atualizações dos serviços sejam bem testadas para que não haja incompatibilidades que possam cessar o funcionamento do repositório ou, ainda, corromper os pacotes existentes.

A ferramenta adota a estrutura de pacotes de informação do OAIS nativamente, tratando a ingestão com o SIP (*Submission Information Package* – Pacote de Submissão de Informações), o arquivamento com o AIP (*Archival Information Package* – Pacote de Arquivamento de Informações) e a difusão do conteúdo com o DIP (*Dissemination Information Package* – Pacote de Disseminação de Informações). Todas as operações de preservação realizadas no material sob sua custódia são documentadas em um arquivo PREMIS a fim de que sejam conhecidos todos os da preservação, corroborando a autenticidade dos materiais.

Ao passo em que o DSpace já incorpora uma plataforma de consulta e acesso aos documentos custodiados, o Archivematica não possui nenhum tipo de interface pública com essa finalidade. Para que os documentos custodiados possam ser acessados, deve-se conectar uma plataforma específica como, por exemplo, o Atom¹⁰³ que recebe os pacotes DIP, os interpreta e disponibiliza as derivadas de acesso e metadados pertinentes à comunidade-alvo.

A utilização do *software* SophiA (PRIMA, 2022) foi relatado por 1 repositório como sendo a solução para a preservação dos seus materiais. O software se destina ao gerenciamento de bibliotecas universitárias, possibilitando o recebimento de materiais em diversos formatos como PDF, imagens, vídeos e sons. Dentre os padrões atendidos pelo software estão o MARC 21 (LIBRARY OF CONGRESS, 2022), padrão difundido mundialmente para catalogação do acervo, e o padrão OAI-PMH (*Open Archive Initiative – Protocol for Metadata Harvesting*) que disponibiliza os metadados dos materiais para serviços especializados na colheita desses metadados, utilizado principalmente para a integração entre repositórios (OPEN ARCHIVES INITIATIVE, 2023). Não foi possível identificar se o software SophiA adere ao padrão OAIS

¹⁰³ <https://www.accesstomemory.org/pt-br/>

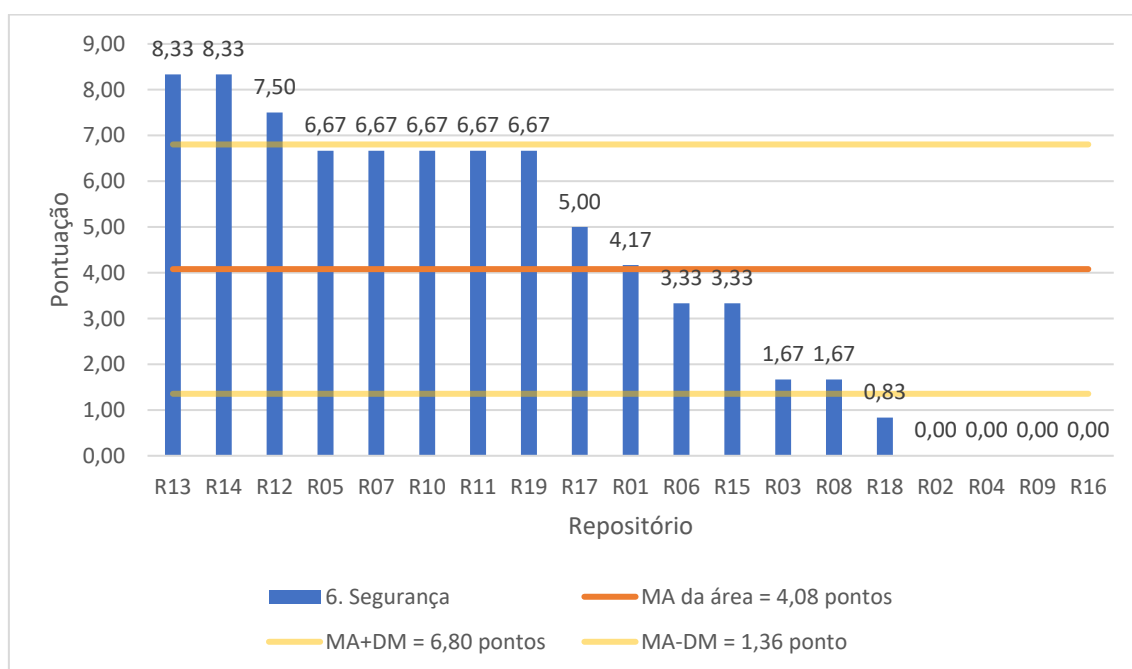
utilizando pacotes de informações e rotinas destinadas à preservação da autenticidade de integridade dos materiais sob sua custódia.

A análise dos resultados da área 5, permite identificar que os repositórios apresentam, de um modo geral, boas infraestruturas físicas e tecnológicas, todos contando com ferramentas de preservação digital, por exemplo, e registrando-se o desempenho máximo em 1 repositório. No entanto, recomenda-se que os repositórios priorizem a elaboração e prática de planos de atualização de *software* e *hardware*, com o objetivo de que não haja perdas e corrompimentos de dados ou que os materiais percam interpretabilidade ao longo do tempo.

5.4.6 Resultados da Área 6: Segurança

A Área 6: Segurança buscou identificar a forma como os repositórios lidam com a segurança das suas instalações físicas e equipamentos de TI envolvidos na preservação digital. As análises das respostas são apresentadas a seguir.

Figura 47 – Pontuações individuais dos repositórios na Área 6: Segurança.



Fonte: elaboração própria.

A Figura 47 exhibe as pontuações totais dos repositórios nas perguntas da área 6, que teve como Média Aritmética, 4,08 pontos, Desvio Médio de 2,72 pontos em relação à MA e Coeficiente de Variação de 75,50%, indicando uma altíssima

variabilidade das pontuações. Os repositórios R13 e R14 obtiveram a maior pontuação, com 8,33 pontos, enquanto a menor pontuação, 0,00 ponto, foi registrada para os repositórios R02, R04, R09 e R16, uma diferença de 8,33 pontos entre os dois extremos.

O Quadro 53 exibe a classificação dos repositórios na área 6, de acordo com as suas faixas de pontuação. Pode-se notar que 3 repositórios se classificaram com desempenho excelente, enquanto 11 repositórios foram classificados com desempenho mediano. Um total de 5 repositórios foi classificado com desempenho inferior.

Quadro 53 – Classificação dos repositórios pelos seus desempenhos Área 6: Segurança.

Desempenho excelente ($\geq 6,80$ pontos)	Desempenho mediano ($\geq 1,36$ e $< 6,80$ pontos)	Desempenho inferior ($< 1,36$ pontos)
1º) R13 e R14	3º) R05, R07, R10, R11 e R19	8º) R18
2º) R12	4º) R17	9º) R02, R04, R09 e R16
	5º) R01	
	6º) R06 e R15	
	7º) R03 e R08	
Total: 3	Total: 11	Total: 5

Fonte: elaboração própria.

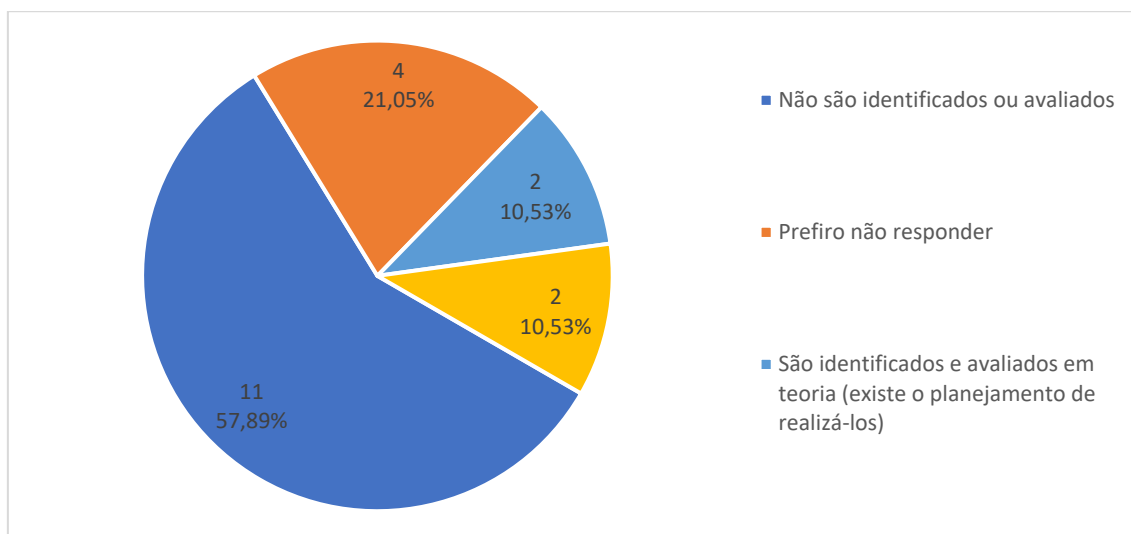
A pergunta “6.1) Os riscos que afetam a preservação dos objetos digitais são identificados e avaliados?” buscou identificar se os repositórios possuem processos para a identificação e avaliação dos riscos que podem comprometer a preservação dos seus respectivos materiais. As respostas são exibidas na Figura 48, sendo possível identificar que 4 repositórios, correspondente a 21,05% do total, preferiram não responder à pergunta.

Dentre os 19 repositórios participantes, 11, equivalente a 57,89%, responderam que não há nenhum processo ou metodologia aplicada para a identificação e avaliação de riscos relativos à preservação digital, representando a maioria dos repositórios. No caso de 2 repositórios, equivalente a 10,53%, existe um processo de identificação e avaliação de riscos, porém a execução do planejamento ainda não foi totalmente concluída. Outros 2 repositórios relataram que existe o planejamento para identificação e avaliação dos riscos, porém ele não foi implementado na prática, existindo atualmente somente em teoria.

Nenhum dos repositórios relatou a existência na prática de um processo totalmente aplicado voltado à identificação dos riscos relativos à preservação digital. Em relação à maioria dos repositórios, 57,89%, não há informações suficientes para se concluir se esse processo é relevante a ponto de já haver alguma iniciativa para a

construção de um planejamento ou se há uma decisão institucional de não realizar o tratamento de riscos. A situação desses repositórios pode ser reflexo de ordem de pessoal, devido à carência de conhecimentos sobre a área de riscos ou a falta de um quantitativo adequado de pessoas destinado a tratar do assunto. Outra possibilidade pode ser a falta de orçamento do repositório que poderia ser destinado à capacitação na área ou à contratação de serviços especializados para identificação e avaliação dos riscos.

Figura 48 – Análise dos riscos da preservação digital.



Fonte: elaboração própria.

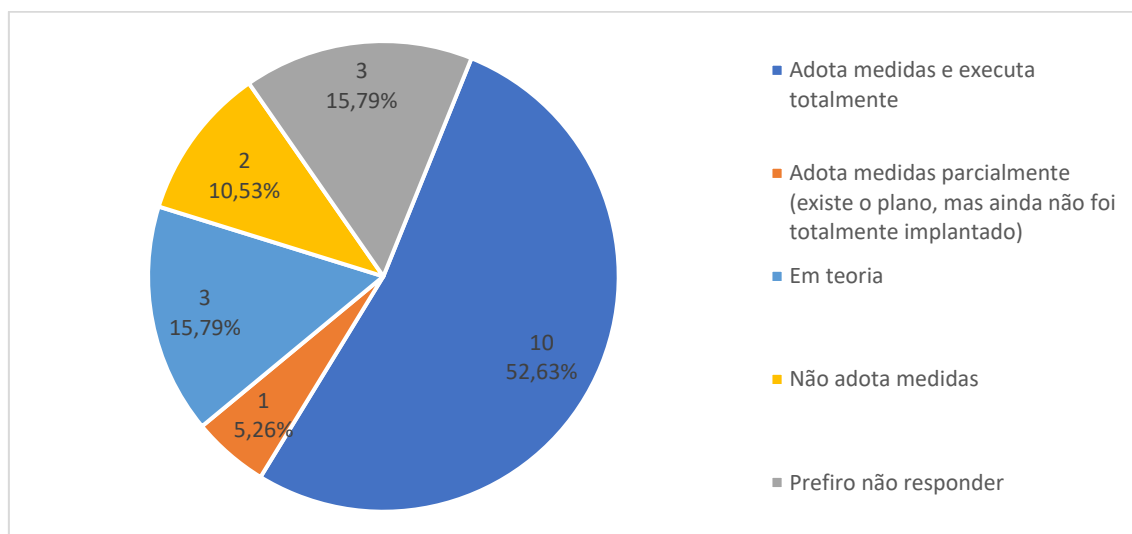
A pergunta “6.2) Caso os riscos tenham sido identificados, quais as ferramentas de avaliação de riscos utilizadas?” buscou identificar a forma como os riscos são identificados e avaliados e só poderia ser respondida caso o repositório tivesse sinalizado alguma forma de tratamento dos riscos na pergunta anterior. Somente uma resposta foi registrada para essa pergunta, referindo-se ao repositório que declarou a existência de um planejamento de identificação e avaliação de riscos, porém com execução parcial.

O repositório em questão declarou utilizar as ferramentas Big IP, Palo Alto Networks e Qualys para a gestão dos riscos. A solução Big IP é fornecida pela empresa F5 e possui diversos softwares destinados a aumentar a segurança da infraestrutura de TI como um todo, possuindo soluções para a otimização de tráfego de rede, segurança das aplicações, serviços em nuvem e softwares instalados em nuvem (F5, 2022). Não foi possível precisar quais metodologias utilizadas pela empresa para a análise efetiva dos riscos do repositório que contratou os seus serviços.

A companhia Palo Alto Networks fornece diversas aplicações que permitem que o ambiente de TI permaneça livre de ameaças e se previna de ataques de diferentes naturezas (PALO ALTO NETWORKS, 2022), porém o repositório em questão não forneceu detalhes acerca das soluções ou serviços contratados. A empresa Qualys é responsável por fornecer uma solução de análise e prevenção de riscos especializada para serviços em nuvem (QUALYS, 2022), no entanto, não foi possível identificar quais serviços ou produtos o repositório em questão possui contratados para se precaver de possíveis riscos que podem comprometer a preservação digital.

A pergunta “6.3) A estrutura de TI (*software* e *hardware*) do repositório adota medidas de segurança contra ataques?” buscou identificar se os repositórios possuem mecanismos para se proteger de eventuais ataques que possam comprometer o funcionamento do repositório e a preservação digital em si. As respostas à pergunta são exibidas na Figura 49, registrando-se que 3 repositórios, 15,79%, preferiram não fornecer uma resposta.

Figura 49 – Adoção de medidas de segurança contra ataques.



Fonte: elaboração própria.

No caso de 10 repositórios, o que equivale à 52,63%, existem procedimentos de prevenção a ataques e as medidas previstas são adotadas totalmente na prática. Em 2 repositórios, equivalente a 10,53%, não existe nenhuma medida de segurança contra ataques implementada na prática. Em 3 repositórios, existe um planejamento para a adoção de medidas contra ataques, porém elas existem somente em teoria, não existindo na prática, ações voltadas a esse fim. No caso de 1 repositório,

5,26%, o planejamento contra ataques existe, mas a sua execução ainda não é total na prática.

Uma vez que a pergunta anterior, 6.2, registrou que 11 repositórios, 57,89% dos repositórios participantes, ou seja, a maioria, não adota nenhuma forma de identificação de riscos nos repositórios, a pergunta 6.3 registrou que 10 repositórios, 52,63%, adota e executa na prática medidas de prevenção de ataques aos repositórios. Esse fato permite inferir que as medidas adotadas contra ataques nos 10 repositórios indicados na pergunta 6.3 não obedecem a uma metodologia definida, o que pode levar a execução de processos que não contemplem realmente os riscos os quais os repositórios estão expostos.

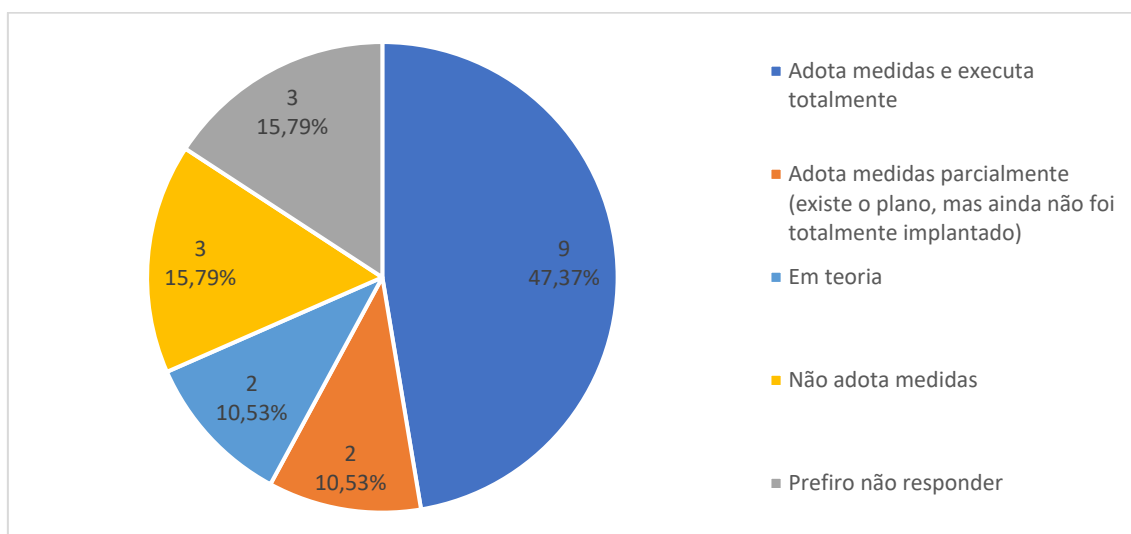
Sem haver a identificação e avaliação dos riscos, não é possível conhecer quais ameaças os repositórios estão mais propensos, assim como o grau de severidade de cada uma das ameaças, fazendo com que se crie uma falsa sensação de segurança, porém deixando falhas que podem ser exploradas em eventos de ataques e que podem comprometer a preservação digital dos seus materiais.

As respostas fornecidas não permitem conhecer exatamente as medidas adotadas nesses 10 repositórios, que podem englobar desde a verificação da força das senhas dos usuários, instalação de firewalls, por exemplo. Essas medidas são úteis e válidas, porém, a infraestrutura de TI de cada repositório é única e requer estudos individuais para que se conheçam as suas vulnerabilidades e seja possível eliminá-las ou diminuir as chances de acontecimentos.

A pergunta “6.4) A estrutura física do repositório adota medidas de segurança para evitar acesso indevido?” teve como objetivo identificar se as instalações físicas do repositório, como prédios e salas, são protegidas de acessos indevidos que podem levar a ataques que comprometam a preservação digital. As respostas à pergunta são exibidas na Figura 50, sendo possível notar que 3 repositórios, equivalente a 15,79% dos repositórios participantes, preferiram não fornecer uma resposta.

No caso de 9 repositórios, 47,37%, são adotadas na prática medidas de controle de acesso às instalações do repositório, enquanto 2 repositórios, 10,53%, declarou que as medidas estão implantadas parcialmente, ou seja, o planejamento sobre quais medidas adotar existe, porém nem todas as ações foram executadas ainda. Na conclusão do planejamento das medidas desses 2 repositórios, poderia se alcançar um total de 57,89% de repositórios com controle de acesso ao repositório, representando a maioria dos repositórios participantes.

Figura 50 – Adoção de medidas na estrutura física contra acessos indevidos.



Fonte: elaboração própria.

Em 2 repositórios, 10,53%, foi relatado haver somente em teoria um plano de controle de acesso às instalações do repositório, significando que na prática, não existe nenhum controle de acesso. Em 3 repositórios, 15,79%, não há o controle de acesso assim como não há um plano para implementar algum tipo de controle de acesso ao repositório. A soma desses dois casos permite identificar que 5 repositórios, 26,31%, não possuem nenhum tipo de controle de acesso ao repositório, deixando-os propensos a acessos que podem comprometer a preservação dos materiais custodiados.

Caso haja intenção, pessoas alheias à equipe do repositório podem acessar as instalações do repositório e executar ataques que podem corromper os dados a ponto de não haver meios para se recuperar o estado anterior. Ainda, pessoas mal intencionadas podem modificar documentos a seu próprio favor, assim como os mecanismos utilizados para a verificação de alterações, fazendo com que o repositório seja fonte de documentos que não sejam autênticos. Outro reflexo dessa fragilidade é o vazamento de informações sigilosas ou protegidas por propriedade intelectual e pelas leis reguladoras de dados pessoais (LGPD¹⁰⁴ e GDPR¹⁰⁵), por exemplo, causando prejuízos, muitas vezes incalculáveis aos interessados, além de tornar o próprio repositório um alvo para ações judiciais indenizatórias.

¹⁰⁴ (BRASIL, 2018)

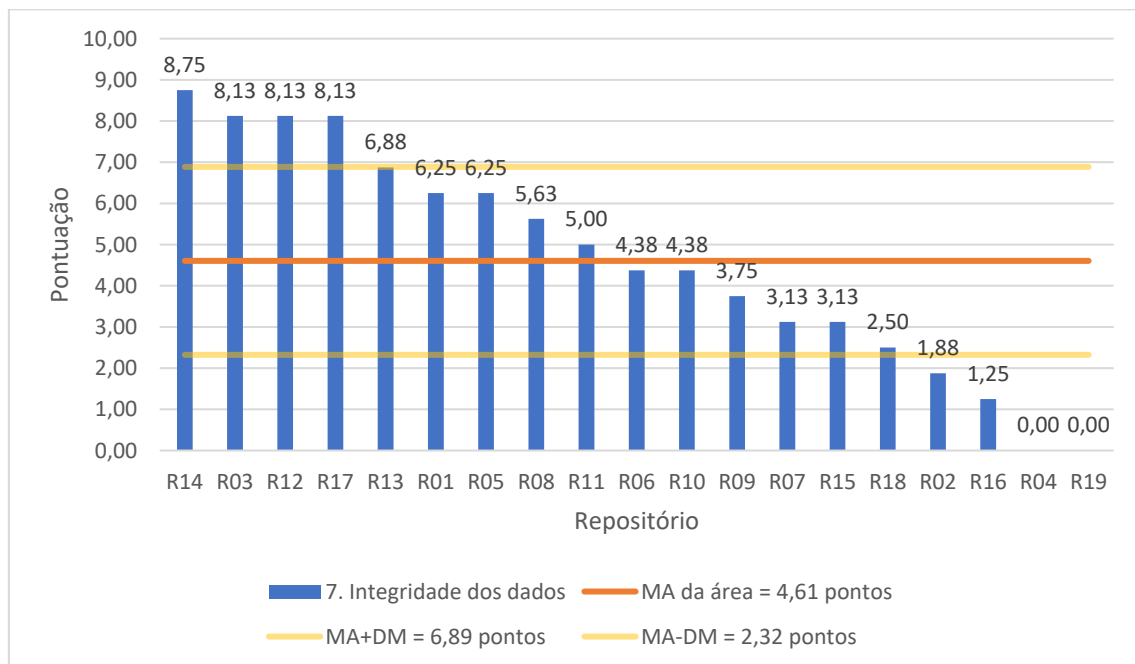
¹⁰⁵ (EUROPEAN COMISSION, 2018)

5.4.7 Resultados da Área 7: Integridade dos Dados

Nessa seção são analisadas as respostas das perguntas da Área 7: Integridade dos Dados, que analisa os mecanismos adotados pelos repositórios para garantir a integridade dos seus dados. Ao todo, a seção apresenta quatro perguntas que analisam desde o tratamento de formatos recebidos até as estratégias utilizadas para *backup*.

A Figura 51 exibe as pontuações totais dos repositórios nas perguntas da área 7 que teve como Média Aritmética, 4,61 pontos, Desvio Médio de 2,28 pontos em relação à MA e Coeficiente de Variação de 58,28%, indicando uma alta variabilidade das pontuações. O repositório R14 obteve a maior pontuação, com 8,75 pontos, enquanto a menor pontuação, 0,00 ponto, foi registrada para os repositórios R04 e R19, uma diferença de 8,75 pontos entre os dois extremos.

Figura 51 – Pontuações individuais dos repositórios na Área 7: Integridade dos Dados.



Fonte: elaboração própria.

O Quadro 54 exibe a classificação dos repositórios na área 7, de acordo com as suas faixas de pontuação. Pode-se notar que 3 repositórios se classificaram com desempenho excelente, enquanto 11 repositórios foram classificados com desempenho mediano. Um total de 5 repositórios foi classificado com desempenho inferior.

Quadro 54 – Classificação dos repositórios pelos seus desempenhos Área 7: Integridade dos Dados.

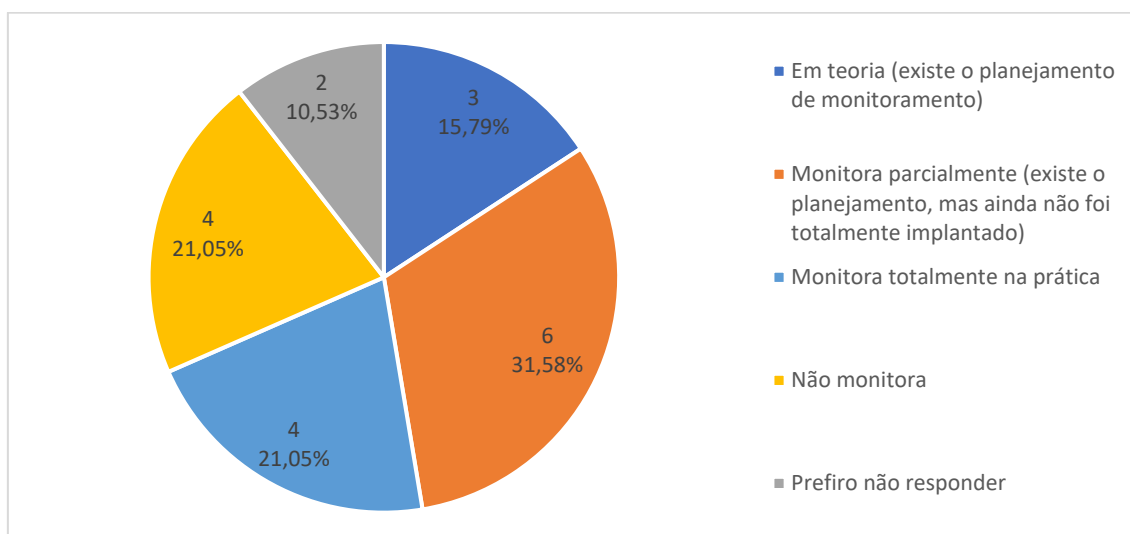
Desempenho excelente ($\geq 6,89$ pontos)	Desempenho mediano ($\geq 2,32$ e $< 6,89$ pontos)	Desempenho inferior ($< 2,32$ pontos)
1º) R14	3º) R05, R07, R10, R11 e R19	8º) R18
2º) R03, R12 e R17	4º) R17	9º) R02, R04, R09 e R16
	5º) R01	
	6º) R06 e R15	
	7º) R03 e R08	
Total: 3	Total: 11	Total: 5

Fonte: elaboração própria.

A pergunta “7.1) O repositório implementa o monitoramento e tratamento dos formatos de arquivos recebidos?” buscou identificar se os repositórios participantes possuem procedimentos para controle dos formatos de arquivos recebidos assim como o monitoramento para adaptação dos formatos à demanda da sua Comunidade-Alvo. O plano de preservação do repositório deve prever quais formatos de arquivos são aceitos, além dos procedimentos a serem executados para a migração, atualização e geração de derivadas de acessos dos documentos recebidos, tornando essas diretivas disponíveis à sua comunidade-alvo.

As respostas são exibidas na Figura 52, sendo possível notar que 2 repositórios, 10,53%, preferiram não responder à pergunta. Em 4 repositórios, 21,05%, nenhum procedimento é realizado para a verificação de formatos de arquivo, podendo-se concluir que quaisquer arquivos de quaisquer formatos são aceitos na durante o processo de ingestão. Esse comportamento pode levar o repositório a um alto grau de heterogeneidade de arquivos tornando difícil a gestão dos seus documentos, uma vez que qualquer envio é aceito. Uma situação similar deve ocorrer no caso de 3 repositórios, 15,79%, que declaram que o monitoramento de formatos existe somente em teoria, sem a aplicação prática.

Figura 52 – Monitoramento e tratamento dos formatos de arquivos.



Fonte: elaboração própria.

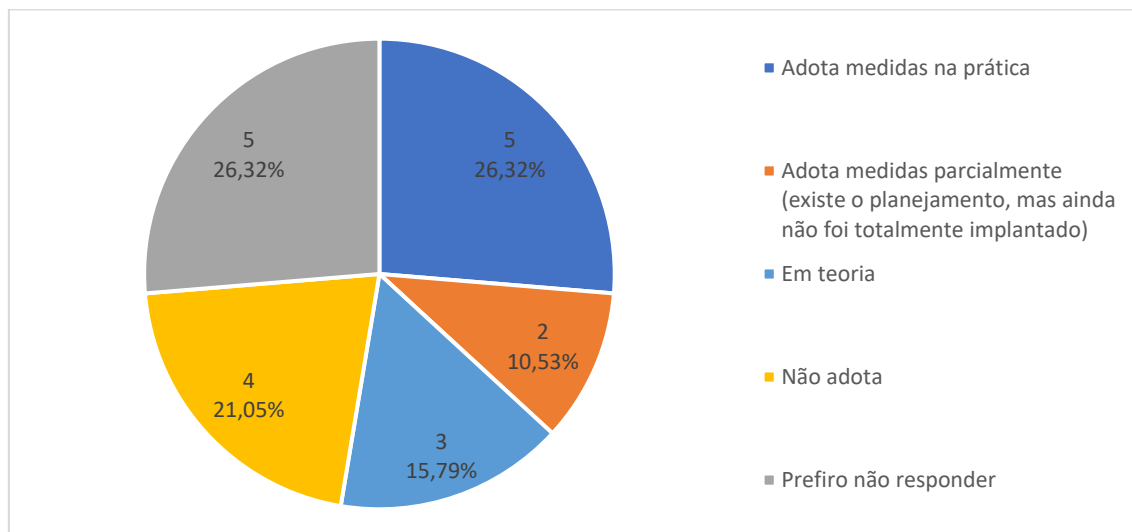
No caso de 6 repositórios, 31,58%, existe o monitoramento e tratamento dos formatos de arquivos recebidos, porém os planejamentos elaborados não se encontram totalmente aplicados na prática, não sendo possível se conhecer quais são efetivamente os procedimentos adotados nesses repositórios. Em 4 repositórios existe o monitoramento e tratamento de formatos totalmente aplicado na prática, representando 21,05% do total de repositórios participantes. Somando-se essas duas parcelas, tem-se um total de 10 repositórios que possuem processos voltados aos formatos de arquivos recebidos durante a ingestão de materiais, alcançando-se 52,63% dos repositórios.

Um fato interessante a ser notado é que 95% dos repositórios participantes declararam não possuir uma política de preservação ou um plano estratégico de preservação, que são os documentos que permitem um delineamento da Comunidade-Alvo do repositório, de onde se extraem as suas demandas e necessidades, os formatos de arquivos suportados e estratégias de preservação a serem empregadas. No entanto, 52,63% dos repositórios declararam monitorar e tratar os formatos de arquivos ingeridos no repositório, de onde se conclui que esses processos são realizados de maneira empírica e sem a devida formalização e, talvez, centrados nas pessoas que atualmente são responsáveis pelos repositórios.

A pergunta “7.2) O repositório adota medidas para prevenir o corrompimento dos dados?” buscou compreender como os repositórios se previnem contra o corrompimento dos dados sob sua custódia, a fim de que não haja perda de informações. As respostas são exibidas na Figura 53, sendo possível identificar que 5

repositórios preferiram não responder à questão, o que representa 26,32% do total de participantes. Dentre todas as perguntas do formulário da pesquisa, essa foi a que mais registrou o desejo dos participantes de não fornecer uma resposta.

Figura 53 – Adoção de medidas contra o corrompimento de dados.



Fonte: elaboração própria.

Em 4 repositórios, 21,05%, foi relatado que não existe nenhum tipo de medida voltada a prevenção de corrompimento de dados ao passo em que 3 repositórios, 15,79%, declararam que as medidas existem somente em teoria. Ao todo, 7 repositórios, equivalente a 36,84% dos repositórios participantes, não possuem aplicados procedimentos que impedem que os seus dados se percam, o que representa uma grave ameaça à preservação digital dos seus acervos. Sem a adoção dessas medidas, os materiais desses repositórios podem ser perdidos a partir de diversas causas como, por exemplo, falhas nos HDs dos servidores ou até mesmo durante a manipulação dos pacotes de informação como SIP, AIP e DIP.

Apenas 5 repositórios, 26,32% dos repositórios, adotam totalmente na prática medidas para prevenir que os seus dados se percam por corrompimento, enquanto 2 repositórios, 10,53%, possuem o planejamento dessas medidas, porém ainda não as aplicam totalmente na prática. Ainda que a implantação dessas medidas seja concluída, 7 repositórios teriam medidas de prevenção de corrompimento, o que representa apenas 36,84% do total de participantes.

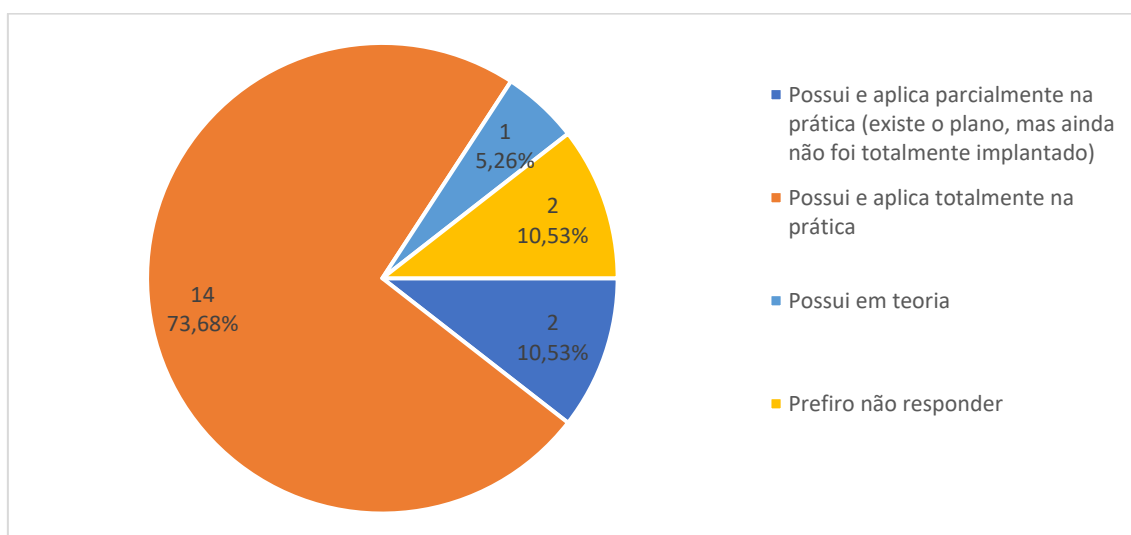
Esse fato é de extrema preocupação, uma vez que os arquivos, juntamente com os seus respectivos metadados, são o objeto principal da preservação digital e, caso eles sejam corrompidos, toda a estrutura, processos e equipamentos do repositório se

tornam inúteis. Os *softwares* utilizados pelos repositórios devem possuir ferramentas que gerem códigos de checagem (*checksum*) a partir dos bytes armazenados, assim como deve-se adotar procedimentos periódicos que verificam a consistência de todos os dados armazenados, fazendo com que se tenha conhecimento de eventuais corrompimentos o mais rápido possível e que as medidas de recuperação sejam acionadas imediatamente.

A pergunta “7.3) O repositório possui um plano de *backup*?” buscou identificar a presença de processos bem estabelecidos de *backups* que permitam que os dados estejam a salvo e possam ser recuperados plenamente na ocorrência de eventos que exijam esses esforços. Ao todo, 2 repositórios, 10,53% dos repositórios participantes, preferiram não responder à pergunta, a Figura 54 exibe os resultados.

Apenas 1 repositório, 5,26%, declarou que o plano de *backup* existe somente em teoria, sem aplicação na prática. A maioria dos repositórios, 14 ao todo, o que representa 73,68% dos repositórios participantes declarou existir um plano de *backup* que é aplicado totalmente na prática, enquanto 2 repositórios, 10,53%, sinalizaram que o plano de *backup* existe, mas não se encontra totalmente aplicado na prática.

Figura 54 – Presença de planos de *backup* nos repositórios.



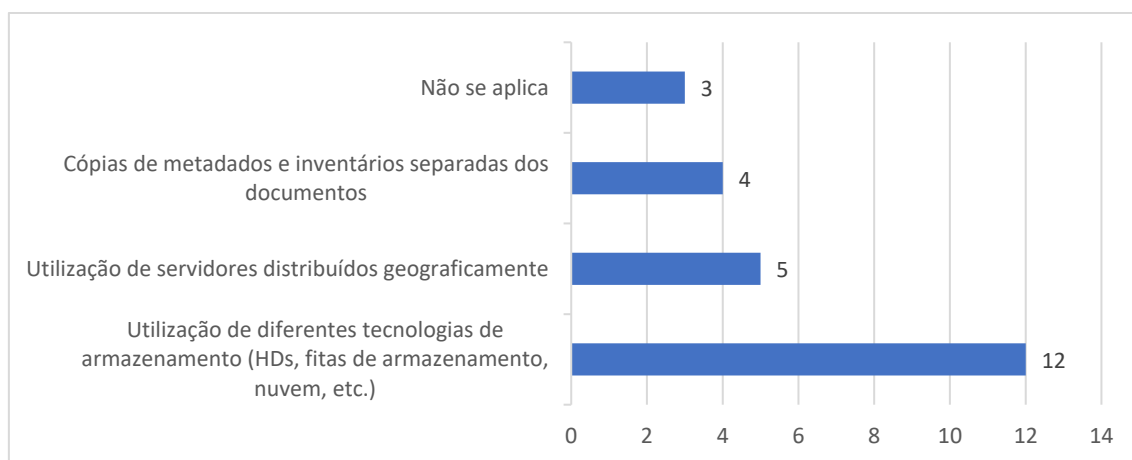
Fonte: elaboração própria.

Caso a implantação dos planos de *backup* desses 2 repositórios estivesse concluída, seria alcançado um total de 16 repositórios com aplicação prática total dos seus planos de *backup*, o que seria equivalente a 84,21% do total de repositórios. Ainda,

caso o repositório que possui o plano de *backup* somente em teoria o aplicasse na prática, se chegaria a 89,47% de repositórios que executam os seus planos de *backup* na prática totalmente. No entanto, cabe ressaltar que a pesquisa atual não possui detalhes sobre os motivos ou restrições que impedem a finalização ou a implantação dos planos de *backup* dos repositórios que relataram pendências quanto a aplicação prática, podendo haver restrições da estrutura organizacional, infraestrutura de TI ou, até mesmo, restrições orçamentárias, por exemplo.

A pergunta “7.4) Caso o repositório possua um plano de *backup*, quais estratégias são empregadas?” buscou identificar quais estratégias para salvaguardar os seus dados os repositórios utilizam, permitindo que os respondentes selecionassem mais de uma resposta ao mesmo tempo. A Figura 55 exibe o gráfico das 24 diferentes respostas recebidas. Tanto os repositórios que declararam não existir um plano de *backup* quanto o repositório que declarou existir o plano de *backup* somente em teoria selecionaram a resposta “Não se aplica” para essa pergunta, totalizando 3 repositórios, o que equivale a 15,79% dos participantes.

Figura 55 – Estratégias de *backup* adotadas pelos repositórios.



Fonte: elaboração própria.

No caso de 4 repositórios, 21,05%, é utilizada a estratégia de separação dos metadados dos documentos e o inventário dos próprios documentos, fazendo com que o repositório possa implementar de detecção de corrompimento nos documentos a partir dos seus metadados ou a partir do inventário que possui a relação de documentos sob a custódia do repositório. A utilização dessa estratégia permite uma independência em relação aos próprios documentos, uma vez que, caso eles sejam corrompidos, utiliza-se

os metadados e/ou os inventários para a descoberta das falhas e auxílio na recuperação dos dados.

A utilização de servidores distribuídos geograficamente é uma estratégia utilizada em 5 repositórios, 26,31%. Essa estratégia tem a tendência de aumentar a resiliência do repositório na ocorrência de eventos catastróficos que possam comprometer o local de guarda dos documentos. A matriz NDSA de níveis de preservação (LEVELS OF PRESERVATION REVISIONS WORKING GROUP, 2019, p. 1) possui a área funcional Armazenamento que avalia o nível em que os repositórios se encontram. No primeiro nível, avalia-se apenas a presença de cópias dos dados em localizações distintas, enquanto no nível 4, o mais sofisticado, avalia se os dados são armazenados em três cópias com diferentes localizações geográficas sujeitas as diferentes ameaças de desastres.

O rol de respostas fornecidas para a pergunta não permitiu a identificação da quantidade de cópias de backup utilizadas ou da verificação de localizações com tipos diferentes de ameaça, sendo possível aferir que, no mínimo, são utilizadas duas cópias que são armazenadas em locais distintos.

Em 12 repositórios, 63,16%, utiliza mais de uma tecnologia para ao armazenamento dos backups, fazendo com que esses repositórios possam contar com diferentes estratégias de recuperação dos dados em casos de perdas. A dependência em somente um tipo de tecnologia pode fazer com que o repositório seja suscetível a problemas de obsolescência, podendo fazer com que os dados sejam irrecuperáveis em determinados casos como o corrompimento de HDs, ou o comprometimento de equipamentos de escrita e leitura de fitas.

Utilizando-se a matriz de níveis de preservação NDSA para comparação dos resultados em relação à área funcional **Armazenamento**, no Nível 1 identifica-se o critério deles “**Possuir duas cópias completas em localizações separadas**” e no Nível 3 identifica-se o critério “**Possuir no mínimo uma cópia em um tipo de mídia de armazenamento diferente**”.

O Quadro 55 exibe o grau de alcance ou cumprimento desses critérios a partir das respostas dos repositórios. Para se determinar o cumprimento dos critérios, levou-se em consideração quantos repositórios declararam realizar procedimentos que os satisfizessem minimamente. Em ambos os níveis 1 e 3, a matriz especifica 3 critérios ao todo e, assumindo-se pesos iguais entre eles, se determinou o grau de cumprimento do nível em si.

Quadro 55 - Cumprimento de critérios da área funcional “Armazenamento”.

Área funcional	Nível 1	Nível 3
Armazenamento	“Possuir duas cópias completas em localizações separadas”	“Possuir no mínimo uma cópia em um tipo de mídia de armazenamento diferente”
Cumprimento do critério	18,75% (3/16 repositórios)	75% (12/16 repositórios)
Cumprimento do nível	6,25%	25%
Cumprimento REBIUN	89%	60%

Fonte: elaboração própria.

Ao se analisar os resultados na área funcional “Armazenamento”, identificados por Azorín *et al* sobre os repositórios REBIUN, 89% alcançaram o Nível 1, enquanto 60% dos repositórios alcançou o Nível 3 (2018, p. 3). Identifica-se assim um melhor desempenho dos repositórios espanhóis em relação à matriz de preservação NDSA, ainda que a conclusão final do estudo não tenha sido favorável (2018, p. 6, tradução nossa):

São detectados problemas graves quanto ao número de cópias de segurança e sua localização, verificações da integridade dos arquivos e capacidade de realizar auditorias em caso de detecção de incidentes de segurança. Não se aplica a nível técnico o compromisso que promete ter com a preservação dos materiais armazenados.

A partir das respostas da seção 7, é possível identificar que a grande maioria dos repositórios possuem planos de backup e utilizam diferentes estratégias visando salvaguardar os seus dados, por outro lado, a grande maioria não monitora os formatos dos arquivos armazenados e não adota medidas na prática para evitar o corrompimento dados. Essa situação pode levar a um cenário em que os bytes dos arquivos podem estar a salvo de catástrofes, por exemplo, porém serem inúteis por não serem mais acessíveis, tornando a preservação digital inútil. Dessa forma, recomenda-se a priorização dessas questões pelos repositórios.

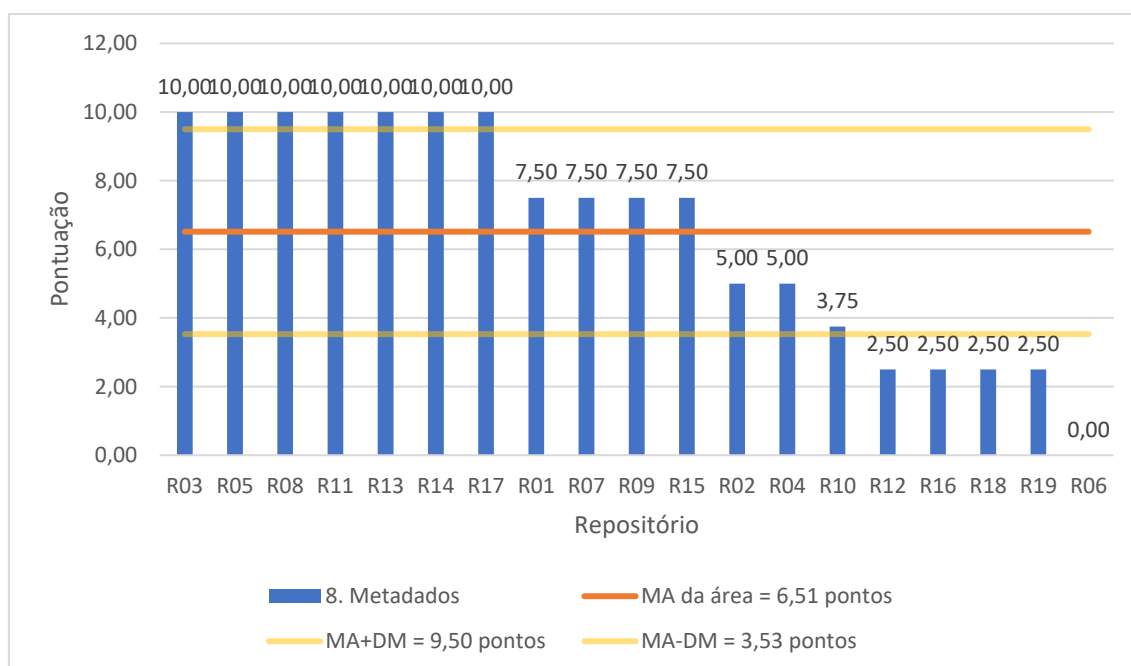
5.4.8 Resultados da Área 8: Metadados

Nessa seção, são analisados os resultados das perguntas relativas à seção 9 do formulário de perguntas, que trata do recolhimento de metadados. A seção possui duas perguntas e os resultados são mostrados a seguir.

A Figura 56 exibe as pontuações totais dos repositórios nas perguntas da área 8 que teve a maior Média Aritmética de todas as áreas, 6,51 pontos, Desvio Médio de 2,98 pontos em relação à MA e Coeficiente de Variação de 50,81%, indicando uma

alta variabilidade das pontuações. Os repositórios R03, R05, R08, R11, R13, R14 e R17 obtiveram a maior pontuação, com 10,00 pontos, sendo a área que apresentou mais repositórios com notas máximas. Esse fato indica que a área de metadados é bastante desenvolvida nos repositórios, caracterizando-se como um ponto forte em geral. A menor pontuação, 0,00 ponto, foi registrada para o repositório R06, uma diferença de 10,00 pontos entre os dois extremos.

Figura 56 – Pontuações individuais dos repositórios na Área 8: Metadados.



Fonte: elaboração própria.

O Quadro 56 exibe a classificação dos repositórios na área 8, de acordo com as suas faixas de pontuação.

Quadro 56 – Classificação dos repositórios pelos seus desempenhos Área 8: Metadados.

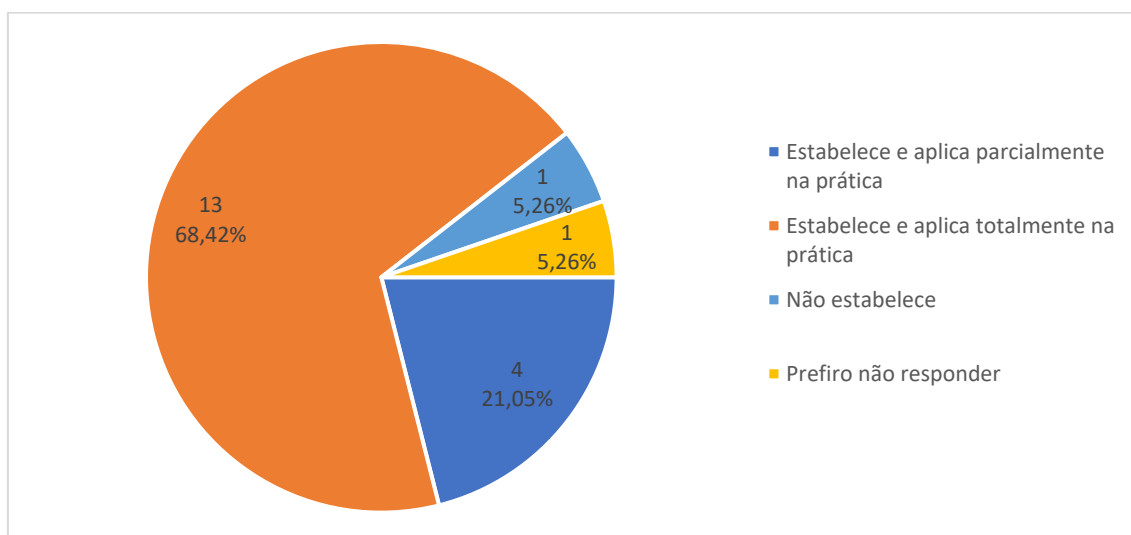
Desempenho excelente ($\geq 9,50$ pontos)	Desempenho mediano ($\geq 3,53$ e $< 9,50$ pontos)	Desempenho inferior ($< 3,53$ pontos)
1º) R03, R05, R08, R11, R13, R14 e R17	2º) R01, R07, R09 e R15	5º) R12, R16, R18 e R19
	3º) R02, R04	6º) R06
	4º) R10	
Total: 7	Total: 7	Total: 5

Fonte: elaboração própria.

Pode-se notar que 7 repositórios se classificaram com desempenho excelente, enquanto 7 repositórios foram classificados com desempenho mediano. Um total de 5 repositórios foi classificado com desempenho inferior.

A pergunta “8.1) O repositório estabelece um padrão de metadados reconhecido a ser adotado pelos documentos recebidos?” procurou identificar se existe uma padronização de metadados dos documentos recebidos pelos repositórios, fazendo com que os materiais possuam uma certa uniformização dos seus metadados. A Figura 57 exibe as respostas da pergunta, sendo que 1 repositório, equivalente a 5,26% dos repositórios participantes, preferiu não responder.

Figura 57 – Adoção de um padrão de metadados na ingestão dos documentos.



Fonte: elaboração própria.

Dentre os 19 repositórios participantes, 1 declarou não haver um padrão de metadados a ser seguido pelos produtores durante a ingestão dos documentos, equivalendo a 5,26% dos participantes. No caso de 4 repositórios, 21,05%, foi respondido que existe um padrão de metadados, porém ele é aplicado parcialmente na prática. Nesses casos, os repositórios podem ter estabelecido um padrão de metadados a ser seguido, no entanto, na prática, nem todos os valores dos metadados são fornecidos ou estão disponíveis durante o recebimento dos materiais.

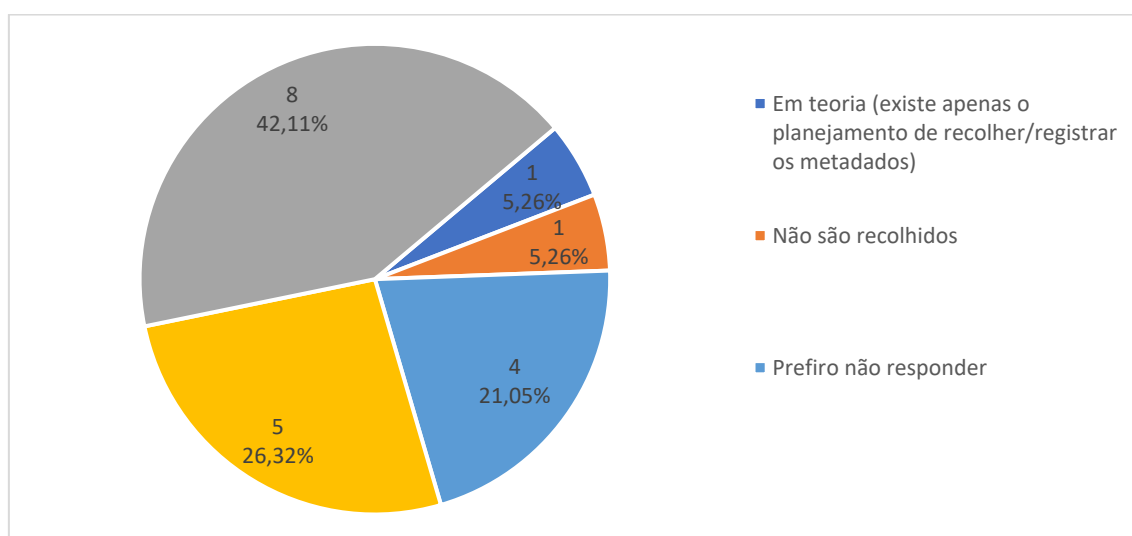
No caso de 13 repositórios, 68,42%, a grande maioria, existe um padrão de metadados a ser seguido e ele é implantando totalmente na prática, concluindo-se que todos os materiais que são recebidos nesses repositórios possuem uma uniformização nos conjuntos de metadados aceitos que são preservados juntos aos respectivos documentos. O fato de a grande maioria adotar um padrão de metadados na ingestão se

relaciona com as respostas da pergunta 4.3, que analisou os padrões, de um modo geral, utilizados pelo repositório.

A pergunta 3.3 registrou que 18 repositórios, quase a totalidade dos 19 repositórios participantes, utiliza o padrão Dublin Core de metadados para os seus materiais. Além desse, os padrões ISAD(G) e NOBRADE também registraram respostas, sinalizando uma grande adesão dos repositórios aos padrões de metadados e descrição de documentos. No entanto, com os dados recolhidos pelo estudo atual não há meios de se confirmar a real qualidade dos metadados fornecidos ou recolhidos nesses casos.

A pergunta “8.2) São recolhidos/registrados metadados que permitam identificar o contexto, referência, direitos e autenticidade dos documentos?”, tendo os resultados apresentados na Figura 58. Foi registrado que 4 repositórios, 21,05% do total de participantes, preferiram não responder à pergunta.

Figura 58 – Recolhimento de metadados de contexto dos documentos.



Fonte: elaboração própria.

Em 8 repositórios, 42,11%, foi relatado haver o recolhimento de metadados referentes aos contextos dos documentos recebidos, sendo possível identificar as respectivas proveniências, direitos e informações que permitem atestar as suas autenticidades. Em 5 repositórios, 26,32%, existe um planejamento para recolhimento desses metadados, porém os procedimentos práticos ainda não foram totalmente implantados. Pode-se chegar a um cenário de 13 repositórios, 68,42%, com recolhimento dos metadados necessários, caso a implantação dos planejamentos seja

concluída nos 5 repositórios que relataram a existência parcialmente aplicada dos planejamentos.

No caso de 1 repositório, 5,26%, existe apenas em teoria o planejamento de recolhimento desses metadados, assim como 1 repositório relatou não haver o recolhimento dos metadados. No caso desses 2 repositórios, os documentos recebidos não possuem informações suficientes para a análise de suas proveniências a partir dos metadados recolhidos. Sendo assim, torna-se difícil a presunção de autenticidade dos materiais desses repositórios, fazendo com que eles não possam ser usados como fontes de prova, por exemplo.

A situação desses repositórios pode correr o risco de estar usando uma estrutura organizacional e tecnológica para preservar documentos que não podem ter a sua autenticidade comprovada. Os gestores dos repositórios que recolhem documentos sem os metadados necessários à preservação devem ser informados dessa situação para que as devidas providências sejam aplicadas e o prejuízo de seja minimizado no acervo já presente nos repositórios digitais.

Ao se analisar os resultados das respostas de ambas as perguntas em relação à matriz de níveis de preservação NDSA, pode-se utilizar a área funcional “**Metadados**”, que possui no Nível 2 somente o critério: “**Armazene metadados suficientes para saber o que é o conteúdo (isso pode incluir alguma combinação de metadados administrativos, técnicos, descritivos, de preservação e estruturais)**”. No Nível 3, utiliza-se o critério: “**Determine quais padrões de metadados aplicar**” de 2 existentes. Por fim, no Nível 4, o critério: “**Implemente os padrões de metadados escolhidos**” de 2 existentes (LEVELS OF PRESERVATION REVISIONS WORKING GROUP, 2019).

O Quadro 57 exhibe o grau de alcance ou cumprimento desses critérios a partir das respostas dos repositórios. Para se determinar o cumprimento dos critérios, levou-se em consideração quantos repositórios declararam realizar procedimentos que os satisfaçam minimamente. No Nível 2, existe somente 1 critério, enquanto os níveis 3 e 4, possuem ambos 2 critérios cada um, assumindo-se pesos iguais para os seus cumprimentos entre si. Assim, a partir do total de 19 repositórios que responderam às perguntas, 42,10% cumprem o critério do Nível 2, enquanto 47,74% cumprem o Nível 3 e 34,21% cumprem o Nível 4.

Quadro 57 – Cumprimento de critérios da área funcional “Metadados”.

Nível	Metadados	Cumprimento do critério	Cumprimento do nível	Cumprimento REBIUN
Nível 2	“Armazene metadados suficientes para saber o que é o conteúdo (isso pode incluir alguma combinação de metadados administrativos, técnicos, descritivos, de preservação e estruturais)”	42,10% (8/19 repositórios)	42,10%	59%
Nível 3	“Determine quais padrões de metadados aplicar”	94,74% (18/19 repositórios)	47,37%	81%
Nível 4	“Implemente os padrões de metadados escolhidos”	68,42% (13/19 repositórios)	34,21%	44%

Fonte: elaboração própria.

Ao se comparar os resultados com a avaliação dos repositórios REBIUN em relação à área funcional Metadados, no Nível 2, 59% dos repositórios espanhóis cumprem com os requisitos do nível. No Nível 3, 81% repositórios cumprem com os requisitos e no Nível 4, 44% declararam cumprir os requisitos (AZORÍN, BARRUECO, *et al.*, 2018). No caso dos repositórios REBIUN, notam-se resultados bastante superiores aos do presente estudo, ressaltando-se a grande disparidade do Nível 3, enquanto a diferença se torna menor no Nível 4, sendo inferior a 10 pontos percentuais. No geral, em relação à área funcional Metadados, nota-se um desempenho inferior dos repositórios brasileiros em relação aos espanhóis.

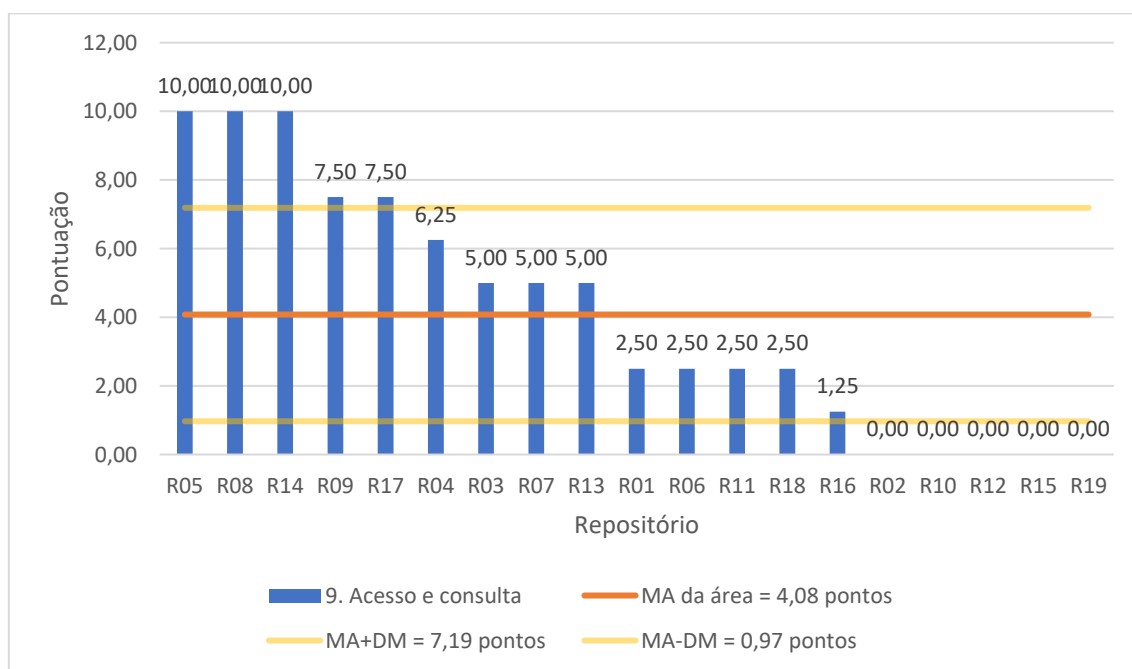
A partir das respostas, é possível determinar que a área 8 foi a que obteve maior pontuação média geral dos repositórios, com 6,51 pontos, tendo 7 repositórios atingido a pontuação máxima. Trata-se, portanto, de um ponto forte dos repositórios no que se refere à utilização de padrões de metadados e recolhimento de metadados de contexto dos seus materiais digitais preservados. Sugere-se o estabelecimento de trabalhos de cooperação entre os repositórios com desempenho excelente e os demais repositórios, a fim de que haja uma melhoria geral da pontuação.

5.4.9 Resultados da Área 9: Acesso e Consulta

Nessa seção são analisadas as respostas da área 9, do formulário da pesquisa que avalia critérios de acesso e consulta aos documentos preservados nos respectivos repositórios. Ao todo, foram disponibilizadas três perguntas aos respondentes sobre o tema.

A Figura 59 exibe as pontuações totais dos repositórios nas perguntas da área 8 que teve a maior Média Aritmética de todas as áreas, 4,08 pontos, Desvio Médio de 3,11 pontos em relação à MA e Coeficiente de Variação de 86,59%, o segundo maior CV de todas as áreas indicando uma altíssima variabilidade das pontuações. Os repositórios R05, R08, e R14 obtiveram a maior pontuação, com 10,00 pontos, sendo a segunda área que apresentou mais repositórios com notas máximas. A menor pontuação, 0,00 ponto, foi registrada para os repositórios R02, R10, R12, R15 e R19, uma diferença de 10,00 pontos entre os dois extremos. A área 9 foi a segunda que mais registrou pontuações zero dentre todas as áreas avaliadas, caracterizando-a como uma deficiência nos repositórios, recomendando-se uma priorização de tratamento e atendimento aos critérios avaliados nessa área, que trata das políticas de acesso e o monitoramento das necessidades das Comunidades-Alvo dos repositórios.

Figura 59 – Pontuações individuais dos repositórios na Área 9: Acesso e Consulta.



Fonte: elaboração própria.

O Quadro 58 exibe a classificação dos repositórios na área 9, de acordo com as suas faixas de pontuação. Pode-se notar que 5 repositórios se classificaram com desempenho excelente, enquanto 9 repositórios foram classificados com desempenho mediano. Um total de 5 repositórios foi classificado com desempenho inferior.

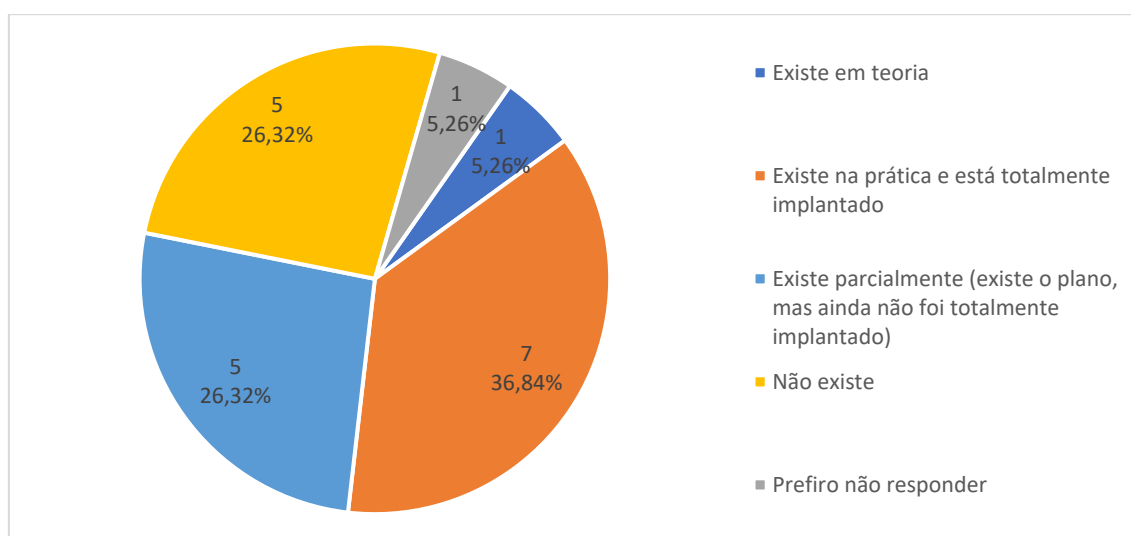
Quadro 58 – Classificação dos repositórios pelos seus desempenhos Área 8: Metadados.

Desempenho excelente ($\geq 7,19$ pontos)	Desempenho mediano ($\geq 0,97$ e $< 7,19$ pontos)	Desempenho inferior ($< 0,97$ pontos)
1º) R05, R08 e R14	3º) R04	7º) R02, R10, R12, R15 e R19
2º) R09 e R17	4º) R03, R07 e R13	
	5º) R01, R06, R11 e R18	
	6º) R16	
Total: 5	Total: 9	Total: 5

Fonte: elaboração própria.

A pergunta “9.1) Existe uma política de acesso aos objetos digitais?” buscou analisar se os repositórios estabelecem políticas de acesso, disponibilizando os materiais de acordo com os interesses das suas comunidades-alvo. Da mesma forma, a política de acesso deve conter as regras de restrição de documentos, quando houverem, e os respectivos procedimentos a serem realizados para que as pessoas acessem esses documentos. A Figura 60 exibe os resultados, sendo registrado que 1 repositório, o equivalente a 5,26% dos repositórios participantes, preferiu não responder à pergunta.

Figura 60 – Existência de política de acesso nos repositórios.



Fonte: elaboração própria.

Em 7 repositórios, 36,84%, foi relatada a existência de uma política de acesso totalmente implementada na prática, enquanto 5 repositórios, 26,32%, declararam que a política de acesso existe, porém não se encontra totalmente implantada na prática. No caso de 1 repositório, 5,26%, foi relatado que a política de acesso foi elaborada, porém não é aplicada na prática, existindo somente em teoria. Em 5

repositórios, 26,32%, não existe nenhuma política de acesso aos materiais digitais preservados.

A elaboração de uma política de acesso aos materiais deve prever quais tipos de materiais, e seus formatos, são disponibilizados à comunidade-alvo do repositório, trazendo utilidade ao acervo custodiado. Somente através de uma regulamentação de acesso pode-se certificar de que os propósitos do repositório são cumpridos. A política de acesso deve ainda especificar sob quais condições os materiais podem ser consultados ou quais sob quais justificativas o repositório não permite a consulta aos seus materiais.

Quando for o caso, a política pode ainda estabelecer o conjunto de instruções e procedimentos a serem seguidos por usuários que desejam acessar materiais que sejam protegidos, restritos ou sigilosos, por exemplo. À critério de cada instituição, a política de acesso pode, ainda, estabelecer custos a serem aplicados sobre as consultas realizadas no repositório.

Outro aspecto passível de ser regulamentado na política de acesso é a integração com outros *softwares*, especificando quais os requisitos a serem seguidos para eventuais interessados possam realizar integrações com o repositório. A utilização de APIs pode promover uma interoperabilidade entre sistemas de TI, que podem permitir, por exemplo, a comunicação entre o sistema de gestão de documentos adotado por uma instituição e o seu repositório para que a submissão de materiais ocorra de maneira automatizada. Dependendo da ferramenta de preservação implantada no repositório, pode se decidir pela adoção de protocolos de colheita de metadados como, por exemplo, o OAI-PMH¹⁰⁶, especificando-se na política de acesso quais dados estarão disponíveis para a colheita e sob quais condições essa tarefa pode ocorrer.

A política de acesso desenvolvida pelos repositórios deve, preferencialmente, se basear nos princípios FAIR (WILKINSON, DUMONTIER, *et al.*, 2016). Os materiais custodiados pelos repositórios devem poder ser localizados, acessíveis, interoperáveis, no sentido de procurar não ser dependente de uma tecnologia exclusiva e pouco conhecida, além de promover o reuso dos materiais por outras pessoas ou outros propósitos diferentes dos quais os originaram. É possível perceber que para se alcançar o propósito de disponibilizar o acesso dos seus materiais, é fundamental que a Comunidade-Alvo do repositório esteja definida e bem delimitada, caso contrário,

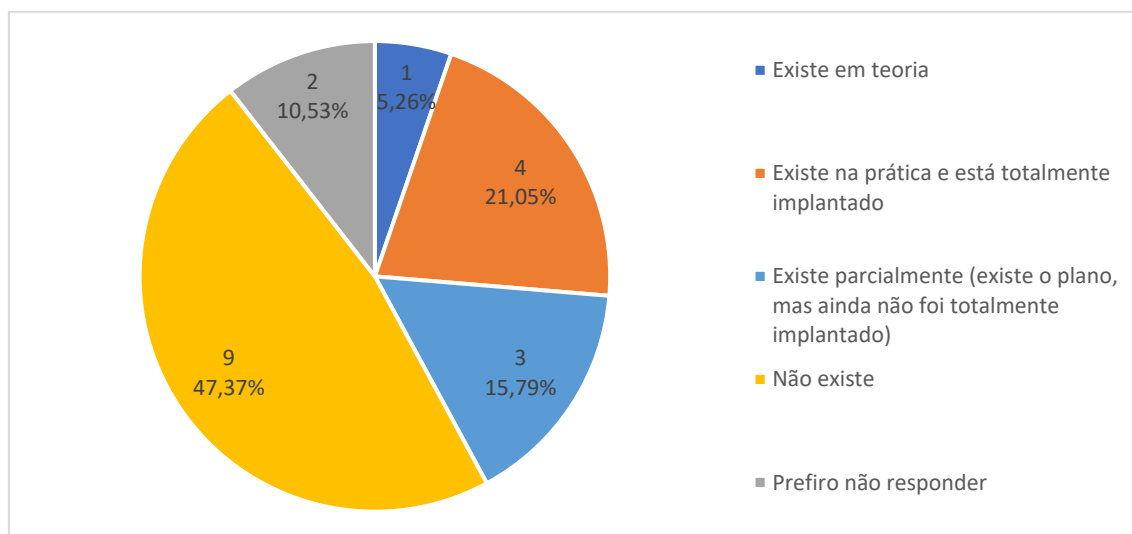
¹⁰⁶ Open Archive Initiative – Protocol for Metadata Harvesting (OPEN ARCHIVES INITIATIVE, 2023).

corre-se o risco de o repositório não disponibilizar os materiais que sejam de seu interesse ou disponibilizá-los de maneira ou formatos inapropriados.

A pergunta “9.2) Existe uma política de monitoramento das necessidades da sua Comunidade-Alvo a fim de disponibilizar os documentos da forma correta e nos formatos de arquivo mais adequados?” buscou compreender como são feitas as escolhas de formatos para a disponibilização dos materiais dos repositórios. As respostas são apresentadas na Figura 61, sendo possível perceber que 2 repositórios, equivalente a 10,53% dos participantes, preferiram não responder à pergunta.

Em 4 repositórios, 21,05%, foi relatada a existência de um plano totalmente implantado de monitoramento das necessidades da Comunidade-Alvo. Em 3 repositórios, 15,79%, o plano existe, porém a sua aplicação não se encontra executada totalmente na prática. No caso de 1 repositório, o planejamento para monitoramento das necessidades da sua Comunidade-Alvo existe, mas as ações práticas não se encontram executadas.

Figura 61 – Existência de um plano de monitoramento das necessidades das Comunidades-Alvo.



Fonte: elaboração própria.

Em 9 repositórios, 47,37%, os respondentes relataram a não existência de qualquer plano de monitoramento das necessidades das suas Comunidades-Alvo. Logo, em 10 repositórios, 52,63%, ou seja, a maioria, os gestores não buscam compreender a forma como os seus usuários utilizam os materiais disponibilizados. Há que se ressaltar que, em um determinado instante de tempo, os repositórios podem ter identificado as melhores formas e formatos de arquivos a serem disponibilizados, no entanto, não

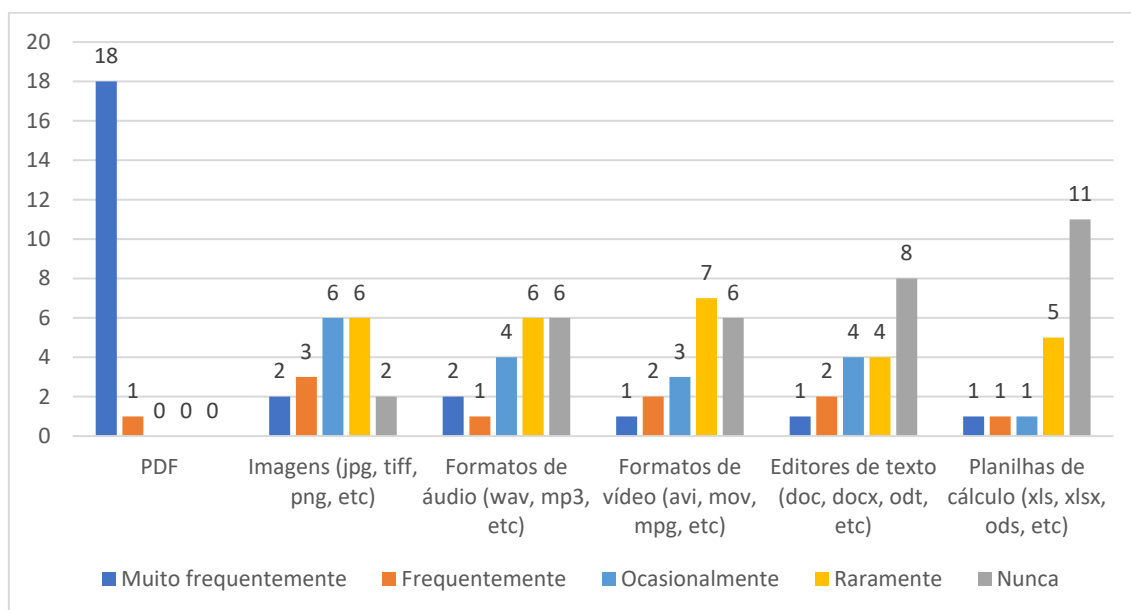
existe um processo contínuo de monitoramento das mudanças que podem ter ocorrido desde então.

À medida em que a tecnologia avança, novos formatos de arquivo ou novas versões de formatos são disponibilizados, fazendo com que haja a necessidade de o repositório se atualizar, buscando incluir essas evoluções nos requisitos da ingestão, preservação e disseminação dos seus materiais. Há que se dedicar atenção especial aos nichos de usuários que podem fazer parte da Comunidade-Alvo dos repositórios, que utilizam formatos de arquivo específicos para as realizações das suas pesquisas necessitam preservá-los. Dentre essas peculiaridades, pode-se citar os materiais utilizados para a realização de cálculos gráficos e estatísticos do *software* R, que possuem um formato de arquivo próprio, vastamente utilizado e conhecido (R FOUNDATION, 2022).

Outro aspecto que pode ser identificado através do monitoramento das necessidades da Comunidade-Alvo são os formatos dos arquivos disponibilizados no acesso e consulta do acervo. Através dessa ferramenta, pode-se identificar que em determinados casos o público não necessita de imagens com altíssima resolução para o acesso a documentos digitalizados, sendo possível a geração de derivadas de acesso que possuam alta qualidade, porém em um formato de arquivo compactado que consome menos espaço e menor consumo de banda de rede, por exemplo.

A pergunta “9.3) Quais os formatos de arquivo mais habituais que o repositório costuma disponibilizar aos usuários?” apresentou uma matriz de grupos de formatos mais comuns de arquivos e solicitou aos respondentes que informassem a frequência com que cada arquivo é disponibilizado à sua Comunidade-Alvo. O objetivo dessa pergunta foi identificar quais os tipos de arquivos mais utilizados pelos repositórios, procurando compreender a forma como os documentos são preservados. Foram registradas respostas para todos os repositórios participantes da pesquisa, os resultados são exibidos na Figura 62.

Figura 62 – Frequência de formatos de arquivos disponibilizados



Fonte: elaboração própria.

O formato de arquivo mais disponibilizado por todos os repositórios é o PDF, observando-se que em 18 repositórios esse formato é disponibilizado muito frequentemente, enquanto 1 repositório sinalizou que disponibiliza esse formato frequentemente. Somando-se as duas respostas, 100% dos repositórios participantes disponibilizam frequentemente documentos em formato PDF à sua Comunidade-Alvo. Dentre as vantagens desse formato de arquivo está o alto grau de presença de programas e aplicativos nos dispositivos para a sua leitura, facilitando o seu acesso e utilização. A pesquisa atual não verificou o formato específico de PDF utilizado pelos repositórios, no entanto, é sabido que o PDF/A é o formato mais adequado à preservação, uma vez que traz embutidas informações que permitem que o programa leitor exiba o documento de uma maneira mais fiel possível àquele gerado e submetido pelo produtor.

A disponibilização pelos repositórios por documentos de imagens, em seus mais variados formatos, foi registrada com ocorrência ocasional ou rara, tendo 12 repositórios registrado essas duas respostas, o que equivale a 63,16% dos repositórios participantes. A pesquisa atual não focou em identificar exatamente os formatos de arquivos das imagens disponibilizadas ou as suas adequações à preservação digital. Em 2 repositórios, 10,53%, foi registrado que arquivos de imagens nunca são disponibilizados.

Os documentos do tipo arquivos de áudio, em seus mais diversos formatos específicos, são pouco disponibilizados pelos repositórios, tendo 6 repositórios, 31,58%,

relatado que nunca disponibilizam esses tipos de arquivo enquanto outros 6 repositórios relataram uma rara disponibilização desses arquivos.

Os documentos de vídeo registraram apenas 1 repositório, 5,26%, com disponibilização muito frequente de materiais com os formatos desse tipo de arquivo. Em 7 repositórios, 36,84%, foi relatada uma rara disponibilização desses materiais, enquanto em 6 repositórios, 31,58%, os materiais de vídeo nunca são disponibilizados.

Os arquivos em formato de editores de texto como, por exemplo, Microsoft Word¹⁰⁷ (doc e docx) ou LibreOffice Writer¹⁰⁸ (odt) nunca são disponibilizados em 8 dos repositórios participantes, representando 42,10%, do total de participantes. Outros 4 repositórios, 21,05%, relataram uma rara disponibilização desses materiais e outros 4 repositórios informaram que ocasionalmente disponibilizam materiais nesses formatos.

O último formato analisado foi o de planilhas de cálculo, geradas por programas como o Microsoft Excel¹⁰⁹ (xls exlsx) e LibreOffice Calc¹¹⁰ (ods), por exemplo. Nesse caso, 11 repositórios, o equivalente a 57,89% dos participantes, relatou que nunca disponibiliza arquivos nesses formatos. Outros 5 repositórios, 26,31%, informaram que raramente disponibilizam planilhas de cálculo, nos seus formatos originais, às suas Comunidades-Alvo.

Mesmo que todos os tipos de arquivos sugeridos pela pesquisa possuam formatos adequados à preservação, conhecidos como formatos sustentáveis¹¹¹, o formato de arquivo PDF é o mais utilizado para a disponibilização dos materiais preservados. Acredita-se que, muito provavelmente, os materiais disponibilizados dessa forma sejam do gênero textual, sobressaindo-se aos outros formatos de editores de texto que também englobam esse gênero. Ao todo, 1 repositório os disponibiliza muito frequentemente e 2 repositórios os disponibilizam frequentemente, somando-se os casos, tem-se apenas 15,79% do total de participantes.

Dentre os 19 repositórios participantes, 6, o equivalente a 31,58%, registraram que disponibilizam todos os tipos de formatos de arquivo sugeridos, ainda que em graus de frequência diferentes. Os outros 13 repositórios, 68,42%, relataram não disponibilizar nunca algum dos tipos de formatos sugeridos nas respostas da pergunta.

¹⁰⁷ <https://www.microsoft.com/pt-br/microsoft-365/word>

¹⁰⁸ <https://pt-br.libreoffice.org/descubra/writer/>

¹⁰⁹ <https://www.microsoft.com/pt-br/microsoft-365/excel>

¹¹⁰ <https://pt-br.libreoffice.org/descubra/calc/>

¹¹¹ (NATIONAL DIGITAL STEWARDSHIP ALLIANCE, 2022, p. 8)

A pergunta “9.3.1) Existem outros formatos habituais disponibilizados aos usuários?” buscou permitir aos respondentes informarem a disponibilização de outros formatos de arquivos diferentes dos sugeridos na pergunta anterior. Foram registradas respostas de 2 repositórios para essa pergunta, indicando das suas particularidades.

A resposta de um repositório declarou que disponibiliza materiais em formato HTML. Esse formato é utilizado para a visualização de informações através de um programa navegador web, popularizando-se com a grande adoção da internet pelo público geral (W3C, 2022). A renderização de arquivos nesse formato utiliza arquivos CSS (*Cascading Style Sheets* – Folhas de Estilo em Cascata) para a decoração dos elementos criados e a sua interpretação é fortemente dependente do dispositivo e programa navegador utilizados. Essas características se adaptam à natureza do ambiente web por trazer dinamicidade e flexibilidade ao conteúdo, no entanto, a forma como o conteúdo é exibido não é fixa, variando conforme as características do dispositivo que lê os arquivos. Assim, os pilares para o estabelecimento da autenticidade dos documentos, que é a presença de forma fixa e conteúdo estável (DURANTI, MACNEIL, *et al.*, 2005, p. 1), ficam comprometidos ao se utilizar arquivos nesse formato.

O segundo repositório que declarou preservar outros formatos de arquivos citou os seguintes formatos:

- a) epub: formato aberto para a publicação de e-books baseado em XML que, entre outras possibilidades, permite que a visualização do conteúdo se adapte ao dispositivo de leitura (W3C, 2022);
- b) tex: formato de arquivo em texto voltado principalmente para a redação de textos como livros e artigos, permitindo a inclusão de fórmulas matemáticas e a escolha de tipografia mais adequada, por exemplo (VULIS, 1992);
- c) tar, xz e zip: formatos de arquivos que empacotam outros arquivos e permitem compressão de tamanho, diminuindo o espaço físico ocupado e facilitando a troca de informações.

Assim como os arquivos em formato HTML não possuem uma forma fixa e conteúdo estável, os arquivos em formato epub e tex dependem de estilização para a renderização dos seus conteúdos. O formato epub é fortemente dependente do software e dispositivos utilizados para a sua leitura, enquanto o formato tex depende da aplicação de estilos de formatação da tipografia e criação de gráficos que pode variar entre os pacotes de dependências usados e entre plataformas. Cabe lembrar que a preservação

nos casos desses arquivos pode preservar o conteúdo original, uma vez que se baseiam em texto e marcações de formatações, no entanto, a representação e interpretação desses conteúdos podem se tornar um desafio ao longo do tempo.

Os formatos de arquivos compactados permitem uma menor ocupação do espaço físico de armazenamento do repositório, permitindo a otimização da infraestrutura ocupada. No entanto, a preservação deve se centrar nos arquivos contidos dentro dos pacotes comprimidos, utilizando procedimentos que permitam a extração dos seus conteúdos para posterior manipulação execução de verificações de integridade, varreduras de vírus, etc. Ainda, o repositório deve se certificar que as compressões efetuadas não resultam em perdas de dados que possam comprometer a interpretabilidade dos arquivos, assim como as respectivas autenticidades.

Os resultados da área 9 demonstram uma baixa preocupação dos repositórios em regular o acesso aos seus materiais, delimitando a política, requisitos e regras a serem utilizadas, por mais que todos os repositórios forneçam acesso dos seus materiais às suas Comunidades-Alvo. Ainda, somente 4 repositórios declarou possuir um plano de monitoramento das necessidades das Comunidades-Alvo, o que pode significar que os demais repositórios não disponibilizam os formatos mais adequados ou não estão atentos às possíveis mudanças de requisitos nesse sentido, distanciando os repositórios da realidade dos seus usuários. Assim, sugere-se que os repositórios apliquem esforços para atender a essa área. Como aspecto positivo identificado, está o fato de 100% dos repositórios utilizarem o formato PDF para difusão dos seus materiais, uma vez que esse formato de arquivo possui fácil utilização e uma grande gama de *softwares* que o leem.

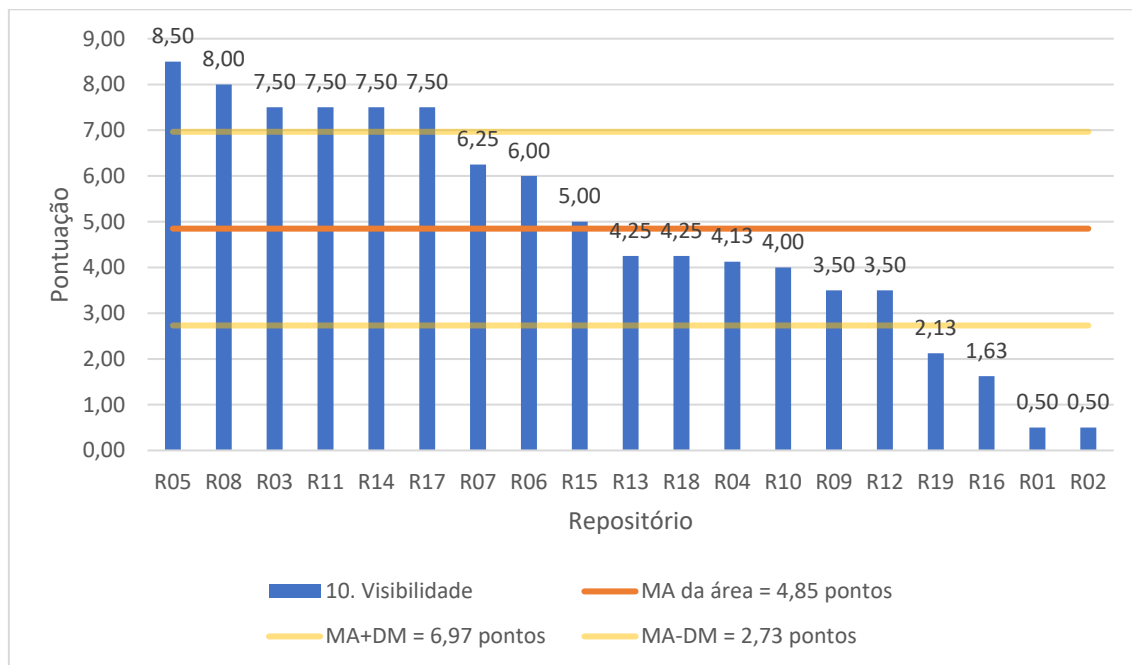
5.4.10 Resultados da Área 10: Visibilidade dos Repositórios

A seguinte seção analisa os resultados obtidos pelas respostas da área 10 do formulário de perguntas, que trata sobre a visibilidade do repositório mediante a sua comunidade-alvo. A todo, a seção conta com quatro perguntas que verificam as estratégias utilizadas para a promoção do uso do repositório e sua forma de comunicação com a sua Comunidade-Alvo.

A Figura 63 exhibe as pontuações totais dos repositórios nas perguntas da área 10 que teve Média Aritmética de 4,85 pontos, Desvio Médio de 2,12 pontos em relação à MA e Coeficiente de Variação de 50,94%, indicando uma variabilidade média

das pontuações. O repositório R05 obteve a maior pontuação, com 8,50 pontos, a menor pontuação, 0,50 ponto, foi registrada para os repositórios R01 e R02, uma diferença de 8,00 pontos entre os dois extremos.

Figura 63 – Pontuações individuais dos repositórios na Área 10: Visibilidade.



Fonte: elaboração própria.

O Quadro 59 exibe a classificação dos repositórios na área 10, de acordo com as suas faixas de pontuação. Pode-se notar que 6 repositórios se classificaram com desempenho excelente, enquanto 9 repositórios foram classificados com desempenho mediano. Um total de 4 repositórios foi classificado com desempenho inferior.

Quadro 59 – Classificação dos repositórios pelos seus desempenhos Área 10: Visibilidade.

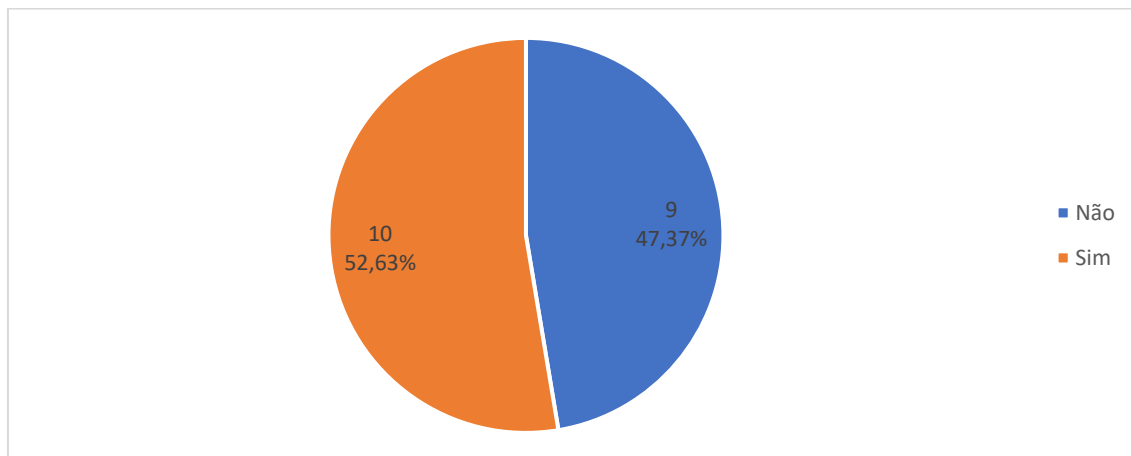
Desempenho excelente ($\geq 6,97$ pontos)	Desempenho mediano ($\geq 2,73$ e $< 6,97$ pontos)	Desempenho inferior ($< 2,73$ pontos)
1º) R05	4º) R07	11º) R19
2º) R08	5º) R06	12º) R16
3º) R03, R11, R14 e R17	6º) R15	13º) R01 e R02
	7º) R13 e R18	
	8º) R04	
	9º) R10	
	10º) R09 e R12	
Total: 6	Total: 9	Total: 4

Fonte: elaboração própria.

Os resultados da pergunta “10.1) É possível localizar o acesso ao repositório a partir da página principal da instituição?” são exibidos na Figura 64. No caso de 10

repositórios, o equivalente a 52,63% dos repositórios participantes, é possível encontrar na página principal da instituição alguma forma de acessar o repositório de materiais. Por outro lado, em 9 repositórios, 47,37%, nas páginas institucionais não é possível encontrar uma forma que permita ao usuário se direcionar ao repositório.

Figura 64 – Acesso ao repositório a partir da página institucional.



Fonte: elaboração própria.

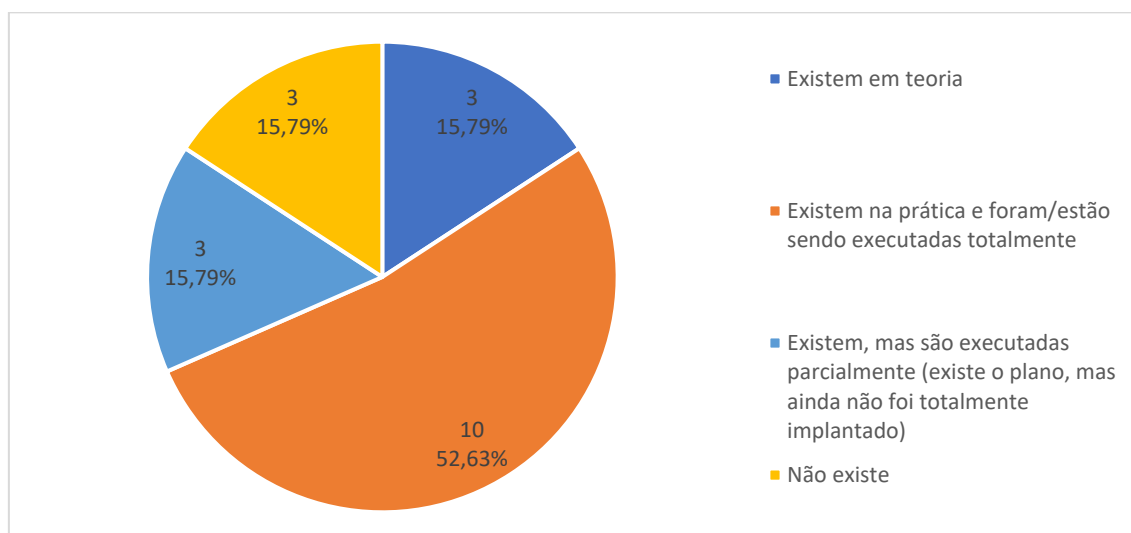
É de interesse das instituições, principalmente públicas, que os seus materiais sejam encontrados pelas Comunidades-Alvo, favorecendo o seu uso, além de ser uma forma de a instituição os trabalhos desenvolvidos. Caso os usuários não consigam encontrar os repositórios na página principal das suas próprias instituições, o seu acesso pode se tornar prejudicado, fazendo com que o seu uso seja diminuído ou até mesmo impedido. Assim, recomenda-se que as páginas das instituições apresentem alguma forma de comunicação através de links, banners, menus ou botões, por exemplo, que favoreça a navegação do usuário ao repositório para a consulta e acesso aos materiais custodiados.

A pergunta “10.2) Existem iniciativas para a promoção do repositório institucional?” buscou verificar se os repositórios e as suas respectivas instituições promovem o uso do repositório institucional, no qual são encontrados os materiais preservados. Os resultados são apresentados na Figura 65.

Em 10 repositórios, o que representa 52,63% dos repositórios participantes, existe um planejamento de iniciativas para a promoção dos repositórios e esse planejamento foi ou está sendo executado na prática. Em 3 repositórios, 15,79%, algumas ações para a promoção dos repositórios são executadas, porém não por completo. No entanto, nesses casos, existe um planejamento estabelecido para essas

ações. O estudo atual não possui dados que permitam identificar as razões pelas quais esses planejamentos não foram totalmente implantados, havendo a possibilidade de que esses planos ainda possam se encontrar em execução.

Figura 65 – Existência de iniciativas de promoção dos repositórios.



Fonte: elaboração própria.

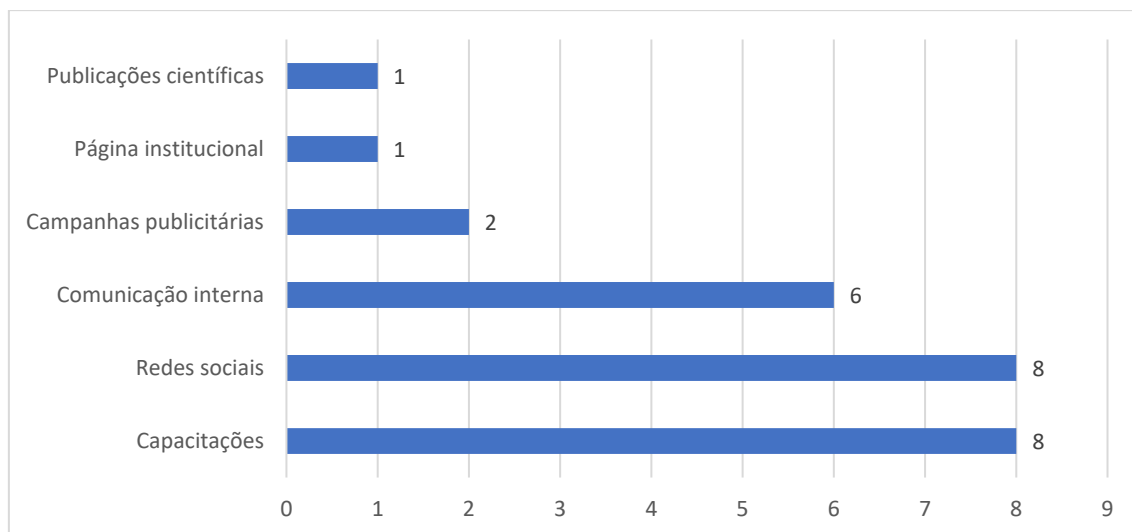
Em 3 repositórios, 15,79%, foi relatado haver um planejamento de promoção dos repositórios somente em teoria, ou seja, na prática não ocorrem ações que promovam ou aumentem a visibilidade dos repositórios frente às suas Comunidades-Alvo. No caso de 3 repositórios, 15,79%, foi relatado que não há qualquer planejamento para que o repositório seja visível e disponível aos seus usuários.

A pergunta “10.3) Caso haja iniciativas de promoção do repositório, quais estratégias e instrumentos utilizados?” buscou compreender as iniciativas adotadas pelos repositórios para a promover as suas visibilidades e respectivos usos. A pergunta elaborada é do tipo livre, permitindo que o respondente informasse todas as estratégias e instrumentos utilizados, assim os termos informados foram agrupados por similaridade. Ao todo, foram registradas 26 respostas, com os resultados sendo exibidos na Figura 66. Somente foram computadas as respostas cujos repositórios sinalizaram na pergunta 10.2 haver iniciativas de promoção dos próprios repositórios.

Dentre os 19 repositórios participantes, 1 repositório, 5,26% do total de participantes, utiliza publicações científicas como forma de promoção, assim como 1 repositório utiliza estratégias de promoção na página da sua instituição. No caso de 2 repositórios, 10,53%, fazem uso de campanhas publicitárias para as suas divulgações,

no entanto, através das respostas fornecidas não foi possível compreender quais os veículos de comunicação utilizados para essas campanhas publicitárias.

Figura 66 – Iniciativas de promoção dos repositórios.



Fonte: elaboração própria.

A comunicação interna é utilizada por 6 repositórios, 31,58%, para a divulgação das ações dos repositórios, voltando-se à sua comunidade de alunos, professores e demais servidores no caso dos repositórios mantidos por instituições de ensino. A utilização de redes sociais foi relatada por 8 repositórios, 42,10%, como sendo um meio para a realização das suas iniciativas de promoção, gerando publicações aos usuários dessas redes para a divulgação de materiais disponíveis e novas funcionalidades do repositório, por exemplo. Em 8 repositórios, 41,10%, existe a cultura de realização de capacitações, palestras, treinamentos e reuniões a respeito dos repositórios e seus usos, gerando como reflexo, a própria promoção dos repositórios.

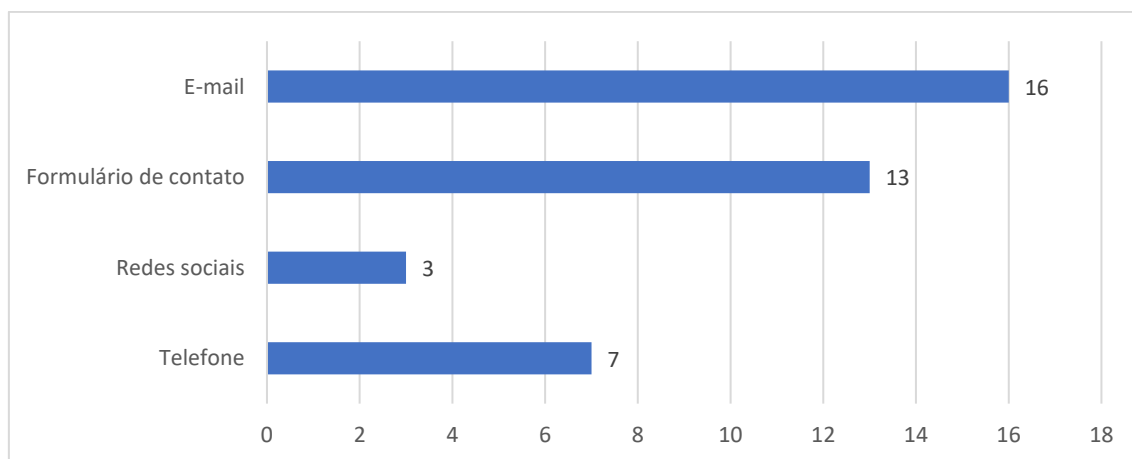
A utilização de capacitações da Comunidade-Alvo pelo repositório favorece uma melhor utilização, permitindo que os usuários tenham conhecimento sobre os tipos de materiais custodiados e as formas de interação existentes com as ferramentas, como filtros e ferramentas de busca, por exemplo. As capacitações podem, ainda, divulgar a política de preservação e acesso adotadas pelo repositório, fazendo com que as características, funcionamento e restrições do repositório tornem-se conhecidas pela Comunidade-Alvo.

As ações de promoção dos repositórios, como um todo, devem ressaltar o compromisso assumido com a preservação digital a longo prazo, difundindo-se as normas e padrões utilizados voltados a preservar a autenticidade e, por consequência,

usabilidade dos materiais depositados. Dessa forma, os produtores e consumidores dos documentos tomam conhecimento do funcionamento dos repositórios que permite impulsionar os seus graus de confiança nos repositórios institucionais.

A pergunta “10.4) Na página web do repositório, são utilizadas quais as formas de contato com os responsáveis pelo repositório?” foi solicitado ao respondente que informasse como os usuários dos repositórios podem se comunicar com os seus respectivos responsáveis. A pergunta, do tipo múltipla escolha, permitiu que os respondentes selecionassem todas as respostas aplicáveis às suas realidades. Todos os 19 repositórios participantes responderam que possuem no mínimo uma forma de contato com o público, apresentando-se os resultados na Figura 67.

Figura 67 – Formas de contato dos repositórios



Fonte: elaboração própria.

Em 16 repositórios, 84,21% dos repositórios participantes, nas páginas dos repositórios é possível encontrar o endereço de e-mail com o qual os interessados podem se comunicar com os responsáveis pelos repositórios. Em 13 repositórios, 68,42%, as páginas web dos repositórios possuem um formulário de contato no qual os interessados podem redigir diretamente as suas mensagens que são destinadas aos responsáveis pelos repositórios.

Em 3 repositórios, 15,79%, as páginas dos repositórios possuem um redirecionamento para as suas respectivas redes sociais permitindo que as pessoas, caso sejam usuárias dessas redes sociais, possam interagir com o perfil do repositório. Em 7 repositórios, 36,84%, as respectivas páginas disponibilizam um número de telefone no qual as pessoas podem se comunicar com os responsáveis pelos repositórios.

Todos os repositórios participantes da pesquisa relataram possuir alguma forma de comunicação entre os usuários e os responsáveis pelos repositórios. Em todos, é possível encontrar um endereço de e-mail ou um formulário de contato para que os usuários consigam enviar os seus questionamentos. A utilização das redes sociais para a comunicação com os usuários pode ser vista como uma maneira de aproximar a interação entre esses dois papéis, facilitando que as pessoas interessadas tenham contato com o perfil do repositório na rede social e interaja com as publicações realizadas, propiciando um aumento no nível do conhecimento a respeito do repositório e sua instituição.

Dentre as formas de comunicação disponibilizadas pelos repositórios, o telefone pode ser considerada um meio síncrono, no qual os usuários contatam o repositório e obtêm uma resposta instantânea, no entanto, o atendimento, normalmente, está restrito aos horários de atendimentos dos repositórios, o que pode ser um impeditivo em certos casos como a diferença entre fusos horários, por exemplo. Por outro lado, o telefone pode ser um meio bastante eficaz para o público interno institucional que necessite se comunicar com o repositório, uma vez que os horários de trabalho ou de realização de aulas pode coincidir muitas vezes.

Embora as respostas obtidas forneçam meios para o conhecimento da forma de comunicação dos repositórios, a pesquisa não utilizou técnicas para mensurar a taxa de resposta e o tempo médio de resposta dos repositórios, não sendo possível determinar se as formas de comunicação são realmente eficazes.

Para fins de comparação, a terceira edição do guia de avaliação de repositórios REBIUN possui os critérios “1.1 – Menção do repositório na página principal da instituição” e “2.5 – Existe uma oferta de contato e assessoramento visível” (AZORÍN MILLARUELO, CÍVICO MARTÍN, *et al.*, 2015, p. 11, 14, tradução nossa) que possuem propósitos similares às perguntas elaboradas no presente estudo e possuem resultados publicados (RED DE BIBLIOTECAS UNIVERSITARIAS ESPAÑOLAS, 2018).

O Quadro 60 exhibe o grau de cumprimento dos repositórios frente a esses requisitos, observando-se um desempenho superior dos repositórios brasileiros, principalmente quanto a disponibilização de acesso ao repositório na página principal das instituições, identificando-se 21,43% dos repositórios espanhóis e 52,63% dos repositórios brasileiros. Quanto à disponibilização de formas de contato, em ambos os

casos, pode-se considerar um alto grau de cumprimento, registrando-se 97,62% dos repositórios espanhóis e 100% dos repositórios brasileiros.

Quadro 60 – Cumprimento de critérios de visibilidade dos repositórios.

Critério	REBIUN	Rede Cariniana
Acesso ao repositório a partir da página inicial da instituição	21,43% (9/42 repositórios)	52,63% (10/19 repositórios)
Existência de uma forma de contato com os responsáveis pelo repositório	97,62% (41/42 repositórios)	100% (19/19 repositórios)

Fonte: elaboração própria.

De um modo geral, a falta de visibilidade de um repositório não implica diretamente na sua perda de confiabilidade, no entanto, o emprego de técnicas e instrumentos para a sua promoção melhoram a sua reputação e criam um ambiente de confiança, impactando positivamente a sua Comunidade-Alvo. Ainda, caso os usuários do repositório não consigam utilizá-lo, por desconhecimento ou falta de acesso, a preservação digital falha com um dos seus princípios que é a utilização dos materiais preservados. Observa-se um bom desempenho nessa área, uma vez que os usuários dos repositórios encontram forma de contato com as suas equipes, no entanto, recomenda-se que o acesso aos repositórios seja disponibilizado a partir das páginas iniciais de todas as instituições, facilitando e promovendo os seus usos.

5.5 ANÁLISES DE FATORES EXTERNOS

Nessa seção são apresentados os resultados e respectivas análises sobre as respostas recebidas dos repositórios, baseando-se nas suas pontuações gerais, sem analisar as especificidades de cada área. O intuito é demonstrar os desempenhos individuais e traçar similaridades e disparidades entre os repositórios para que se compreenda a realidade de cada um. A partir desse panorama, é possível encontrar justificativas que embasam os desempenhos encontrados.

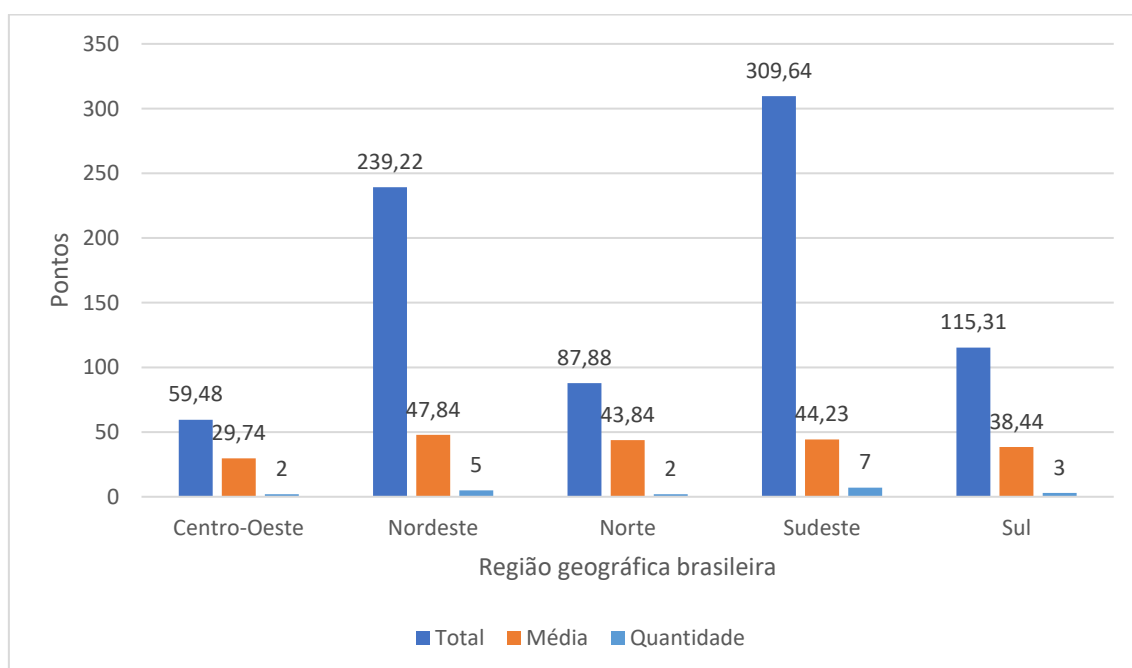
5.5.1 Resultados por Região Geográfica Brasileira

Nessa seção do trabalho são analisados os dados de desempenho dos repositórios participantes referentes às pontuações ponderadas e médias em relação às regiões geográficas brasileiras. O cálculo das pontuações totais absolutas levou em

consideração as somas das pontuações ponderadas individuais, enquanto as médias consideraram as quantidades de repositórios em cada região geográfica.

Uma vez que a quantidade de repositórios participantes entre as regiões não é uniforme, escolheu utilizar as médias para fins comparativos entre os repositórios, buscando uma equalização das pontuações. A Figura 68 exibe o gráfico comparativo entre as regiões geográficas brasileiras, suas pontuações absolutas e respectivas médias. Quando se observam as pontuações absolutas de modo isolado, pode-se perceber uma grande disparidade dos resultados, destacando-se a região Sudeste com 309,64 pontos ao todo, devido ao fato da região possuir 7 repositórios ao todo. Quando se observam as médias das pontuações pelas respectivas quantidades de repositórios, percebe-se uma menor discrepância entre os desempenhos das regiões, sem haver grandes distorções. Assim, embora as análises seguintes demonstrem as pontuações absolutas obtidas, a comparação pelas médias se faz mais igualitária e apropriada.

Figura 68 - Pontuações absolutas e médias entre as regiões geográficas brasileiras.



Fonte: elaboração própria.

O Quadro 61 exibe o desempenho de pontuação dos 19 repositórios por área de avaliação e região geográfica brasileira. Pode-se observar que a quantidade de repositórios participantes entre as regiões não é igual, registrando-se 2 repositórios na região Centro-Oeste, 5 repositórios na região Nordeste, 2 repositórios na região Norte, 7 repositórios na região Sudeste e 3 repositórios na região Sul. Para fins de registro, as

pontuações totais obtidas pelos repositórios em cada região geográfica e área foram incluídas, no entanto, as análises seguintes consideram as médias de pontuação, uma vez que poderiam haver distorções ao se considerar as pontuações absolutas, como citado anteriormente.

Quadro 61 - Desempenho dos repositórios por área e região geográfica brasileira.

Área		Centro-Oeste	Nordeste	Norte	Sudeste	Sul	Totais
Quantidade de repositórios		2	5	2	7	3	19
1. Contextualização do repositório	Total	8,67	31,33	12,00	45,33	16,67	114,00
	Média	4,33	6,27	6,00	6,48	5,56	6,00
2. Estratégia e responsabilidade	Total	2,50	9,17	2,50	20,83	5,83	40,83
	Média	1,25	1,83	1,25	2,98	1,94	2,15
3. Certificações e normas	Total	2,22	7,22	1,67	17,22	2,22	30,56
	Média	1,11	1,44	0,83	2,46	0,74	1,61
4. Infraestrutura organizacional	Total	1,00	19,00	8,50	20,50	7,50	56,50
	Média	0,50	3,80	4,25	2,93	2,50	2,97
5. Infraestrutura técnica e tecnológica	Total	15,00	33,13	11,25	36,25	15,63	111,25
	Média	7,50	6,63	5,63	5,18	5,21	5,86
6. Segurança	Total	13,33	22,50	8,33	25,00	8,33	77,50
	Média	6,67	4,50	4,17	3,57	2,78	4,08
7. Integridade dos dados	Total	4,38	25,63	10,63	31,25	15,63	87,50
	Média	2,19	5,13	5,31	4,46	5,21	4,61
8. Metadados	Total	6,25	37,50	15,00	47,50	17,50	123,75
	Média	3,13	7,50	7,50	6,79	5,83	6,51
9. Acesso e consulta	Total	0,00	25,00	10,00	35,00	7,50	77,50
	Média	0,00	5,00	5,00	5,00	2,50	4,08
10. Visibilidade	Total	6,13	28,75	8,00	30,75	18,50	92,13
	Média	3,06	5,75	4,00	4,39	6,17	4,85
Pontuação total absoluta (1)		59,48	239,22	87,88	309,64	115,31	811,51
Média da região (2)		29,74	47,85	43,94	44,24	38,44	42,72
Média das médias (3)		2,97	4,79	4,39	4,42	3,84	4,27
Desvio Médio (DM) (4)		1,96	1,51	1,49	1,16	1,75	1,29
Desvio Padrão (DP) (5)		2,41	1,87	1,94	1,40	1,84	1,56
Coeficiente de Variação (CV) (%) (6)		81,01%	39,13%	44,23%	31,67%	47,99%	36,51%

Fonte: elaboração própria.

As últimas linhas do Quadro 61, apresentam cálculos referente às pontuações das regiões, como a Pontuação total absoluta (1), que representa o somatório das pontuações absolutas de cada área e região, a Média Aritmética da região (2), que representa a Média Aritmética da pontuação total da região pela quantidade de repositórios dessa região e a Média das médias (3) que representa a Média Aritmética das pontuações médias das 10 áreas de cada região. Com o objetivo de medir a

variabilidade das pontuações apresentadas, utilizou-se o Desvio Médio (DM) (4), que é a média das diferenças das pontuações por área em relação à média da própria área, o Desvio Padrão (DP) (5) e o Coeficiente de Variação (CV) expresso em porcentagem que demonstra o grau de variabilidade da amostra de valores apresentados e serve como medida de comparação da variabilidade entre os repositórios.

Observando-se o CV (6) das médias das regiões, a região Centro-Oeste registra o maior valor, com 81,01% de variação entre as pontuações das áreas de avaliação. De maneira análoga, a região também apresenta o maior Desvio Médio entre todas as regiões, 1,96 pontos. Em parte, essa alta variação se justifica por se registrar 0,00 ponto de média na área “9. Acesso e consulta”, que também foi a menor pontuação da área entre as regiões, e 7,50 pontos de média na área “5. Infraestrutura técnica e tecnológica”, que foi a maior pontuação da área entre as regiões. Conclui-se que os repositórios da região Centro-Oeste atribuem importâncias e aplicam tratamentos muito distintos entre as áreas avaliadas. Opostamente, a região Sudeste registrou o menor CV, com 31,67% de variação, e menor Desvio Médio, com 1,16 pontos, registrando sua menor pontuação média, 2,46 pontos, na área “3. Certificações e normas” e sua maior pontuação média, 6,79 pontos, na área “8. Metadados”. Assim, conclui-se que os repositórios da região Sudeste apresentam um tratamento mais uniforme entre as áreas avaliadas.

A partir das pontuações médias de cada região, é possível estabelecer-se um ranking por ordem decrescente de pontuação: 1º lugar região Nordeste (4,79 pontos), 2º lugar região Sudeste (4,42 pontos), 3º lugar região Norte (4,39 pontos), 4º lugar região Sul (3,84 pontos) e 5º lugar região Centro-Oeste (2,97 pontos). Uma vez que a média das pontuações médias das áreas se estabeleceu com valor 4,27 pontos, é possível identificar que as regiões Nordeste, Sudeste e Norte se posicionaram acima da média, enquanto as regiões Sul e Centro-Oeste se posicionaram abaixo da média. Conclui-se que a maioria das regiões se apresenta acima da pontuação média das áreas e regiões.

O Quadro 62 fornece uma análise comparativa das pontuações, apresentando as áreas, a média da área das regiões, a maior média registrada e a região à qual pertence, a menor média registra e a região à qual pertence e a diferença entre a maior e a menor das regiões por área.

Observando-se os dados, é possível notar duas áreas com as maiores pontuações médias. A área “5. Infraestrutura técnica e tecnológica”, apresenta pontuação máxima de 7,50 pontos na região Centro-Oeste, 1,64 pontos acima da média

da área (5,86), concluindo-se que os repositórios dessa região possuem um tratamento mais direcionado às questões de estrutura física e de equipamentos relacionados à preservação digital. A segunda área com maior pontuação é a “8. Metadados”, com as maiores pontuações (7,50 pontos) registradas para as regiões Nordeste e Norte, 0,99 ponto acima da média da área (6,51), concluindo-se que os repositórios dessas regiões apresentam maiores preocupações com a adoção de um padrão de metadados e um conjunto mínimo de metadados que devem estar presentes na ingestão dos materiais.

Quadro 62 - Maiores e menores médias por área e região geográfica.

Área	Média	Maior média (A)	Região	Menor média (B)	Região	Diferença (A- B) (A/B)%	Desvio Médio
1. Contextualização do repositório	6,00	6,48	Sudeste	4,33	Centro-Oeste	2,15 (49,65%)	0,57
2. Estratégia e responsabilidade	2,15	2,98	Sudeste	1,25	Centro-Oeste Norte	1,73 (138,40%)	0,63
3. Certificações e normas	1,61	2,46	Sudeste	0,74	Sul	1,72 (232,43%)	0,63
4. Infraestrutura organizacional	2,97	4,25	Norte	0,50	Centro-Oeste	3,75 (750,00%)	1,02
5. Infraestrutura técnica e tecnológica	5,86	7,50	Centro-Oeste	5,18	Sudeste	2,32 (44,79%)	0,79
6. Segurança	4,08	6,67	Centro-Oeste	2,78	Sul	3,89 (139,93%)	0,98
7. Integridade dos dados	4,61	5,31	Norte	2,19	Centro-Oeste	3,12 (142,47%)	0,88
8. Metadados	6,51	7,50	Nordeste Norte	3,13	Centro-Oeste	4,37 (139,62%)	1,26
9. Acesso e consulta	4,08	5,00	Nordeste Norte Sudeste	0,00	Centro-Oeste	5,00 (500%)	1,68
10. Visibilidade	4,85	6,17	Sul	3,06	Centro-Oeste	3,11 (101,63%)	1,06
Médias	4,27	5,43	-	2,32	-	3,11 (134,05%)	

Fonte: elaboração própria.

Opostamente, a área que possui menores pontuações médias é “9. Acesso e consulta”, com média de 0 ponto registrada para a região Centro-Oeste. Pode-se perceber um alto grau de deficiência no que se refere à elaboração e aplicação de regras de acesso e monitoramento para a disponibilização dos materiais preservados. A melhoria de desempenho nessa área pode ser alcançada mediante esforços das equipes dos repositórios em direção a estabelecer e aprovar as suas políticas de acesso, por

exemplo, restringindo-se a aspectos e trâmites, em geral, internos das próprias instituições.

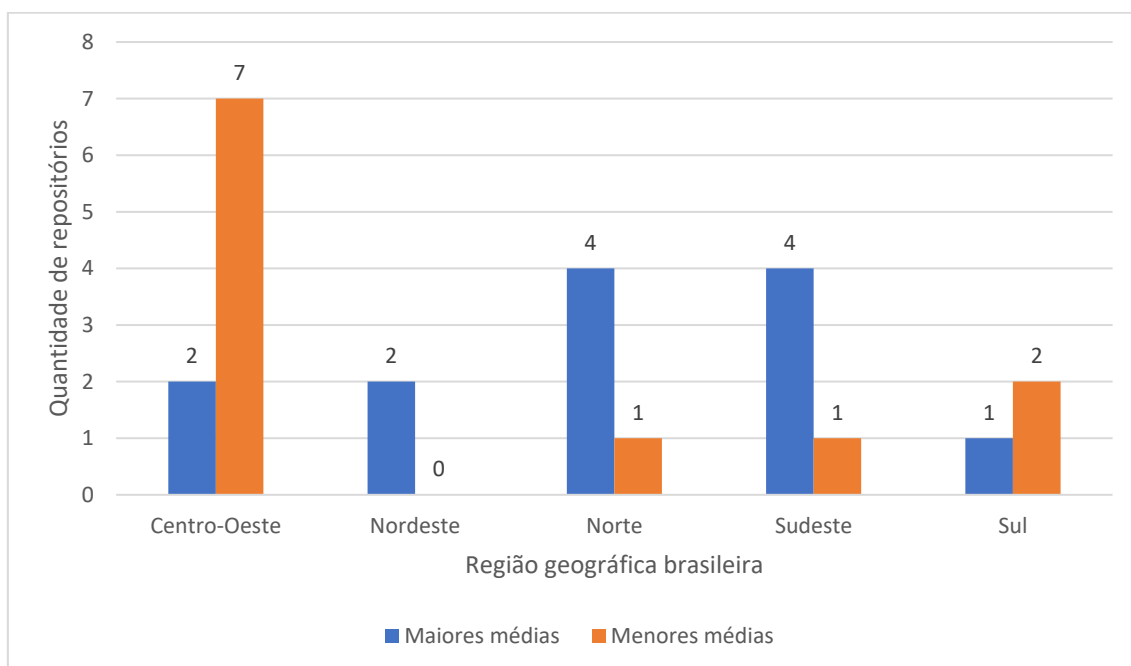
A segunda menor média de pontuação das áreas registrada foi a “4. Infraestrutura organizacional”, com 0,50 ponto na região Centro-Oeste. Pode-se perceber que nesse caso, os repositórios sofrem com a inexistência ou falta de apoio para o estabelecimento de uma estrutura que permita a realização da preservação digital. Sabe-se, no entanto, que os desafios nessa área envolvem diferentes aspectos que muitas vezes dependem de fatores externos ao repositório ou até mesmo às suas instituições, como alinhamento político nos casos dos repositórios públicos e disponibilidade orçamentária, por exemplo. Assim, a melhoria desses resultados pode necessitar um planejamento que leve em consideração a colaboração e negociação com órgãos governamentais superiores, visando desenvolver melhores condições para a efetivação da preservação digital.

O fato de as duas piores médias das áreas pertencerem à região Centro-Oeste desperta atenção e apreensão frente à situação atual e aos desafios de melhoria que os repositórios dessa região necessitam. No entanto, essa concentração geográfica pode beneficiar a realização de iniciativas locais conjuntas entre os repositórios voltadas a atender as suas deficiências e necessidades.

Analisando-se as disparidades entre as maiores e menores pontuações médias por área e região, observou-se a área “4. Infraestrutura organizacional” com diferença de 750%, o que evidencia um alto contraste nos quesitos que tratam de se estabelecer estruturas com competências e responsabilidades bem delimitadas entre as regiões Centro-Oeste e Norte, nesse caso. A área “9. Acesso e consulta” apresentou uma disparidade de 500% entre a maior média das regiões Nordeste, Norte e Sudeste, e a menor média da região Centro-Oeste, indicando um alto grau de diversidade nos aspectos que se relacionam com o estabelecimento de regras para acesso e consulta dos materiais preservados.

Em contrapartida, as áreas “5. Infraestrutura técnica e tecnológica” e “1. Contextualização do repositório” apresentaram as menores diferenças entre as maiores e menores médias, com 44,79% e 49,65%, respectivamente. Assim, evidencia-se que essas áreas possuem um tratamento mais uniforme entre os repositórios no que se refere às suas estruturas tecnológicas e à existência de parâmetros que guiam os funcionamentos dos repositórios.

Figura 69 – Regiões geográficas frente às maiores e menores médias.



Fonte: elaboração própria.

A Figura 69 resume a incidência das regiões geográficas brasileiras em relação à frequência de maiores e menores médias nas áreas de avaliação do formulário. Percebe-se que as regiões Norte e Sudeste possuem desempenhos similares, apresentando maiores médias em 4 áreas cada uma, enquanto ambas registram 1 área com menor média de pontuação. Assim, destacam-se como as regiões que possuem melhores resultados frente às médias de pontuação das avaliações realizadas. Opostamente, a região Centro-Oeste apresenta melhor desempenho das médias em 2 áreas de avaliação, enquanto registra 7 áreas com as menores médias de pontuação. Assim, nota-se que os repositórios dessa região possuem grandes desafios a serem transpostos em relação às demais regiões.

A região Nordeste obteve maior média em 2 áreas e não registrou menores médias em nenhuma área. Em contrapartida, a região Sul apresentou-se com maior média em apenas 1 área enquanto em 2 áreas foi identificada como as menores médias.

Embora os resultados expressem diferenças entre as regiões geográficas do Brasil, devido à baixa quantidade de repositórios em certas regiões, seria necessária uma amostra mais significativa para estabelecer uma relação entre a localização dos repositórios e os seus desempenhos.

5.5.2 Resultados por Tipo de Parceria

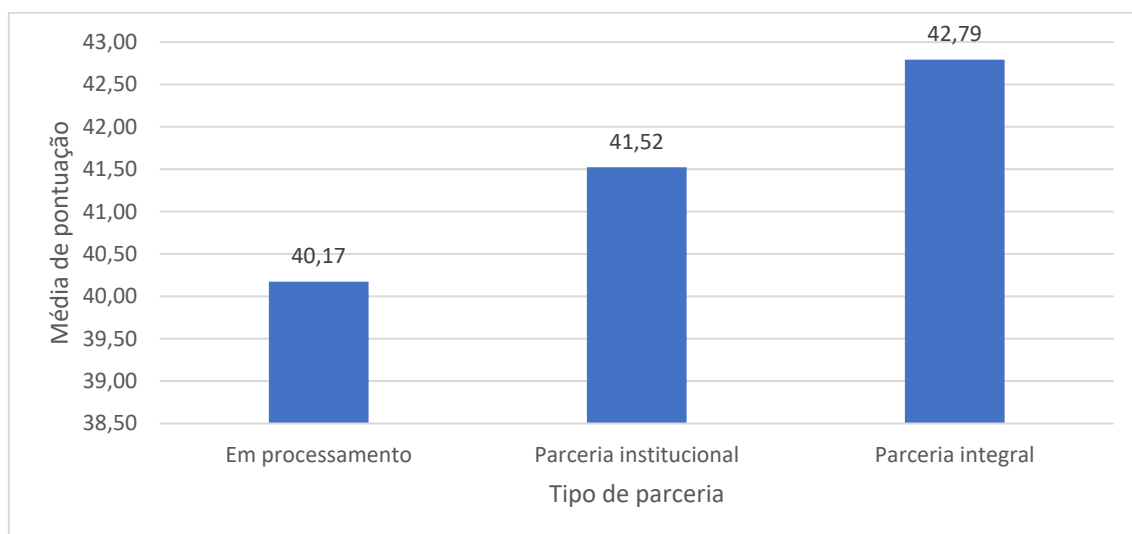
Nessa seção do trabalho são analisados os dados de desempenho dos repositórios participantes referentes aos tipos de parceria estabelecidos com a Rede Cariniana. O Quadro 63 apresenta os tipos de parcerias, as respectivas quantidades de repositórios pertencentes aos tipos, a pontuação geral total do tipo de parceria e a Média Aritmética (MA) da pontuação do tipo de parceria. É possível notar que a “Parceria integral” é o tipo de parceria que possui mais repositórios participantes, com 8 repositórios ao todo, equivalente a 42,11% do total de participantes, seguido pelo tipo “Em processamento”, com 7 repositórios, equivalente a 36,84% dos repositórios. Conforme especificado na seção 4.2.1, no período de coleta dos dados do presente estudo, esses repositórios estavam em tratativas com o IBICT, coordenador da Rede Cariniana, para definição do tipo específico de parceria e estabelecimento dos planos de trabalho.

Quadro 63 – Pontuações dos repositórios por tipo de parceria com a Rede Cariniana.

Tipo de parceria	Repositórios	Pontuação total	MA
Acordo de cooperação	1	63,40	63,40
Parceria integral	8	342,33	42,79
Parceria institucional	3	124,57	41,52
Em processamento	7	281,21	40,17
Total:	19	811,51	42,71

Fonte: elaboração própria.

A partir das médias de pontuações, se obteve um Desvio Médio de 8,22 pontos em relação à média, 42,71 pontos, assim como um Coeficiente de Variação (CV) de 22,32% entre as médias dos tipos de parcerias. No entanto, é possível perceber que o tipo de parceria “Acordo de cooperação” é o que possui maior pontuação, ocupando o primeiro lugar, porém, possui somente 1 repositório participante do estudo o que acaba por distorcer, de certa forma, o *ranking*. Assim, escolheu-se elaborar um *ranking* secundário descartando-se a pontuação e a MA desse tipo de parceria e recalculando a MA geral. Esse ranking, exibido na Figura 70, considera a pontuação total de 748,11 pontos, 18 repositórios e MA geral de 41,56 pontos.

Figura 70 – *Ranking* dos tipos de parceria com a Rede Cariniana.

Fonte: elaboração própria.

Como a pontuação do tipo de parceria “Acordo de cooperação” foi removida, os cálculos se modificaram obtendo-se 41,56 pontos de MA geral, 0,88 ponto de DM e 2,57% de CV. A partir desses indicadores, é possível perceber um baixo grau de variabilidade das pontuações por tipo de parceria, o que indica uma uniformidade de tratamento nas questões relativas aos aspectos identificados no presente estudo.

Enquanto a Parceria integral se destina às instituições com infraestrutura de armazenamento digital e documentos aptos à preservação digital, a Parceria institucional requer somente a existência de projetos institucionais ligados à preservação digital, sem a obrigatoriedade de uma iniciativa a âmbito geral da instituição, como explicado na seção 4.2.1. A partir dos resultados, percebe-se que o tipo de parceria não influencia o desempenho de pontuação dos repositórios participantes, ao menos enquanto as instituições que ainda estão “Em processamento” não formalizam o tipo de parceria a ser firmada.

5.5.3 Resultados por Tipo de Instituição

Nessa seção do trabalho são analisados os dados de desempenho dos repositórios participantes referentes aos seus tipos de instituição. O Quadro 64 exhibe os repositórios com os respectivos tipos de instituição e suas pontuações totais. Optou-se pela criação de 3 categorias de tipo de instituição: 1) **Universidade**: atribuído aos repositórios que funcionam dentro de universidades particulares ou públicas de qualquer

esfera administrativa; 2) **Arquivo, Biblioteca e Museu**: atribuído aos repositórios que representam especificamente esses tipos de instituição, e; 3) **Outro tipo de instituição**: atribuído aos repositórios que não se encaixam nos tipos anteriores.

Quadro 64 – Repositórios, seus tipos de instituição e pontuações totais.

Repositório	Tipo de instituição	Pontuação total	Repositório	Tipo de instituição	Pontuação total
R01	Outro tipo de instituição	46,03	R11	Universidade	53,94
R02	Arquivo, Biblioteca ou Museu	21,01	R12	Universidade	37,65
R03	Universidade	47,94	R13	Arquivo, Biblioteca ou Museu	63,40
R04	Arquivo, Biblioteca ou Museu	27,43	R14	Universidade	66,86
R05	Universidade	64,65	R15	Universidade	31,26
R06	Universidade	36,10	R16	Universidade	17,90
R07	Universidade	47,24	R17	Universidade	66,74
R08	Universidade	51,11	R18	Universidade	23,49
R09	Universidade	49,28	R19	Universidade	24,85

Fonte: elaboração própria.

O Quadro 65 agrupa as pontuações dos repositórios pelos respectivos tipos de instituição, calculando-se as Médias Aritméticas (MA) das pontuações de cada tipo, além do Desvio Médio (DM) e do Coeficiente de Variância (CV). A partir das MA, se obteve um Desvio Médio (DM) de 3,34 pontos e um Coeficiente de Variação (CV) de 8,62%, indicando uma uniformidade dos desempenhos dos repositórios em relação aos tipos de suas instituições.

Quadro 65 – Pontuações totais dos repositórios por tipo de instituição.

Tipo de instituição	Repositórios	Pontuação total	MA
Universidade	15	653,64	43,58
Arquivo, Biblioteca ou Museu	3	111,84	37,28
Outro tipo de instituição	1	46,03	46,03
Total:	19	811,51	
MA:	42,71		
DM:	3,34		
CV (%):	8,62		

Fonte: elaboração própria.

Com o intuito de possibilitar uma análise alternativa, os mesmos indicadores foram recalculados desconsiderando a pontuação do tipo “Outro tipo de instituição”, que possui apenas 1 repositório. Assim, obteve-se uma MA de 42,53 pontos, um DM de 2,56 pontos e um CV de 4,7%, notando-se uma maior uniformidade das pontuações, ainda que a MA geral registrada se apresente 0,18 ponto abaixo do cálculo anterior.

É possível notar uma grande predominância de repositórios de universidades, totalizando 15 repositórios, o que equivale a 78,95% do total de participantes e, por isso, acumulam a maior pontuação total. Dentre as 3 categorias, 2 se posicionam acima da MA, 42,71 pontos, indicando um alto desempenho dos repositórios.

Ao se analisar a maior e a menor MA, 46,03 e 37,28 pontos respectivamente, observa-se uma diferença de 8,75 pontos, equivalendo a uma diferença de 23,47% da menor para a maior MA. Conclui-se que, devido à uniformidade das pontuações indicada pela CV e à baixa diferença de pontuação entre as pontuações extremas, o tipo de instituição não influencia de maneira significativa nos desempenhos dos repositórios nos quesitos avaliados no presente estudo.

5.5.4 Resultados por Tempo de Criação dos Repositórios

Nessa seção do trabalho são analisados os dados de desempenho dos repositórios participantes com relação ao tempo de suas criações e estabelecimentos oficiais. As informações referentes às datas de criação dos repositórios foram extraídas a partir dos sítios dos próprios repositórios ou em documentos normativos que estabelecem e oficializam a implantação dos repositórios das instituições. No caso de quatro repositórios, não foi possível encontrar essa informação em nenhum meio, tendo sido feito contato através de mensagens de e-mail, com respostas positivas de dois deles. Nos dois repositórios restantes, não houve resposta, não sendo possível afirmar as suas datas de criação.

Quadro 66 – Repositórios e seus anos de criação.

Repositório	Criação	Repositório	Criação	Repositório	Criação
R03	2005	R14	2010	R11	2017
R17	2005	R15	2010	R13	2019
R19	2008	R08	2011	R01	2021
R02	2009	R06	2013	R12	Não disponível
R10	2009	R04	2015	R16	Não disponível
R07	2010	R05	2015		
R09	2010	R18	2015		

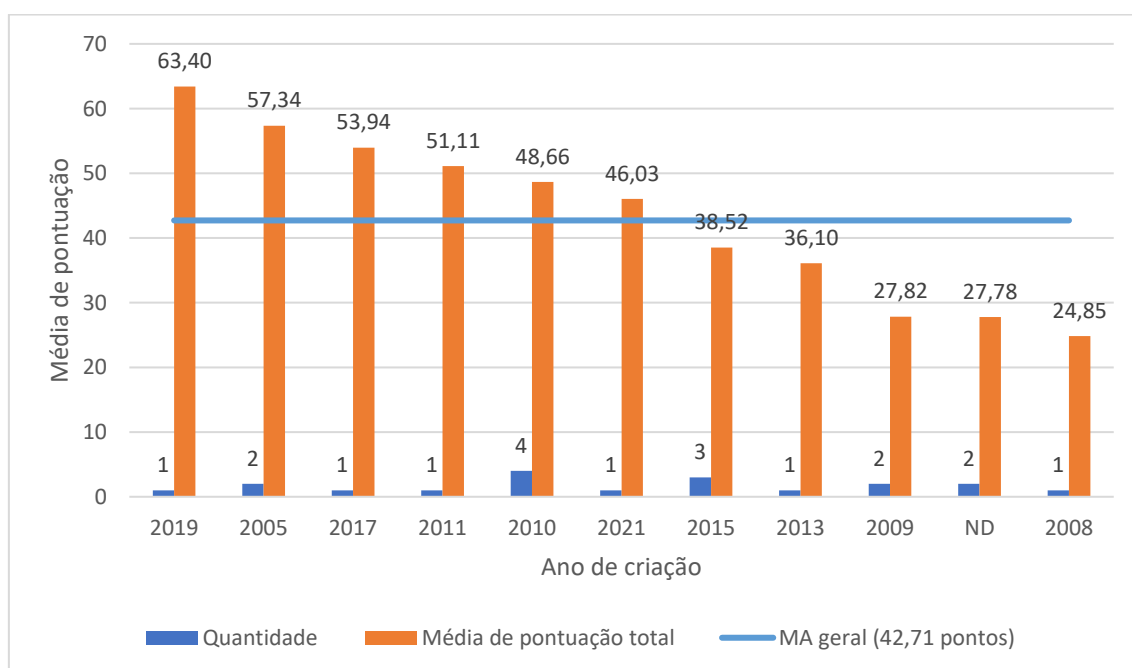
Fonte: elaboração própria.

O Quadro 66 exhibe os repositórios e os seus respectivos anos de criação, sendo possível constatar a existência de 2 repositórios mais antigos, criados em 2005 e

tendo 18 anos atualmente, enquanto o mais recente possui 2 anos, tendo sido criado em 2021.

A Figura 71 exibe o gráfico que agrupa os anos de criação dos repositórios com as suas respectivas médias de pontuação e a linha da Média Aritmética (MA) geral de pontuação dos repositórios, 42,71 pontos. É possível notar que a pontuação mais alta, 63,40 pontos, é ocupada por um repositório criado em 2019 e que possui 4 anos desde a sua criação, enquanto a menor pontuação registrada, 24,85 pontos, é de um repositório criado em 2008, tendo, portanto, 15 anos de criação. Acima da MA, é possível encontrar 10 repositórios, com criações desde 2005 até 2021, possuindo tempo médio de criação de 9,17 anos, equivalente a aproximadamente 9 anos e 2 meses. Abaixo da MA, encontram-se 9 repositórios, com tempo médio de criação de 11,75 anos, equivalente a 11 anos e 9 meses, desconsiderando-se os 2 repositórios sem a informação de data de criação.

Figura 71 – Média de pontuação dos repositórios em relação ao ano de criação.



Fonte: elaboração própria.

Enquanto é possível verificar 2 repositórios com 18 anos desde a sua criação com média de pontuação acima da MA geral, é possível verificar 1 repositório com 15 anos desde a sua criação com a menor pontuação de todos os repositórios, não sendo possível estabelecer uma relação entre a idade do repositório e seu desempenho.

5.5.5 Resultados dos Repositórios Universitários

Como a grande maioria dos repositórios participantes do estudo pertencem à Instituições de Ensino Superior (IES), somando 15 repositórios de um total de 19 participantes, equivalente a 78,95%, escolheu-se analisar os seus desempenhos em relação a aspectos específicos de suas naturezas como o orçamento, tamanho da comunidade, *ranking* de universidades do Ministério da Educação e Cultura (MEC) e o *ranking* Scimago. Ainda que a análise por orçamento não seja uma dimensão exclusiva das universidades, esses dados relativos às instituições públicas federais estão disponíveis para consulta e acesso no Portal da Transparência do governo brasileiro, enquanto que no caso dos repositórios privados ou de outras naturezas, não foi possível encontrar dados confiáveis para realizar a análise. As análises dessa seção utilizam os repositórios categorizados como Universidade no Quadro 64 da seção 5.5.3 anterior.

Isolando-se as pontuações dos repositórios das IES, obteve-se uma MA de 43,58 pontos, um DM de 13,22 pontos, DP de 15,25 pontos e CV de 35%. Percebe-se que a MA desse conjunto de repositórios apresenta-se acima da MA geral, de 42,71, enquanto as demais variáveis apresentaram-se muito próximas das observadas em todos os repositórios. Esse fato é devido à grande maioria dos repositórios pertencerem a IES.

5.5.5.1 Por Orçamento

Nessa seção são analisados os desempenhos dos repositórios em relação aos seus orçamentos executados e publicamente declarados relativos ao ano de 2022. No caso dos repositórios de instituições federais, a fonte de dados foi o Portal da Transparência do Governo Federal (CONTROLADORIA-GERAL DA UNIÃO, 2023). No caso de 1 repositório estadual, a fonte de dados utilizada foi o respectivo Portal da Transparência estadual, porém para fins de anonimização, essa referência será omitida.

O Quadro 67 exhibe os repositórios com as respectivas naturezas, orçamento de 2022 em Reais (R\$) e a pontuação total obtida. Cabe ressaltar que os orçamentos não se referem exclusivamente aos repositórios, mas sim, às universidades às quais pertencem. O processo de detalhamento e discriminação dos orçamentos das instituições buscando identificar o montante destinado exclusivamente aos repositórios pode apresentar uma alta complexidade, uma vez que custos com TI, pessoal e demais equipamentos podem ser compartilhados entre diferentes setores da organização e

podem beneficiar diferentes processos dentro das próprias instituições, além do repositório.

Quadro 67 – Repositórios, seus respectivos orçamentos e pontuações totais.

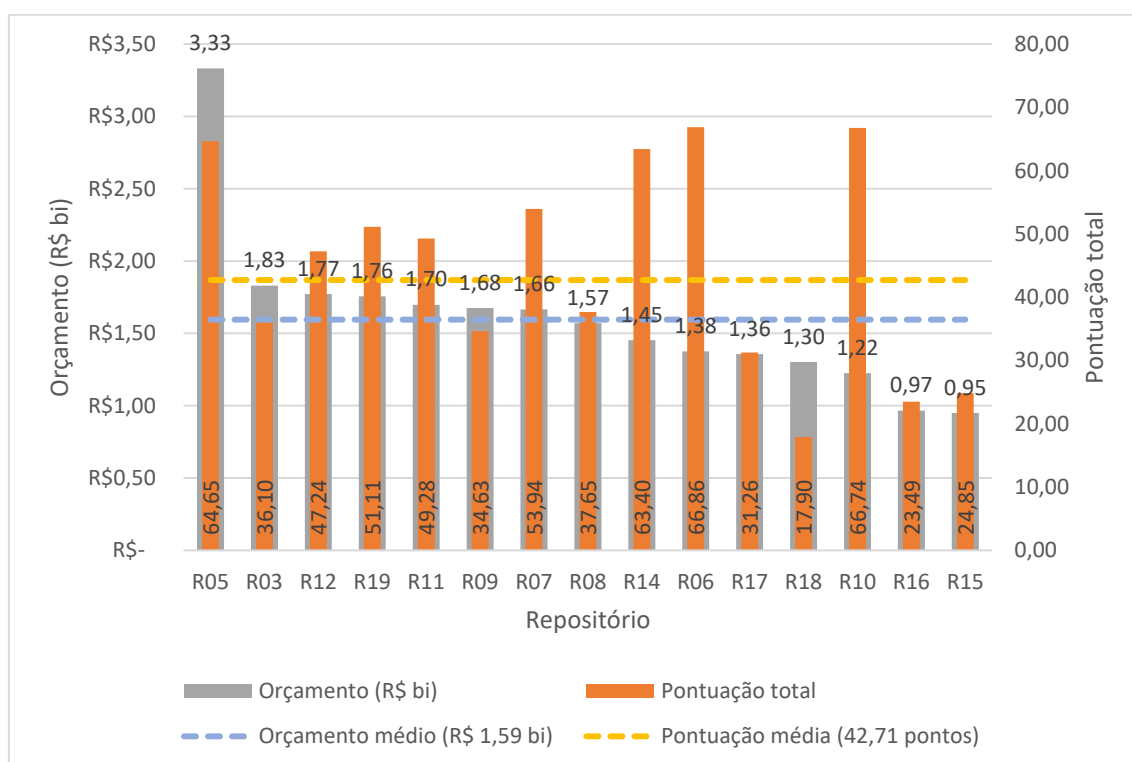
Repositório	Natureza	Orçamento 2022 (R\$)	Pontuação total
R01	Privado	Não se aplica	46,03
R02	Público federal	Não se aplica	21,01
R03	Público federal	R\$ 1.830.307.263,08	47,94
R04	Público federal	Não se aplica	27,43
R05	Público estadual	R\$ 3.331.656.507,00	64,65
R06	Público federal	R\$ 1.375.126.032,62	36,10
R07	Público federal	R\$ 1.664.907.673,57	47,24
R08	Público federal	R\$ 1.570.412.153,13	51,11
R09	Público federal	R\$ 1.675.642.032,81	49,28
R10	Público federal	R\$ 1.224.866.970,13	34,63
R11	Público federal	R\$ 1.697.438.781,62	53,94
R12	Público federal	R\$ 1.772.065.579,36	37,65
R13	Público federal	Não se aplica	63,40
R14	Público federal	R\$ 1.452.654.474,29	66,86
R15	Público federal	R\$ 949.459.505,42	31,26
R16	Público federal	R\$ 965.943.791,45	17,90
R17	Público federal	R\$ 1.356.253.238,31	66,74
R18	Público federal	R\$ 1.301.851.498,03	23,49
R19	Público federal	R\$ 1.756.301.978,85	24,85
	Média Aritmética:	R\$ 1.594.992.498,64	42,71

Fonte: elaboração própria.

Analisando-se os orçamentos em relação à MA, percebe-se 7 repositórios acima da MA e 8 repositórios abaixo da MA, comportamento condizente com a Média Aritmética. No entanto, a soma dos orçamentos dos repositórios acima da MA é de R\$ 13.728.319.816,29, enquanto a soma dos orçamentos dos repositórios abaixo da MA é de R\$ 10.196.567.663,38, uma diferença de 34,64% que sinaliza uma grande disparidade de valores destinados aos orçamentos das universidades.

A Figura 72 exibe os orçamentos, em bilhões de Reais (R\$) dos repositórios universitários com a Média Aritmética (MA) geral dos orçamentos, de R\$ 1.594.992.498,64 (R\$ 1,59 bi). É possível verificar uma grande distorção do orçamento do repositório R05, sendo 2,08 vezes superior à MA dos orçamentos e 3,51 vezes superior ao orçamento do repositório R15, o menor registrado.

Figura 72 – Orçamento das universidades com repositórios.



Fonte: elaboração própria.

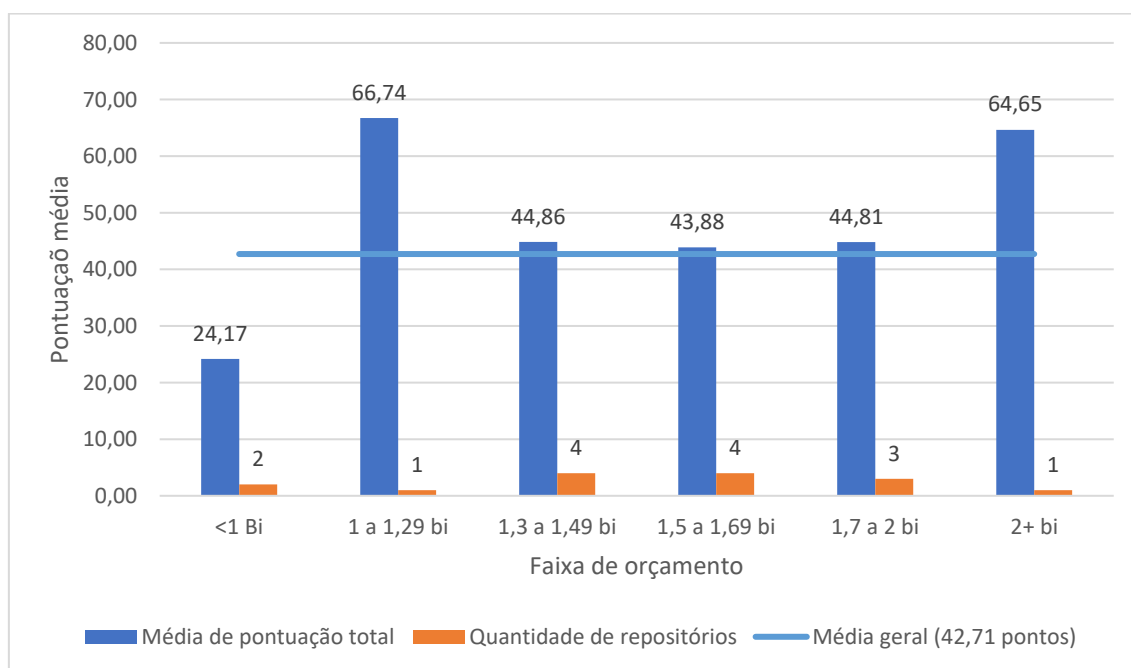
Outra análise possível é em relação às pontuações. A MA das pontuações dos repositórios com orçamentos superiores à MA dos orçamentos (R05, R03, R12, R19, R11, R09, R07) é de 50,19 pontos, estando 7,48 pontos acima da MA geral de 42,71 pontos, enquanto a MA das pontuações dos repositórios com orçamentos inferiores à MA dos orçamentos é de 41,01 pontos, 1,70 ponto abaixo da MA geral. Assim, ainda que os orçamentos possam influenciar pontuações maiores, a MA de pontuação repositórios com orçamentos abaixo da MA não se distanciam consideravelmente da própria MA.

Em termos relativos, a pontuação média dos repositórios superiores, é 22,38% maior que a pontuação média dos repositórios inferiores, percentual menor que a diferença entre os somatórios dos orçamentos dos dois estratos, de 34,63%. Nota-se assim, uma certa assimetria entre orçamento e desempenho de pontuação.

A Figura 73 exibe a pontuação média dos repositórios agrupada por 7 faixas de orçamento. Pode-se observar que todas as pontuações médias dos repositórios das universidades com orçamentos superiores a R\$ 1 bilhão são superiores à MA de pontuação de 42,71 pontos, podendo servir como um limiar mínimo de orçamento para as instituições. A partir dessa faixa, de orçamentos acima de R\$ 1 bilhão, não se nota

uma relação direta entre o incremento de orçamento e a melhoria de desempenho dos repositórios, com exceção dos repositórios R10, que alcançou 66,74 pontos e está na faixa de orçamento de R\$ 1 a 1,29 bilhão e do repositório R05 que alcançou 64,65 pontos e está na faixa de maiores de R\$ 2 bilhões.

Figura 73 – Pontuação média dos repositórios por faixa de orçamento.



Fonte: elaboração própria.

Ressalta-se, novamente, sobre a limitação existente de identificação das verbas direcionadas exclusivamente e diretamente aos repositórios. Assim, a realização de aumentos dos orçamentos das universidades pode não influenciar diretamente na melhoria de seus desempenhos e serviços caso as suas administrações internas não realizem a destinação de verbas adequadas a apoiar a estrutura de preservação de cada instituição.

O Quadro 68 exhibe os orçamentos das IES divididos por região geográfica brasileira com as respectivas pontuações médias regionais, com a média geral de orçamento de R\$ 1.594.992.498,64 (R\$ 1,59 bi). Pode-se notar que a região Norte possui apenas um repositório pertencente à uma IES, o que, de certa forma, pode distorcer os resultados dessa região. A Figura 74 exhibe gráfico para análise dos dados dos orçamentos regionais, exibidos no eixo principal, com relação à média dos orçamentos e das pontuações regionais, exibidas no eixo secundário, com relação à média de pontuação geral.

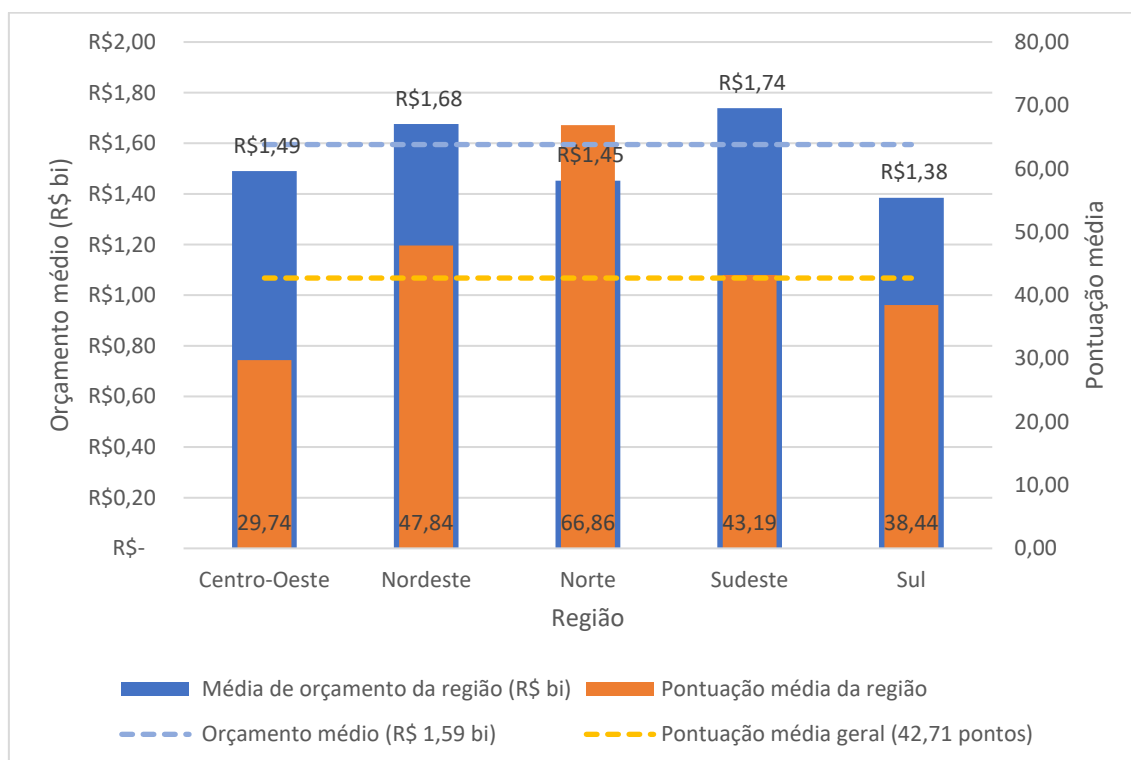
Quadro 68 – Orçamentos das instituições por região geográfica brasileira e suas pontuações médias.

Região	Quantidade	Orçamento médio da região	Pontuação média da região
Centro-Oeste	2	R\$ 1.490.584.474,49	29,74
R10	1	R\$ 1.224.866.970,13	34,63
R19	1	R\$ 1.756.301.978,85	24,85
Nordeste	5	R\$ 1.676.093.244,10	47,84
R07	1	R\$ 1.664.907.673,57	47,24
R08	1	R\$ 1.570.412.153,13	51,11
R09	1	R\$ 1.675.642.032,81	49,28
R11	1	R\$ 1.697.438.781,62	53,94
R12	1	R\$ 1.772.065.579,36	37,65
Norte	1	R\$ 1.452.654.474,29	66,86
R14	1	R\$ 1.452.654.474,29	66,86
Sudeste	4	R\$ 1.738.926.258,70	43,19
R05	1	R\$ 3.331.656.507,00	64,65
R16	1	R\$ 965.943.791,45	17,90
R17	1	R\$ 1.356.253.238,31	66,74
R18	1	R\$ 1.301.851.498,03	23,49
Sul	3	R\$ 1.384.964.267,04	38,44
R03	1	R\$ 1.830.307.263,08	47,94
R06	1	R\$ 1.375.126.032,62	36,10
R15	1	R\$ 949.459.505,42	31,26
Total:	15	R\$ 1.594.992.498,64	MA: 43,58

Fonte: elaboração própria.

É possível notar que as regiões que possuem orçamentos acima da média geral (R\$ 1,59 bi) também possuem pontuações acima da pontuação média geral (42,71 pontos), com exceção da região Norte, que possui 66,86 pontos e orçamento abaixo da média (R\$ 1,45 bi). Nas regiões Centro-Oeste e Sul, tanto as pontuações quanto os orçamentos se posicionam abaixo das respectivas médias. De onde se pode estabelecer um tipo de relação em que as verbas regionais destinadas às instituições contribuem com os desempenhos dos respectivos repositórios.

Figura 74 – Pontuações médias das regiões em relação aos orçamentos médios.



Fonte: elaboração própria.

Como explicado anteriormente, a região Norte possui somente 1 repositório ligado a uma IES, de onde foram extraídos os dados de orçamentos, o que pode ocasionar uma distorção das análises dessa região. Ainda assim, ao se remover os dados da região Norte, observa-se um orçamento médio de R\$1.605.159.500,38, valor 0,64% acima do orçamento médio geral e uma MA de 41,91 pontos, 0,80 ponto acima da média geral de todos os repositórios, de 42,71 pontos. Ainda assim, as mudanças ocasionadas por esse cenário não impactam as análises realizadas com todas as regiões e orçamentos, sendo possível estabelecer um tipo de relação entre as verbas regionais destinadas às instituições com os desempenhos dos respectivos repositórios.

5.5.5.2 Por Tamanho das Comunidades

Nessa seção do trabalho são analisados os desempenhos dos repositórios de IES em relação aos tamanhos das suas comunidades de estudantes de graduação, docentes e servidores, se baseando nos números detalhados na seção 4.2.1 e que utilizam como fonte de dados se a Sinopse Estatística da Educação Superior realizada no ano de 2020 (INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2022). O Quadro 69 exhibe os repositórios com

os respectivos tamanhos das suas comunidades, expressos em faixas de valores, e as suas pontuações totais obtidas. Ressalta-se que os tamanhos das comunidades só se aplicam nos casos das Instituições de Ensino Superior.

Quadro 69 - Repositórios, tamanhos das suas comunidades e pontuações totais.

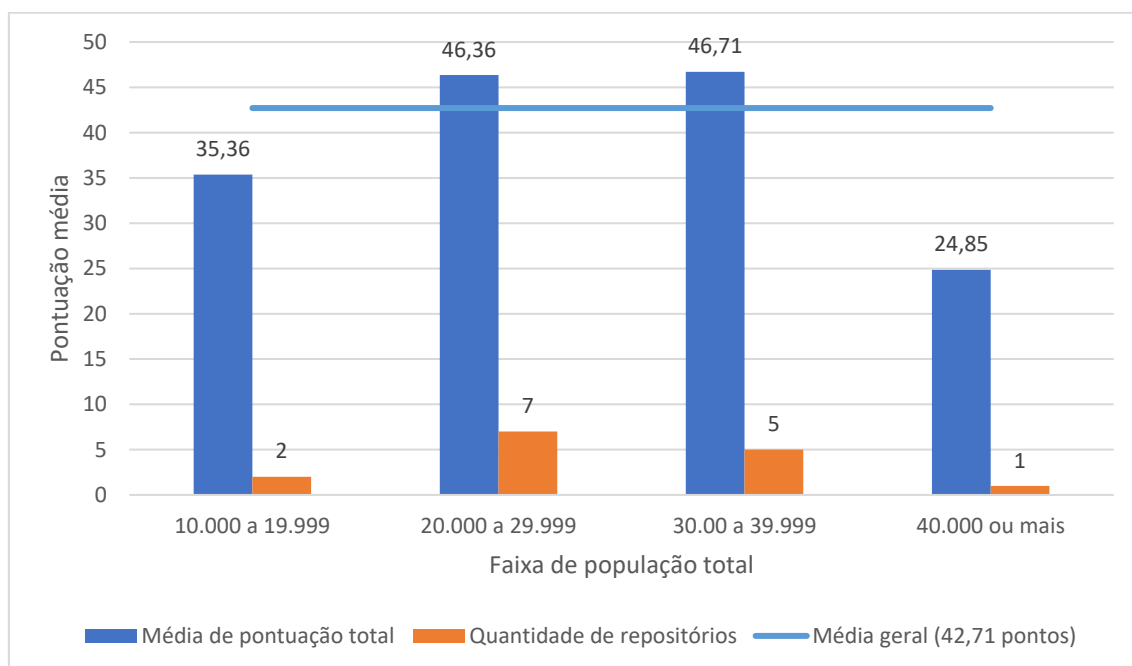
Repositório	Tamanho	Pontuação total	Repositório	Tamanho	Pontuação total
R07	10.000 a 19.999	47,24	R08	30.000 a 39.999	51,11
R18	10.000 a 19.999	23,49	R09	30.000 a 39.999	49,28
R05	20.000 a 29.999	64,65	R11	30.000 a 39.999	53,94
R06	20.000 a 29.999	36,10	R15	30.000 a 39.999	31,26
R10	20.000 a 29.999	34,63	R19	40.000 ou mais	24,85
R12	20.000 a 29.999	37,65	R01	Não se aplica	46,03
R14	20.000 a 29.999	66,86	R02	Não se aplica	21,01
R16	20.000 a 29.999	17,90	R04	Não se aplica	27,43
R17	20.000 a 29.999	66,74	R13	Não se aplica	63,40
R03	30.000 a 39.999	47,94			
Tamanho total:	427.188		Tamanho médio:	28.479,20	

Fonte: elaboração própria.

A Figura 75 exibe a pontuação média dos repositórios com as suas respectivas populações gerais expressas em faixas de valores e a média geral de pontuação. Pode-se perceber que a maioria dos repositórios universitários, 12 de um total de 15 repositórios, possui uma população geral entre 20.000 e 39.999 pessoas. A pontuação média dos repositórios com população abaixo de 20.000 pessoas, 35,36 pontos, se apresenta abaixo da média geral de pontuações dos repositórios, 42,71 pontos, assim como o único repositório que possui mais de 40.000 pessoas em sua comunidade obteve 24,85 pontos.

Assim, percebe-se uma evolução dos desempenhos dos repositórios a partir de uma população de 20.000 pessoas, que passa a ter uma piora de pontuação quando essa população ultrapassa as 40.000 pessoas.

Figura 75 – Pontuação média dos repositórios e suas populações totais.



Fonte: elaboração própria.

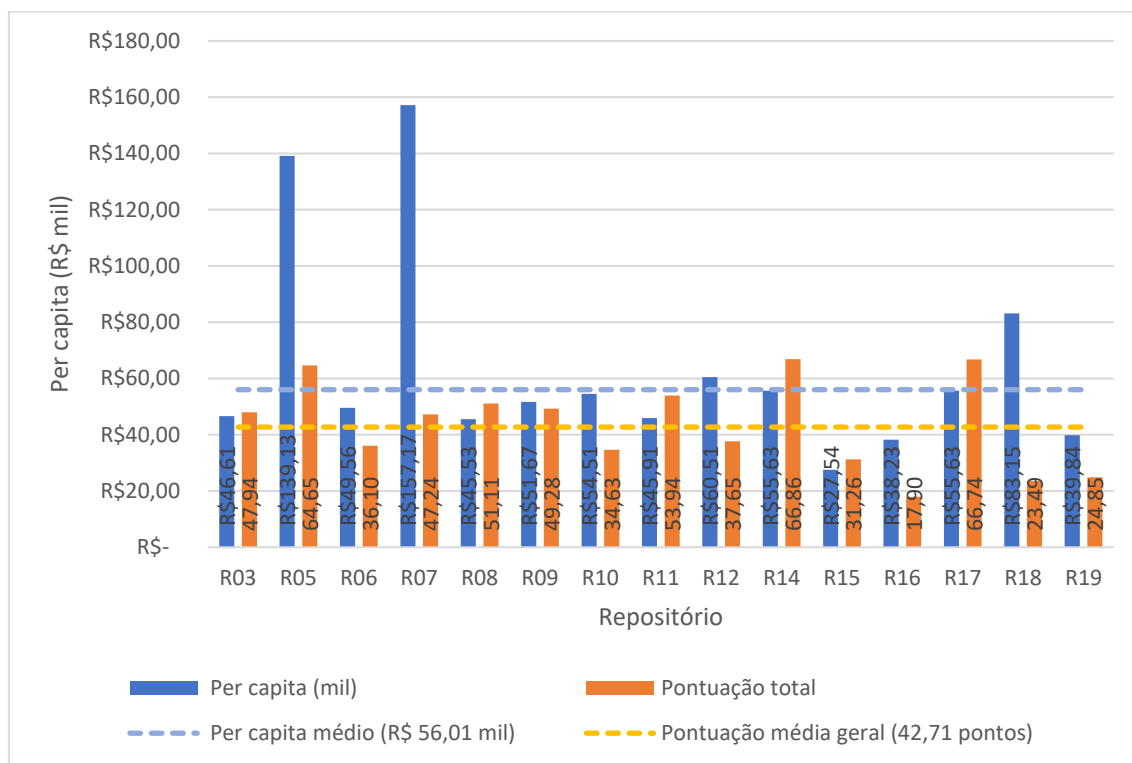
Na faixa de população de 10.000 a 19.999 encontram-se os repositórios R07 e R18. Ao se analisar as pontuações específicas das áreas desses repositórios, verifica-se que o R07 possui 0,00 ponto na área “2. Estratégia e responsabilidade” e 1,11 ponto na área “3. Certificações e normas”, enquanto o repositório R18 apresenta 0,83 ponto na área “6. Segurança” e 1,00 ponto na área “4. Infraestrutura organizacional”. Na faixa de população acima de 40.000 pessoas, encontra-se o repositório R19, com 0,00 ponto nas áreas “2. Estratégia e responsabilidade”, “4. Infraestrutura organizacional”, “7. Integridade dos dados” e “9. Acesso e consulta”.

Pode-se perceber a que as áreas “2. Estratégia e responsabilidade” e “4. Infraestrutura organizacional” possuem pontuações baixas em ambas faixas de população dos repositórios, de onde se pode concluir que essas áreas sofrem em desempenho tanto em instituições de menor porte quanto nas de grande porte. Embora a primeira área não possua características diretamente ligadas ao tamanho da comunidade atendida, uma vez que trata do estabelecimento de políticas e instrumentos de preservação, a segunda área é diretamente afetada pelo tamanho da comunidade, de onde se verifica que nos repositórios com menor população, a infraestrutura organizacional pode não estar totalmente implementada, com formação de equipes e aquisição de talentos, por exemplo. Por outro lado, no caso de grandes populações atendidas, pode ser o caso de a infraestrutura organizacional estar subdimensionada e

sobrecarregada, não possuindo mais meios para atender as demandas da sua comunidade sem que melhorias e adaptações sejam realizadas.

A Figura 76 exibe o orçamento médio per capita das IES em relação às suas populações totais de estudantes de graduação e pontuações totais obtidas.

Figura 76 – Orçamento per capita dos repositórios e pontuações totais.



Fonte: elaboração própria.

Inicialmente, se destacam os repositórios R05, com per capita de R\$ 139,13 mil, e R07, com per capita de R\$ 157,17 mil, respectivamente 148,40% e 180,61% acima do per capita médio de R\$ 56,01 mil, sendo os únicos repositórios que possuem ambos orçamentos e pontuações acima das respectivas médias. Dentre os repositórios que possuem pontuação acima da média geral de 42,71 pontos e orçamento per capita abaixo da média de R\$ 56,01 mil, pode-se citar R03, R08, R09, R11, R14 e R17, sendo que o orçamento per capita desses dois últimos, ambos com R\$ 55,63 mil, se aproxima bastante do per capita médio. A análise desses casos permite concluir que o desempenho dos repositórios não está diretamente relacionado ao orçamento per capita das respectivas instituições, confirmada pelo fato de que o 1º lugar no ranking geral de repositórios, R14 com 66,86 pontos, possui orçamento per capita abaixo da média.

5.5.5.3 Por Ranking do MEC

Nessa seção do trabalho são analisados os desempenhos dos repositórios de IES em relação Índice Geral de Cursos (IGC) do Ministério da Educação e Cultura, realizado pela última vez em 2021. O IGC é utilizado para avaliar a qualidade das IES brasileiras e utiliza a pontuação de avaliação dos cursos de graduação, as médias dos conceitos de avaliação dos programas de pós-graduação e a distribuição dos estudantes entre os diferentes níveis de ensino (INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2020). O cálculo do IGC considera sempre o ano atual de realização dos cálculos em conjunto com os dois anos anteriores, referindo-se assim a um triênio e possui pontuação mínima de 1 ponto e máxima de 5 pontos, sendo a maior pontuação um sinônimo de excelência da instituição.

O Quadro 70 apresenta os repositórios referentes a IES com os respectivos IGC e suas pontuações gerais. Inicialmente, pode-se verificar que nenhuma instituição possui o menor IGC possível, 3 pontos, enquanto a grande maioria, 11 instituições, equivalente a 73,33%, possui valor 4 de IGC e 4 instituições, equivalente a 26,67%, possui valor 5 de IGC.

Quadro 70 – Repositórios, seus IGC (*ranking* do MEC) e pontuações totais.

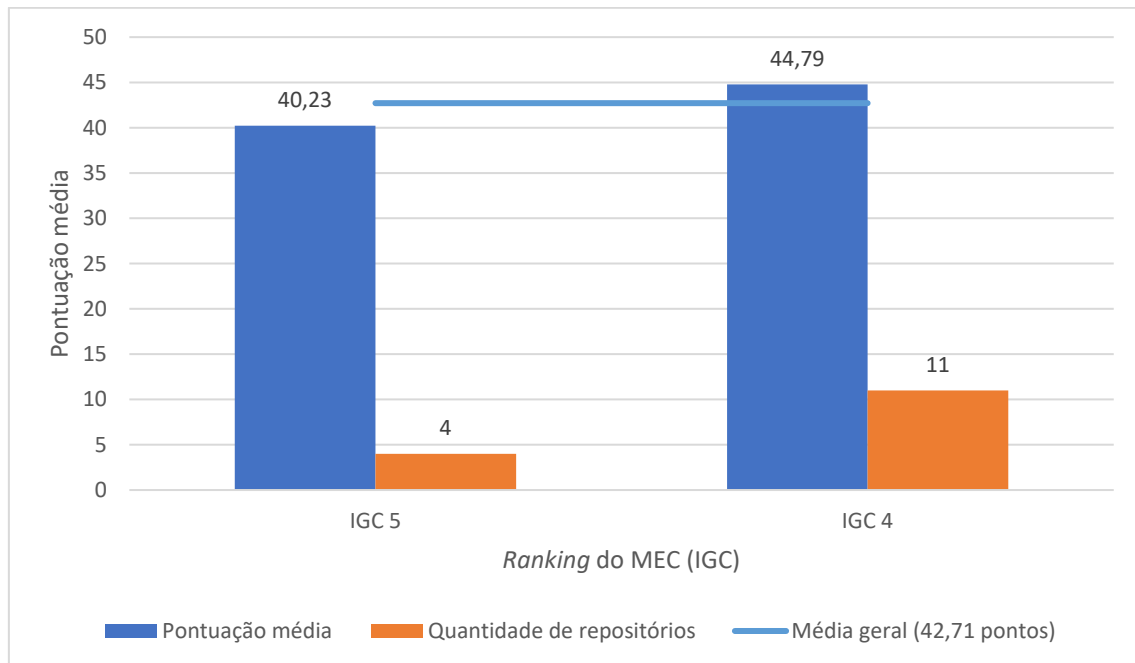
Repositório	IGC 2021	Pontuação total	Repositório	IGC 2021	Pontuação total
R06	4	36,10	R17	4	66,74
R07	4	47,24	R03	5	47,94
R08	4	51,11	R05	5	64,65
R09	4	49,28	R18	5	23,49
R10	4	34,63	R19	5	24,85
R11	4	53,94	R01	Não se aplica	46,03
R12	4	37,65	R02	Não se aplica	21,01
R14	4	66,86	R04	Não se aplica	27,43
R15	4	31,26	R13	Não se aplica	63,40
R16	4	17,90			

Fonte: elaboração própria.

A Figura 77 exibe o gráfico das pontuações médias dos repositórios em relação ao *ranking* MEC das suas instituições e a pontuação média geral dos repositórios, de 42,71 pontos. Pode-se perceber um melhor desempenho dos repositórios cujas instituições possuem IGC 4, alcançando pontuação média de 44,79 pontos, 2,08 pontos acima da média geral dos repositórios de 42,71 pontos, enquanto os repositórios

de instituições com IGC 5 alcançaram pontuação média de 40,23 pontos, 2,48 pontos abaixo da média geral dos repositórios. Esse dado pode revelar que as instituições promoveram mudanças internas de seus cursos para atingir a progressão de IGC, porém as mudanças não se refletiram nos repositórios dessas instituições, levando-os a um baixo desempenho.

Figura 77 – Pontuação média dos repositórios pelo *ranking* MEC das suas instituições.



Fonte: elaboração própria.

5.5.5.4 Por Ranking do SCImago

Nessa seção do trabalho são analisados os desempenhos dos repositórios de IES em relação ao *SCImago Institutions Ranking* (SIR) (*Ranking* de Instituições SCImago) 2023, tido como (SCIMAGO, 2023, p. tradução nossa):

uma classificação de instituições acadêmicas e relacionadas à pesquisa, classificadas por um indicador composto que combina três conjuntos diferentes de indicadores baseados no desempenho da pesquisa, nos resultados da inovação e no impacto social medido por sua visibilidade na Web.

O cálculo do índice é composto por indicadores que avaliam o impacto das instituições em relação à Pesquisa (*Research*), Inovação (*Innovation*) e Impacto social (*Societal*), em uma proporção de 50%, 30% e 20%, respectivamente. Essa metodologia permite a comparação entre diferentes instituições, não somente universidades, a nível global, permitindo monitorar os seus desempenhos ao longo dos anos, além de permitir

identificar quais áreas as instituições se destacam ou necessitam melhorias. O índice possui uma escala de 0 a 100 e, como se trata de um *ranking*, quanto menor o valor, melhor a colocação da instituição e mais impacto ela proporciona nas três dimensões avaliadas. Para fins de exemplificação, no *ranking* mundial das universidades, a Universidade de Harvard (*Harvard University*) se encontra em 1º lugar, a Universidade de Salamanca (*Universidad de Salamanca*) se encontra em 471º lugar, 15º lugar na Espanha, e a Universidade de São Paulo em 22º lugar, sendo a 1ª colocada no *ranking* de universidades brasileiras (SCIMAGO, 2023).

O Quadro 71 exibe os repositórios, suas colocações no *ranking* SCImago a nível de universidades brasileiras e geral mundial, a faixa de avaliação e as respectivas pontuações totais.

Quadro 71 – Repositórios, ranking SIR, faixa de avaliação e pontuações totais.

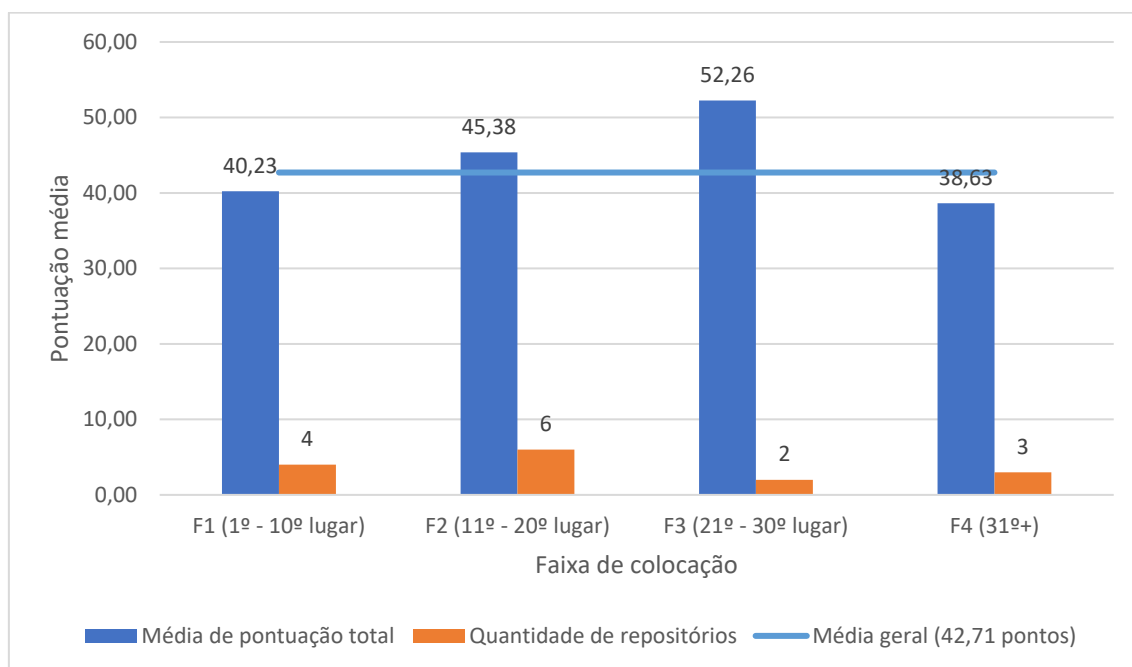
Repositório	SCImago universidades brasileiras	Faixa de avaliação	SCImago geral	Pontuação total
R05	3º	F1	600º	64,65
R03	5º	F1	733º	47,94
R18	9º	F1	1.681º	23,49
R19	10º	F1	2.210º	24,85
R08	13º	F2	3.029º	51,11
R10	15º	F2	3.109º	34,63
R07	16º	F2	3.173º	47,24
R06	17º	F2	3.183º	36,10
R09	18º	F2	3.327º	49,28
R11	18º	F2	3.327º	53,94
R12	26º	F3	4.030º	37,65
R14	29º	F3	4.377º	66,86
R17	31º	F4	4.457º	66,74
R16	35º	F4	4.891º	17,90
R15	41º	F4	5.266º	31,26
R01	Não se aplica	-	-	46,03
R02	Não se aplica	-	-	21,01
R04	Não se aplica	-	-	27,43
R13	Não se aplica	-	-	63,40

Fonte: elaboração própria a partir de (SCIMAGO, 2023).

Com o intuito de agrupar as colocações dos repositórios em estratos para a realização de análises, adotou-se 4 faixas de avaliação de acordo com as posições dos repositórios no SIR: 1) F1: colocações entre 1º e 10º lugar; 2) F2: colocações entre 11º e

20º lugar; 3) F3: colocações entre 21º e 30º lugar, e; 4) F4: colocações a partir do 31º lugar. A Figura 78 exibe o gráfico das pontuações médias dos repositórios agrupadas pelas faixas de avaliação criadas com a quantidade de repositório em cada faixa e a média geral de pontuação dos repositórios. É possível perceber que os 4 repositórios mais bem colocados no SIR possuem pontuação média, 40,23 pontos, abaixo da pontuação média geral dos repositórios, 42,71 pontos, assim com os 3 repositórios que ocupam a partir da 31ª colocação no SIR. Nas 2 faixas intermediárias, que compreende entre a 11ª e a 30ª colocação, a pontuação média dos repositórios supera a pontuação média geral.

Figura 78 – Pontuação média dos repositórios de acordo com suas colocações no SIR.



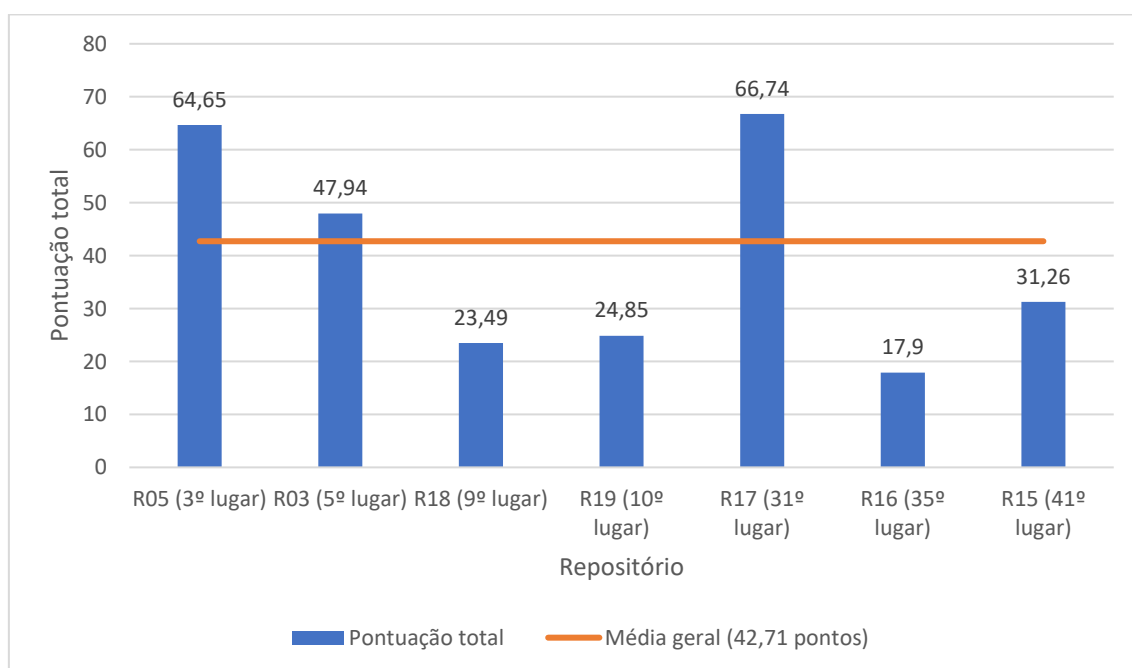
Fonte: elaboração própria.

Buscando-se analisar o desempenho dos 4 primeiros e dos 3 últimos colocados no SIR, a Figura 79 exibe as respectivas pontuações totais e a média geral dos repositórios. Pode-se perceber que na primeira faixa de pontuação, os repositórios R05 e R03, ocupantes respectivamente das 3ª e 5ª colocações do SIR possuem pontuações totais acima da média geral dos repositórios, no entanto, os repositórios R18 e R19, ocupantes respectivamente das 9ª e 10ª colocações possuem pontuações totais bem abaixo da média geral. Do lado oposto, o repositório R17, ocupante da 31ª colocação no SIR possui pontuação muito acima da média geral e ocupa a 2ª colocação

no ranking geral do presente estudo, enquanto os demais repositórios R16 e R15 apresentam pontuações totais abaixo da média geral.

A partir dessa análise, conclui-se que os indicadores que medem a pesquisa, inovação e impacto social utilizados para o cálculo do SIR não possuem relações diretas com as áreas e critérios analisados no presente estudo. Ainda assim, recomenda-se que os repositórios que obtiveram desempenhos abaixo da média no presente trabalho dediquem esforços em direção à excelência na preservação digital confiável, acompanhando os desempenhos obtidos nos rankings mundiais como o SIR.

Figura 79 – Pontuações totais dos repositórios e suas colocações no SIR.



Fonte: elaboração própria.

A seguir são apresentadas as conclusões da seção atual de capítulos, analisando-se os desempenhos de pontuação dos repositórios.

5.6 AVALIAÇÃO GERAL DOS RESULTADOS

A seção atual apresenta um panorama geral do desempenho dos repositórios pesquisados frente às perguntas elaboradas no presente estudo. Os resultados da primeira seção “1. Informações sobre o respondente” foram omitidos pois serviram para caracterizar o perfil das pessoas que forneceram as respostas para o questionário e não são comparáveis entre si. As respostas das perguntas das demais seções que possuem

características mensuráveis e que não são de natureza descritiva foram incluídas na análise exibida no Quadro 72.

A primeira coluna do quadro “**Seção do questionário/Pergunta**” exibe as seções e respectivas perguntas, a coluna “**Respostas**” exibe o total de repositórios que forneceram alguma resposta à pergunta, sendo 19 o total de repositórios participantes do presente estudo. A coluna “**Cumprem na prática**” representa o percentual e a fração de repositórios que cumpre e aplica na prática os requisitos do critério da pergunta.

A coluna “**Em teoria**”, representa o percentual e a fração de repositórios que sinalizaram que possuem ações planejadas somente em teoria ou parcialmente em implantação, relativas aos requisitos da pergunta. Em alguns casos, as perguntas não possuíam respostas para a identificação da existência em teoria dos critérios, não possuindo resultados preenchidos nessa coluna.

A coluna “**Aplicação da teoria**” sumariza os resultados dos repositórios que aplicam na prática os requisitos da pergunta com os resultados dos repositórios que declararam aplicar parcialmente ou somente em teoria os mesmos requisitos. Os valores dessa coluna permitem realizar uma projeção dos resultados alcançados caso os planejamentos atualmente em teoria ou parcialmente em execução fossem aplicados totalmente na prática, permitindo uma visualização da situação dos repositórios como um todo.

Como forma de destacar os resultados observados, adotou-se uma escala de cores, utilizando-se cor vermelha quando o percentual está na faixa entre 0% e 24,99%, a cor amarela quando o percentual está entre 25% e 49,99%, a cor verde clara para os percentuais entre 50% e 74,99% e a cor verde escura para a faixa entre 75% e 100%.

Quadro 72 – Desempenho dos repositórios em relação ao alcance efetivo ou em teoria dos requisitos das perguntas.

Seção do questionário/Pergunta	Respostas	Cumprem na prática	Em teoria	Aplicação da teoria
1. Contextualização do repositório				
1.1 O repositório institucional possui missão, visão e valores elaborados?	18	61,11% 11/18	N.A. ¹¹²	
1.2 O repositório institucional possui a sua Comunidade-Alvo delimitada?	19	78,95% 15/19	N.A.	
2. Estratégia e responsabilidade				
2.1 Existe uma Política de Preservação com os requisitos mínimos dos documentos recebidos, tratamento de preservação e acesso?	18	5,55% 1/18	61,11% 11/18	66,67% 12/18
2.2 Existe um Plano de Continuidade que garante que os documentos estarão preservados e acessíveis a longo prazo, caso ocorra algum desastre ou descontinuidade do funcionamento?	17	5,88% 1/17	29,41% 5/17	35,29% 6/17
3. Certificações e normas				
3.1 O repositório possui certificações alcançadas?	19	0% 0/19	N.A.	
3.2 Quais certificações o repositório tem a intenção de alcançar? (O repositório possui intenções de alcançar certificações?)	16	75% 12/16	N.A.	
3.3 O repositório adota padrões e normas estabelecidos?	19	100% 19/19	N.A.	
4. Infraestrutura organizacional				
4.1 Existe uma estrutura organizacional adequada ao funcionamento do repositório para a preservação a longo prazo?	18	27,78% 5/18	33,33% 6/18	61,11% 11/18
4.2 Existe orçamento adequado ao funcionamento do repositório para a preservação a longo prazo?	17	11,76% 2/17	35,29% 6/17	47,06% 8/17
4.3 Em relação à quantidade de pessoas, a equipe responsável pelo repositório é adequada?	19	15,79% 3/19	15,79% 3/19	31,58% 6/19
4.4 Em relação às capacitações e conhecimentos, a equipe responsável pelo repositório é adequada?	18	38,89% 7/18	11,11% 2/18	50% 9/18
4.5 Existe um plano de capacitação permanente da equipe do repositório?	17	17,65% 3/17	29,41% 5/17	47,06% 8/17
5. Infraestrutura técnica e tecnológica				
5.1 O repositório possui estrutura física de prédios e/ou salas adequada à preservação e garantia de preservação a longo prazo?	18	50% 9/18	16,67% 3/18	66,67% 12/18
5.2 O repositório possui estrutura de equipamentos (servidores de aplicação, armazenamento, rede, etc) adequada à preservação e garantia de preservação a longo prazo?	17	47,06% 8/17	23,56% 4/17	70,59% 12/17
5.3 Existe um plano de atualização periódico de <i>hardware</i> e <i>software</i> ?	17	17,65%	52,94%	70,59%

¹¹² Não se aplica.

		3/17	9/17	12/17
6. Segurança				
6.1) Os riscos que afetam a preservação dos objetos digitais são identificados e avaliados	15	0% 0/15	26,67% 4/15	26,67%
6.3) A estrutura de TI (<i>software</i> e <i>hardware</i>) do repositório adota medidas de segurança contra ataques?	16	62,50% 10/16	25% 4/16	87,50% 14/16
6.4) A estrutura física do repositório adota medidas de segurança para evitar acesso indevido?	16	56,25% 9/16	25% 4/16	81,25% 13/16
7. Integridade dos dados				
7.1) O repositório implementa o monitoramento e tratamento dos formatos de arquivos recebidos?	17	23,53% 4/17	52,94% 9/17	76,47% 13/17
7.2) O repositório adota medidas para prevenir o corrompimento dos dados?	14	35,71% 5/14	21,43% 3/14	57,14% 8/14
7.3) O repositório possui um plano de backup?	17	82,35% 14/17	17,65% 3/17	100% 17/17
8. Metadados				
8.1) O repositório estabelece um padrão de metadados reconhecido a ser adotado pelos documentos recebidos?	18	72,22% 13/18	22,22% 4/18	94,44% 17/18
8.2) São recolhidos/registrados metadados que permitam identificar o contexto, referência, direitos e autenticidade dos documentos?	15	53,33% 8/15	40% 6/15	93,33% 14/15
9. Acesso e consulta				
9.1) Existe uma política de acesso aos objetos digitais?	18	38,89% 7/18	33,33% 6/18	72,22% 13/18
9.2) Existe uma política de monitoramento das necessidades da sua comunidade-alvo a fim de disponibilizar os documentos da forma correta e nos formatos de arquivo mais adequados?	17	23,53% 4/17	23,53% 4/17	47,05% 8/17
10. Visibilidade				
10.1) É possível localizar o acesso ao repositório a partir da página principal da instituição?	19	52,63% 10/19	47,37% 9/19	100% 19/19
10.2) Existem iniciativas para a promoção do repositório institucional?	19	52,63% 10/19	31,58% 6/19	84,21% 16/19
10.4) Na página web do repositório, são utilizadas quais as formas de contato com os responsáveis pelo repositório? (Existe uma forma de contato?)	19	100% 19/19		100% 19/19

Fonte: elaboração própria.

Pode-se perceber que os percentuais de alcance de determinadas perguntas se modificam consideravelmente de forma positiva caso haja a aplicação prática total das medidas planejadas pelos repositórios. Como exemplo, cita-se a pergunta 3.1 que trata da existência de um plano de preservação nos repositórios, que registra 5,55% de aplicação prática total, porém caso os repositórios que declararam possuir o plano somente em teoria ou parcialmente aplicados realizassem esforços para a aplicação total, se alcançaria 66,67% dos 18 repositórios que declararam respostas, ou ainda 63,16% do total de 19 repositórios participantes do estudo.

Outro caso semelhante pode ser observado com a pergunta 8.1 que verifica se os repositórios monitoram e tratam os formatos de arquivos recebidos na ingestão. Na prática, 23,53% aplicam esse monitoramento, enquanto 52,94% declararam que existe o planejamento de monitoramento somente em teoria ou parcialmente aplicado. Em caso de efetivação prática desses monitoramentos, se alcançaria 76,47% do total de 17 repositórios que responderam à pergunta, ou 68,42% do total de 19 repositórios participantes do estudo.

Outra situação de evolução nos resultados pode ser observada na pergunta 11.1 que trata de disponibilizar um acesso ao repositório institucional na página principal da respectiva instituição. Atualmente, foi registrado que 52,63% possuem esse acesso implementado na prática, porém 47,37% possuem o planejamento em teoria para essa disponibilização de acesso. Na efetivação desses planejamentos, se chegaria a 100% dos repositórios possuindo acesso a partir das páginas das suas instituições, contribuindo para a visibilidade dos repositórios e promovendo o uso de seus materiais.

Do lado oposto, cita-se como exemplo a pergunta 7.1 que trata sobre a identificação e avaliação de riscos dos repositórios, onde foi registrado que nenhum repositório realiza essas tarefas, enquanto 26,67%, dos 15 repositórios que responderam, registraram que existe o planejamento em teoria, porém não aplicados. Mesmo em uma eventual efetivação desses planejamentos, o percentual de repositórios que possuiria identificação e avaliação de riscos seria muito baixo, representando 21,05% do total de 19 repositórios pesquisados. A aplicação de processos de identificação e avaliação de impacto dos riscos permite um autoconhecimento das fragilidades e pontos fortes do repositório, permitindo planejar ações para evitar ou mitigar eventos que possam comprometer a preservação confiável a longo prazo dos seus materiais. Como consequência, a existência desse mapeamento de riscos impacta positivamente na confiança do repositório frente à sua Comunidade-Alvo.

Os resultados permitem identificar grandes disparidades entre as áreas, demonstrando um alto grau de evolução de alguns repositórios ao mesmo tempo em que revela grandes desafios a serem transpostos por outros repositórios rumo a melhores desempenhos. Uma vez que a maioria dos repositórios pertence a instituições públicas, sabe-se que as decisões e iniciativas de mudanças requerem sensibilizações, diálogos e negociações em esferas superiores das instituições, tornando o processo de mudança lento e, muitas vezes, burocrático. Além desse fato, os investimentos necessários às adaptações e melhorias das instituições requerem previsão legal, planejamento orçamentário e direcionamento das verbas, possibilitando que outras áreas da instituição sejam beneficiadas em detrimento dos repositórios. No entanto, é possível que se criem redes de cooperação entre as instituições e parcerias com os repositórios privados, como é o caso da Rede Cariniana, favorecendo o compartilhamento de conhecimentos e experiências, visando diminuir as diferenças existentes em busca da excelência na preservação digital confiável.

6 CONCLUSÕES

A realização do presente estudo permitiu identificar a forma como os repositórios participantes da Rede Cariniana realizam a preservação de seus materiais, partindo-se de aspectos mais gerais com a existência de planos e políticas de preservação até a identificação de normas e padrões utilizados nos repositórios. Em termos de participação, 19 dos 31 repositórios participaram do estudo, possibilitando um grande alcance de instituições, no entanto, somente 1 repositório privado forneceu respostas, impossibilitando um conhecimento mais abrangente sobre a dinâmica de preservação dos repositórios dessa natureza. Assim, ainda que o conhecimento de aspectos específicos de instituições privadas possa ter sido prejudicado, a grande participação de instituições públicas permitiu conhecer o estado atual da preservação digital dos repositórios.

A forma como a comunicação com os repositórios foi realizada, em sua maioria, ocorreu através de manifestações de Pedido de Acesso à Informação na plataforma Fala.BR, gerando um certo grau de formalidade e burocracia no fornecimento das respostas, oposto ao que se pretendia no presente estudo ao realizar as primeiras comunicações via mensagens de e-mail. Mesmo assim, as manifestações foram respondidas e permitiram a realização das análises necessárias, sendo fornecidas pelos responsáveis pelos repositórios e integrantes das suas equipes, agregando confiabilidade e conhecimento nas respostas recebidas.

Buscando realizar uma comparação objetiva entre os repositórios, foi elaborada uma ferramenta de pontuação dos critérios avaliados no estudo, com pesos iguais entre as áreas e perguntas dentro das áreas. Assim, foi possível identificar pontos fortes e aspectos a serem melhorados nos repositórios que permitem um diagnóstico da situação atual, mas também podem servir para acompanhar a evolução dos aspectos no futuro. Além de servir para a criação do *ranking* de repositórios, a ferramenta pode ser usada para o intercâmbio de soluções entre os repositórios, incentivando a troca de experiências que possibilitem um amadurecimento e evolução da preservação digital na Rede Cariniana.

O presente capítulo apresenta efetivamente as conclusões extraídas do estudo. A primeira seção apresenta o nível de cumprimento dos objetivos propostos, em seguida, são apresentadas as conclusões gerais, as confirmações ou refutações das hipóteses e as linhas futuras de pesquisa sugeridas.

6.1 NÍVEL DE CUMPRIMENTO DOS OBJETIVOS

A presente seção recapitula os objetivos propostos pelo presente trabalho, identificando as seções que apontam os seus respectivos cumprimentos. O objetivo geral **OG1**, “definir um modelo geral de análise e avaliação de repositórios digitais com base em fontes normativas internacionais”, caracteriza-se como cumprido, uma vez que todos os objetivos específicos que o compõe foram cumpridos, conforme se apresenta em seguida. O Quadro 73 resume os objetivos específicos desse objetivo geral e as seções do trabalho onde se encontram os respectivos cumprimentos.

Quadro 73 – Cumprimento dos objetivos específicos do Objetivo Geral 1.

OG1: Definir um modelo geral de análise e avaliação de repositórios digitais com base em fontes normativas internacionais	Seção do trabalho
OE1.1: identificar e analisar as ferramentas de avaliação de repositórios digitais usadas no contexto internacional, a fim de obter uma visão geral dos elementos comuns e das diferenças mais significativas entre elas.	Seção 4.2.2.1.
OE1.2: identificar, selecionar e classificar os critérios de avaliação para a certificação de repositórios confiáveis identificados nas ferramentas mencionadas no ponto anterior, a fim de definir as áreas de gestão de repositórios digitais nas quais um processo de avaliação deve se concentrar.	Seção 4.2.2.2.
OE1.3: elaborar um questionário estruturado para a avaliação de repositórios digitais que possa ser aplicado em qualquer ambiente institucional, tecnológico e normativo.	Seção 4.2.3.
OE1.4: projetar uma ferramenta de avaliação quantitativa	Seção 4.2.5.
OE1.5: testar o modelo de avaliação em um ambiente delimitado no contexto brasileiro, a fim de comprovar a aplicação do modelo e validar sua utilidade.	Capítulo 5.

Fonte: elaboração própria.

O cumprimento do objetivo específico **OE1.1** “identificar e analisar as ferramentas de avaliação de repositórios digitais usadas no contexto internacional, a fim de obter uma visão geral dos elementos comuns e das diferenças mais significativas entre elas” é demonstrado na seção 4.2.2.1. O cumprimento do objetivo específico

OE1.2 “identificar, selecionar e classificar os critérios de avaliação para a certificação de repositórios confiáveis identificados nas ferramentas mencionadas no ponto anterior, a fim de definir as áreas de gestão de repositórios digitais nas quais um processo de avaliação deve se concentrar” é demonstrado na seção 4.2.2.2. O cumprimento do objetivo específico **OE1.3** “elaborar um questionário estruturado para a avaliação de repositórios digitais que possa ser aplicado em qualquer ambiente institucional, tecnológico e normativo” é demonstrado na seção 4.2.3. O cumprimento do objetivo específico **OE1.4** “projetar uma ferramenta de avaliação quantitativa, associada ao questionário indicado no ponto anterior, que permita fazer comparações objetivas entre as diferentes áreas de avaliação do mesmo repositório e entre diferentes repositórios” é demonstrado na seção 4.2.5. O cumprimento do objetivo específico **OE1.5** “testar o modelo de avaliação em um ambiente delimitado no contexto brasileiro, a fim de comprovar a aplicação do modelo e validar sua utilidade” é demonstrado no capítulo 5 de resultados, uma vez que a aplicação do modelo foi realizada nos repositórios da Rede Cariniana.

Quadro 74 – Cumprimento dos objetivos específicos do Objetivo Geral 2.

OG2: Analisar o estado da preservação digital dos repositórios digitais da Rede Cariniana usando o modelo definido anteriormente	
OE2.1: avaliar aspectos relacionados ao planejamento e gestão dos repositórios digitais da Rede Cariniana	Seções 5.4.1, 5.4.2, 5.4.3 e 5.4.4.
OE2.2: avaliar os aspectos técnicos dos repositórios e a utilização de metadados de preservação	Seções 5.4.5 e 5.4.8.
OE2.3: avaliar as técnicas e mecanismos utilizados de modo a garantir a segurança e integridade dos dados	Seções 5.4.6 e 5.4.7.
OE2.4: avaliar de que forma os repositórios difundem seus materiais e se fazem acessíveis ao público	Seções 5.4.9 e 5.4.10.

Fonte: elaboração própria.

O objetivo geral **OG2** “analisar o estado da preservação digital dos repositórios digitais da Rede Cariniana usando o modelo definido anteriormente” considera-se cumprido pois todos os objetivos específicos que o compõe foram cumpridos, conforme se demonstra a seguir. O Quadro 74 resume os objetivos específicos desse objetivo geral e as seções do trabalho onde se encontram os respectivos cumprimentos.

O cumprimento do objetivo específico **OE2.1** “avaliar aspectos relacionados ao planejamento e gestão dos repositórios digitais da Rede Cariniana” é demonstrado nas seções 5.4.1, 5.4.2, 5.4.3 e 5.4.4. O cumprimento do objetivo específico **OE2.2** “avaliar os aspectos técnicos dos repositórios e a utilização de metadados de preservação” é demonstrado nas seções 5.4.5 e 5.4.8. O cumprimento do objetivo específico **OE2.3** “avaliar as técnicas e mecanismos utilizados de modo a garantir a segurança e integridade dos dados” é demonstrado nas seções 5.4.6 e 5.4.7. O cumprimento do objetivo específico **OE2.4** “avaliar de que forma os repositórios difundem seus materiais e se fazem acessíveis ao público” é demonstrado nas seções 5.4.9 e 5.4.10.

O objetivo geral **OG3** “determinar a existência de fatores externos que podem influenciar na qualidade dos repositórios digitais, estudando o caso da Rede Cariniana” considera-se cumprido pois todos os objetivos específicos que o compõe foram cumpridos, conforme se demonstra a seguir. O Quadro 75 resume os objetivos específicos desse objetivo geral e as seções do trabalho onde se encontram os respectivos cumprimentos.

Quadro 75 – Cumprimento dos objetivos específicos do Objetivo Geral 3.

OG3: Determinar a existência de fatores externos que podem influenciar na qualidade dos repositórios digitais, estudando o caso da Rede Cariniana	
OE3.1: determinar a influência do entorno geográfico na qualidade dos repositórios	Seção 5.5.1.
OE3.2: determinar a influência dos aspectos relacionados com o entorno institucional na qualidade dos repositórios	Seções 5.5.2, 5.5.3, 5.5.5.1, 5.5.5.2, 5.5.5.3 e 5.5.5.4
OE3.3: determinar como o tempo de criação do repositório (idade) influencia na sua qualidade.	Seção 5.5.4.

Fonte: elaboração própria.

O cumprimento do objetivo específico **OE3.1** “determinar a influência do entorno geográfico na qualidade dos repositórios” é demonstrado na seção 5.5.1. O cumprimento do objetivo específico **OE3.2** “determinar a influência dos aspectos relacionados com o entorno institucional na qualidade dos repositórios” é demonstrado nas seções 5.5.2, 5.5.3 e 5.5.5 e suas devidas subseções. O cumprimento do objetivo específico **OE3.3** “determinar como o tempo de criação do repositório (idade) influencia na sua qualidade” é demonstrado na seção 5.5.4.

6.2 AVALIAÇÃO DAS HIPÓTESES

A partir das hipóteses propostas no presente estudo, nessa seção são apresentadas as suas confirmações ou refutações conforme se demonstra a seguir. De modo a auxiliar a compreensão, o Quadro 76 apresenta um resumo das confirmações ou refutações das hipóteses propostas no estudo.

Quadro 76 – Hipóteses e suas avaliações.

Hipótese	Confirmação
H1: os repositórios institucionais da Rede Cariniana possuem estruturas organizacionais, físicas e tecnológicas para a realização da preservação digital a longo prazo	Confirmada com ressalvas.
H2: os repositórios institucionais brasileiros utilizam diretrizes estabelecidas para a realização da preservação de seus materiais digitais	Parcialmente confirmada.
H3: os repositórios institucionais brasileiros não aplicam ferramentas de avaliação de confiabilidade	Confirmada
H4: a elaboração de um conjunto de critérios de confiabilidade mínimos, adaptados de padrões internacionais, é útil para a avaliação dos repositórios brasileiros e para a formulação de planos de melhoria	Confirmada
H5: a qualidade dos repositórios é influenciada por fatores externos, relacionados aos seus entornos	Confirmada
H6: os repositórios mais antigos, criados há mais tempo, devem possuir mais qualidade que os repositórios mais recentes	Refutada

Fonte: elaboração própria.

A hipótese **H1** “os repositórios institucionais da Rede Cariniana possuem estruturas organizacionais, físicas e tecnológicas para a realização da preservação digital a longo prazo” é considerada **confirmada com ressalvas**. Quando se observam os resultados dos critérios relacionados à infraestrutura organizacional, disponibilidade orçamentária e adequação das equipes, apresentados na seção 5.4.4, a hipótese H1 poderia ser refutada, uma vez que as porcentagens não ultrapassam 50% de aplicação prática. No entanto, quando se considera a aplicação parcial e a existência em teoria dos critérios, os números ultrapassam 50%, daí a confirmação da hipótese com a ressalva de que os repositórios se comprometam a investirem em esforços para a aplicação prática total dos devidos critérios. As equipes dos repositórios se propuseram a desenvolver soluções para a preservação dos materiais digitais, mesmo com todas as dificuldades que essas deficiências impõem a essa atividade. No entanto, é vital que haja uma evolução em busca da excelência da preservação digital, buscando a criação ou

ampliação de estruturas organizacionais com cargos e atribuições bem definidas, orçamentos bem dimensionados que permitam a manutenção das atividades e investimentos em equipamentos e capacitações.

A hipótese **H2** “os repositórios institucionais brasileiros utilizam diretrizes estabelecidas para a realização da preservação de seus materiais digitais” é **parcialmente confirmada**, a partir dos resultados demonstrados nas seções 5.4.2 e 5.4.7. Analisando-se as diretrizes utilizadas pelos repositórios, a maioria possui Políticas de Acesso e Planos de *Backup* aplicados total ou parcialmente na prática, no entanto, a maioria dos repositórios não possui Política de Preservação e Plano de Continuidade, de onde se origina o fato de a hipótese ser parcialmente confirmada. Cabe alertar para o fato de que poucos repositórios declararam possuir uma política de monitoramento das suas Comunidades-Alvo, buscando compreender e identificar as suas necessidades e assim como reagir a elas, conforme demonstrado na seção 5.4.9.

A quase totalidade da ausência de planos de continuidade nos repositórios se torna um fato alarmante, uma vez que os repositórios desconhecem as ameaças, por não identificar e avaliar os riscos e esperam que eventos que impeçam os seus funcionamentos não ocorram. Ressalta-se, aqui o fato de que não somente eventos catastróficos podem cessar as atividades dos repositórios, mas também eles podem se tornar alvos de erros gerenciais e ingestões políticas que coloquem em risco a autenticidade dos seus materiais. Uma possível razão levantada para a deficiência de planos de continuidade é a de que a natureza pública da grande maioria das instituições dos repositórios cria um cenário de segurança, seja pela estabilidade da carreira das equipes como pela natureza das verbas advindas do governo, desnecessitando a elaboração do plano de continuidade. O presente estudo não permite identificar em que grau essa situação seria similar em repositórios de natureza privada, devido ao baixo índice de participação desses repositórios, no entanto acredita-se que o estabelecimento de contratos e acordos de nível de serviço fomentem o estabelecimento de planos de continuidade maduros e eficazes. Uma vez estabelecido o plano de continuidade, estabelece-se um constante monitoramento de gatilhos que possam ameaçar o repositório e seus materiais, buscando nunca cessar a preservação digital confiável.

A deficiência de diretrizes de preservação permite concluir que a preservação digital dos repositórios é realizada de forma empírica, ainda que a *expertise* das equipes dos repositórios permita um conhecimento do funcionamento das suas instituições e Comunidades-Alvo. As diretrizes de preservação formam um corpus

teórico que baliza todo o funcionamento dos repositórios, assim como respalda e delimita as suas responsabilidades e os protege de eventuais situações legais e jurídicas que venham a surgir. Assim, recomenda-se um tratamento emergencial dessas questões por parte dos repositórios, a fim de sanar essa fragilidade que pode comprometer as suas atividades.

A hipótese **H3** “os repositórios institucionais brasileiros não aplicam ferramentas de avaliação de confiabilidade” é **confirmada** de acordo com o exposto na seção 5.4.3. Comprova-se a inexistência de uma ferramenta de avaliação ligada a práticas, normas e auditorias internacionais, porém voltada à realizada brasileira, de modo a dificultar a aplicação de instrumentos dessa natureza nos repositórios brasileiros. Ainda, comprova-se um baixo grau de conhecimento de avaliações e certificações de confiabilidade pelas equipes dos repositórios, fato que pode ser originado por razões como o desconhecimento ou a falta de planejamento de estudos da área ou ainda por restrições de equipe e/ou orçamentárias. Os repositórios ao alcançarem certificações ou aplicar métodos de auditoria contribuem para o estabelecimento do sentimento de confiança frente às suas comunidades, uma vez que constatarem cumprir requisitos mínimos necessários para a preservação digital a longo prazo. Conclui-se que há uma baixa priorização dos repositórios sobre o conhecimento e aquisição de certificações e auditorias como um todo, sinalizando um baixo embasamento teórico sobre os critérios que avaliam a confiabilidade da preservação digital. Recomenda-se que haja o estudo da área de certificações, modelos de maturidade e auditorias para que a confiança seja construída, se estabeleça e se concretize frente as suas comunidades de usuários.

A hipótese **H4** “a elaboração de um conjunto de critérios de confiabilidade mínimos, adaptados de padrões internacionais, é útil para a avaliação dos repositórios brasileiros e para a formulação de planos de melhoria” é **confirmada** uma vez que a aplicação da ferramenta de avaliação desenvolvida, seus resultados e análise são apresentados no Capítulo 5, e a seção 5.6.

A hipótese **H5** “a qualidade dos repositórios é influenciada por fatores externos, relacionados aos seus entornos” é **confirmada**, conforme as análises realizadas na seção 5.5. Foi possível constatar a influência direta que determinadas variáveis externas aos repositórios possuem nos seus desempenhos de pontuação identificados. Esses fatores permitem identificar uma forte desigualdade entre os repositórios quando se trata de questões alheias às suas alçadas. No caso dos repositórios públicos, essa desigualdade poderia ser diminuída através do

estabelecimento de políticas públicas que, de certa forma, priorizem os repositórios com menor desempenho, visando melhorar a preservação como um todo a nível nacional.

A hipótese **H6** “os repositórios mais antigos, criados há mais tempo, devem possuir mais qualidade que os repositórios mais recentes” é **refutada**, conforme as análises apresentadas na seção 5.5.4. Os resultados demonstraram repositórios criados mais recentemente com pontuações gerais muito superiores aos mais antigos que, em tese, deveriam estar mais organizados em relação a estruturas, diretrizes, etc. Assim, não foi possível estabelecer uma relação direta entre a idade dos repositórios e os seus desempenhos de pontuação.

6.3 CONCLUSÕES GERAIS E RECOMENDAÇÕES

A partir de uma abordagem mais geral, buscou-se identificar a existência dos documentos que guiam a estratégia e responsabilidade dos repositórios como as políticas e planos de preservação, políticas de acesso e planos de continuidade. Em seguida, buscou-se identificar os aspectos relacionados à adequação das estruturas organizacionais, físicas e técnicas/tecnológicas para a preservação confiável a longo prazo. Utilizando-se uma visão mais específica, buscou-se identificar como os repositórios se relacionam com as certificações de confiança e quais mecanismos e processos são adotados para garantir que os materiais preservados sejam dignos de confiança e não sejam corrompidos ou adulterados ao longo do tempo. Por fim, como forma de analisar como os repositórios se promovem e promovem o uso de seus materiais custodiados, foram verificados aspectos relacionados à visibilidade dos repositórios frente às suas Comunidades-Alvo.

Uma visão global dos resultados, apresentados na seção 5.3, constatou que, mesmo o repositório com pontuação mais alta registrada, 66,86 pontos, possui uma pontuação relativamente baixa de uma escala com máximo de 100 pontos, assim como a média geral calculada foi de 42,71 pontos, abaixo de 50% da pontuação máxima possível. Opostamente, o repositório com menor pontuação alcançou 17,90 pontos, apresentando um grande distanciamento da média geral e um distanciamento bastante preocupante do primeiro colocado. Assim, conclui-se que, de um modo geral, os repositórios devem investir em iniciativas e projetos que visem melhorar as questões relacionadas à confiabilidade da preservação digital. Outra conclusão é a de que existe um abismo de diferenças entre as realidades dos repositórios, sinalizando que a própria

Rede Cariniana pode desenvolver iniciativas desde o âmbito da conscientização até o fomento de verbas para as melhorias dos repositórios.

A maioria dos repositórios declarou possuir Missão, Visão e Valores, conforme apresentado na seção 5.4.1, o que pode ser visto como um aspecto positivo, pois essas informações norteiam o estabelecimento de políticas dos repositórios e os seus funcionamentos, além de transmitir às suas Comunidades-Alvo os valores pelos quais os repositórios se pautam. No entanto, no caso dos repositórios que não possuem suas missões declaradas, essa ausência pode ser um fator impeditivo para a realização de auditorias e certificações, uma vez que os propósitos principais dos repositórios não são identificados e, portanto, a avaliação de critérios pode ser comprometida. Assim, se recomenda que todos os repositórios elaborem e aprovem suas missões, visões e valores e os tornem públicos, permitindo que as suas comunidades tomem conhecimento dos propósitos que movem os repositórios.

De maneira similar, a maioria dos repositórios, 15 ao todo, declarou possuir uma delimitação das suas Comunidades-Alvo, conforme apresentado na seção 5.4.1, o que favorece um olhar direcionado e específico para as necessidades dos seus usuários, buscando atender aos seus requisitos. Conclui-se que os repositórios possuem a noção de quem são os seus usuários e suas necessidades, ainda que falte uma materialização e uma certa oficialização dessa definição, a partir da qual outras ações podem ser tomadas como o monitoramento das necessidades da Comunidade-Alvo. Cabe sinalizar que os 4 repositórios restantes devem tratar essa questão com um certo grau de importância e urgência, buscando compreender para quem preservam e quais são suas necessidades.

Um fator preocupante identificado e apresentado na seção 5.4.2 foi a ausência de Políticas de Preservação ou Planos Estratégicos de Preservação na grande maioria dos repositórios. Ambos documentos podem ser tidos como essenciais para o estabelecimento de um repositório, pois é através deles que se sabe o seu propósito e se tem um ponto inicial do qual partem todas as ações, são estabelecidos processos e atividades voltados à realização da preservação confiável do repositório. A existência desses documentos pode ser tida como obrigatória nos repositórios, trazendo à tona o fato alarmante de que 17 repositórios que responderam às perguntas relacionadas a esses temas relataram não possuir os documentos. Conclui-se que os repositórios participantes realizam a preservação dos seus materiais de maneira empírica e com pouco embasamento, uma vez que faltam as políticas que a orientam, direcionam e estabelecem o escopo da própria preservação.

A inexistência de um Plano de Continuidade, apresentado na seção 5.4.2, nos repositórios foi um fato alarmante identificado a partir das respostas dos participantes, nas quais apenas 1 repositório, dos 19 participantes, relatou possuir esse documento. Como reflexo desse cenário, não se têm garantias de que os materiais depositados nos repositórios serão de fato preservados e acessíveis na ocorrência de eventos que impossibilitem a operação desses repositórios. Ainda, de acordo com a Lei nº 8.159 de 8 de janeiro de 1991, o Arquivo Nacional é o órgão máximo responsável pela preservação e acesso dos documentos dos órgãos do poder executivo federal (BRASIL, 1991), logo, faz-se necessária a existência de um acordo e elaboração de um plano para que os repositórios institucionais federais possam transferir os seus materiais para o Arquivo Nacional na ocorrência de eventos de falha.

Sendo o Arquivo Nacional o responsável final pela preservação e acesso dos documentos próprios e das demais instituições, o próprio órgão deve possuir um plano de continuidade elaborado e robusto que garanta que os materiais de pesquisa, documentos arquivísticos e históricos sejam corretamente preservados e mantenham-se acessíveis, independente de quaisquer eventos que possam vir a ocorrer. Cabe lembrar que a descontinuação de um repositório pode ocorrer devido a eventos catastróficos naturais, dificuldades financeiras ou até mesmo por mudança de escopo, finalidade e cessação de interesses, o que torna os repositórios públicos vulneráveis aos interesses dos governantes em mandato. Conclui-se que os repositórios atualmente não compreendem a necessidade, importância e urgência de se estabelecer um Plano de Continuidade e que os seus gestores vislumbram um horizonte pleno de prosseguimento das suas operações, sem contarem com imprevistos, catástrofes ou mudanças de políticas que os afetem. Assim, se recomenda que os repositórios atendam a essa questão o mais urgentemente possível.

Analisando-se os resultados da seção 5.4.3 relativos à presença e interesse dos repositórios por certificações de confiança, verifica-se um baixo percentual alcançado, que pode ser originado por diversas razões como o desconhecimento ou a falta de planejamento de estudos da área ou ainda por restrições de equipe e/ou orçamentárias. Os repositórios ao alcançarem certificações ou aplicar métodos de auditoria contribuem para o estabelecimento de sentimento de confiança frente às suas comunidades, uma vez que constatarem cumprir requisitos mínimos necessários para a preservação digital a longo prazo. Conclui-se que há uma baixa priorização dos repositórios sobre o conhecimento e aquisição de certificações, sinalizando um baixo

embasamento teórico sobre os critérios que avaliam a confiabilidade da preservação digital. Recomenda-se que haja o estudo da área de certificações, modelos de maturidade e auditorias para que a confiança seja construída, se estabeleça e se concretize frente as suas comunidades de usuários.

De um modo geral, tanto as estruturas organizacionais, físicas, técnicas e tecnológicas dos repositórios pesquisados não são adequadas à preservação digital quando se considera a aplicação prática existente. No entanto, caso as instituições que possuem aspectos parcialmente aplicados ou somente teoria, os aplicassem totalmente, o cenário se modificaria, alcançando-se a maioria e, portanto, essas estruturas estariam adequadas.

A partir dos resultados da seção 5.4.4, foram identificados problemas graves quanto à inexistência de equipes capacitadas atuando nos repositórios assim como a inexistência de recursos financeiros que permitam o funcionamento e investimentos nos repositórios. Em relação aos orçamentos, a partir dos 17 repositórios que forneceram respostas, somente 2 relataram possuir orçamentos adequados e aplicados na prática, enquanto 6 relataram que o orçamento existe somente em teoria ou não está totalmente aplicado, enquanto em 9 repositórios, a maioria, foi relatado que o orçamento não é adequado.

Sendo a maioria dos repositórios pertencentes a instituições públicas federais, elas dependem de verbas da administração federal que as disponibiliza de acordo com os seus interesses e prioridades. Internamente, essas instituições alocam os devidos percentuais dos recursos para a manutenção das suas atividades e realizações de investimentos, notando-se uma baixa prioridade na alocação de verbas adequadas às atividades de preservação digital. Essa baixa priorização, precarização e escassez de investimentos perpassa décadas e é relatada por Maher (1978), Beagrie (2003), Lavoie e Dempsey (2004) e Matlala, Ncube e Parbanath (2022). Esse último estudo faz uma revisão de literatura sobre o estado da preservação digital na África do Sul, observando situações semelhantes às encontradas no presente estudo (2022, p. 206, tradução nossa):

A preservação de documentos digitais no setor público da África do Sul é inibida principalmente pela falta ou escassez das ferramentas e habilidades necessárias. Essas ferramentas incluem políticas, diretrizes e padrões adequados, infraestrutura, apoio da alta administração, orçamento dedicado e equipamentos adequados, além de uma força de trabalho qualificada. A ausência de políticas, diretrizes e padrões afeta a eficácia da preservação de documentos digitais. Além disso, o estudo revelou que a preservação digital de documentos na África do Sul continua a enfrentar uma escassez de equipamentos de armazenamento; opera em condições ambientais precárias;

prática práticas inadequadas de manuseio de arquivos, tanto por parte dos funcionários quanto dos usuários; as instalações de armazenamento são inadequadas; trabalha com instalações subdesenvolvidas, recebe financiamento inadequado para a preservação; e carece de funcionários treinados e experientes.

A falta de orçamentos adequados aos repositórios impacta diretamente na forma como a preservação digital é realizada, limitando investimentos que vão desde a aquisição e manutenção de equipamentos, passa pela contratação de equipes qualificadas e culmina na escassez de verbas para os investimentos em capacitações, por exemplo. A baixa prioridade destinada à preservação digital nos repositórios pode ser reflexo do não imediatismo da percepção de retorno das verbas aplicadas, uma vez que os investimentos realizados constroem uma base para o futuro, quando muitos gestores e governantes não estarão mais ocupando seus cargos.

Torna-se evidente a importância e urgência com as quais os gestores das instituições dos repositórios devem tratar a preservação digital, de maneira a favorecer a criação e estabelecimento de estruturas sólidas que permitam preservar confiavelmente e a longo prazo os materiais para usos futuros. De forma complementar, cabe aos gestores e equipes dos repositórios a tarefa de conscientizar e sensibilizar os gestores de suas instituições, comunidade e sociedade como um todo sobre a importância das tarefas que desenvolvem, evidenciando tanto os benefícios de uma preservação digital realizada adequadamente quanto os prejuízos, impactos e punições que podem vir a acontecer essa preservação não se realiza plenamente.

Conclui-se que, apesar das deficiências de orçamento, de estruturas técnicas e organizacionais e de pessoal, as equipes dos repositórios realizam a preservação utilizando os recursos possíveis e disponíveis, adaptando-se às necessidades e demandas que a preservação digital exige, acumulando responsabilidades e se apoiando em esforços individuais, sinalizando áreas que devem ter alta prioridade para os seus atendimentos nos repositórios e suas instituições.

A análise das respostas sobre as estruturas técnicas e tecnológicas dos repositórios apresentou resultados satisfatórios, no sentido de haver equipamentos e ferramentas de preservação instaladas em todos os repositórios, conforme demonstrado na seção 5.4.5. Esse fato revela um alto grau de direcionamento e importância para os aspectos tecnológicos envolvidos na preservação digital em detrimento dos outros aspectos existentes. Os resultados ainda permitem concluir que a existência de uma ferramenta de preservação aderente ao modelo de referência OAIS, baseada em pacotes

de informação, traz a crença de que o repositório se torna um Arquivo OAIS ou que se estabelece um RDC-Arq. Ao se olhar o panorama geral da ISO 14721 e da Resolução n. 43, que define o RDC-Arq, percebe-se que ambas englobam muito mais aspectos que os puramente tecnológicos, possuindo requisitos organizacionais e políticas de preservação, por exemplo. Daí a caracterização de que a crença dos repositórios quanto a serem aderentes à OAIS ou serem RDC-Arqs é errônea.

Ainda, a inexistência de planos de atualização de *hardware* e *software* que foi relatada em 14 repositórios que responderam à questão. É essencial que exista um monitoramento para controle da obsolescência da tecnologia envolvida na preservação digital, observando-se desde o tempo de vida útil das mídias de armazenamento até as versões de *software* utilizadas, por exemplo, evitando que haja perdas ou corrompimento de dados. Conclui-se que os materiais preservados na maioria dos repositórios correm sérios riscos de se tornarem inúteis em algum momento futuro, uma vez que os equipamentos, programas e formatos de arquivos podem se tornarem obsoletos.

A partir das análises, se identificou que todas as soluções de preservação, desde o nível gerencial até o operacional, são realizadas por equipes e setores internos das instituições às quais os repositórios pertencem, sem haver terceirização da preservação dos materiais digitais. Essa situação pode ser causada por diferentes fatores como a falta de conhecimentos sobre a terceirização desses serviços ou a possíveis restrições orçamentárias, baseadas na crença de que as instituições não possuiriam verbas para sustentar os contratos firmados. Uma vez que os relatos apontam carências nas estruturas organizacionais, equipes e orçamentos dos repositórios, sugere-se como alternativa estudos de viabilidade de contratação de serviços e empresas voltadas à prestação de serviços de preservação digital, utilizando-se como base a ISO 17068 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012) que trata sobre Repositórios de Terceiros Confiáveis para Documentos de Arquivo (*Trusted Third Party Repository for Digital Records*). Optando por essa direção, os repositórios atuariam estabelecendo as suas políticas e acordos de níveis de serviço, por exemplo, a serem utilizados na fiscalização dos serviços contratados, atuando mais em demandas de gestão e menos em demandas operacionais.

Em relação a metadados, aspecto analisado na seção 5.4.8, a maioria dos repositórios declarou possuir um padrão estabelecido assim como declararam existir metadados que apoiem a presunção de autenticidade dos materiais digitais preservados,

sejam eles totalmente ou parcialmente recolhidos na prática. Essa situação é corroborada pelo fato de que a área que avaliou critérios de metadados foi a que obteve maior desempenho médio dentre todas as demais áreas. A análise da qualidade dessas informações recolhidas para a comprovação da autenticidade requer a aplicação de uma metodologia específica e direcionada, concluindo-se que requer um estudo minucioso e detalhado, no entanto, ultrapassa o escopo do presente trabalho.

Em relação à salvaguarda dos materiais preservados, destaca-se positivamente a existência de planos de *backup* em execução na maioria dos repositórios, contando, inclusive, com tecnologias diferentes entre as cópias e localizações geográficas distintas, conforme é apresentado na seção 5.4.7. Conclui-se que as soluções de cópias de segurança dos repositórios, em geral, estão bem estruturadas e possuem certo grau de resiliência a falhas e acontecimentos de catástrofes.

Constata-se que a preservação dos repositórios se estrutura fortemente na utilização de tecnologias com ferramentas de preservação e estratégias de *backup*, ainda que essas situações não constem em planos de preservação ou planos de *backup*, por exemplo. Ao que se pode concluir que esse foco estreito em tecnologia advém do fato de se associar a preservação digital apenas ao aspecto da garantia de não corrompimento de dados, em detrimento dos demais aspectos envolvidos. Recomenda-se que os repositórios realizem estudos mais amplos da preservação digital, seja através de projetos internos ou via colaborações entre instituições, para que haja um entendimento mais amplo das características e responsabilidades envolvidas na preservação digital, de modo a aproveitar e direcionar corretamente as soluções tecnológicas já desenvolvidas.

Uma vez que o DSpace é a ferramenta de preservação utilizada pela grande maioria dos repositórios e ela já possui embutida funcionalidades de acesso e consulta dos materiais, destaca-se positivamente o alto grau de disponibilização dos acervos como um todo. No entanto, os repositórios possuem carências no estabelecimento das regras sobre quais materiais podem ser acessados e sob quais condições específicas, além do fato de que uma minoria possui um monitoramento das necessidades das suas Comunidades-Alvo. Conclui-se que os repositórios difundem e dão acesso aos seus materiais sem se basear em regras plenamente estabelecidas, correndo riscos de infringirem direitos ou disponibilizar formatos de arquivos não apropriados. Os repositórios devem destacar esforços para o estabelecimento das suas políticas de acesso, incluindo planos de monitoramento das Comunidades-Alvo, buscando compreender os seus requisitos e anseios assim como reagir às mudanças identificadas.

Em relação à visibilidade dos repositórios, constatou-se que, em praticamente metade dos casos, as páginas iniciais das suas instituições possuem formas de direcionamento aos repositórios, conforme apresentado na seção 5.4.10. Acredita-se que a situação dos repositórios que não possuem esse acesso possa ser resolvida com poucos esforços tanto das equipes dos repositórios quanto das equipes responsáveis pelas páginas institucionais, no entanto, essa melhoria poderia facilitar e aumentar a utilização dos repositórios. De um modo geral, todos os repositórios disponibilizam formas de contato com a comunidade, facilitando a comunicação entre a comunidade e as equipes responsáveis. Conclui-se que os repositórios possuem ferramentas implementadas de relacionamento com as comunidades, necessitando de poucos ajustes para que as páginas institucionais possuam meios de direcionar os usuários aos repositórios em si.

Ao final, a execução do presente estudo permitiu identificar as ferramentas, instrumentos, métodos e normas existentes para a avaliação de repositórios digitais, de onde se elaborou uma ferramenta com ampla cobertura de critérios desses referencias existentes. Quando se analisa um repositório e se identifica o atendimento a um critério, é possível inferir o atendimento ao mesmo tempo a diversos critérios dos instrumentos utilizados como referência, favorecendo e facilitando a aplicação do instrumento desenvolvido. Assim, os repositórios podem mapear os seus pontos fortes e debilidades que necessitam ser tratadas, servindo como embasamento para a tomada de decisão das gestões dos repositórios, por exemplo.

A aplicação do instrumento nos repositórios brasileiros possibilitou a sua própria homologação, ao mesmo tempo em que permitiu identificar o estado atual da preservação desses repositórios. A partir dos dados coletados, foram feitas análises, sendo possível concluir que existem muitos desafios a serem transpassados por esses repositórios em direção à excelência na preservação digital, que vão desde a reformulação de suas estruturas organizacionais, priorização de investimentos, capacitações de equipes, até o monitoramento das necessidades das Comunidades-Alvo. No entanto, sabe-se que o caminho para essas mudanças requer primeiramente a sensibilização dos gestores para que as prioridades sejam compreendidas, de onde o presente estudo pode vir a servir como apoio.

6.4 LINHAS FUTURAS DE PESQUISA

Dentre as possibilidades identificadas para trabalhos futuros, está a diferenciação de pesos e valores entre as áreas e perguntas de acordo com os seus graus de importância para a preservação digital. Para a realização desse trabalho, deve-se elencar quais áreas e perguntas são mais críticas e essenciais para a preservação digital e quais podem ser consideradas melhorias ou aperfeiçoamentos de processos, utilizando-se tanto as normas de certificação e auditoria disponíveis como também a opinião de pesquisadores da área e a realização de testes-piloto.

Junto à coordenação da Rede Cariniana, ainda é possível elaborar uma espécie de calendário de avaliação dos repositórios integrantes, executada no momento do ingresso dos repositórios e depois periodicamente, buscando compreender a forma como a participação na Rede afeta as suas evoluções das suas preservações digitais. Assim, é possível diagnosticar áreas carentes comuns aos repositórios que necessitariam de ações de melhoria ou ainda estabelecer um nível mínimo exigido para o ingresso e permanência na Rede.

A automatização da coleta de informações dos repositórios, no que diz respeito aos critérios avaliados, pode vir a ser realizada através do uso de APIs gerenciadas pela coordenação da Rede Cariniana, que teria acesso aos dados automaticamente. Para o caso de determinadas áreas e critérios, a API poderia coletar informações das ferramentas de preservação enviando-as à coordenação e facilitando a tarefa de fornecer respostas ao formulário de avaliação. Nesse caso, cabe analisar a viabilidade técnica, intenção da Rede e repositórios participantes para a implementação dessa solução.

Uma vez que a ferramenta desenvolvida avaliou especificamente os repositórios da Rede Cariniana, a sua aplicação pode se estender para outros grupos de repositórios a nível nacional ou internacional, uma vez que a avaliação utiliza critérios baseados em normas de certificação e auditoria internacionais. Essa ampliação contribuiria para o avanço da preservação digital mundial, identificando pontos fortes existentes praticados e deficiências a serem corrigidas, além de permitir a realização de *benchmarks* com outros países. A viabilização desse estudo pode requerer a coordenação um órgão central que normatize a realização do estudo e se responsabilize pela coleta e análise dos dados, sugerindo ou demandando mudanças identificadas nos

repositórios, de acordo com as suas respectivas esferas administrativas. Assim, seria necessária a criação de futuras parcerias para a efetivação da ampliação desse estudo.

REFERÊNCIAS

- ABRAMS, S.; KUNZE, J.; LOY, D. An Emergent Micro-Services Approach to Digital Curation Infrastructure. **The International Journal of Digital Curation**, v. 5, n. 1, 21 Jul 2010. 172-186. Disponível em: <<https://doi.org/10.2218/ijdc.v5i1.151>>. Acesso em: 19 out. 2022.
- ALI, M.; LOAN, F. A.; MUSHATQ, R. Open Access Scientific Digital Repositories : An Analytical Study of the Open DOAR. **5th International Symposium on Emerging Trends and Technologies in Libraries and Information Services (ETTLIS)**, Noida, India, 2018. 213-216. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/8485265>>. Acesso em: 05 maio 2023.
- ALONSO LÓPEZ, A. L. Archivos sólidos en un sociedad líquida: La archivística en el marco del eGov. **Tábula**, v. 25, 2023. 101-121.
- ALONSO, V. C. Los documentos y su tratamiento archivístico. **Boletín de la ANABAD**, v. 31, n3, 1981. 365-381. Disponível em: <<https://dialnet.unirioja.es/descarga/articulo/965337.pdf>>. Acesso em: 30 nov. 2022.
- ALTMAN, D. G.; BLAND, J. M. Standard deviations and standard errors. **BMJ**, v. 331, 15 October 2005. 903. Disponível em: <<https://doi.org/10.1136/bmj.331.7521.903>>. Acesso em: 01 ago. 2023.
- ANDRADE, F. L. D.; CHAGAS, C. A. Repositórios digitais confiáveis: a verificação de compatibilidade entre modelos internacionais de critérios de preservação digital no longo prazo e o RDC-Arq. **Ágora: Arquivologia em Debate**, v. 33, n. 6, 2023. 1-22. Disponível em: <<https://agora.emnuvens.com.br/ra/article/view/1177>>. Acesso em: 31 jan. 2023.
- ARCHIVES NATIONALES. Archives Nationales (France), 2023. Disponível em: <<https://www.archives-nationales.culture.gouv.fr/en/web/guest/histoire-de-l-institution>>.
- ARELLANO, M. Á. M.; OLIVEIRA, A. F. D. Gestão de Repositórios de Preservação Digital. **RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, SP, v. 14, n. 3, 30 set. 2016. 465-483. Disponível em: <<https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/864634>>.
- ARQUIVO NACIONAL. **Dicionário brasileiro de terminologia arquivística**. Rio de Janeiro: Arquivo Nacional, 2005.
- ARQUIVO NACIONAL. **Código de Classificação de Documentos de Arquivo Relativos às Atividades-Fim das Instituições Federais de Ensino Superior - IFES**. Rio de Janeiro: Arquivo Nacional, 2011. Disponível em: <https://www.gov.br/arquivonacional/pt-br/servicos/gestao-de-documentos/orientacao-tecnica-1/codigo-de-classificacao-e-tabela-de-temporalidade-e-destinacao-de-documentos-de-arquivo/copy_of_portaria_n0922011_codigo_de_classificacao_de_documentos_.pdf>. Acesso em: 22 dez. 2022.
- ARQUIVO NACIONAL. **Tabela de Temporalidade e Destinação Final de Documentos de Arquivo Relativo às Atividades-Fim das Instituições Federais de Ensino Superior - IFES**. Rio de Janeiro: Arquivo Nacional, 2011. Disponível em: <https://www.gov.br/arquivonacional/pt-br/servicos/gestao-de-documentos/orientacao-tecnica-1/codigo-de-classificacao-e-tabela-de-temporalidade-e-destinacao-de-documentos-de-arquivo/copy_of_portaria_n0922011_tabela_de_temporalidade_e_destinacao_.pdf>. Acesso em: 22 dez. 2022.

- ARQUIVO NACIONAL. **Código de classificação e tabela de temporalidade e destinação de documentos relativos às atividades-meio do Poder Executivo Federal**. Rio de Janeiro: Arquivo Nacional, 2020. Disponível em: <https://www.gov.br/arquivonacional/pt-br/servicos/gestao-de-documentos/orientacao-tecnica-1/codigo-de-classificacao-e-tabela-de-temporalidade-e-destinacao-de-documentos-de-arquivo/copy_of_cod_classif_-e_tab_temp_2019_m_book_digital_25jun2020.pdf>. Acesso em: 22 dez. 2022.
- ARQUIVO NACIONAL. **Recomendações para elaboração da listagem de eliminação de documentos de arquivo pelos órgãos e entidades do poder executivo federal. Versão atualizada em junho 2022**. Rio de Janeiro: Arquivo Nacional, 2020. Disponível em: <https://www.gov.br/arquivonacional/pt-br/servicos/gestao-de-documentos/orientacao-tecnica-1/recomendacoes-tecnicas-1/recomendacao_led_2020_03_24_1.pdf>. Acesso em: 07 fev. 2023.
- ARTEFACTUAL SYSTEMS INC. Archivematica: open-source digital preservation system, 2022. Disponível em: <<https://www.archivematica.org/pt-br/>>. Acesso em: 26 ago. 2022.
- ARTEFACTUAL SYSTEMS INC. AtoM: Open Source Archival Description Software, 2022. Disponível em: <<https://www.accesstomemory.org/pt-br/>>. Acesso em: 17 ago. 2022.
- ARTEFACTUAL SYSTEMS INC. Archivematica: open-source digital preservation system. Documentation, 2023. Disponível em: <<https://archivematica.org/pt-br/docs/archivematica-1.13/>>. Acesso em: 10 maio 2023.
- ARTEFACTUAL SYSTEMS INC. PREMIS Metadata in Archivematica. Documentation. **Archivematica**, 2023. Disponível em: <<https://www.archivematica.org/en/docs/archivematica-1.13/user-manual/metadata/premis/#premis-template>>. Acesso em: 14 mar. 2023.
- ARXIV. Arxiv e-Print Archive. **Arxiv**, 2023. Disponível em: <<https://arxiv.org/>>. Acesso em: 30 abr. 2023.
- ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN. **UNE-ISO 14721. Sistemas de transferencia de datos e información espaciales. Sistema abierto de información de archivo (OAIS). Modelo de referencia**. Madrid: UNE, 2015.
- ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN. **UNE-ISO 30302. Información y documentación. Sistemas de gestión para los documentos. Guía de implantación**. Madrid: AENOR, 2015.
- ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN. **UNE-ISO 15489-1. Información y documentación. Gestión de documentos. Parte 1: Generalidades**. Madrid: AENOR, 2016.
- ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN. **UNE-ISO 16363. Sistemas de transferencia de información y datos espaciales. Auditoría y certificación de repositorios digitales de**. Madrid: AENOR, 2017.
- ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN. **UNE-ISO 30301. Información y documentación. Sistemas de gestión para los documentos. Requisitos**. Madrid: UNE, 2019.
- ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN. **UNE-ISO 17068. Información y documentación. Repositorio de tercero de confianza para documentos electrónicos**. Madrid: AENOR, 2020.
- ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN. **UNE-ISO 30300. Información y documentación. Gestión de documentos. Conceptos fundamentales y vocabulario**. Madrid: UNE, 2021.

- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR ISO 15489-1**: Informação e documentação - Gestão de documentos de arquivo. Parte 1: Conceitos e princípios. Rio de Janeiro: ABNT, 2018.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR ISO 19011**: Diretrizes para auditoria de sistemas de gestão. Rio de Janeiro: ABNT, 2018.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR ISO 23081-1**: Informação e documentação - Processos de gestão de documentos de arquivo - Metadados para documentos de arquivo - Parte 1: Princípios. Rio de Janeiro: ABNT, 2019.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR ISO 23081-2**: Informação e documentação — Gerenciamento de metadados para documentos de arquivo - Parte 2: Problemas conceituais e implementação. Rio de Janeiro: ABNT, 2020.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR ISO 14721**: Sistemas espaciais de transferência de dados e informação - Sistema Aberto de Arquivamento de Informação (SAAI) - Modelo de referência. Rio de Janeiro: ABNT, 2021.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR ISO 17068**: Informação e documentação. Repositório de terceiros confiável para documentos de arquivo digital. Rio de Janeiro: ABNT, 2021.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR ISO/TR 23081-3**: Informação e documentação - Gerenciamento de metadados para documentos de arquivo - Parte 3: Método de autoavaliação. Rio de Janeiro: ABNT, 2022.
- AZORÍN MILLARUELO, C. et al. **Guía para la evaluación de repositorios institucionales de investigación. Modificación de la segunda versión**. Madrid: FECYT, 2015.
- AZORÍN, C. et al. **Evaluación del estado de la preservación de los repositorios REBIUN 2018**. XVII Workshop REBIUN de Proyectos Digitales. Sevilla, España: [s.n.]. 2018.
- AZORÍN, C. et al. **Guía para la evaluación de repositorios institucionales de investigación. Cuarta versión**. Madrid: FECYT, 2021. Disponível em: <https://repositoriorebiun.org/bitstream/handle/20.500.11967/809/2021GuiaEvaluacionRecolecta_vf%281%29.pdf?sequence=1&isAllowed=y>. Acesso em: 25 ago. 2022.
- BAILEY, C. W. J. **SPEC Kit 292**: Institutional Repositories. July 2006. University of Houston Libraries. Institutional Repository Task Force. Association of Research Libraries. Washington, DC, USA: ARL, 2006. Disponível em: <<https://doi.org/10.29242/spec.292>>. Acesso em: 04 maio 2023.
- BALL, A. **Web Archiving. Version 1.1. DCC State of the Art Report**. Edinburgh, UK: Digital Curation Centre, 2010. Disponível em: <<https://www.dcc.ac.uk/sites/default/files/documents/reports/sarwa-v1.1.pdf>>. Acesso em: 10 maio 2023.
- BARNARD, A.; DELGADO, A.; VOUTSSÁS, J. **Los Archivos Digitales. Una visión integradora. Tomo I**. Primera edición. ed. Puebla, México: Archivo Histórico Universitario, 2014. Disponível em: <https://iibi.unam.mx/voutssasmt/documentos/archivos_digitales_3_corto.pdf>. Acesso em: 21 set. 2023.
- BARRUECO CRUZ, J. M. et al. **Guía para la evaluación de repositorios institucionales de investigación. V 1.0**. Madrid: FECYT, 2010. Disponível em: <<http://hdl.handle.net/10261/35735>>. Acesso em: 01 mar. 2023.
- BARRUECO, J. M. **La Guía de Evaluación de Repositorios**. 15 años Recolecta. Alcobendas, España: FECYT. 2022.
- BAUCOM, E. et al. **Using the Levels of Digital Preservation as an Assessment Tool**. [S.l.]: NDSA, 2019. Disponível em: <<https://osf.io/m6j4q>>. Acesso em: 19 ago. 2022.

- BEAGRIE, N. **National Digital Preservation Initiatives: An Overview of Developments in Australia, France, the Netherlands, and the United Kingdom and of Related international Activity.** Washington, DC, USA: Council on Library and Information Resources, 2003. Disponível em: <<https://www.clir.org/wp-content/uploads/sites/6/pub116.pdf>>. Acesso em: 20 jun. 2023.
- BEAGRIE, N. et al. **Digital Preservation Policies Study. Part 1: final report october 2008.** [S.l.]: Charles Beagrie Limited, 2008.
- BEAGRIE, N.; POTHEN, P. **Digital Curation: Digital Archives, Libraries and e-Science Seminar.** Digital Preservation Day. London, UK: [s.n.]. 2001.
- BEARMAN, D. Record-Keeping Systems. *Archivaria*, v. 36, 1993. 16-36.
- BEDEIAN, A. G.; MOSSHOLDER, K. W. On the Use of the Coefficient of Variation as a Measure of Diversity. *Organizational Research Methods*, v. 3, n. 3, July 2000. 285-297. Disponível em: <<https://doi.org/10.1177/109442810033005>>. Acesso em: 01 ago. 2023.
- BJÖRK, B.-C. Open Access Subject Repositories: An Overview. *Journal of the Association for Information Science and Technology*, v. 65 n. 4, apr 2014. 698-706.
- BONAL ZAZO, J. L.; DE LORENZO-CÁCERES, M. D. P. O. Criterios de certificación y auditoría de repositorios digitales seguros en archivos. In: VAQUINHA, N.; CAIXAS, M.; VINAGRE, H. **Da produção à preservação informacional: desafios e oportunidades.** Évora: Publicações do Cidehus, 2017. p. 529-550. ISBN 9782821882676. Disponível em: <<http://books.openedition.org/cidehus/2835>>. Acesso em: 08 mar. 2023.
- BRAGA, T. E. N.; HOLANDA, A. P. D.; CANELHAS, T. Resolução RDC-Arq CONARQ: uma análise dos novos requisitos informacionais propostos. *Revista Brasileira de Preservação Digital (RBDP)*, Campinas, SP, 2022. 1-10. Disponível em: <<https://econtents.bc.unicamp.br/inpec/index.php/rebpred/article/view/16583/11385>>. Acesso em: 10 maio 2023.
- BRASIL. **Lei Nº 8.159, de 8 de janeiro de 1991. Dispõe sobre a política nacional de arquivos públicos e privados e dá outras providências.** Brasília. 1991.
- BRASIL. **Lei Nº 12.527 de 18 de novembro de 2011. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei nº 8.112, de 11 de dezembro de 1990.** Brasília, DF. 2011.
- BRASIL. **Lei Nº 13.460 de 26 de junho de 2017. Dispõe sobre participação, proteção e defesa dos direitos do usuário dos serviços públicos da administração pública.** Brasília, DF. 2017.
- BRASIL. **Lei Nº 13.709 de 14 de agosto de 2018. Lei Geral de Proteção de Dados (LGPD).** Brasília. 2018.
- BRASIL. **Decreto Nº 4.915, de 12 de dezembro de 2003. Dispõe sobre o Sistema de Gestão de Documentos e Arquivos da administração pública federal.** Brasília: [s.n.], 2019. Disponível em: <https://www.planalto.gov.br/ccivil_03/decreto/2003/D4915.htm>. Acesso em: 30 mar. 2023.
- BROWN, A. **Practical digital preservation. A how-to guide for organizations of any size.** London, England: Facet publishing, 2013.
- BROWN, P. et al. **Bethesda Statement on Open Access Publishing. Released June 20, 2003.** Maryland: [s.n.], 2003. Disponível em: <<https://www.ouvrirlascience.fr/wp-content/uploads/2019/03/Bethesda-Statement-on-Open-Access-Publishing.pdf>>. Acesso em: 01 maio 2023.
- BUDAPEST OPEN ACCESS INITIATIVE. Read the Declartaion. **Budapest Open Access Initiative**, 2002. Disponível em: <<https://www.budapestopenaccessinitiative.org/read/>>. Acesso em: 01 maio 2023.

- BUSTELO-RUESTA, C. La serie de normas iso 30300 - management system for records: la gestión de los documentos integrada en la gestión de las organizaciones. **Acervo - Revista do Arquivo Nacional**, Rio de Janeiro, v. 28, n. 2, jul/dez 2015. 72-83. Disponível em: <<https://www.brapi.inf.br/index.php/res/v/42723>>. Acesso em: 17 ago. 2022.
- CAFÉ, L. et al. Repositórios institucionais: nova estratégia para publicação científica na Rede. **XXVI Congresso Brasileiro de Ciências da Comunicação. Trabalho apresentado no XIII ENDOCOM, 03 de setembro de 2003, Belo Horizonte/ MG**, Belo Horizonte, MG, 2-6 set 2003. Disponível em: <http://www.intercom.org.br/papers/nacionais/2003/www/pdf/2003_endocom_trabalho_cafe.pdf>. Acesso em: 05 maio 2023.
- CAPLAN, P. **Entendendo o PREMIS. Publicado originalmente em 2009. Revisado em 2017.** Tradução de Amarilis Montagnolli Gomes Corrêa e Laerte Pereira Silva Júnior. [S.l.]: PREMIS Editorial Committee, 2009. Disponível em: <https://www.loc.gov/standards/premis/understandingPREMIS_portuguese_2021.pdf>. Acesso em: 13 mar. 2023.
- CENTER FOR RESEARCH LIBRARIES. **Ten Principles**. Chicago, USA: CRL, 2007. Disponível em: <<https://www.crl.edu/archiving-preservation/digital-archives/metrics-assessing-and-certifying/core-re>>. Acesso em: 22 fev. 2023.
- CENTRO UINVERSITÁRIO INTERNACIONAL UNINTER. Institucional. **Centro Universitário Uninter**, 2022. Disponível em: <<https://www.uninter.com/centro-universitario-internacional/>>. Acesso em: 24 ago. 2022.
- CHRISTIANE, P. **The Myth of Immateriality - Presenting & Preserving New Media**. REFRESH! THE FIRST INTERNATIONAL CONFERENCE ON THE HISTORIES OF ART, SCIENCE AND TECHNOLOGY. [S.l.]: [s.n.], 2019.
- COMISIÓN DE NORMAS ESPAÑOLAS DE DESCRIPCIÓN ARCHIVÍSTICA. **Modelo conceptual de descripción archivística: Entidades, Relaciones y Atributos**. 2017. ed. Madrid: Secretaría General Técnica, 2017. Acesso em: 28 junho 2022.
- CONFEDERATION OF OPEN ACCESS REPOSITORIES. Repository Toolkit. **Confederation of Open Access Repositories**, 2019. Disponível em: <<https://coartraining.gitbook.io/coar-repository-toolkit/>>. Acesso em: 10 maio 2023.
- CONFEDERATION OF OPEN ACCESS REPOSITORIES. **COAR Annual Report 2022-2023. April 2023. Version for members**. Göttingen, Germany: COAR Office, 2023. Disponível em: <<https://www.coar-repositories.org/files/COAR-Annual-Report-2022-2023.pdf>>. Acesso em: 10 maio 2023.
- CONSELHO INTERNACIONAL DE ARQUIVOS. **ISAD(G): Norma Internacional de Descrição Arquivística: segunda edição**, adotada pelo Comitê de Normas de Descrição, Estocolmo, Suécia, 19-22 de setembro de 1999, versão aprovada pelo CIA. Versão brasileira. Rio de Janeiro: Arquivo Nacional, 2000. ISBN 85-7009-032-3.
- CONSELHO INTERNACIONAL DE ARQUIVOS. **ISDF: Norma Internacional para Descrição de Funções**. Primeira edição. Elaborada pelo Comitê de Boas Práticas e Normas. Dresden, Alemanha, 2-4 de maio de 2007. Tradução brasileira. Rio de Janeiro: Arquivo Nacional, 2008.
- CONSELHO INTERNACIONAL DE ARQUIVOS. **ISDIAH: Norma Internacional para Descrição de Instituições com Acervo Arquivístico**. Primeira edição. Elaborada pelo Comitê de Boas Práticas e Normas. Londres, Reino Unido, 10-11 de março de 2008. Tradução brasileira. Rio de Janeiro: Arquivo Nacional, 2009.
- CONSELHO NACIONAL DE ARQUIVOS. **NOBRADE: Norma Brasileira de Descrição Arquivística**. Rio de Janeiro: Arquivo Nacional, 2006. ISBN 85-60207-02-3. Disponível em: <<https://www.gov.br/conarq/pt-br/centrais-de-conteudo/publicacoes/nobrade.pdf>>. Acesso em: 28 set. 2022.

CONSELHO NACIONAL DE ARQUIVOS. **e-ARQ Brasil: Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos. Versão 1.1.** Adotada pelo Conselho Nacional de Arquivos em dezembro de 2009. Rio de Janeiro: CONARQ, 2011. ISBN 978-85-60207-30-5.

CONSELHO NACIONAL DE ARQUIVOS. **Resolução Nº 37, de 19 de dezembro de 2012. Aprova as Diretrizes para a Presunção de Autenticidade de Documentos Arquivísticos Digitais.** Rio de Janeiro: Arquivo Nacional, 2012.

CONSELHO NACIONAL DE ARQUIVOS. **Resolução N 43, de 04 de setembro de 2015. Diretrizes para a implementação de repositórios arquivísticos digitais confiáveis - RDC-Arq.** Rio de Janeiro: Arquivo Nacional, 2015. Disponível em: <https://www.gov.br/conarq/pt-br/centrais-de-conteudo/publicacoes/conarq_diretrizes_rdc_arq_resolucao_43.pdf>. Acesso em: 19 ago. 2022.

CONSELHO NACIONAL DE ARQUIVOS. **RIC-CM: consultation draft v0.1. Comments by Brazilian Rio de Janeiro Group and Câmara Técnica de Documentos Eletrônicos of Conselho Nacional de Arquivos.** Conselho Nacional de Arquivos. Rio de Janeiro. 2016.

CONSELHO NACIONAL DE ARQUIVOS. **Glossário dos Documentos Arquivísticos Digitais.** 8ª. ed. Rio de Janeiro: Arquivo Nacional, 2020.

CONSELHO NACIONAL DE ARQUIVOS. **e-ARQ Brasil: Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos. Versão 2.** Adotada pelo Conselho Nacional de Arquivos em junho de 2021. Rio de Janeiro: CONARQ, 2022. ISBN 978-85-7009-007-2. Disponível em: <<https://www.gov.br/conarq/pt-br/centrais-de-conteudo/publicacoes/EARQV203MAI2022.pdf>>. Acesso em: 18 jan. 2023.

CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS. **Reference Model for an Open Archival Information System (OAIS). Recommendation for Space Data System Standards. CCSDS 650.0-B-1. Blue Book.** Washington, DC, USA: CCSDS Secretariat, 2002. Disponível em: <<https://public.ccsds.org/Pubs/650x0b1s.pdf>>. Acesso em: 15 fev. 2023.

CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS. **Audit and Certification of Trustworthy Digital Repositories (ACTDR). Recommendation for Space Data System Practices. Recommended Practice CCSDS 652.0-M-1. Magenta Book.** Washington, DC, USA: CCSDS Secretariat, 2011.

CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS. **Reference Model for an Open Archival Information System (OAIS). 650.0-M-2. Magenta Book.** Washington, DC, USA: CCSDS Secretariat, 2012.

CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS. **Requirements for Bodies Providing Audit and Certification of Candidate Trustworthy Digital Repositories. Recommended Practice. CCSDS 652.1-M-2. Magenta Book.** Washington, DC, USA: CCSDS Secretariat, 2014. Disponível em: <<https://public.ccsds.org/Pubs/652x1m2.pdf>>. Acesso em: 27 fev. 2023.

CONTROLADORIA-GERAL DA UNIÃO. Fala.BR. Plataforma Integrada de Ouvidoria e Acesso à Informação, 2022. Disponível em: <<https://falabr.cgu.gov.br/>>. Acesso em: 08 set. 2022.

CONTROLADORIA-GERAL DA UNIÃO. Portal da Transparência. **Portal da Transparência**, 2023. Disponível em: <<https://portaldatransparencia.gov.br/>>. Acesso em: 10 ago. 2023.

COOK, T. The Concept of the Archival Fonds in the Post-Custodial Era: Theory, Problems and Solutions. **Archivaria. Proceedings of the ACA Seventeenth Annual Conference, Montreal 12-15 September 1992**, n. 35, january 1993. Disponível em: <<https://www.archivaria.ca/index.php/archivaria/article/view/11882>>. Acesso em: 03 fev. 2023.

COOK, T. The Impact of David Bearman on Modern Archival Thinking: An Essay of Personal Reflection and Critique. **Archives and Museum Informatics**, 11, march 1997. 15-37. Disponível em: <<https://doi.org/10.1023/A:1009035121019>>. Acesso em: 19 jan. 2023.

- COOK, T. Archival Science and postmodernism: new formulations for old concepts. **Archival Science**, v. 1, p. 3-24, march 2001.
- CORETRUSTSEAL. Data Seal of Approval Synopsis (2008-2018), 2023. Disponível em: <<https://www.coretrustseal.org/about/history/data-seal-of-approval-synopsis-2008-2018/>>. Acesso em: 24 fev. 2023.
- CORETRUSTSEAL BOARD. About, 2022. Disponível em: <<https://www.coretrustseal.org/about/>>. Acesso em: 05 ago. 2022.
- CORETRUSTSEAL BOARD. Core Certified Repositories, 2022. Disponível em: <<https://www.coretrustseal.org/why-certification/certified-repositories/>>. Acesso em: 08 ago. 2022.
- CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD. **CoreTrustSeal Trustworthy Data Repositories Requirements 2020–2022**. [S.l.]: [s.n.], 2019. Acesso em: 22 jan. 2022.
- CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD. **CoreTrustSeal Trustworthy Data Repositories Requirements 2020-2022 (v02.00-2020-2022)**. [S.l.]: [s.n.], 2019. Acesso em: 05 ago. 2022.
- CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD. **CoreTrustSeal Trustworthy Data Repositories Requirements: Glossary 2023-2025. V01.00**. [S.l.]: [s.n.], 2022. Disponível em: <<https://doi.org/10.5281/zenodo.7051125>>. Acesso em: 02 mar. 2023.
- CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD. **CoreTrustSeal Trustworthy Digital Repositories Requirements 2023-2025. Extended Guidance. V01.00**. [S.l.]: [s.n.], 2022. Disponível em: <<https://doi.org/10.5281/zenodo.7051096>>. Acesso em: 02 mar. 2023.
- CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD. **CoreTrustSeal Trustworthy Digital Repositories Requirements 2023-2025. V01.00**. [S.l.]: [s.n.], 2022. Disponível em: <<https://doi.org/10.5281/zenodo.7051011>>. Acesso em: 02 mar. 2023.
- CORUJO, L. **Introdução aos Repositórios Digitais**. 1ª. ed. Lisboa, Portugal: Edições Colibri, 2022. ISBN 978-989-566-148-0.
- CORUJO, L. M. N. **Repositórios Digitais e Confiança: um exemplo de repositório de Preservação Digital: o RODA**. Dissertação de mestrado. Faculdade de Letras da Universidade de Lisboa. Lisboa: Universidade de Lisboa, 2014. Disponível em: <<http://hdl.handle.net/10760/30375>>. Acesso em: 05 maio 2023.
- CORUJO, L. M. N. **Avaliação Arquivística de Informação Eletrónica. Da Emergência Teórica ao Consenso sobre um Modelo. Tese de doutorado**. Coimbra, Portugal: Universidade de Coimbra, 2020.
- COSTA, M. P. D.; LEITE, F. C. L. **Repositórios institucionais da América Latina e o acesso aberto à informação científica**. Brasília: IBICT, 2017. Disponível em: <<http://hdl.handle.net/10760/31109>>. Acesso em: 01 maio 2023.
- CROW, R. **The Case for Institutional Repositories: A SPARC Position Paper**. Prepared by Raym Crow, SPARC Senior Consultant. Release 1.0. Washington, DC, USA: The Scholarly Publishing & Academic Resources Coalition, 2002. Disponível em: <https://ils.unc.edu/courses/2014_fall/inls690_109/Readings/Crow2002-CaseforInstitutionalRepositoriesSPARCPaper.pdf>. Acesso em: 02 maio 2023.
- DAPPERT, A.; PEYRARD, S.; GUENTHER, R. S. An Introduction to Implementing Digital Preservation Metadata. In: DAPPERT, A.; GUENTHER, R. S.; PEYRARD, S. **Digital Preservation Metadata for Practitioners. Implementing PREMIS**. Cham, Switzerland: Springer, 2016. p. 1-9. ISBN 978-3-319-43763-7 (eBook). Disponível em: <https://doi.org/10.1007/978-3-319-43763-7_3>. Acesso em: 13 mar. 2023.

- DATA SEAL OF APPROVAL. **Guidelines version 1**. [S.l.]: [s.n.], 2010. Disponível em: <<https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:116038/tab/2>>. Acesso em: 05 ago. 2022.
- DATA SEAL OF APPROVAL. **Guidelines version 2014-2017**. [S.l.]: [s.n.], 2013. Disponível em: <<https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:116038/tab/2>>. Acesso em: 05 ago. 2022.
- DE GIUSTI, M. Calidad en los repositorios digitales. Los principios TRUST para repositorios de datos. **Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología**, n. 29, ene-jun 2021. 55-59.
- DE GIUSTI, M. R. et al. **Las actividades y el planeamiento de la preservación en un repositorio institucional**. BIREDIAL - Conferencia Internacional Acceso Abierto, Comunicación Científica y Preservación Digital. Barranquilla, Colombia: [s.n.]. 2012.
- DE GIUSTI, M. R.; VILLARREAL, G. L. Revision of different implementations for digital preservation: towards a methodological proposal for preserving and auditing IR reliability. **Revista Digital Biblioteconomia e Ciência da Informação**, Campinas, SP, v. 16, n. 2, May/Aug 2018. p. 273-292.
- DEUTSCHES INSTITUT FÜR NORMUNG. **DIN 31644**: Information and documentation. Criteria for trustworthy digital archives. Berlin: DIN, 2012.
- DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE. **Digital Repository Audit Method Based on Risk Assessment (DRAMBORA). Draft for Public Testing & Comment. v1.0**. [S.l.]: [s.n.], 2007. Disponível em: <<http://www.repositoryaudit.eu/download>>. Acesso em: 08 ago. 2022.
- DIGITAL PRESERVATION COALITION. **Digital Preservation Coalition Rapid Assessment Model. Version 2**. [S.l.]: DPC, 2021. Acesso em: 02 set. 2022.
- DIGITAL PRESERVATION COALITION. Digital Preservation Awards 2022 Winners Announced. **Digital Preservation Coalition**, 2022. Disponível em: <<https://www.dpconline.org/events/digital-preservation-awards/the-winners>>. Acesso em: 14 mar. 2023.
- DIGITALPRESERVATIONEUROPE. **Repository Planning Checklist and Guidance DPE-D3.2**. [S.l.]: [s.n.], 2008. Disponível em: <https://digitalpreservationeuropa.eu/publications/reports/Repository_Planning_Checklist_and_Guidance.pdf>. Acesso em: 03 mar. 2023.
- DLM FORUM FOUNDATION. **MoReq2010**: Modular Requirements for Records Systems. Volume 1: Core Services & Plug-in Modules. [S.l.]: [s.n.], v. Version 1.1, 2011. Disponível em: <<http://moreq2010.eu/>>. Acesso em: 21 nov. 2022.
- DOBRATZ, S.; SCHÖGER, A.; STRATHMANN, S. The nestor Catalogue of Criteria for Trusted Digital Repository Evaluation and Certification. **Digital Curation & Trusted Repositories**, v. 8. n.2, set. 2007. 1-6. Disponível em: <<https://jodi-ojs-tdl.tdl.org/jodi/article/view/199>>. Acesso em: 16 fev. 2023.
- DUBLIN CORE METADATA INITIATIVE. DCMI Metadata Terms, 2020. Disponível em: <<https://www.dublincore.org/specifications/dublin-core/dcmi-terms/>>. Acesso em: 28 set. 2022.
- DURANTI, L. Registros documentais contemporâneos como provas de ação. Tradução de Adelina Novaes e Cruz. **Revista Estudos Históricos**, Rio de Janeiro, v. 7, n. 13. CPDOC 20 anos, 01 jul. 1994. 50-64. Disponível em: <<https://bibliotecadigital.fgv.br/ojs/index.php/reh/article/view/1976>>. Acesso em: 10 maio 2023.
- DURANTI, L. Reliability and Authenticity: The Concepts and Their Implications. **Archivaria**, 39, May 1995. 5-10. Disponível em: <<https://archivaria.ca/index.php/archivaria/article/view/12063>>. Acesso em: 08 dez. 2022.

- DURANTI, L. Concepts and principles for the management of electronic records, or records management theory is archival diplomatics. **Records Management Journal**, 9, December 1999. 149-171. Disponível em: <<https://doi.org/10.1108/EUM0000000007248>>. Acesso em: 12 dez. 2022.
- DURANTI, L. et al. Strategy Task Force Report. In: FORCE, I. S. T. **The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project**. San Miniato, Italy: Archilab, 2005. Disponível em: <http://www.interpares.org/book/interpares_book_g_part4.pdf>. Acesso em: 18 ago. 2022.
- DURASPACE. About DSpace. **DSpace**, 2022. Disponível em: <<https://dspace.lyrasis.org/about/>>. Acesso em: 25 ago. 2022.
- DURASPACE. DuraSpace registry. **DuraSpace**, 2022. Disponível em: <https://duraspace.org/registry/?gv_search=&filter_10=&filter_4_6=Brazil&filter_3=academic&filter_20=&filter_28=&mode=all>. Acesso em: 14 out. 2022.
- DURASPACE. About DSpace. **DSpace**, 2023. Disponível em: <<https://dspace.lyrasis.org/about/>>. Acesso em: 10 maio 2023.
- DURASPACE. **DSpace. An open source dynamic digital repository. Diagram**. [S.l.]: [s.n.], 2023. Disponível em: <https://dspace.lyrasis.org/wp-content/uploads/2022/01/DSpace_Diagram.pdf>. Acesso em: 12 maio 2023.
- ESPANHA. **Ley 16/1985, de 25 de junio, del Patrimonio Historico Español. BOE 155, de 29 de junio de 1985**. Madrid: BOE, 1985. Disponível em: <<https://www.boe.es/buscar/pdf/1985/BOE-A-1985-12534-consolidado.pdf>>. Acesso em: 31 mar. 2023.
- ESPANHA. **Ley 11/2007, de 22 de junio, de acceso electrónico de los ciudadanos a los Servicios Públicos. «BOE» núm. 150, de 23 de junio de 2007. Referencia: BOE-A-2007-12352**. Madrid: [s.n.], 2007. Disponível em: <<https://www.boe.es/eli/es/l/2007/06/22/11/con>>. Acesso em: 20 dez. 2022.
- ESPANHA. **Orden CUL/1524/2007, de 25 de mayo, por la que se crea la Comisión de Normas Españolas de Descripción Archivística**. [S.l.]: Boletín Oficial del Estado, 2007. 23899-23900 p. ISBN 131. Disponível em: <<https://www.boe.es/eli/es/o/2007/05/25/cul1524>>. Acesso em: 28 junho 2022.
- ESPANHA. **Real Decreto 1671/2009, de 6 de noviembre, por el que se desarrolla parcialmente la Ley 11/2007, de 22 de junio, de acceso electrónico de los ciudadanos a los servicios públicos. BOE núm. 278, de 18 de noviembre de 2009. Referencia: BOE-A-2009-18358**. Madrid: BOE, 2009. Disponível em: <<https://www.boe.es/buscar/pdf/2009/BOE-A-2009-18358-consolidado.pdf>>. Acesso em: 31 mar. 2023.
- ESPANHA. **Real Decreto 4/2010, de 8 de enero, por el que se regula el Esquema Nacional de Interoperabilidad en el ámbito de la Administración Electrónica. BOE núm. 25, de 29 de enero de 2010**. Madrid: BOE, 2010. Disponível em: <<https://www.boe.es/buscar/pdf/2010/BOE-A-2010-1331-consolidado.pdf>>. Acesso em: 20 dez. 2022.
- ESPANHA. **Real Decreto 1708/2011, de 18 de noviembre, por el que se establece el Sistema Español de Archivos y se regula el Sistema de Archivos de la Administración General del Estado y de sus Organismos Públicos y su régimen de acceso. BOE núm. 284**. Madrid: Ministerio de la Presidencia, 2011. Disponível em: <<https://www.boe.es/buscar/pdf/2011/BOE-A-2011-18541-consolidado.pdf>>. Acesso em: 29 mar. 2023.
- ESPANHA. **Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas. BOE n. 236**. Madrid: BOE, 2015. Disponível em: <<https://boe.es/boe/dias/2015/10/02/pdfs/BOE-A-2015-10565.pdf>>. Acesso em: 29 mar. 2023.

ESPAÑA. **Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público**. Madrid: BOE, 2015. Disponível em: <<https://boe.es/boe/dias/2015/10/02/pdfs/BOE-A-2015-10566.pdf>>. Acesso em: 29 mar. 2023.

EUROPEAN COMMISSION. **MoReq2 Specification: Model Requirements for the Management of Electronic Records. Update and Extension**. Paris, France: Office for Official Publications of the European Communities, 2008. ISBN 978-92-79-09772-0. Disponível em: <https://irms.org.uk/resource/resmgr/moreq/moreq2_typeset_version.pdf>. Acesso em: 21 nov. 2022.

EUROPEAN COMMISSION. **Reform of EU data protection rules**. [S.l.]. 2018.

F5. F5 Application Services, 2022. Disponível em: <<https://www.f5.com/products>>. Acesso em: 20 out. 2022.

FARIA, M. I.; PERICÃO, M. D. G. **Dicionário do Livro. Da escrita ao livro electrónico**. Coimbra, Portugal: Edições Almedina, SA, 2008. ISBN 978-972-40-3499-7.

FEDORA. About Us. **FEDORA**, 2023. Disponível em: <<https://fedora.lyrasis.org/about/>>. Acesso em: 01 maio 2023.

FERREIRA, M. **Introdução à Preservação Digital: conceitos, estratégias e actuais consensos**. Guimarães, Portugal: Escola de Engenharia da Universidade do Minho, 2006. ISBN 972-8692-30-7.

FLORES, D.; ROCCO, B.; SANTOS, H. Cadeia de custódia para documentos arquivísticos digitais. **Acervo - Revista do Arquivo Nacional**, Rio de Janeiro, v. 29, n. 2, 2016. 117-132. Disponível em: <<http://hdl.handle.net/20.500.11959/brapci/40511>>. Acesso em: 18 ago. 2022.

FLORES-FERNÁNDEZ, C. et al. Estrategias de preservación digital de archivos sonoros. Revisión sistematizada. **Revista Española de Documentación Científica**., Santiago, Chile, abril-junio 2022. 1-10. Disponível em: <<https://doi.org/10.3989/redc.2022.2.1864>>. Acesso em: 10 out. 2023.

FRANK, R. D. Risk in trustworthy digital repository audit and certification. **Archival Science**, 22, 2022. 43-73.

FUNDACIÓN ESPAÑOLA PARA LA CIENCIA Y LA TECNOLOGÍA. Recolecta. Recolector de Ciencia Abierta. **Recolecta**, 2023. Disponível em: <<https://recolecta.fecyt.es/>>. Acesso em: 10 maio 2023.

GARDEREN, P. V. Archivematica: using micro-services and open-source software to deliver a comprehensive digital curation solution. **Proceedings of the 7th International Conference on Digital Preservation, iPRES**, Viena, Áustria, 19-24 September 2010. Disponível em: <<https://hdl.handle.net/11353/10.185511>>. Acesso em: 19 out. 2022.

GAVA, T. B. S.; FLORES, D. Problematizando a Pós-Custódia com a contemporaneidade da Cadeia de Custódia Digital Arquivística compartilhada e distribuída na Preservação Digital Sistêmica. **InCID: Revista de Ciência da Informação e Documentação**, v. 13, n. 2, 2022. 222-243. Disponível em: <<https://www.revistas.usp.br/incid/article/view/191654>>. Acesso em: 03 fev. 2023.

GAVA, T.; FLORES, D. Auditoria e Certificação ao Longo da Cadeia de Custódia Digital Arquivística. **Informação & Informação**, v. 26, n. 4, 2021. 424-449. Disponível em: <<https://www.uel.br/revistas/uel/index.php/informacao/article/view/44504>>. Acesso em: 18 ago. 2022.

GIARETTA, D.; HARMSSEN, H.; KEITEL, C. Memorandum of Understanding, 2010. Disponível em: <<http://www.trusteddigitalrepository.eu/Memorandum%20of%20Understanding.html>>. Acesso em: 18 ago. 2022.

GIL, A. C. **Métodos e Técnicas de Pesquisa Social**. 6ª edição. ed. São Paulo: Editora Atlas, 2008. ISBN 978-85-224-5142-5.

- GILLEAN, D. Artefactual response to RiC draft, 25 jan 2017. Disponível em: <<https://groups.google.com/g/ica-atom-users/c/QwSor7OQ90U/m/DpsDM1P5BAAJ>>. Acesso em: 06 jul. 2022.
- GINSPARG, P. Scholarly Information Architecture, 1989-2015. **Data Science Journal**, v. 3 n. 4, 4 feb 2004. 29-37. Disponível em: <<https://storage.googleapis.com/jnl-up-j-dsj-files/journals/1/articles/273/submission/proof/273-1-529-1-10-20150416.pdf>>.
- GODOY, A. S. Pesquisa qualitativa: tipos fundamentais. **Revista de Administração de Empresas. Fundação Getúlio Vargas, Escola de Administração de Empresas de S.Paulo**, São Paulo, SP, v. 35, n. 3, 1995. 20-29. Disponível em: <<http://hdl.handle.net/11449/20595>>. Acesso em: 01 maio 2023.
- GOMES, W. D. S.; AUTRAN, M. D. M. M. Análise dos aspectos de confiabilidade do Repositório Digital Arquivístico Archivematica à luz da Resolução nº 43 do Conselho Nacional de Arquivos. **Ciência da Informação em Revista**, Maceió, v. 7, n. 1, jan/abr 2020. 105-120. Disponível em: <<https://doi.org/10.28998/cirev.2020v7n1g>>. Acesso em: 20 out. 2022.
- GRANGER, S. Digital Preservation and Deep Infrastructure. **D-Lib Magazine**, v. 8, n. 2, feb 2002. Disponível em: <<http://www.dlib.org/dlib/february02/granger/02granger.html>>. Acesso em: 20 maio 2023.
- HACKETT, Y. Preserver guidelines. Preserving Digital Records: Guidelines for Organizations. In: DURANTI, L.; PRESTON, R. **International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2: Experiential, Interactive and Dynamic Records**. Padova, Italy: Associazione Nazionale Archivistica Italiana, 2008. p. 699-712. Disponível em: <http://www.interpares.org/ip2/display_file.cfm?doc=ip2_book_complete.pdf>. Acesso em: 17 ago. 2022.
- HANNIGAN, K. et al. **InterPARES 1 Project Book: Glossary Committee Report**. [S.l.]: [s.n.], 2001. Disponível em: <http://www.interpares.org/book/interpares_book_h_part5.pdf>. Acesso em: 11 nov. 2022.
- HEDSTROM, M. Descriptive practices for electronic records: deciding what is essential and imagining what is possible. **Archivaria**, v. 36, 01 jan. 1993. 53-63. Disponível em: <<https://archivaria.ca/index.php/archivaria/article/view/11934/12890>>. Acesso em: 01 fev. 2023.
- HEDSTROM, M.; LEE, C. A. Significant properties of digital objects: definitions, applications, implications. **Proceedings of the DLM-Forum 2002. @ccess and preservation of electronic information: best practices and solutions**, Barcelona, Spain, 6-8 May 2002. 218-227. Disponível em: <https://ils.unc.edu/callee/sigprops_dlm2002.pdf>. Acesso em: 31 maio 2023.
- HEDSTROM, M.; MONTGOMERY, S. **Digital Preservation: needs and Requirements in RLG Member Institutions**. A study commissioned by the Research Libraries Group. Mountain View, California, USA: RLG, 1998. Disponível em: <<https://www.webjunction.org/content/dam/research/activities/digpresneeds/digpres.pdf>>. Acesso em: 10 maio 2023.
- HEERY, R.; ANDERSON, S. **Digital Repositories Review. Final version**. UKOLN and The Arts and Humanities Data Service (AHDS). [S.l.], p. 37. 2005.
- HERRERA, A. H. **Archivística General. Teoría y Práctica. Actualizada y Aumentada**. 5ª. ed. Sevilla, España: Diputación Provincial de Sevilla, 1991. Disponível em: <<https://alexavidal.files.wordpress.com/2015/07/archivisticageneralteoriaypractica-antonia-heredia-herrera.pdf>>. Acesso em: 30 nov. 2022.
- HESLOP, H.; DAVIS, S.; WILSON, A. **An Approach to the Preservation of Digital Records**. Canberra, Australia: National Archives of Australia, 2002. Disponível em: <<https://www.naa.gov.au/sites/default/files/2020-01/An-Approach-to-the-Preservation-of-Digital-Records.pdf>>. Acesso em: 12 dez. 2022.

- HIGASHI, A. K. et al. Ambientes digitais confiáveis para preservação holística de documentos arquivísticos. **Informação & Informação**, Londrina, v. 25. n. 4, out/dez 2020. 499-527.
- HIRTLE, P. B. Authenticity in a Digital Environment. In: CULLEN, C. T., et al. **Authenticity in a Digital Environment**. Washington: Council on Library and Information Resources, 2000. p. 8-23. ISBN 1-887334-77-7. Disponível em: <<https://www.clir.org/wp-content/uploads/sites/13/pub92.pdf#page=16>>. Acesso em: 17 ago. 2022.
- HOCKX-YU, H. **Digital preservation in the context of institutional repositories. Preprint. Digital Preservation Programme Manager, Joint Information**. [S.l.]: [s.n.], 2006. Disponível em: <<http://eprints.rclis.org/8189/>>. Acesso em: 02 maio 2023.
- INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA. Parcerias da Rede Cariniana, 2014. Disponível em: <<https://cariniana.ibict.br/index.php/parceiros-da-rede>>. Acesso em: 10 jan. 2022.
- INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA. Open Journal Systems (OJS), 2021. Disponível em: <<https://www.gov.br/ibict/pt-br/assuntos/tecnologias-para-a-informacao/ojs>>. Acesso em: 25 ago. 2022.
- INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. Índice Geral de Cursos (IGC). **INEP**, 2020. Disponível em: <<https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/indicadores-de-qualidade-da-educacao-superior/indice-geral-de-cursos-igc>>. Acesso em: 10 ago. 2023.
- INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Censo da Educação Superior 2020: notas estatísticas**. Brasília, DF: INEP, 2022. Disponível em: <https://download.inep.gov.br/publicacoes/institucionais/estatisticas_e_indicadores/notas_estatisticas_censo_da_educacao_superior_2020.pdf>. Acesso em: 26 jun. 2023.
- INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Sinopse Estatística da Educação Superior 2020**. Brasília: INEP, 2022. Disponível em: <<https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/sinopses-estatisticas/educacao-superior-graduacao>>. Acesso em: 23 ago. 2022.
- INTERNATIONAL COUNCIL ON ARCHIVES. **Guide for Managing Electronic Records from an Archival Perspective. Committee on Electronic Records**. Paris, France: ICA, 1997. ISBN 0-9682361-0-3. Disponível em: <https://www.ica.org/sites/default/files/ICA%20Study%208%20guide_eng.pdf>. Acesso em: 19 jan. 2023.
- INTERNATIONAL COUNCIL ON ARCHIVES. **Dictionary of Archival Terminology**. Draft Third Edition/DAT III. ed. [S.l.]: [s.n.], 1999.
- INTERNATIONAL COUNCIL ON ARCHIVES. **ISAD(G): General International Standard Archival Description: adopted by the Committee on Descriptive Standards**, Stockholm, Sweden, 19-22 September 1999. 2nd. ed. Ottawa: [s.n.], 2000.
- INTERNATIONAL COUNCIL ON ARCHIVES. **ISAAR(CPF): International Standard Archival Authority Record for Corporate Bodies, Persons and Families**. 2nd. ed. Canberra: [s.n.], 2004.
- INTERNATIONAL COUNCIL ON ARCHIVES. **ISDF: International Standard for Describing Functions**. 1st. ed. Desden: [s.n.], 2007.
- INTERNATIONAL COUNCIL ON ARCHIVES. **ISDIAH: International Standard for Describing Institutions with Archival Holdings**. 1st. ed. London: [s.n.], 2008.
- INTERNATIONAL COUNCIL ON ARCHIVES. **Records in Contexts: a conceptual model for archival description**. Consultation Draft v0.1. ed. [S.l.]: [s.n.], 2016.

INTERNATIONAL COUNCIL ON ARCHIVES. **Records in Contexts**: a conceptual model for archival description. Consultation draft v0.2 (preview). ed. [S.l.]: [s.n.], 2019.

INTERNATIONAL COUNCIL ON ARCHIVES. Records in Contexts Ontology (RiC-O), 2021. Disponível em: <https://www.ica.org/standards/RiC/RiC-O_v0-2.html>. Acesso em: 06 jul. 2022.

INTERNATIONAL COUNCIL ON ARCHIVES. **Records in Contexts**: conceptual model. Consultation Draft v0.2. ed. [S.l.]: [s.n.], 2021. Disponível em: <https://www.ica.org/sites/default/files/ric-cm-02_july2021_0.pdf>. Acesso em: 07 jul. 2022.

INTERNATIONAL COUNCIL ON ARCHIVES. **Records in Contexts**: introduction to archival description. Consultation Draft v0.2. ed. [S.l.]: [s.n.], 2021. Disponível em: <https://www.ica.org/sites/default/files/ric-iad-02_0_0.pdf>. Acesso em: 07 jul. 2022.

INTERNATIONAL COUNCIL ON ARCHIVES. Call for comments on ICA RiC-CM 0.2, 2022. Disponível em: <<https://www.ica.org/en/call-for-comments-on-ica-ric-cm-02>>. Acesso em: 03 agosto 2022.

INTERNATIONAL NETWORK FOR THE CONSERVATION OF CONTEMPORARY ART. INCCA Statement of Values. **INCCA**, 2015. Disponível em: <<https://incca.org/incca-statement-values>>. Acesso em: 17 maio 2023.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 15489-1**: Information and documentation - Records Management. Part 1: Concepts and principles. Switzerland: ISO, 2001.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO/IEC 17799**: Information Technology. Security Techniques. Code of Practice for Information Security Management. 2nd. ed. Switzerland: ISO, 2005.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO/TR 18492**: Long-term preservation of electronic document-based information. Switzerland: ISO, 2005.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 23081-1**: Information and documentation - Records Management Processes. Metadata for Records. Part 1: Principles. Switzerland: ISO, 2006.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO/TS 23081-2**: Information and documentation — Records management processes - Metadata for records - Part 2: Conceptual and implementation issues. Switzerland: ISO, 2007.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO/TR 23081-3**: Information and documentation - Managing metadata for records - Part 3: Self-assessment method. Switzerland: ISO, 2011.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 14721**: Space data and information transfer systems. Open archival information system (OAIS). Reference model. Switzerland: ISO, 2012.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 16363**: Space data and information transfer systems. Audit and certification of trustworthy digital repositories. Switzerland: ISO, 2012.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO/TR 17068:2012. Information and documentation - Trusted third party repository for digital records. **ISO**, 2012. Disponível em: <<https://www.iso.org/standard/58087.html>>. Acesso em: 27 fev. 2023.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 16919**: Space data and information transfer system. Requirements for bodies providing audit and certification of candidate trustworthy digital repositories. Switzerland: ISO, 2014.

- INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 9000:** Quality management systems - Fundamentals and vocabulary. Switzerland: ISO, 2015.
- INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 15489-1:** Information and documentation - Records Management. Part 1: Concepts and principles. Switzerland: ISO, 2016.
- INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 17021-1:** Conformity assessment. Requirements for bodies providing audit and certification of management systems. Switzerland: ISO, 2016.
- INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 15836-1:** Information and documentation. The Dublin Core metadata element set. Part 1: Core elements. Switzerland: ISO, 2017.
- INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 23081-1:** Information and documentation - Records management processes - Metadata for records - Part 1: Principles. Switzerland: ISO, 2017.
- INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 14641:** Electronic document management. Design and operation of an information system for the preservation of electronic documents. Specifications. Switzerland: ISO, 2018.
- INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO/IEC 27000:** Information technology - Security techniques - Information security management systems - Overview and vocabulary. Switzerland: ISO, 2018.
- INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 30300:** Information and documentation - Records management - Core concepts and vocabulary. Geneva, Switzerland: ISO, 2020.
- INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 23081-2:** Information and documentation - Managing metadata for records - Part 2: Conceptual and implementation issues. Switzerland: ISO, 2021.
- INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 13008:** Information and documentation. Digital records conversion and migration process. Switzerland: ISO, 2022.
- INTERPARES. **InterPARES 1 Project Book:** The InterPARES Glossary. [S.l.]: [s.n.], 2001. Disponível em: <http://www.interpares.org/book/interpares_book_q_gloss.pdf>. Acesso em: 11 nov. 2022.
- INTERPARES PRESERVATION TASK FORCE. **How to Preserve Authentic Electronic Records. Appendix 6.** [S.l.]: [s.n.], 2001. Disponível em: <http://interpares.org/display_file.cfm?doc=interpares_book_o_app06.pdf>. Acesso em: 17 ago. 2022.
- INTERPARES TRUST. InterPARES Trust response to EGAD-RiC, 2016. Disponível em: <<http://interparestrust.org/trust/article/interpares-trust-response-to-egadric>>. Acesso em: 06 jul. 2022.
- JENKINSON, H. **A Manual of Archive Administration. New and revised edition. First edition published 1922.** 2. ed. London, UK: Percy Lund, Humphries & Co. Ltd., 1937.
- JOINT INFORMATION SYSTEMS COMMITTEE. About OpenDOAR. **OpenDOAR**, 2023. Disponível em: <<https://v2.sherpa.ac.uk/opendoar/>>. Acesso em: 15 maio 2023.
- JOINT INFORMATION SYSTEMS COMMITTEE. OpenDOAR: Browse by Country and Region. Brazil. **OpenDOAR**, 2023. Disponível em: <https://v2.sherpa.ac.uk/view/repository_by_country/Brazil.html>. Acesso em: 26 jun. 2023.
- JONES, M.; BEAGRIE, N. **Preservation Management of Digital Materials:** a handbook. The Handbook was first compiled by Neil Beagrie and Maggie Jones and is now maintained and updated by the DPC. London: British Library, 2001. Disponível em: <<https://www.dpconline.org/docs/digital-preservation-handbook/299-digital-preservation-handbook/file>>. Acesso em: 01 abr. 2023.

- KEEP SOLUTIONS. **RODA 4. Long-Term Digital Preservation. Characteristics and Technical Requirements.** [S.l.]: Keep Solutions, 2021. Disponível em: <https://www.keep.pt/wp-content/uploads/2021/11/WP181216.2-whitepaper-roda-4_EN.pdf>. Acesso em: 14 mar. 2023.
- KENNEY, A.; MCGOVERN, N. The Five Organizational Stages of Digital Preservation. In: HODGES, P., et al. **Digital Libraries: A Vision for the 21st Century.** A Festschrift in Honor of Wendy Lougee on the Occasion of Her Departure from the University of Michigan. Michigan, USA: Michigan Publishing, University of Michigan Library, 2003. Disponível em: <<https://quod.lib.umich.edu/s/spobooks/bbv9812.0001.001/1:11?rgn=div1;view=fulltext>>. Acesso em: 15 fev. 2023.
- LAVOIE, B.; DEMPSEY, L. Thirteen Ways of Looking at Digital Preservation. **DLib Magazine**, v. 10 n. 7/8, July/August 2004. Disponível em: <https://www.researchgate.net/profile/Lorcan-Dempsey-2/publication/240949457_Thirteen_Ways_of_Looking_at_Digital_Preservation/links/5611d13308ae6b29b49e338b/Thirteen-Ways-of-Looking-at-Digital-Preservation.pdf>. Acesso em: 20 jun. 2023.
- LAVOIE, B.; GARTNER, R. **Technology Watch Report. Preservation Metadata. DPC Technology Watch Series Report 05-01.** [S.l.]: OCLC Online Computer Library Center Inc., 2005. Disponível em: <https://d1wqtxts1xzle7.cloudfront.net/6723337/o050912l-libre.pdf?1390847763=&response-content-disposition=inline%3B+filename%3DPreservation_Metadata_Technology_Watch_R.pdf&Expires=1682793051&Signature=UFnQ5pPCsP8~d745Yg6qcjZSqZPeIbZ-D2LP0raWpQVvW5mmjgtKWu>. Acesso em: 31 mar. 2023.
- LEVELS OF PRESERVATION REVISIONS WORKING GROUP. **NDSA Levels of Digital Preservation Matrix V2.0.** [S.l.]: NDSA, 2019. Disponível em: <<https://osf.io/qd54c>>. Acesso em: 19 ago. 2022.
- LEVELS OF PRESERVATION REVISIONS WORKING GROUP. **Using the Levels of Digital Preservation: an overview for v2.0.** [S.l.]: NDSA, 2019. Disponível em: <<https://osf.io/vnc32>>. Acesso em: 19 ago. 2022.
- L'HOURS, H.; KLEEMOLA, M.; LEEUW, L. CoreTrustSeal: from academic collaboration to sustainable services. **IASSIST Quarterly**, v. 43, n. 1, p. 1-17, 2019.
- LIBRARY OF CONGRESS. MARC 21 Format for Bibliographic Data. Update No. 34 (July 2022), 2022. Disponível em: <<https://www.loc.gov/marc/bibliographic/>>. Acesso em: 20 out. 2022.
- LIN, D. et al. The TRUST Principles for digital repositories. **Scientific Data**, v. 7, n. 144, 2020. 1-5. Disponível em: <<https://doi.org/10.1038/s41597-020-0486-7>>. Acesso em: 18 ago. 2022.
- LINDLAR, M.; SCHWAB, F. 203.5 All that work ... for what? Return on investment for trustworthy archive certification processes – a case study, 29 July 2022. Disponível em: <<https://osf.io/8a3sc>>.
- LLANES-PADRÓN, D. **La descripción archivística en los tiempos posmodernos: conceptos, principios y normas.** Marília: Oficina Universitária, 2016. ISBN 978-85-7983-777-7.
- LUHMANN, N. Familiarity, Confidence, Trust: Problems and Alternatives. In: GAMBETTA, D. **Trust. Making and Breaking Cooperative Relations.** Great Britain: Basil Blackwell, 1988. Cap. 6, p. 94-107. Disponível em: <https://www.nuffield.ox.ac.uk/users/gambetta/Trust_making%20and%20breaking%20cooperative%20relations.pdf>. Acesso em: 30 maio 2023.
- LYNCH, C. Institutional Repositories: Essential Infrastructure For Scholarship In The Digital Age. **portal: Libraries and the Academy**, v. 3. n. 2, 2003. 327-336.
- MAEMURA, E.; MOLES, N.; BECKER, C. Organizational Assessment Frameworks for Digital Preservation: A Literature Review and Mapping. **Journal of the Association for Information Science and Technology**, v. 68, n. 7, 2017. 1619-1637. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1002/asi.23807>>.

- MAHER, W. J. The importance of financial analysis of archival programs. **The Midwestern Archivist**, vol. 3, n. 2, 1978. 3-24. Disponível em: <<https://www.jstor.org/stable/41057898>>. Acesso em: 20 jun. 2023.
- MARCUM, D. B. The preservation of digital information. **The Journal of Academic Librarianship**, 22, n. 6, 1996. 451-454. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0099133396900063>>.
- MÁRQUEZ, J. V. **Confianza e información digital**: bibliotecas, archivos y web. Colección Tecnologías de la Información. Instituto de Investigaciones Bibliotecológicas y de la Información. Mexico: UNAM, 2017.
- MATLALA, M. E.; NCUBE, T. R.; PARBANATH,. The state of digital records preservation in South Africa's public sector in the 21st century: a literature review. **Records Management Journal**, v. 32, n. 2, 2022. 198-212. Disponível em: <<https://www.emerald.com/insight/content/doi/10.1108/RMJ-02-2021-0004/full/html>>. Acesso em: 20 jun. 2023.
- METS BOARD. **METS. Metadata Encoding and Transmission Standard**: primer and reference manual. Version 1.6. Revised 2010. United States: LOC, 2010. Disponível em: <<https://www.loc.gov/standards/mets/METSPrimer.pdf>>. Acesso em: 25 ago. 2022.
- MINISTERIO DE CULTURA. **Diccionario de Terminología Archivística**. 2ª. ed. Madrid: [s.n.], 1995. Disponível em: <<https://www.culturaydeporte.gob.es/cultura/areas/archivos/mc/dta/diccionario.html>>. Acesso em: 29 nov. 2022.
- MINISTERIO DE HACIENDA Y ADMINISTRACIONES PÚBLICAS. **Guía de aplicación de la Norma Técnica de Interoperabilidad de Documento electrónico (2ª edición electrónica)**. Madrid: Ministerio de Hacienda y Administraciones Públicas, 2016. Disponível em: <https://administracionelectronica.gob.es/pae_Home/dam/jcr:5881e773-6d5d-48b6-b4a6-7760e63fcfef/Guia_NTI_documento_electronico_PDF_2ed_2016.pdf>. Acesso em: 29 mar. 2023.
- MINISTERIO DE POLÍTICA TERRITORIAL Y ADMINISTRACIÓN PÚBLICA. **Resolución de 19 de julio de 2011, de la Secretaría de Estado para la Función Pública, por la que se aprueba la Norma Técnica de Interoperabilidad de Documento Electrónico**. BOE num. 182, de 30 de julio de 2011, páginas 87094 a 87107. Madrid: BOE, 2011. Disponível em: <https://www.boe.es/diario_boe/txt.php?id=BOE-A-2011-13169>. Acesso em: 31 mar. 2023.
- MORO-CABERO, M.; LLANES-PADRÓN, D. Records in contexts: Un nuevo modelo para la representación de la información archivística en el entorno de la web semántica. **El profesional de la información**, v. 26, n. 3, p. 525-533, junio 2017.
- MORO-CABERO, M.; MARTÍN-POZUELO, M.; BONAL-ZAZO, J. L. ISO 15489 and other standardized management systems: analogies and synergies. **Records Management Journal**, 21, n. 2, 2011. 104-121. Disponível em: <<https://doi.org/10.1108/09565691111152044>>. Acesso em: 17 ago. 2022.
- MULLER, S. F.; FEITH, J. A.; FRUIN, R. **Manual for the Arrangement and Description of Archives. Translation of the second edition with new introductions**. Tradução de Arthur H. Leavitt. Chicago: The Society of American Archivists, 2003. Disponível em: <<https://babel.hathitrust.org/cgi/pt?id=mdp.39015057022447&view=1up&seq=5>>. Acesso em: 22 março 2022.
- NATIONAL ARCHIVES. **Statement of Public Task**. London, England: [s.n.], 2015.
- NATIONAL ARCHIVES. National Archives History and Mission. **National Archives**, 2021. Disponível em: <<https://www.archives.gov/about/history/about/history/history-and-mission>>. Acesso em: 31 mar. 2023.
- NATIONAL DIGITAL STEWARDSHIP ALLIANCE. Levels of Digital Preservation, 2022. Disponível em: <<https://nds.a.org/publications/levels-of-digital-preservation/>>. Acesso em: 19 ago. 2022.

NESTOR CERTIFICATION WORKING GROUP. **Explanatory Notes on the Nestor Seal for Trustworthy Digital Archives. Nestor Materials 17.** [S.l.]: Deutsche Nationalbibliothek, 2013. Disponível em: <<https://nbn-resolving.org/process-urn-form?identifier=urn%3Anbn%3Ade%3A0008-2013100901&verb=FULL>>. Acesso em: 24 fev. 2023.

NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION. **Catalogue of Criteria for Trusted Digital Repositories. Version 1. Draft for public comment. Nestor Materials 8.** Frankfurt, Germany: Deutsche Nationalbibliothek, 2006. Disponível em: <<http://d-nb.info/1047607387/34>>. Acesso em: 16 fev. 2023.

NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION. **Nestor criteria: Catalogue of Criteria for Trusted Digital Repositories. Version 2.** Nestor Materials 8. Frankfurt, Germany: Deutsche Nationalbibliothek, 2009. Disponível em: <<http://nbn-resolving.de/urn:nbn:de:0008-2010030806>>. Acesso em: 22 ago. 2022.

NETWORK OF EXPERTISE IN LONG-TERM STORAGE OF DIGITAL RESOURCES. Nestor siegel, 2022. Disponível em: <https://www.langzeitarchivierung.de/Webs/nestor/EN/Services/nestor_Siegel/nestor_siegel_node.html>. Acesso em: 22 ago. 2022.

OFFICE OF SCIENCE AND TECHNOLOGY POLICY. **Memorandum for the Heads of Executive Departments and Agencies. Subject: Increasing Access to the Results of Federally Funded Scientific Research.** Washington, DC, USA: Executive Office of the President, 2013. Disponível em: <https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf>. Acesso em: 01 maio 2023.

OFFICE OF SCIENCE AND TECHNOLOGY POLICY. **Memorandum for the Heads of Executive Departments and Agencies. Subject: Ensuring Free, Immediate, and Equitable Access to Federally Funded Research.** Washington, DC, USA: Executive Office of the President, 2022. Disponível em: <<https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-Access-Memo.pdf>>. Acesso em: 01 maio 2023.

ONLINE COMPUTER LIBRARY CENTER. **A Metadata Framework to Support the Preservation of Digital Objects. Preservation Metadata and the OAIS Information Model. A report by the OCLC/RLG Working Group on Preservation Metadata.** Dublin, Ohio, USA: OCLC, 2002. Disponível em: <https://www.oclc.org/content/dam/research/activities/pmwg/pm_framework.pdf>. Acesso em: 10 mar. 2023.

ONLINE COMPUTER LIBRARY CENTER. **Trustworthy Repositories Audit & Certification: Criteria and Checklist (TRAC).** Version 1.0. Online Computer Library Center (OCLC). The Center for Research Libraries (CRL). Chicago, Illinois, USA: CRL, 2007. Disponível em: <https://www.crl.edu/sites/default/files/d6/attachments/pages/trac_0.pdf>. Acesso em: 16 fev. 2023.

OPEN ACCESS MOVEMENT. **Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities.** Berlin: OAM, 2003. Disponível em: <https://openaccess.mpg.de/67605/berlin_declaration_engl.pdf>. Acesso em: 01 maio 2023.

OPEN ARCHIVES INITIATIVE. The Open Archives Initiative Protocol for Metadata Harvesting. Protocol Version 2.0 of 2002-06-14, 2023. Disponível em: <<https://www.openarchives.org/OAI/openarchivesprotocol.html>>. Acesso em: 10 maio 2023.

ORGANIZATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. **OECD Principles and Guidelines for Access to Research Data from Public Funding.** [S.l.]: OECD, 2007. Disponível em: <<https://www.oecd.org/science/inno/38500813.pdf>>. Acesso em: 30 abr. 2023.

PALO ALTO NETWORKS. Customer Success Tools, 2022. Disponível em: <<https://www.paloaltonetworks.com/services/customer-success-tools>>. Acesso em: 20 out. 2022.

- PAYETTE, S.; LAGOZE, C. Flexible and Extensible Digital Object and Repository Architecture (FEDORA). **Research and Advanced Technology for Digital Libraries. ECDL 1998. Lecture Notes in Computer Science**, Heraklion, Crete, Greece, v. 1513, 21-23 sep 1998. 41-59.
- PEARCE-MOSES, R. **A Glossary of Archival and Records Terminology**. Chicago, USA: The Society of American Archivists, 2005. ISBN 1-931666-14-8.
- PENNOCK, M. **Web-Archiving. DPC Technology Watch Report 13-01 March 2013**. Great Britain: Digital Preservation Coalition, 2013. Disponível em: <<http://dx.doi.org/10.7207/twr13-01>>. Acesso em: 18 maio 2023.
- PHILLIPS, M. et al. **The NDSA Levels of Digital Preservation: An Explanation and Uses**. NDSA. [S.l.]. 2013.
- PREMIS EDITORIAL COMMITTEE. **Conformant Implementation of the PREMIS Data Dictionary**. [S.l.]: PREMIS, 2015. Disponível em: <<https://www.loc.gov/standards/premis/premis-conformance-20150429.pdf>>. Acesso em: 14 mar. 2023.
- PREMIS EDITORIAL COMMITTEE. **PREMIS Data Dictionary for Preservation Metadata. Version 3.0**. [S.l.]: Library of Congress, 2015. Disponível em: <<https://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf>>. Acesso em: 28 set. 2022.
- PREMIS EDITORIAL COMMITTEE. **Guidelines for Using PREMIS with METS for Exchange. Revised January 2017**. [S.l.]: PREMIS, 2017. Disponível em: <<https://www.loc.gov/standards/premis/guidelines2017-premismets.pdf>>. Acesso em: 13 mar. 2023.
- PRIMA. Biblioteca Universitária - SophiA, 2022. Disponível em: <<https://sophia.com.br/solucoes/bibliotecas/biblioteca-universitaria/>>. Acesso em: 20 out. 2022.
- PRIMARY TRUSTWORTHY DIGITAL REPOSITORY AUTHORISATION BODY. ISO 16363 certification of Trustworthy Digital Repositories, 2022. Disponível em: <<http://www.iso16363.org/>>. Acesso em: 22 ago. 2022.
- QUALYS. IT Security and Compliance Platform, 2022. Disponível em: <<https://www.qualys.com/>>. Acesso em: 20 out. 2022.
- R FOUNDATION. What is R?, 2022. Disponível em: <<https://www.r-project.org/about.html>>. Acesso em: 26 out. 2022.
- RAMALHO, J. C. et al. RODA and Crib. A Service-Oriented Digital Repository. **International Conference on Preservation of Digital Objects. iPress.**, London, England, 2008. Disponível em: <<https://hdl.handle.net/1822/8226>>. Acesso em: 04 maio 2023.
- RAMALHO, J. C. et al. RODA: Repositório de Objetos Digitais Autênticos: passado, presente e futuro. **Encontro Internacional de Arquivos (EIA)**, Évora, 3-4 Outubro 2014. Disponível em: <<https://hdl.handle.net/1822/31335>>. Acesso em: 02 set. 2022.
- RAMPAZZO, L. **Metodologia Científica. Para alunos dos cursos de graduação e pós-graduação**. 3ª. ed. São Paulo: Edições Loyola, 2005.
- RED DE BIBLIOTECAS UNIVERSITARIAS ESPAÑOLAS. **Selección de criterios utilizados para la evaluación de repositorios**. [S.l.]: CRUE, 2015. Disponível em: <https://www.rebiun.org/sites/default/files/Documentos/GT-REP/EvaluacionREP/Seleccion_Criterios_Evaluacion_Repositorios.pdf>. Acesso em: 27 mar. 2023.
- RED DE BIBLIOTECAS UNIVERSITARIAS ESPAÑOLAS. **Formulario de Evaluación de Repositorios. GREDOS. Universidad de Salamanca**. [S.l.]: [s.n.], 2018. Disponível em: <https://www.rebiun.org/sites/default/files/Documentos/GT-REP/EvaluacionREP/Resultado_Evaluacion_Repositorio_GREDOS_USAL.pdf>. Acesso em: 28 mar. 2023.

- RED DE BIBLIOTECAS UNIVERSITARIAS ESPAÑOLAS. Evaluación de repositorios. **Red de Bibliotecas Universitarias Españolas**, 2022. Disponível em: <<https://www.rebiun.org/grupos-trabajo/repositorios/evaluacion>>. Acesso em: 28 mar. 2023.
- RESEARCH LIBRARIES GROUP. **Trusted Digital Repositories: Attributes and Responsibilities**. An RLG-OCLC Report. Mountain View, CA, USA: RLG, 2002. Disponível em: <<https://www.oclc.org/content/dam/research/activities/trustedrep/repositories.pdf>>. Acesso em: 19 ago. 2022.
- ROCHA, R. P. D. et al. Análise dos sistemas DSpace e Dataverse para repositórios de dados de pesquisa com acesso aberto. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v. 17, jun 2021, p. 1-25. Disponível em: <<https://rbbd.febab.org.br/rbbd/article/view/1572/1261>>. Acesso em: 11 out. 2022.
- ROGERS, C. Diplomats of born digital documents – considering documentary form in a digital environment. **Records Management Journal**, 25, 16 março 2015. 6-20.
- ROSS, S.; MCHUGH, A. The Role of Evidence in Establishing Trust in Repositories. **D-Lib Magazine**, v. 12, n. 7/8, jul/ago 2006. Disponível em: <<http://www.dlib.org/dlib/july06/ross/07ross.html>>. Acesso em: 30 maio 2023.
- SAMPAIO, T. B. **Metodologia da Pesquisa**. 1ª. ed. Santa Maria, RS: UFSM, CTE, UAB, 2022. Disponível em: <https://repositorio.ufsm.br/bitstream/handle/1/26138/MD_Metodologia_da_Pesquisa.pdf?sequence=1&isAllowed=y>. Acesso em: 03 jul. 2023.
- SANTOS, H. M. **Auditoria de repositórios arquivísticos digitais confiáveis: uma análise das normas SO 14721 e ISO 16363**. Dissertação de mestrado. Santa Maria: Universidade Federal de Santa Maria, 2018.
- SANTOS, H. M. **Manual para Auditoria de Repositórios Arquivísticos Digitais Confiáveis**. [S.l.]: [s.n.], 2018. Disponível em: <<http://furg.academia.edu/HenriqueMachadosSantos>>.
- SANTOS, H. M. D.; FLORES, D. Responsabilidade de um Repositório Arquivístico Digital Confiável na perspectiva do Open Archival Information System. **Páginas a&b: Arquivos e Bibliotecas**, Série 3, n. 11, 2019. 116-132. Disponível em: <<https://ojs.letras.up.pt/index.php/paginasaeab/article/view/5459>>. Acesso em: 10 maio 2023.
- SANTOS, H. M. D.; FLORES, D. Modelo Lógico da Informação no Open Archival Information System: uma reflexão arquivística sobre o Pacote de Informação de Arquivamento. **Perspectivas em Gestão & Conhecimento**, João Pessoa, v. 10, n. 1, jan/abr 2020. 23-38. Disponível em: <<http://dx.doi.org/10.21714/2236-417X2020v10n1p23>>. Acesso em: 20 maio 2023.
- SCELLENBERG, T. R. **Modern Archives. Principles and Techniques. With a new introduction by H. G. Jones. Originally published in 1956**. Reissued in 2003. ed. Chicago, USA: The Society of American Archivists, 2003. ISBN 0-931666-02-4.
- SCIMAGO. SCImago methodology. **SCImago Institutions Ranking (SIR)**, 2023. Disponível em: <<https://www.scimagoir.com/methodology.php>>. Acesso em: 10 ago. 2023.
- SCIMAGO. SCImago University Rankings 2023. **SCImago Institutions Rankings**, 2023. Disponível em: <<https://www.scimagoir.com/rankings.php?sector=Higher%20educ.>>. Acesso em: 10 ago. 2023.
- SOUZA, M. V. B. D. Análise comparativa de modelos conceituais de descrição arquivística: RiC e NEDA. **Scire: representación y organización del conocimiento**, v. 26, n. 2, p. 39-46, jul-dic 2020. ISSN 1135-3716. Disponível em: <<https://doi.org/10.54886/scire.v26i2.4697>>. Acesso em: 28 jun 2022.
- SOUZA, M. V. B.; FLORES, D. Possíveis impactos do modelo Records in Contexts para os usuários de arquivos. **Acervo**, Rio de Janeiro, v. 33, n. 3, p. 49-67, set/dez 2020.

- TENNIS, J. T.; ROGERS, C. **Authenticity Metadata and the IPAM**: Progress toward the InterPARES. Proceedings of International Conference on Dublin Core and Metadata Application 2012. Kuching, Sarawak, Malaysia: Dublin Core Metadata Initiative. 2012. p. 38-45.
- TÉRMENS, M. **Preservación Digital. Colección El profesional de la información**. Barcelona, España: Editorial UOC, 2014. ISBN 978-84-9064-082-1.
- TESCH, R. **Qualitative Research**: Analysis Types and Software. 1st. ed. London, UK: Routledge, 1990.
- THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS. **IEEE Standard Glossary of Software Engineering Terminology. IEEE Std 610.12-1990**. New York, USA: IEEE, 1990.
- THIBODEAU, K. Overview of technological approaches to digital preservation and challenges in coming years. **The state of digital preservation: an international perspective**, Washington, DC, USA, 24-25 April 2002. 4-31. Disponível em: <<http://www.clir.org/wp-content/uploads/sites/6/pub107.pdf#page=10>>. Acesso em: 17 nov. 2022.
- UNESCO. **Declaração REA de Paris em 2012**. Congresso Mundial sobre Recursos Educacionais Abertos (REA) de 2012. Paris, França: UNESCO. 2012.
- UZWYSHYN, R. Online Research Data Repositories: The What, When, Why and How. **Computers in Libraries**, v. 36, n. 3, 2016. 18-21. Disponível em: <<https://digital.library.txstate.edu/handle/10877/7597>>. Acesso em: 01 maio 2023.
- VOUTSSÁS, J.; AMOZORRUTIA, A. B. **Glosario de preservación archivística digital versión 4.0**. Primera edición. ed. México: UNAM, Instituto de Investigaciones, 2014. Disponível em: <<https://sistemas.iibi.unam.mx/publica20/conmutarl.php?arch=1&idx=288>>. Acesso em: 21 set. 2023.
- VULIS, M. **Modern TEX and its Applications**. Boca Raton, Florida, USA: CRC Press, 1992.
- W3C. EPUB Format, 2022. Disponível em: <<https://www.w3.org/AudioVideo/ebook/>>. Acesso em: 04 nov. 2022.
- W3C. HTML Standard, 2022. Disponível em: <<https://html.spec.whatwg.org/multipage/>>. Acesso em: 04 nov. 2022.
- WATERS, D.; GARRET, J. **Preserving Digital Information. Report of the Task Force on Archiving of Digital Information**. Commission on Preservation and Access. Washington, DC, p. 68. 1996. (1-887334-50-5).
- WHEATLEY, P. Institutional Repositories in the Context of Digital Preservation. **Microform & Digitization Review**, v. 33 n.3, 7 dec 2007. 135-146.
- WILKINSON, M. D. et al. The FAIR Guiding Principles for scientific data. **Scientific Data**, 3, n. 1, 2016. 1-9.
- XIA, J. A Comparison of Subject and Institutional Repositories in Self-archiving Practices. **The Journal of Academic Librarianship**, v. 34, n. 6, nov 2008. 489-495.
- YEO, G. Fijeza y fluidez en el mundo digital. Traducción: Verónica Fernández de Cabo. **Tábula**, v. 25, 2022. 23-38.
- YOON, A. End users' trust in data repositories: definition and influences on trust development. **Archival Science**, v. 14, 2014. 17-37. Disponível em: <<https://doi.org/10.1007/s10502-013-9207-8>>. Acesso em: 30 maio 2023.
- ZALAMA, J. R. **Documentos en contexto**: las nuevas normas de descripción archivística y su aplicación en la administración digital. 23 Jornadas de Archivos Universitarios. A Coruña: [s.n.]. 2018. p. 27-41.

ANEXO A
QUESTIONÁRIO DE AVALIAÇÃO DOS REPOSITÓRIOS

No presente Anexo, é apresentado o questionário de avaliação aplicado aos repositórios na sua totalidade, com as perguntas, respectivas respostas e justificativas para as suas elaborações.

Quadro 77 – Questionário completo de avaliação dos repositórios com perguntas e justificativas.

Área 0: Informações sobre o respondente	
Pergunta	
0.1. Qual o seu cargo ou função desempenhada no órgão que gerencia o repositório?	
Respostas	
☐ Diretor(a) ☐ Vice-diretor(a) ☐ Responsável pelo setor ☐ Prefiro não responder ☐ Outros: _____	
Pergunta	
0.2. Qual a sua formação acadêmica?	
Respostas	
☐ Arquivista ☐ Bibliotecário ☐ Administrador ☐ Técnico de Arquivo ☐ Historiador ☐ Museólogo ☐ Prefiro não responder ☐ Outros: _____	
Área 1: Contextualização do repositório	
Pergunta	
1.1 O repositório institucional possui missão, visão e valores elaborados?	
Respostas	
☐ Sim ☐ Não ☐ Prefiro não responder ☐ Outros: _____	
Justificativas	
O estabelecimento da missão, visão e valores do repositório devem descrever o compromisso da organização para administrar os objetos digitais sob sua custódia, conforme o critério ISO 16363:2012 3.1.1 (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 23). As declarações tornam claro aos usuários do repositório que a preservação e acesso a longo prazo são responsabilidades explícitas do repositório, conforme consta no critério CoreTrust Seal R1 (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019, p. 10). A ferramenta DRAMBORA identifica como um risco conforme, a ausência de comprometimento com a preservação digital na missão do repositório. Nos modelos disponibilizados, esses riscos são identificados como T1, T2 e R04 (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007, p. 109, 138). A análise da missão e objetivos do repositório digital é analisada no critério Nestor Seal 1 (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 10).	
Pergunta	
1.1.1 Em caso positivo, qual a missão, visão e valores do repositório institucional?	

Respostas
Pergunta aberta, podendo-se preencher com os valores solicitados ou indicando uma URL onde é possível acessar as informações.
Pergunta
1.2 O repositório institucional possui a sua Comunidade-Alvo delimitada?
Respostas
☐ Sim ☐ Não ☐ Prefiro não responder ☐ Outros
Justificativas
A Comunidade-Alvo é o “grupo identificado de potenciais Consumidores, que convém que sejam capazes de entender, em particular, um conjunto de informações” (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2021, p. 19). O repositório deve voltar as suas ações à sua Comunidade-Alvo procurando atendê-la e monitorando as suas necessidades. A definição da comunidade designada do repositório delimita o seu alcance assim como delimita as necessidades que deve atender, correspondendo ao critério ISO 16363 3.3.1 (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 27), critério CoreTrust Seal R0 (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019, p. 6) e critério Nestor Seal 1.3 (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 11).
Pergunta
1.2.1 Em caso positivo qual a Comunidade-Alvo do repositório institucional?
Respostas
Pergunta aberta, podendo-se preencher com os valores solicitados ou indicando uma URL onde é possível acessar as informações.
Pergunta
1.3 Particularidades da Comunidade-Alvo. A Comunidade-Alvo do repositório possui especificidades em suas demandas de uso dos documentos? Caso julgue que não haja particularidades, deixe a resposta em branco.
Respostas
Pergunta aberta
Justificativas
O fato de o repositório identificar particularidades em sua Comunidade-Alvo, mesmo que não ela não tenha sido formalmente delimitada, conforme solicitado na pergunta 2.1, sinaliza que há uma certa delimitação e identificação de necessidades da comunidade designada do repositório.
Pergunta
1.4 Qual a natureza do material depositado no repositório?
Respostas
Pergunta semiaberta que possibilita a seleção de várias alternativas ao mesmo tempo, permitindo também o fornecimento de uma opção que não tenha sido listada: ☐ Dados de pesquisa (informações geradas ou obtidas a partir de investigação científica, como: gravações de áudio e vídeo, planilhas, questionários, etc.) ☐ Documentos arquivísticos ☐ Documentos históricos ☐ Materiais objetos de pesquisa (artigos, monografias, dissertações e teses) ☐ Outros:
Justificativas

A definição dos tipos de materiais recebidos pelo repositório e disponibilizados à sua Comunidade-Alvo permite o estabelecimento de requisitos específicos, sendo analisada pelos critérios CoreTrust Seal R0 (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019, p. 6) e Nestor Seal 1.1 (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 10).

Área 1: Estratégia e responsabilidade

Pergunta

2.1 Existe uma Política de Preservação com os requisitos mínimos dos documentos recebidos, tratamento de preservação e acesso?

Respostas

- ☐ Existe em teoria
- ☐ Existe na prática, mas ainda não foi totalmente implantado
- ☐ Existe na prática e está totalmente implantado
- ☐ Não existe
- ☐ Prefiro não responder

Justificativas

A Política de Preservação é o documento que descreve o “[...] enfoque adotado pelo repositório para a preservação dos objetos depositados” (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 18, tradução nossa). A política estabelecida deve estar em conformidade com o Plano Estratégico de Preservação.

Pergunta

2.1.1 Caso não exista uma Política de Preservação, existe um Plano Estratégico de Preservação?

Respostas

- ☐ Existe em teoria
- ☐ Existe na prática, mas ainda não foi totalmente implantado
- ☐ Existe na prática e está totalmente implantado
- ☐ Não existe
- ☐ Prefiro não responder

Justificativas

O Plano Estratégico de Preservação é o documento que “[...] fixa as metas e objetivos para alcançar a parte da missão do repositório concernente à preservação”. (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 18, tradução nossa).

Pergunta

2.2 Existe um Plano de Continuidade que garante que os documentos estarão preservados e acessíveis a longo prazo, caso ocorra algum desastre ou descontinuidade do funcionamento?

Respostas

- ☐ Existe em teoria
- ☐ Existe na prática, mas ainda não foi totalmente implantado
- ☐ Existe na prática e está totalmente implantado
- ☐ Não existe
- ☐ Prefiro não responder

Justificativas

O plano de continuidade, também conhecido como Plano de Herança é “o plano de como e quando o gerenciamento, a propriedade e/ou o controle dos conteúdos do arquivo OAIS serão transferidos a um subsequente arquivo OAIS, com o propósito de garantir a efetiva preservação daqueles conteúdos” (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2021, p. 22).

O plano de continuidade é o documento que norteará os processos a serem executados em caso de desastres e demais motivos que levem à descontinuidade do repositório, como a falta de financiamento, ou mudança de prioridades de órgãos superiores, por exemplo. Documentos de guarda permanente são preservados para sempre e, por isso, o plano de continuidade deve antever as potenciais ameaças que

coloquem em risco a preservação.
Área 3: Certificações e normas
Pergunta
3.1 O repositório possui certificações alcançadas?
Respostas
Pergunta aberta para que o respondente informe quais certificações o repositório já alcançou
Justificativas
Busca-se identificar quais certificações, de qualquer natureza, o repositório possui juntamente com o seu ano de obtenção. A pergunta se refere ao critério ISO 16363:2012 3.3.6 (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 30) e ao critério Nestor Seal 5 (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 17).
Pergunta
3.2 Quais certificações o repositório tem a intenção de alcançar?
Respostas
Pergunta semiaberta que possibilita ao respondente selecionar mais de uma resposta ao mesmo tempo, possibilitando fornecer uma opção que não tenha sido contemplada: ☐ ISO 16363 ☐ CoreTrust Seal ☐ Nestor Seal ☐ Existe intenção de obter certificações, mas ainda não foram decididas quais ☐ O repositório já possui todas as certificações que projetou alcançar ☐ Prefiro não responder ☐ Outras:
Justificativas
A pergunta visa identificar o conhecimento quanto a algumas das certificações existentes, informando se o repositório tem interesse em se submeter à aplicação dos respectivos critérios. Em alguns casos, pode ser que o repositório tenha interesse em obter uma certificação, porém ainda não houve o processo para decidir qual exatamente.
Pergunta
3.3 O repositório adota padrões e normas estabelecidos?
Respostas
☐ Dublin Core ☐ ISAD(G) ☐ ISO 14721:2012 (OAIS) ☐ METS ☐ NOBRADE ☐ PREMIS ☐ Outros: _____
Justificativas
Buscar-se analisar se o repositório usa normas consolidadas (internacionais ou padrões de fato estabelecidos) na sua estrutura como um todo. O repositório digital ao utilizar a ISO 14721:2012, por exemplo, automaticamente se beneficia de toda a estrutura voltada à preservação a longo prazo, objeto principal da norma. A adoção de normas tanto para a descrição arquivística como para o empacotamento dos documentos e seus metadados possui entre as suas vantagens, o favorecimento da realização de integrações com outras ferramentas.
Área 4: Infraestrutura organizacional

Pergunta
4.1 Existe uma estrutura organizacional adequada ao funcionamento do repositório para a preservação a longo prazo?
Respostas
Pergunta do tipo semiaberta, com a possibilidade de escolha de uma única resposta: ☐ Existe em teoria ☐ Existe parcialmente (existe o planejamento de estrutura, mas ainda não foi totalmente implantado) ☐ Existe na prática e está totalmente implantado ☐ Não existe ☐ Prefiro não responder ☐ Outros: _____
Justificativas
É de extrema importância que o repositório possua estrutura adequada prevendo subdivisões com atribuições e responsabilidades bem especificadas. O mesmo se aplica ao repositório em si, caso faça parte de uma organização maior. A formalização dessa estrutura favorece o estabelecimento de relações bem delimitadas entre as unidades, facilitando a comunicação interna e o restante da estrutura. (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 25). A pergunta se relaciona ao critério ISO 16363:2012 3.2.1 (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 25), ao critério CoreTrust Seal R5 (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019, p. 14) e ao critério Nestor Seal 4.3 (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 16).
Pergunta
4.2 Existe orçamento adequado ao funcionamento do repositório para a preservação a longo prazo?
Respostas
Pergunta do tipo semiaberta, com a possibilidade de escolha de uma única resposta: ☐ Existe em teoria ☐ Existe parcialmente (o orçamento está previsto, mas ainda não foi totalmente executado) ☐ Existe na prática e está totalmente implantado (o orçamento previsto foi totalmente executado) ☐ Não existe ☐ Prefiro não responder ☐ Outros: _____
Justificativas
O orçamento deve contemplar a aquisição de equipamentos (servidores e rede, por exemplo), aquisição de <i>softwares</i> (quando for o caso), eventuais adaptações da estrutura física que impactem a preservação e capacitação da equipe. A pergunta se relaciona a todos os critérios ISO 16363 3.4, que tratam sobre a sustentabilidade financeira do repositório (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 31), ao critério CoreTrust Seal R5 (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019, p. 14) e ao critério Nestor Seal 4.1 (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 15).
Pergunta
4.3 Em relação à quantidade de pessoas, a equipe responsável pelo repositório é adequada?
Respostas
Pergunta do tipo semiaberta, com a possibilidade de escolha de uma única resposta: ☐ É adequada em teoria ☐ É adequada na prática ☐ Não é adequada ☐ Prefiro não responder ☐ Outros: _____

Justificativas
A formação de uma equipe bem dimensionada em número e com pessoas com os conhecimentos adequados permite ao repositório a divisão de trabalho sem sobrecargas, além de torná-lo mais seguro, uma vez que as pessoas compreendem as responsabilidades assumidas (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 26).
Pergunta
4.4 Em relação às capacitações e conhecimentos, a equipe responsável pelo repositório é adequada?
Respostas
Pergunta do tipo semiaberta, com a possibilidade de escolha de uma única resposta: ☐ É adequada em teoria ☐ É adequada na prática ☐ Não é adequada ☐ Prefiro não responder ☐ Outros: _____
Justificativas
A formação de uma equipe bem dimensionada em número e com pessoas com os conhecimentos adequados permite ao repositório a divisão de trabalho sem sobrecargas, além de torná-lo mais seguro, uma vez que as pessoas compreendem as responsabilidades assumidas (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 26).
Pergunta
4.5 Existe um plano de capacitação permanente da equipe do repositório?
Respostas
Pergunta do tipo semiaberta, com a possibilidade de escolha de uma única resposta: ☐ Existe em teoria ☐ Existe parcialmente (existe o plano, mas ainda não foi totalmente implantado) ☐ Existe na prática e está totalmente implantado ☐ Não existe ☐ Prefiro não responder ☐ Outros: _____
Justificativas
O repositório deve se comprometer a acompanhar a evolução de normas, legislações e tecnologias a fim de se adaptar às mudanças e garantir a preservação a longo prazo (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 26).
Área 5: Infraestrutura técnica e tecnológica
Pergunta
5.1 O repositório possui estrutura física de prédios e/ou salas adequada à preservação e garantia de preservação a longo prazo?
Respostas
Pergunta do tipo semiaberta, com a possibilidade de escolha de uma única resposta: ☐ Possui em teoria ☐ Possui na prática ☐ Não possui ☐ Prefiro não responder ☐ Outro
Justificativas
A disponibilização adequada de prédios e salas permite a instalação dos servidores, equipamentos de rede e equipe necessários à realização da preservação.
Pergunta
5.2 O repositório possui estrutura de equipamentos (servidores de aplicação, armazenamento, rede, etc) adequada à preservação e garantia de preservação a longo prazo?

Respostas
Pergunta do tipo semiaberta, com a possibilidade de escolha de uma única resposta: ☐ Possui em teoria ☐ Possui na prática ☐ Não possui ☐ Prefiro não responder ☐ Outro
Justificativas
O orçamento deve contemplar a aquisição de equipamentos (servidores e rede, por exemplo), aquisição de <i>softwares</i> (quando for o caso), eventuais adaptações da estrutura física que impactem a preservação e capacitação da equipe. A pergunta se relaciona a todos os critérios ISO 16363 3.4, que tratam sobre a sustentabilidade financeira do repositório (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2012, p. 31), ao critério CoreTrust Seal R5 (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019, p. 14) e ao critério Nestor Seal 4.1 (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009, p. 15).
Pergunta
5.3 Existe um plano de atualização periódico de hardware e software?
Respostas
Pergunta do tipo semiaberta, com a possibilidade de escolha de uma única resposta: ☐ Existe em teoria ☐ Existe parcialmente (existe o plano, mas ainda não foi totalmente implantado) ☐ Existe na prática e está totalmente implantado ☐ Não existe ☐ Prefiro não responder ☐ Outros
Justificativas
Os equipamentos e softwares usados no repositório se tornam obsoletos com o passar do tempo, necessitando assim, que haja o monitoramento periódico com o planejamento de substituições e atualizações necessárias.
Pergunta
5.4 Qual plataforma (<i>software</i>) utilizada pelo repositório?
Respostas
Pergunta do tipo semiaberta, com a possibilidade de escolha de uma única resposta: ☐ DSpace ☐ Archivematica ☐ RODA ☐ Nenhuma ☐ Prefiro não responder ☐ Outros
Justificativas
Busca-se identificar primeiramente se o repositório utiliza alguma plataforma de preservação digital e, em caso positivo, identificar as características dessas plataformas e conformidades com padrões e normas internacionais.
Área 6: Segurança
Pergunta
6.1 Os riscos que afetam a preservação dos objetos digitais são identificados e avaliados?
Respostas
Pergunta do tipo semiaberta, com a possibilidade de escolha de uma única resposta: ☐ São identificados e avaliados em teoria (existe o planejamento de realizá-los)

☐ São identificados e avaliados parcialmente (existe o planejamento, mas ainda não foi totalmente implantado) ☐ São identificados e avaliados ☐ Não são identificados ou avaliados ☐ Prefiro não responder ☐ Outros: _____
Justificativas
Deve-se avaliar os riscos relativos às tecnologias, objetos e <i>softwares</i> que permitam estimar as suas obsolescências e tratar a matriz de riscos.
Pergunta
6.2 Caso os riscos tenham sido identificados, quais as ferramentas de avaliação de riscos utilizadas?
Respostas
Pergunta do tipo aberta
Justificativas
Busca-se conhecer as ferramentas e/ou metodologias de avaliação de riscos utilizadas pelos repositórios.
Pergunta
6.3 A estrutura de TI (<i>software</i> e <i>hardware</i>) do repositório adota medidas de segurança contra ataques?
Respostas
Pergunta do tipo semiaberta, com a possibilidade de escolha de uma única resposta: ☐ Em teoria ☐ Adota medidas parcialmente (existe o plano, mas ainda não foi totalmente implantado) ☐ Adota medidas e executa totalmente ☐ Não adota medidas ☐ Prefiro não responder ☐ Outros: _____
Justificativas
Os equipamentos do repositório devem estar protegidos de acessos indevidos que possam levar a invasões, vazamentos ou corrompimento dos dados armazenados.
Pergunta
6.4 A estrutura física do repositório adota medidas de segurança para evitar acesso indevido?
Respostas
Pergunta do tipo semiaberta, com a possibilidade de escolha de uma única resposta: ☐ Em teoria ☐ Adota medidas parcialmente (existe o plano, mas ainda não foi totalmente implantado) ☐ Adota medidas e executa totalmente ☐ Não adota medidas ☐ Prefiro não responder ☐ Outros: _____
Justificativas
As instalações físicas do repositório devem ser protegidas de acessos indevidos, evitando assim, o acontecimento de ataques que possam comprometer a autenticidade dos documentos.
Área 7: Integridade dos dados
Pergunta
7.1 O repositório implementa o monitoramento e tratamento dos formatos de arquivos recebidos?
Respostas
Pergunta do tipo semiaberta, com a possibilidade de escolha de uma única resposta: ☐ Em teoria (existe o planejamento de monitoramento) ☐ Monitora parcialmente (existe o planejamento, mas ainda não foi totalmente implantado)

☐ Monitora totalmente na prática ☐ Não monitora ☐ Prefiro não responder ☐ Outros: _____
Justificativas
O repositório deve se certificar que os formatos dos arquivos recebidos estão de acordo com o seu plano de preservação a fim de garantir suporte e prover tratamento adequado aos arquivos.
Pergunta
7.2 O repositório adota medidas para prevenir o corrompimento dos dados?
Respostas
Pergunta do tipo semiaberta, com a possibilidade de escolha de uma única resposta: ☐ Em teoria ☐ Adota medidas parcialmente (existe o planejamento, mas ainda não foi totalmente implantado) ☐ Adota medidas na prática ☐ Não adota ☐ Prefiro não responder ☐ Outros: _____
Justificativas
Uma das principais funcionalidades do repositório é garantir que os documentos e demais arquivos utilizados não se corrompam, caso contrário toda a preservação se torna comprometida.
Pergunta
7.3 O repositório possui um plano de <i>backup</i> ?
Respostas
Pergunta do tipo semiaberta, com a possibilidade de escolha de uma única resposta: ☐ Possui em teoria ☐ Possui e aplica parcialmente na prática (existe o plano, mas ainda não foi totalmente implantado) ☐ Possui e aplica totalmente na prática ☐ Não possui ☐ Prefiro não responder ☐ Outros: _____
Justificativas
Busca-se verificar se o repositório possui um plano que gerencia a construção de suas cópias de segurança visando não perder dados e possibilitando uma pronta recuperação em casos de catástrofes.
Pergunta
7.4 Caso o repositório possua um plano de <i>backup</i> , quais estratégias são empregadas?
Respostas
Pergunta do tipo semiaberta, com a possibilidade de escolha de várias respostas: ☐ Não se aplica ☐ Utilização de servidores distribuídos geograficamente ☐ Utilização de diferentes tecnologias de armazenamento (HDs, fitas de armazenamento, nuvem, etc) ☐ Cópias de metadados e inventários separadas dos documentos ☐ Outros: _____
Justificativas
Distribuir o <i>backup</i> dos dados do repositório em diferentes localizações geográficas e mediante uso de tecnologias diferentes diminui os riscos envolvidos nos casos em que uma das instalações ou tecnologia específica empregada é comprometida (LEVELS OF PRESERVATION REVISIONS WORKING GROUP, 2019).
Área 8: Metadados
Pergunta

8.1 O repositório estabelece um padrão de metadados reconhecido a ser adotado pelos documentos recebidos?
Respostas
Pergunta do tipo semiaberta, com a possibilidade de escolha de uma única resposta: ☐ Em teoria ☐ Estabelece e aplica parcialmente na prática ☐ Estabelece e aplica totalmente na prática ☐ Não estabelece ☐ Prefiro não responder ☐ Outros: _____
Justificativas
O repositório deve estabelecer o conjunto de metadados esperado nos documentos recebidos assim como as suas obrigatoriedades a fim de que haja uma homogeneidade entre os documentos sob sua guarda.
Pergunta
8.2 São recolhidos/registrados metadados que permitam identificar o contexto, referência, direitos e autenticidade dos documentos?
Respostas
Pergunta do tipo semiaberta, com a possibilidade de escolha de uma única resposta: ☐ Em teoria (existe apenas o planejamento de recolher/registrar os metadados) ☐ São recolhidos/registrados parcialmente (existe o planejamento, mas ainda não foi totalmente implantado) ☐ São recolhidos/registrados totalmente ☐ Não são recolhidos ☐ Prefiro não responder ☐ Outros: _____
Justificativas
A utilização de metadados adequados permite ao repositório identificar o contexto, proveniência, referência e integridade dos documentos.
Área 9: Acesso e consulta
Pergunta
9.1 Existe uma política de acesso aos objetos digitais?
Respostas
Pergunta do tipo semiaberta, com a possibilidade de escolha de uma única resposta: ☐ Existe em teoria ☐ Existe parcialmente (existe o plano, mas ainda não foi totalmente implantado) ☐ Existe na prática e está totalmente implantado ☐ Não existe ☐ Prefiro não responder ☐ Outros: _____
Justificativas
Devem ser implementadas políticas de acesso à Comunidade-Alvo informando as possibilidades oferecidas registrando-se as consultas realizadas, de acordo com o critério ISO 16363 4.6.1 (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2017, p. 60).
Pergunta
9.2 Existe uma política de monitoramento das necessidades da sua Comunidade-Alvo a fim de disponibilizar os documentos da forma correta e nos formatos de arquivo mais adequados?
Respostas
Pergunta do tipo semiaberta, com a possibilidade de escolha de uma única resposta: ☐ Existe em teoria

☐ Existe parcialmente (existe o plano, mas ainda não foi totalmente implantado) ☐ Existe na prática e está totalmente implantado ☐ Não existe ☐ Prefiro não responder ☐ Outros: _____
Justificativas
Há que se manter um monitoramento ativo das necessidades da Comunidade-Alvo do repositório para que a disponibilização dos dados ocorra da maneira mais adequada, além de utilizar os formatos de arquivos apropriados (ASOCIACIÓN ESPAÑOLA DE NORMALIZACIÓN Y CERTIFICACIÓN, 2015, p. 51). Deve-se priorizar a utilização de formatos abertos de arquivos, evitando-se assim a dependência de <i>softwares</i> proprietários que podem mudar a maneira como os dados são lidos ou serem descontinuados, por exemplo.
Pergunta
9.3 Quais os formatos de arquivo mais habituais que o repositório costuma disponibilizar aos usuários?
Respostas
Pergunta fechada, com a possibilidade de escolha de várias respostas. ☐ PDF ☐ Imagens (jpg, tiff, png, etc.) ☐ Formatos de áudio (wav, mp3, etc.) ☐ Formatos de vídeo (avi, mov, mpg, etc.) ☐ Editores de texto (doc, docx, odt, etc.) ☐ Planilhas de cálculo (xls, xlsx, ods, etc.) Para cada uma das respostas, deve-se informar a frequência com que o formato de arquivo é disponibilizado aos usuários de acordo com a escala: ☐ Nunca ☐ Raramente ☐ Ocasionalmente ☐ Frequentemente ☐ Muito frequentemente
Justificativas
Busca compreender os formatos disponibilizados aos usuários com as respectivas frequências.
Pergunta
9.3.1 Existem outros formatos habituais disponibilizados aos usuários?
Respostas
Pergunta do tipo aberta
Justificativas
Possibilita ao respondente complementar formatos de materiais não contemplados nas opções anteriores.
Área 10: Visibilidade
Pergunta
10.1 É possível localizar o acesso ao repositório a partir da página principal da instituição?
Respostas
Pergunta do tipo semiaberta, com a possibilidade de escolha de uma única resposta: ☐ Sim ☐ Não ☐ Prefiro não responder ☐ Outros
Justificativas
A partir da página inicial da instituição deve existir um link ou outro mecanismo (botão, figura,

etc.) que permita ao usuário acessar o repositório institucional (AZORÍN MILLARUELO, CÍVICO MARTÍN, <i>et al.</i> , 2015, p. 11).
Pergunta
10.2 Existem iniciativas para a promoção do repositório institucional?
Respostas
Pergunta do tipo semiaberta, com a possibilidade de escolha de uma única resposta: ☐ Existem em teoria ☐ Existem, mas são executadas parcialmente (existe o plano, mas ainda não foi totalmente implantado) ☐ Existem na prática e foram/estão sendo executadas totalmente ☐ Não existem ☐ Prefiro não responder ☐ Outros
Justificativas
Promover a visibilidade do repositório auxilia os usuários a saber da sua existência e permite que eles tomem conhecimento dos conteúdos armazenados (AZORÍN, BERNAL, <i>et al.</i> , 2021, p. 12)
Pergunta
10.3 Caso haja iniciativas de promoção do repositório, quais estratégias e instrumentos utilizados?
Respostas
Pergunta aberta
Justificativas
Capacitações, palestras, campanhas publicitárias (anúncios, comerciais, etc.).
Pergunta
10.4 Na página web do repositório, são utilizadas quais as formas de contato com os responsáveis pelo repositório?
Respostas
Pergunta do tipo semiaberta, com a possibilidade de escolha de várias respostas: ☐ E-mail ☐ Formulário de contato ☐ Redes sociais (Facebook, Instagram, Twitter, Whatsapp, LinkedIn, etc.) ☐ Telefone ☐ Nenhuma forma de contato ☐ Outros
Justificativas
Disponibilizar uma ferramenta de contato possibilita que o público interessado possa interagir com os responsáveis pelo repositório facilitando a comunicação com os usuários (AZORÍN, BERNAL, <i>et al.</i> , 2021, p. 14)

Fonte: elaboração própria.

ANEXO B
INSTRUMENTO QUANTITATIVO DE AVALIAÇÃO DOS
REPOSITÓRIOS

No presente Anexo, é apresentado o instrumento quantitativo de avaliação dos repositórios com as devidas áreas avaliadas e suas pontuações máximas, assim como as respectivas perguntas, respostas e pontuações. A partir desse instrumento, se obteve as pontuações dos repositórios participantes do presente estudo.

Figura 80 – Instrumento quantitativo de avaliação dos repositórios com perguntas, respostas e pontuações.

Área 1: Contextualização do repositório	
Pergunta 1.1 O repositório institucional possui missão, visão e valores elaborados?	Pontuação máxima: 3,33
Respostas	Valor
Sim	3,33 (100%)
Não	0,00
Prefiro não responder	0,00
Pergunta 1.2 O repositório institucional possui a sua Comunidade-Alvo delimitada?	Pontuação máxima: 3,33
Respostas	Valor
Sim	3,33 (100%)
Não	0,00
Prefiro não responder	0,00
Pergunta 1.4 Qual a natureza do material depositado no repositório?	Pontuação máxima: 3,33
Respostas	Valor
1 Categoria de material	0,67 (20%)
2 Categorias de materiais	1,33 (40%)
3 Categorias de materiais	2,00 (60%)
4 Categorias de materiais	2,67 (80%)
5 ou mais categorias de materiais	3,33 (100%)
Prefiro não responder	0,00
Área 2: Estratégia e responsabilidade	
Pergunta 2.1 Existe uma Política de Preservação com os requisitos mínimos dos documentos recebidos, tratamento de preservação e acesso?	Pontuação máxima: 3,33
Respostas	Valor
Existe em teoria	0,83 (25%)
Existe na prática, mas ainda não foi totalmente implantado	1,67 (50%)
Existe na prática e está totalmente implantado	3,33 (100%)
Não existe	0,00
Prefiro não responder	0,00
Pergunta 2.1.1 Caso não exista uma Política de Preservação, existe um Plano Estratégico de Preservação?	Pontuação máxima: 3,33
Respostas	Valor
Existe em teoria	0,83 (25%)
Existe na prática, mas ainda não foi totalmente implantado	1,67 (50%)
Existe na prática e está totalmente implantado	3,33 (100%)
Não existe	0,00

Prefiro não responder	0,00
Pergunta 2.2 Existe um Plano de Continuidade que garante que os documentos estarão preservados e acessíveis a longo prazo, caso ocorra algum desastre ou descontinuidade do funcionamento?	Pontuação máxima: 3,33
Respostas	Valor
Existe em teoria	0,83 (25%)
Existe na prática, mas ainda não foi totalmente implantado	1,67 (50%)
Existe na prática e está totalmente implantado	3,33 (100%)
Não existe	0,00
Prefiro não responder	0,00
Área 3: Certificações e normas	
Pergunta 3.1 O repositório possui certificações alcançadas (relativas a repositórios)?	Pontuação máxima: 3,33
Respostas	Valor
Sim	3,33 (100%)
Não	0,00
Prefiro não responder	0,00
Pergunta 3.2 Quantas certificações o repositório tem a intenção de alcançar?	Pontuação máxima: 3,33
Respostas	Valor
Existe intenção de obter certificações, mas ainda não foram decididas quais	0,57 (17%)
Existe a intenção de obter 1 certificação	1,10 (33%)
Existe a intenção de obter 2 certificações	1,66 (50%)
Existe a intenção de obter 3 certificações	2,23 (67%)
Existe a intenção de obter 4 ou mais certificações	2,76 (83%)
O repositório já possui todas as certificações que projetou alcançar	3,33 (100%)
Não existe intenção de obter certificações	0,00
Prefiro não responder	0,00
Pergunta 3.3 O repositório adota padrões e normas estabelecidos?	Pontuação máxima: 3,33
Respostas	Valor
Adota 1 padrão ou norma	0,57 (17%)
Adota 2 padrões ou normas	1,10 (33%)
Adota 3 padrões ou normas	1,66 (50%)
Adota 4 padrões ou normas	2,23 (67%)
Adota 5 padrões ou normas	2,76 (83%)
Adota 6 ou mais padrões ou normas	3,33 (100%)
Não adota padrões e normas	0,00
Prefiro não responder	0,00
Área 4: Infraestrutura organizacional	
Pergunta 4.1 Existe uma estrutura organizacional adequada ao funcionamento do repositório para a preservação a longo prazo?	Pontuação máxima: 2,00
Respostas	Valor
Existe em teoria	0,50 (25%)

Existe parcialmente (existe o planejamento de estrutura, mas ainda não foi totalmente implantado)	1,00 (50%)
Existe na prática e está totalmente implantado	2,00 (100%)
Não existe	0,00
Prefiro não responder	0,00
Pergunta 4.2 Existe orçamento adequado ao funcionamento do repositório para a preservação a longo prazo?	Pontuação máxima: 2,00
Respostas	Valor
Existe em teoria	0,50 (25%)
Existe parcialmente (o orçamento está previsto, mas ainda não foi totalmente executado)	1,00 (50%)
Existe na prática e está totalmente implantado	2,00 (100%)
Não existe	0,00
Prefiro não responder	0,00
Pergunta 4.3 Em relação à quantidade de pessoas, a equipe responsável pelo repositório é adequada?	Pontuação máxima: 2,00
Respostas	Valor
É adequada em teoria	0,50 (25%)
É adequada na prática	2,00 (100%)
Não é adequada	0,00
Prefiro não responder	0,00
Pergunta 4.4 Em relação às capacitações e conhecimentos, a equipe responsável pelo repositório é adequada?	Pontuação máxima: 2,00
Respostas	Valor
É adequada em teoria	0,50 (25%)
É adequada na prática	2,00 (100%)
Não é adequada	0,00
Prefiro não responder	0,00
Pergunta 4.5 Existe um plano de capacitação permanente da equipe do repositório?	Pontuação máxima: 2,00
Respostas	Valor
Existe em teoria	0,50 (25%)
Existe parcialmente (existe o plano, mas ainda não foi totalmente implantado)	1,00 (50%)
Existe na prática e está totalmente implantado	2,00 (100%)
Não existe	0,00
Prefiro não responder	0,00
Área 5: Infraestrutura técnica e tecnológica	
Pergunta 5.1 O repositório possui estrutura física de prédios e/ou salas adequada à preservação e garantia de preservação a longo prazo?	Pontuação máxima: 2,50
Possui em teoria	0,62 (25%)
Possui na prática	2,50 (100%)
Não possui	0,00
Prefiro não responder	0,00
Pergunta 5.2 O repositório possui estrutura de equipamentos (servidores de aplicação,	Pontuação máxima: 2,50

armazenamento, rede, etc.) adequada à preservação e garantia de preservação a longo prazo?	
Possui em teoria	0,62 (25%)
Possui na prática	2,50 (100%)
Não possui	0,00
Prefiro não responder	0,00
Pergunta	Pontuação máxima:
5.3 Existe um plano de atualização periódico de <i>hardware</i> e <i>software</i> ?	2,50
Existe em teoria	0,62 (25%)
Existe parcialmente (existe o plano, mas ainda não foi totalmente implantado)	1,25 (50%)
Existe na prática e está totalmente implantado	2,50 (100%)
Não existe	0,00
Prefiro não responder	0,00
Pergunta	Pontuação máxima:
5.4 Possui plataforma (<i>software</i>) de repositório implantada?	2,50
Sim	2,50 (100%)
Não	0,00
Prefiro não responder	0,00
Área 6: Segurança	
Pergunta	Pontuação máxima:
5.1 Os riscos que afetam a preservação dos objetos digitais são identificados e avaliados?	3,33
São identificados e avaliados em teoria (existe o planejamento de realizá-los)	0,83 (25%)
São identificados e avaliados parcialmente (existe o planejamento, mas ainda não foi totalmente implantado)	1,67 (50%)
São identificados e avaliados	3,33 (100%)
Não são identificados ou avaliados	0,00
Prefiro não responder	0,00
Pergunta	Pontuação máxima:
5.3 A estrutura de TI (<i>software</i> e <i>hardware</i>) do repositório adota medidas de segurança contra ataques?	3,33
Em teoria	0,83 (25%)
Adota medidas parcialmente (existe o plano, mas ainda não foi totalmente implantado)	1,67 (50%)
Adota medidas e executa totalmente	3,33 (100%)
Não adota medidas	0,00
Prefiro não responder	0,00
Pergunta	Pontuação máxima:
5.4 A estrutura física do repositório adota medidas de segurança para evitar acesso indevido?	3,33
Em teoria	0,83 (25%)
Adota medidas parcialmente (existe o plano, mas ainda não foi totalmente implantado)	1,67 (50%)
Adota medidas e executa totalmente	3,33 (100%)
Não adota medidas	0,00
Prefiro não responder	0,00
Área 7: Integridade dos dados	

Pergunta	Pontuação máxima:
7.1 O repositório implementa o monitoramento e tratamento dos formatos de arquivos recebidos?	2,50
Em teoria (existe o planejamento de monitoramento)	0,62 (25%)
Adota medidas parcialmente (existe o planejamento, mas ainda não foi totalmente implantado)	1,25 (50%)
Adota medidas na prática	2,50 (100%)
Não adota medidas	0,00
Prefiro não responder	0,00
Pergunta	Pontuação máxima:
7.2 O repositório adota medidas para prevenir o corrompimento dos dados?	2,50
Em teoria	0,62 (25%)
Adota medidas parcialmente (existe o planejamento, mas ainda não foi totalmente implantado)	1,25 (50%)
Adota medidas na prática	2,50 (100%)
Não adota medidas	0,00
Prefiro não responder	0,00
Pergunta	Pontuação máxima:
7.3 O repositório possui um plano de <i>backup</i> ?	2,50
Possui em teoria	0,62 (25%)
Possui e aplica parcialmente na prática (existe o plano, mas ainda não foi totalmente implantado)	1,25 (50%)
Possui e aplica totalmente na prática	2,50 (100%)
Não adota medidas	0,00
Prefiro não responder	0,00
Pergunta	Pontuação máxima:
7.4 Caso o repositório possua um plano de <i>backup</i> , quantas estratégias são empregadas?	2,50
Emprega 1 estratégia	0,62 (25%)
Emprega 2 estratégias	1,25 (50%)
Emprega 3 estratégias	1,87 (75%)
Emprega 4 ou mais estratégias	2,50 (100%)
Não se aplica	0,00
Prefiro não responder	0,00
Área 8: Metadados	
Pergunta	Pontuação máxima:
8.1 O repositório estabelece um padrão de metadados reconhecido a ser adotado pelos documentos recebidos?	5,00
Em teoria	1,25 (25%)
Estabelece e aplica parcialmente na prática	2,50 (50%)
Estabelece e aplica totalmente na prática	5,00 (100%)
Não estabelece	0,00
Prefiro não responder	0,00
Pergunta	Pontuação máxima:
8.2 São recolhidos/registrados metadados que permitam identificar o contexto, referência, direitos e autenticidade dos documentos?	5,00
Em teoria (existe apenas o planejamento de recolher/registrar os metadados)	1,25 (25%)

São recolhidos/registrados parcialmente (existe o planejamento, mas ainda não foi totalmente implantado)	2,50 (50%)
São recolhidos/registrados totalmente	5,00 (100%)
Não são recolhidos	0,00
Prefiro não responder	0,00
Área 9: Acesso e consulta	
Pergunta	Pontuação máxima:
9.1 Existe uma política de acesso aos objetos digitais?	5,00
Em teoria	1,25 (25%)
Existe parcialmente (existe a política, mas ainda não foi totalmente implantada)	2,50 (50%)
Existe na prática e está totalmente implantada	5,00 (100%)
Não existe	0,00
Prefiro não responder	0,00
Pergunta	Pontuação máxima:
9.2 Existe uma política de monitoramento das necessidades da sua Comunidade-Alvo a fim de disponibilizar os documentos da forma correta e nos formatos de arquivo mais adequados?	5,00
Em teoria	1,25 (25%)
Existe parcialmente (existe a política, mas ainda não foi totalmente implantada)	2,50 (50%)
Existe na prática e está totalmente implantada	5,00 (100%)
Não existe	0,00
Prefiro não responder	0,00
Área 10: Visibilidade	
Pergunta	Pontuação máxima:
10.1 É possível localizar o acesso ao repositório a partir da página principal da instituição?	2,50
Sim	2,50 (100%)
Não	0,00
Prefiro não responder	0,00
Pergunta	Pontuação máxima:
10.2 Existem iniciativas para a promoção do repositório institucional?	2,50
Existem em teoria	0,62 (25%)
Existem, mas são executadas parcialmente (existe o plano, mas ainda não foi totalmente implantado)	1,25 (50%)
Existem na prática e foram/estão sendo executadas totalmente	2,50 (100%)
Não existem	0,00
Prefiro não responder	0,00
Pergunta	Pontuação máxima:
10.3 Quantas iniciativas de promoção do repositório, são utilizadas?	2,50
É utilizada 1 tipo de iniciativa	0,50 (20%)
São utilizados 2 tipos de iniciativa	1,00 (40%)
São utilizados 3 tipos de iniciativa	1,50 (60%)
São utilizados 4 tipos de iniciativa	2,00 (80%)
São utilizados 5 ou mais tipos de iniciativa	2,50 (100%)
Não se aplica	0,00

Nenhuma iniciativa é utilizada	0,00
Prefiro não responder	0,00
Pergunta 10.4 Na página web do repositório, quantas formas de contato com os responsáveis pelo repositório são utilizadas?	Pontuação máxima: 2,50
Utiliza 1 forma de contato	0,50 (20%)
Utiliza 2 formas de contato	1,00 (40%)
Utiliza 3 formas de contato	1,50 (60%)
Utiliza 4 formas de contato	2,00 (80%)
Utiliza 5 ou mais formas de contato	2,50 (100%)
Não se aplica	0,00
Nenhuma iniciativa é utilizada	0,00
Prefiro não responder	0,00

Fonte: elaboração própria.

RESUMEN EN ESPAÑOL

UNIVERSIDAD DE SALAMANCA
ESCUELA DE DOCTORADO STUDII SALMANTINI
FORMACIÓN EN LA SOCIEDAD DEL CONOCIMIENTO



**ANÁLISIS DE CONFIABILIDAD EN LOS REPOSITARIOS
DIGITALES BRASILEÑOS: ESTUDIO DE CASO DE LA RED
CARINIANA**

MARCOS VINÍCIUS BITTENCOURT DE SOUZA

Salamanca, España
2023

RESUMEN EN ESPAÑOL

**ANÁLISIS DE CONFIABILIDAD EN LOS REPOSITORIOS
DIGITALES BRASILEÑOS: ESTUDIO DE CASO DE LA RED
CARINIANA**

MARCOS VINÍCIUS BITTENCOURT DE SOUZA

Tesis presentada en la Escuela de Doctorado
“STUDII SALAMANTINI”, en el Programa
de Doctorado Formación en la Sociedad del
Conocimiento, para la obtención del título de
doctor por la Universidad de Salamanca.

Directores de la tesis:
Prof.^a Dra. María Manuela Moro Cabero
Prof. Dr. José Luis Bonal Zazo

Salamanca, España
2023

TABLA DE CONTENIDO

1 CONTEXTUALIZACIÓN Y JUSTIFICATIVA DE LA INVESTIGACIÓN.....	4
2 DEFINICIÓN DEL TEMA, ALCANCE Y LIMITACIONES.....	14
3 OBJETIVOS	16
4 HIPÓTESIS.....	19
5 ESTRUCTURA DEL TEXTO	22
6 CONCLUSIONES	24
6.1 Grado de cumplimiento de los objetivos	25
6.2 Confirmación de las hipótesis	28
6.3 Conclusiones y recomendaciones.....	32
6.4 Líneas futuras de investigación	41

1 CONTEXTUALIZACIÓN Y JUSTIFICATIVA DE LA INVESTIGACIÓN

En las últimas décadas, la creación y el uso de documentos digitales se ha convertido en una realidad en la vida cotidiana de personas, instituciones y empresas, lo que suscita preocupación por la conservación de estos documentos. Dado que la naturaleza de los documentos digitales es efímera y frágil, es necesario planificar y adoptar procesos y herramientas que permitan preservarlos durante el tiempo necesario, garantizando su uso en un horizonte temporal a menudo inalcanzable para la imaginación humana.

La obsolescencia y los fallos de los soportes de almacenamiento, así como la constante desactualización de las tecnologías, exigen a las instituciones ubicarse siempre un paso por delante de la obsolescencia del hardware y, en no pocas ocasiones, son detectados tardíamente algunos efectos, lo que exige una respuesta ágil y eficaz por parte de las organizaciones que preservan documentos digitales. En décadas anteriores, el principal soporte de almacenamiento para la preservación fue la cinta magnética, que ya presentaba problemas de compatibilidad de formatos propietarios y de dispositivos de lectura y escritura. Cuando una organización elegía un determinado estándar de cinta magnética, vinculado a su fabricante, adquiría los conocimientos necesarios y estructuraba sus procesos en función de la solución seleccionada. En caso de cambio de norma, por ejemplo, había que revisar todos los procesos de conservación y acceso, lo que implicaba que las organizaciones se mostraran inflexibles ante los cambios y, en algunos casos, rehenes de los fabricantes de la tecnología.

Con el desarrollo de la tecnología, las cintas magnéticas quedaron obsoletas y los dispositivos de lectura y escritura pasaron a ser objetos en desuso, hechos que incidieron para que las organizaciones migrasen a soportes más modernos. Incluso antes de que se abandonara el uso de cintas, la falta de fabricantes, de asistencia técnica y de

personal cualificado ya suponía un obstáculo para las organizaciones, lo que las llevaba a adoptar nuevas tecnologías y, en ocasiones, a la pérdida de datos almacenados por imposibilidad de acceso.

En el transcurso de los años, han surgido nuevos soportes de almacenamiento, que se han popularizado y han sido adoptados para la conservación de documentos. La aparición de los discos duros magnéticos (HDD) supuso una evolución en el conjunto de la industria informática, al disponer de una mayor densidad de datos, permitiendo almacenar más contenido en un menor espacio físico, además de mejorar la velocidad de acceso y escritura de los dispositivos, en comparación con las cintas magnéticas. Los avances en esta tecnología se basaron en el desarrollo de semiconductores más eficientes y pequeños, entre otras tecnologías, lo cual facilitó y popularizó su uso. Como mejora, los propios dispositivos disponen ahora de mecanismos automáticos de detección y corrección de fallos, lo que aumenta su fiabilidad.

Los soportes ópticos fueron otra tecnología utilizada para el almacenamiento de datos, con mayores capacidades que los disquetes y precios más bajos que los discos duros. Sin embargo, fue necesario implantar procesos adecuadamente establecidos para el almacenamiento, lectura y escritura de datos que, en cierto modo, se asimilaban a los utilizados en las cintas magnéticas.

Además de los discos duros, otras tecnologías también utilizan soportes magnéticos como modo de funcionamiento para la lectura y escritura de datos, a modo de ejemplo, señalamos los discos en estado sólido -SSD (*Solid State Drives*). En este caso, el dispositivo utiliza transistores para almacenar datos de forma persistente, magnética, pero no tiene componentes móviles como los HDD, lo que reduce el riesgo de fallo de los componentes físicos. Los discos duros siguen siendo muy utilizados hoy en día para almacenar datos, ya que tienen una gran capacidad de almacenamiento y una vida útil más larga.

Independientemente del soporte utilizado, la preservación digital requiere un examen minucioso de los soportes de almacenamiento, tratando de identificar las condiciones y los puntos de posibles fallos que provocan corrupciones o alteraciones de los materiales. Desde un enfoque considerando su naturaleza física, deben adoptarse procesos de verificación para identificar con antelación el instante en que sea necesario

sustituir el equipo. Desde un enfoque más abstracto, deben vigilarse las tecnologías empleadas en este almacenamiento, observando los sistemas de ficheros utilizados y, asimismo, tratando de anticiparse a los cambios tecnológicos que pueden beneficiar o poner en peligro la preservación digital.

Deteniéndonos en un nivel más abstracto, destacamos los formatos de archivo utilizados para almacenar textos, imágenes, música, vídeo, etc. A medida que los soportes físicos se deterioran con el tiempo, los formatos de archivo se modifican, ya sea para añadir funcionalidades, mejorar el rendimiento o corregir errores detectados. El software utilizado para producir estos materiales evoluciona con el tiempo y, en ocasiones, deja de existir compatibilidad con los formatos antiguos, lo que obliga a migrar a las versiones más actualizadas. En los casos en que se interrumpe la compatibilidad con un formato, también debe llevarse a cabo la conversión al formato más equivalente posible, evitando la máxima pérdida de datos.

Cabe señalar que, incluso, en un nivel más abstracto de los formatos de archivo existe obsolescencia; la cual debe ser el objetivo de las estrategias de preservación, anticipándose a posibles problemas futuros de acceso, interpretabilidad, utilización y reutilización de los materiales. Por ejemplo, optar por formatos sostenibles se considera una de estas estrategias, ya que se buscan formatos abiertos ampliamente aceptados por la comunidad. Sin embargo, dependiendo de las necesidades y de la base de conocimientos de las comunidades designadas, esto no siempre es posible, lo que requiere procedimientos y esfuerzos detallados para controlar y prevenir la obsolescencia.

Asimismo, existen situaciones en las que para la preservación se opta por el uso de estrategias de emulación, con las que los equipos actuales emulan un entorno ya obsoleto y difícil de reproducir, lo que resulta útil para preservar programas informáticos desarrollados o videojuegos, por ejemplo. En otras ocasiones, son necesarias técnicas de encapsulación, cuando se preserva todo el contexto tecnológico necesario para acceder a los materiales preservados y utilizarlos.

La elección de la estrategia de preservación debe valorar su finalidad, optando por preservar todas las características del contexto original en el que se crearon y utilizaron los objetos digitales o por priorizar la utilización del material mediante

migraciones y conversiones de formato, renunciando al contexto tecnológico original. Thibodeau (2002, p. 15) identifica un filtro de cuatro criterios que deben aplicarse al elegir una estrategia de preservación: ser viable, sostenible, practicable y apropiada. Ser viable significa que existen equipos y programas informáticos capaces de aplicar eficazmente la estrategia. Ser sostenible significa que el método puede seguir utilizándose en un futuro a largo plazo o que ofrece alternativas para sustituirlo. El criterio de practicable implica que el método debe ajustarse a las limitaciones presupuestarias y no ser demasiado difícil de aplicar. Ser apropiada significa que el método debe poder aplicarse a los materiales que el depósito pretende conservar, teniendo en cuenta los requisitos impuestos y las normas empresariales.

Comparando las realidades, para los documentos físicos convencionales, la información está estrechamente vinculada al soporte utilizado, lo que significa que los procesos de preservación se centran, principalmente, en el mantenimiento de estos soportes, ya que la utilización y el acceso al contenido de estos documentos no requieren mecanismos tecnológicos adicionales. En el caso de los materiales digitales, el contenido no está ligado al soporte, lo que facilita su copia y transferencia entre soportes y ubicaciones *ad infinitum*, no obstante, la obsolescencia tecnológica supone una gran preocupación y esfuerzo en su preservación digital, requiriendo conocimientos específicos, técnicas, competencias y habilidades especializadas y, por consiguiente, profesionales con perfiles diferenciados del conservador y restaurador tradicional.

El almacenamiento físico de los archivos requiere cuidados especiales para evitar que los datos se corrompan, ya que soportes y equipos se desgastan a medida que se utilizan, lo que exige rutinas que comprueben periódicamente la integridad de los documentos, así como procedimientos programados para sustituir aquellos equipos más utilizados, evitando cualquier tipo de pérdida de documentos. Planteando un paralelismo con los documentos en papel, éstos también requieren revisiones periódicas para detectar o minimizar el deterioro que igualmente puede llevar a pérdidas irreparables.

Una vez abordada la obsolescencia tecnológica mediante estrategias y procesos adecuados, es necesario preservar la autenticidad de los materiales digitales. La acumulación de materiales digitales sin garantías de autenticidad acaba generando el

almacenamiento de bits y bytes inútiles, dado que no es posible certificar que efectivamente son prueba de los hechos que relatan. Esto incluye tanto los documentos digitales de archivo como los datos de investigación que originan y sustentan la investigación científica, por ejemplo.

La autenticidad se construye desde el momento en que se crean los materiales digitales, analizando los sistemas y procesos que intervienen en su producción. Al capturar elementos que nos permiten comprender su contexto de producción, los actores implicados con sus responsabilidades y funciones, y sus relaciones con otros materiales, podemos obtener una especie de instantánea de ese momento, que proporciona la base para verificar la autenticidad.

La preocupación por la autenticidad es un tema central en la gestión y conservación de documentos. Mientras que la presunción de autenticidad de los documentos en papel puede hacerse analizando el propio documento y sus elementos, en el mundo digital, es necesario capturar información adicional, denominada metadatos, para apoyar el proceso de verificación de la autenticidad. La facilidad con la que pueden crearse y manipularse los documentos digitales supone un reto para la Diplomática y la Archivística, ya que un usuario de un programa de edición de textos, por ejemplo, puede reescribir un documento para que parezca escrito originalmente por otra persona.

De este modo, verificar la autenticidad en el universo digital implica analizar los metadatos de los documentos, que registran detalles del fichero creado, el contexto de producción, los permisos de los usuarios y otras interacciones a lo largo de su vida. El análisis minucioso de los metadatos relativos a los documentos permite confirmar y verificar su autenticidad e identificar las manipulaciones intencionadas o no.

En la realidad de los documentos de archivo digitales, los sistemas de gestión de documentos desempeñan un papel central en la composición de la autenticidad al controlar las funciones y permisos de los actores implicados, permitiendo o denegando operaciones en función de sus atribuciones, generalmente relacionadas con sus cargos en las organizaciones. A medida que se utilizan estos documentos y se accede a ellos, se añaden más elementos a sus bloques de información, construyendo una estructura sólida que permite certificar su autenticidad. Estos elementos detallan la posición de los

documentos dentro de la organización, sus propósitos y funciones y se relacionan con las actividades reflejadas en los Cuadros de Clasificación de Documentos de la organización. Así se construye la relación entre los documentos, los cuales se apoyan mutuamente en su autenticidad, permitiendo certificar que son testigos veraces de los hechos que relatan.

En este contexto, cabe señalar que en las últimas décadas la Organización Internacional de Normalización (ISO) ha desarrollado una amplia gama de normas vinculadas de una u otra forma a los documentos de archivo digitales. En relación con la gestión de documentos, podemos mencionar la norma ISO 15489:2016 y la familia de normas ISO 30300:2020, que establecen procesos y requisitos relativos a la gestión de documentos administrativos. También han sido editadas la ISO 23081:2017, que establece un marco para la creación, gestión y uso de metadatos para documentos de archivo, alineada con la ISO 15489, así como la ISO 13008:2022, que establece las directrices que deben observarse en las conversiones de formatos de archivo y migraciones de documentos o ficheros. Igualmente, se menciona la norma ISO/TR 18492:2005, guía práctica para la conservación a largo plazo de documentos auténticos, que define largo plazo como:

"[. ...] el periodo de conservación (que) supera la vida útil prevista de la tecnología (hardware y software) utilizada para crear y mantener la información" (INTERNATIONAL ORGANISATION FOR STANDARDISATION, 2005).

Si nos centramos en la preservación digital, existen diversas normas, tales como la norma ISO 14721:2012, donde se establecen los requisitos para la implementación de un Sistema Abierto de Información Archivística (OAIS) y se define un modelo de referencia con un flujo de información realizado a través de paquetes, o la norma ISO 16363:2012, que se utiliza para la certificación y establece mecanismos y criterios para certificar que un archivo puede denominarse OAIS. Igualmente, se señala la norma ISO 17068:2012, que trata sobre la contratación de repositorios de terceros.

Además, podemos mencionar la familia de normas ISO 9000, que trata sobre Sistemas de Gestión y Aseguramiento de la Calidad, con un alcance muy amplio que puede aplicarse en los nichos más diversos, y la ISO 16919, que establece requisitos para

la composición y funcionamiento de organismos de auditoría y certificación de repositorios digitales confiables. Asimismo, cabe destacar en materia de seguridad de la información a la familia de normas ISO 27000, que trata sobre los Sistemas de Gestión de la Seguridad de la Información (SGSI) y asume que toda organización, independientemente de su tamaño, recopila, procesa, gestiona y almacena información y, por lo tanto, deben existir normas y mecanismos que garanticen la seguridad de esta, tal y como se define en la norma (INTERNATIONAL ORGANISATION FOR STANDARDIZATION, 2018, p. 11):

"El término seguridad de la información se basa generalmente en el hecho de que la información se considera un activo que tiene un valor que requiere una protección adecuada, por ejemplo, contra la pérdida de disponibilidad, confidencialidad e integridad."

Partiendo de la necesidad de mantener la seguridad de la información de las organizaciones, un SGSI se define como (INTERNATIONAL ORGANISATION FOR STANDARDISATION, 2018, p. 11):

"[...] consiste en políticas, procedimientos, directrices y recursos y actividades asociados, gestionados colectivamente por una organización, con el objetivo de proteger sus activos de información."

Por lo tanto, existe una intersección de intereses en los objetivos de establecer un SGSI y garantizar la autenticidad de los documentos gestionados por estos sistemas, característica fundamental para la conservación fiable de los documentos digitales.

En Brasil, en el ámbito de la gestión de documentos, también existe el Modelo de Requisitos para Sistemas de Gestión de Documentos Informatizados, e-ARQ, que contiene requisitos, procesos y metadatos que deben ser implementados por los sistemas de gestión de documentos de archivo (CONSELHO NACIONAL DE ARQUIVOS, 2022).

La captura/recepción y el almacenamiento de documentos digitales es responsabilidad de las instituciones de Archivo o Repositorios Digitales, que representan a una empresa, organismo u organización que dispone de políticas de preservación y acceso, así como de las estructuras físicas y organizativas adecuadas para llevar a cabo la preservación digital. Esas instituciones tienen el deber de establecer los criterios

necesarios para conservar los documentos que reciben, así como de definir los procesos internos que garanticen su correcta custodia durante el tiempo de preservación necesario, que a menudo puede ser indefinido a largo plazo.

Esta definición es coincidente con el concepto de Archivo OAIS, creado por el CCSDS en 2002 y transformado posteriormente en la norma internacional ISO 14721:2012 (INTERNATIONAL ORGANISATION FOR STANDARDISATION, 2012). La norma describe un conjunto de entidades de archivo con responsabilidades delimitadas, un flujo de datos entre las entidades y los estándares de paquetes de información que forman parte de algunas de estas comunicaciones. La norma también especifica los requisitos y condiciones que deben cumplirse para determinar si un archivo cumple o no la norma. En la medida en que el archivo de una institución se adhiere a la norma ISO 14721, puede decirse que la preservación digital se realiza con un diseño de repositorio confiable.

La confiabilidad del repositorio y, por consiguiente, de sus materiales preservados se basa, a un alto nivel, en la constatación de que la cadena de custodia no ha sido corrompida o comprometida. Si se detectan fallos en esta cadena en el entorno de gestión, toda la secuencia de preservación y acceso se torna frágil, no siendo posible certificar la autenticidad de los materiales. La extracción de los documentos de sus sistemas de gestión para su ingesta en el repositorio es un punto crítico que debe analizarse, asegurándose de que se lleva a cabo mediante sistemas desarrollados y especializados en este proceso y no por personas que puedan poner en riesgo la autenticidad de estos documentos.

Una vez creados los archivos digitales, es necesario reconocer cuáles efectúan con autenticidad la preservación de sus materiales. No es suficiente que los archivos o depósitos digitales se autodenominen confiables, sino que hay que establecer criterios para validar sus procesos y herramientas, creando así un entorno de confianza. Al respecto, han surgido herramientas y metodologías para verificar la capacidad que poseen los Archivos para efectuar una preservación confiable a largo plazo, siendo extensivas para la aplicación de modelos de madurez, el análisis de riesgos, la realización de auditorías internas y externas, hasta la obtención de certificaciones y sellos de aprobación.

Estos procesos permiten a las unidades de Archivo comprobar el cumplimiento de los criterios relacionados con los objetivos de conservación de los documentos, por lo que es importante que estas conozcan su estado actual. El mapeo de los criterios incumplidos identifica las debilidades y riesgos existentes, proporcionando material para su resolución o mitigación, según cada caso.

La comunidad archivística internacional ha desarrollado diversos modelos de certificación para repositorios digitales, que pueden ser aplicados por los propios equipos, realizando el proceso de autoevaluación, o mediante la contratación de una entidad evaluadora acreditada que compruebe el cumplimiento de los criterios.

En este universo, pueden citarse herramientas diversas, tales como la matriz de preservación NDSA, que presenta niveles de preservación que evolucionan en sofisticación, culminando en un estado en el que la preservación digital de una organización muestra signos de confiabilidad para algunos aspectos (NATIONAL DIGITAL STEWARDSHIP ALLIANCE, 2022). Otro ejemplo es DRAMBORA, una metodología de análisis de riesgos, que dispone de una herramienta para evaluar la probabilidad y el impacto de los acontecimientos que podrían comprometer la preservación digital de los repositorios, abarcando desde los aspectos organizativos hasta los tecnológicos, pasando por el tratamiento de los paquetes de información preservada (DIGITAL CURATION CENTRE AND DIGITAL PRESERVATION EUROPE, 2007).

En cuanto a la certificación formal de repositorios (esto es, de tercera parte), existe la norma ISO 16363:2012, que establece una metodología y unos criterios para la realización de auditorías que permitan comprobar si un determinado Archivo cumple con la norma de referencia OAIS (INTERNATIONAL ORGANISATION FOR STANDARDISATION, 2012). En este caso, la certificación se obtiene tras una auditoría realizada por un organismo externo que analiza detalladamente todos los criterios y culmina con la certificación. De forma similar a esta norma, la ISO 17068:2012 analiza el cumplimiento de los criterios en el caso de repositorios externos a la organización, los denominados repositorios de tercera parte, certificando que estos repositorios pueden considerarse confiables (INTERNATIONAL ORGANISATION FOR STANDARDIZATION, 2012).

También existe la opción de que el repositorio obtenga el Sello Nestor, un sello que acredita que el repositorio realiza una preservación digital confiable a través de auditorías internas que posteriormente se presentan a la organización que lo controla. El sello tiene características similares a la norma ISO 16363:2012, ya que su grupo de trabajo participó en la redacción de dicha norma (NESTOR WORKING GROUP TRUSTED REPOSITORIES-CERTIFICATION, 2009). Destinado a ser el primer paso en la verificación de la confiabilidad de los repositorios destacamos el Sello *CoreTrust*, que requiere una auditoría interna y posterior revisión por parte de expertos en el campo de la preservación de documentos, sin necesidad de una entidad auditora formalizada. Este sello se obtiene tras el envío de la documentación soporte a la organización que lo controla y se presenta como un primer esfuerzo y compromiso con la preservación confiable de los repositorios (CORETRUSTSEAL STANDARDS AND CERTIFICATION BOARD, 2019).

En España se ha creado una guía que evalúa criterios de repositorios institucionales y verifica la presencia de buenas prácticas y aplicaciones de estándares de metadatos y servicios de recogida de información. La guía se ha elaborado en el marco de la Red de Bibliotecas Universitarias (REBIUN) y, entre sus objetivos, analiza si los repositorios institucionales españoles contribuyen a la causa de la Ciencia Abierta, que amplía y difunde el acceso al conocimiento científico. Aunque la guía se ha elaborado a partir de la realidad española, sus criterios pueden servir para evaluar repositorios de otros países (AZORÍN, BERNAL, et al., 2021).

2 DEFINICIÓN DEL TEMA, ALCANCE Y LIMITACIONES

Partiendo del contexto descrito anteriormente, caracterizado por los desafíos que implica la preservación digital y la necesidad de que los repositorios digitales respondan a dichos retos, este estudio se basa en el siguiente problema: **¿Cómo es posible afirmar que la preservación digital de los repositorios brasileños es fiable y de calidad? ¿Qué mecanismos se utilizan para evaluar los repositorios? ¿Qué áreas funcionales de los repositorios deben ser evaluadas para determinar su confiabilidad?**

Como referencia y alcance, se utiliza un subconjunto de repositorios brasileños, representado por los repositorios digitales pertenecientes a la *Red Cariniana*, que abarca todas las regiones geográficas del territorio brasileño. Entre los elementos que más impactaron en la investigación se destacan los vinculados a la parte exploratoria, dada la realidad brasileña, con gran diversidad de regiones, de divisiones estatales y federales, lo cual implicaba que abarcar la totalidad de los repositorios se tornase imposible. En este sentido, fue necesario buscar una muestra representativa que explorara la realidad, de ahí la elección de una red con múltiples instituciones de alcance nacional.

Se trata de una investigación de naturaleza descriptiva y analítica cualitativa-cuantitativa que incorpora un estudio de caso y permite desarrollar herramientas adaptadas y propuestas de mejora.

Para el marco teórico se utilizaron técnicas descriptivo-analítico-sintéticas basadas en fuentes bibliográficas, normativas internacionales y nacionales, así como fuentes documentales, que facilitaron el análisis y sistematización de los factores mediadores en la preservación a largo plazo de los objetos digitales, lo que consecuentemente permitió conocer el grado de complejidad funcional y tecnológica que rodea a los repositorios digitales.

Además, se consultaron y analizaron las herramientas de evaluación más relevantes utilizadas a nivel internacional con el fin de elegir los criterios esenciales para evaluar los repositorios. De este modo, los criterios de evaluación se basaron en las

siguientes herramientas y normas: ISO 14721:2012, ISO 16363:2012, matriz de niveles NDSA, Sello Nestor, Sello *CoreTrust*, DRAMBORA y la Guía de Evaluación de Repositorios REBIUN. Sus conjuntos de criterios han servido de base para la elaboración de los criterios y áreas utilizadas en este estudio.

Para la realización de la investigación, de carácter aplicado, se utilizaron los repositorios de la *Red Cariniana*, con los permisos adecuados para la exploración de datos. La Red está formada por 31 instituciones, de las cuales 22 son públicas federales, 4 estatales y 5 privadas. Respondieron 19 instituciones, lo que representa el 61% del total de repositorios. De las instituciones que no respondieron, 7 son privadas o estatales y 5 federales. Se pidió a todos los repositorios que cumplimentaran el formulario para obtener una visión general de cómo se lleva a cabo la preservación digital en sus respectivas instituciones.

Por otro lado, una vez realizada la selección, los problemas surgieron por la falta de respuestas de algunos repositorios de la *Red Cariniana*. También ha sido un factor limitante la protección de los datos obtenidos para no difundir los resultados de aquellos repositorios que mostraban debilidades evidentes. Todo ello se señala en los apartados de esta investigación cuando se muestran y analizan resultados. Las demás limitaciones metodológicas se presentan en el capítulo 4 sobre metodología.

Los resultados esperados de la investigación se centran, en primer lugar, en la creación de una herramienta de evaluación de repositorios para valorar los factores relevantes para la calidad de la preservación de estos repositorios y, en segundo lugar, en disponer de un mapeo del grado de confiabilidad de la muestra seleccionada de repositorios institucionales y la realidad brasileña. Además, se relacionan los principales desafíos de preservación en sus áreas a partir de abordajes heterogéneos: objeto, ciclo de vida, custodia y procesos específicos.

En el siguiente subapartado, se presentan los objetivos previstos para este trabajo.

3 OBJETIVOS

Se presentan los objetivos de este trabajo de investigación, estructurados en dos niveles para una mejor comprensión. En el primer nivel, se enuncian los grandes objetivos generales (**OG**), cada uno de los cuales se divide en una serie de objetivos específicos (**OE**):

- **OG1:** definir un modelo general de análisis y evaluación de repositorios digitales basado en fuentes normativas internacionales.

Aunque, como se ha mencionado anteriormente, existen diversas herramientas relacionadas con la auditoría, evaluación, análisis de riesgos o certificación de repositorios digitales, se ha optado por diseñar un modelo general debido a que las herramientas existentes están vinculadas a la evaluación de repositorios según una norma concreta, lo que condiciona el propio proceso de evaluación, o bien están enfocadas a determinados aspectos concretos (evaluación de repositorios de terceros, evaluación de riesgos, evaluación del impacto de la ocurrencia de eventos que comprometen la preservación o realización de auditorías internas, entre otros). Como puede observarse, se trata de un objetivo metodológico, para el que se han establecido los siguientes objetivos específicos:

- **OE1.1:** identificar y analizar las herramientas de evaluación de repositorios digitales utilizadas en el contexto internacional, con el fin de obtener una visión general de los elementos comunes y las diferencias más significativas entre ellas;
- **OE1.2:** identificar, seleccionar y clasificar los criterios de evaluación para la certificación de repositorios fiables identificados en las herramientas mencionadas en el punto anterior, con el fin de definir las áreas de gestión de repositorios digitales en las que debe centrarse un proceso de evaluación. Definir las áreas de gestión en las que se categorizan los criterios de evaluación es un paso necesario para establecer un modelo de evaluación;
- **OE1.3:** desarrollar un cuestionario estructurado de evaluación de repositorios digitales que pueda aplicarse en cualquier entorno institucional, tecnológico y normativo;

- **OE1.4:** diseñar una herramienta de evaluación cuantitativa, asociada al cuestionario indicado en el punto anterior, que permita: 1) valorar objetivamente el nivel de calidad de un repositorio; 2) realizar comparaciones equitativas entre las diferentes áreas de gestión de un repositorio; 3) realizar comparaciones entre diferentes repositorios, y 4) elaborar un ranking de repositorios con relación a sus cualidades;
 - **OE1.5:** probar el modelo de evaluación en un entorno limitado en el contexto brasileño, con el fin de comprobar la aplicación del modelo y validar su utilidad.
- **OG2:** analizar el estado de preservación digital de los repositorios digitales de la *Red Cariniana* utilizando el modelo definido anteriormente.

Este análisis permitirá construir una imagen de los elementos internos de los repositorios relacionados con la fiabilidad. Por razones que serán explicadas en el capítulo de metodología, la *Red Cariniana* es un buen ejemplo de repositorios brasileños, por lo que las conclusiones extraídas de este análisis pueden ser extrapoladas a otros sistemas de repositorios digitales del país. Para alcanzar este objetivo general, se definieron los siguientes objetivos específicos:

 - **OE2.1:** evaluar aspectos relacionados con la planificación y gestión de los repositorios digitales de la *Red Cariniana*;
 - **OE2.2:** evaluar los aspectos técnicos de los repositorios y el uso de metadatos de preservación;
 - **OE2.3:** evaluar las técnicas y mecanismos utilizados para garantizar la seguridad e integridad de los datos;
 - **OE2.4:** evaluar el modo en que los repositorios difunden sus materiales y los hacen accesibles al público;
 - **OE2.5:** identificar en los repositorios las áreas funcionales más desarrolladas (puntos fuertes) y las que necesitan mejoras.
- **OG3:** determinar la existencia de factores externos que pueden influir en la calidad de los repositorios digitales, estudiando el caso de la *Red Cariniana*.

Además de los factores internos comentados en el objetivo anterior, pueden existir elementos externos que influyan en la calidad de los repositorios, como su entorno geográfico, las características de su contexto institucional o incluso el tiempo transcurrido desde la creación del repositorio. Para alcanzar este objetivo general, se definieron los siguientes objetivos específicos:

- **OE3.1:** determinar la influencia del entorno geográfico en la calidad de los repositorios. Este objetivo se estableció con el fin de evaluar el impacto de las políticas generales desarrolladas en diferentes regiones geográficas sobre el objeto de estudio;
- **OE3.2:** Determinar la influencia de aspectos relacionados con el entorno institucional en la calidad de los repositorios. Se pretende analizar aspectos como el tipo de institución, el tipo de asociación entre los repositorios y la Red Cariniana, el presupuesto, el tamaño de las comunidades de repositorios y determinados índices de calidad institucional.
- **OE3.3:** Determinar cómo influye el tiempo de creación del repositorio (antigüedad) en su calidad.

En el siguiente subapartado, se formulan las hipótesis que guiarán la realización de este estudio y cuya verificación se presentará al final del trabajo.

4 HIPÓTESIS

A partir de la problematización, iniciamos el trabajo intelectual de elaboración de hipótesis que, para Gil, se caracterizan como:

"[...] una supuesta respuesta al problema a investigar. Es una proposición que se forma y que sólo será aceptada o rechazada después de haber sido debidamente comprobada" (2008, p. 41). Rampazzo caracteriza una hipótesis como "la suposición de una causa o de una ley destinada a explicar provisionalmente un fenómeno hasta que los hechos la contradigan o la confirmen" (2005, p. 36).

Atendiendo a dicha caracterización, se elaboraron las seis hipótesis que se presentan a continuación.

La hipótesis 1 (H1) se formula como sigue: **"los repositorios institucionales de la Red Cariniana cuentan con las estructuras organizativas, físicas y tecnológicas para llevar a cabo la preservación digital a largo plazo"**. Una vez presentados los retos de la preservación digital, resulta esencial que existan las estructuras adecuadas para apoyar y llevar a cabo la preservación digital propiamente dicha. Las estructuras organizativas definen las funciones, los cargos y las responsabilidades de los agentes implicados, mientras que la estructura física es necesaria tanto para la asignación de los equipos como para apoyar la estructura tecnológica necesaria y fundamental para la preservación digital. La falta, debilidad o ineficacia de cualquiera de estas estructuras podría poner en peligro la preservación de los materiales digitales.

La hipótesis 2 (H2) se formula del modo siguiente: **"los repositorios institucionales de la Red Cariniana utilizan directrices establecidas para la preservación de sus materiales digitales"**. Es a partir de las directrices que los repositorios definen los requisitos de preservación, proporcionando criterios para la implementación de reglas de depósito y acceso a los materiales preservados. Además, las directrices establecen la forma en que el repositorio puede prevenir eventos que podrían cesar y/o poner en peligro las actividades de los repositorios, sin perjudicar los materiales preservados.

La hipótesis 3 (**H3**) se formula como sigue: "**los repositorios institucionales de la Red Cariniana no aplican herramientas de evaluación de confiabilidad**". Dado que no se encontró documentación sobre procesos de certificación, auditorías o sellos de confianza en los repositorios, esto significa que no se aplican métodos para verificar la confianza en los repositorios. Esta situación puede deberse a la ausencia de una herramienta adaptada a la realidad brasileña que aborde los factores de preservación relevantes y facilite la evaluación de los repositorios. En el ámbito de la gestión documental brasileña, cabe mencionar el e-ARQ (Modelo de Requisitos para Sistemas Informatizados de Gestión de Documentos) (CONSELHO NACIONAL DE ARQUIVOS, 2022), que tiene criterios bien establecidos y detallados que deben cumplir los sistemas para implementar una gestión archivística adecuada. En el campo de la preservación, la Resolución CONARQ Nº 43 establece las directrices a seguir para que un repositorio pueda ser denominado Repositorio Archivístico Digital Confiable (RDC-Arq) (CONSELHO NACIONAL DE ARQUIVOS, 2015). Sin embargo, las directrices se presentan de manera amplia y no establecen un instrumento que pueda ser utilizado de manera efectiva para demostrar su cumplimiento o no.

La hipótesis 4 (**H4**) se formula como: "**el desarrollo de un conjunto de criterios mínimos de confiabilidad, adaptados de las normas internacionales, es útil para evaluar los repositorios brasileños y formular planes de mejora**". Como no existen herramientas de evaluación orientadas a las especificidades de la realidad brasileña, la creación de un instrumento que abarque la evaluación de diferentes áreas funcionales de los repositorios con criterios que tengan la cobertura de instrumentos, normas, auditorías y métodos ya consolidados podría ayudar en la realización de evaluaciones en los repositorios. Una vez realizada la evaluación con este conjunto de criterios, los repositorios podrían demostrar su fiabilidad o detectar puntos de mejora.

La hipótesis 5 (**H5**) se formula del modo siguiente: "**la calidad de los repositorios está influida por factores externos relacionados con su entorno**". Al formar parte de la sociedad, los repositorios están sujetos a influencias e interferencias externas a su gestión, que pueden afectar directamente a su funcionamiento y calidad. Entre estos factores externos se encuentran, por ejemplo, la ubicación geográfica, el contexto institucional y el presupuesto.

La hipótesis 6 (**H6**) se formula como sigue: "**los repositorios más antiguos deberían tener mayor calidad que los repositorios creados más recientemente**". Dado que llevan más tiempo en funcionamiento, los repositorios más antiguos deberían ser de mayor calidad, disponer de políticas y directrices, así como de mayor experiencia y conocimientos a la hora de corregir y reajustar los factores que han influido en su funcionamiento.

5 ESTRUCTURA DEL TEXTO

El trabajo de investigación se estructura en dos partes: un marco teórico y un estudio aplicado.

La primera parte estudia el objeto de preservación, analizando su concepción y los retos que plantea su preservación a lo largo de todo su ciclo de vida. También examina los medios y modelos de preservación de los objetos digitales y las principales herramientas para garantizar y demostrar la confiabilidad de los procesos de preservación.

La segunda parte se centra en el estudio aplicado, que incluye los detalles metodológicos para la creación del modelo de evaluación y el estudio exploratorio vinculado a la evaluación de la *Red Cariniana* de repositorios digitales, que facilita la comprobación del modelo de evaluación propuesto.

El marco teórico investiga, en primer lugar, las particularidades del objeto a preservar, cuáles son sus desafíos, dónde preservarlo y, a continuación, analiza cómo garantizar su preservación. Se divide en tres capítulos. El capítulo 1 detalla el objeto que debe preservarse, los principales retos, las normas y las herramientas de preservación. El Capítulo 2 presenta las particularidades de un Repositorio Digital, evolucionando a partir de los Archivos tradicionales, además de presentar el modelo de referencia para repositorios digitales actualmente en uso, ejemplos de herramientas de preservación, la inserción de la realidad brasileña en este contexto y la contribución de los repositorios al Acceso Abierto. El capítulo 3 aborda cuestiones relacionadas con la evaluación de la confiabilidad de los repositorios, la cronología de los métodos y herramientas de evaluación, la auditoría y certificación y concluye con la presentación de algunas de las herramientas destacadas.

La segunda parte de la investigación se centra en la creación de una herramienta de evaluación y en la verificación de su validez mediante un estudio exploratorio de la *Red Cariniana*. Al respecto, el Capítulo 4 detalla la metodología utilizada en este estudio, presentando el universo de repositorios analizados, la forma

de comunicación utilizada, así como las preguntas y respuestas empleadas para evaluar cada criterio. El Capítulo 5 registra los resultados del estudio exploratorio, presentando los resultados por categoría, considerando los ítems integrales y en general.

Finalmente, las conclusiones y consideraciones finales de este estudio se presentan en el Capítulo 6 y las referencias bibliográficas, de acuerdo con la norma ABNT (Brasil), en el Capítulo de Referencias.

Para la redacción y el cuidado de aspectos formales se ha respetado la norma ABNT.

6 CONCLUSIONES

Este estudio permitió identificar cómo los repositorios participantes de la *Red Cariniana* preservan sus materiales, desde aspectos más generales como la existencia de planes y políticas de preservación hasta la identificación de normas y estándares utilizados en los repositorios. En lo que se refiere a la participación, 19 de los 31 repositorios tomaron parte en el estudio, lo que permitió llegar a una amplia gama de instituciones. Sin embargo, sólo 1 repositorio privado proporcionó respuestas, lo que imposibilitó una comprensión más completa de la dinámica de preservación de los repositorios de esta naturaleza. Así pues, aunque el conocimiento de aspectos concretos de las instituciones privadas puede haberse visto obstaculizado, la amplia participación de instituciones públicas permitió conocer el estado actual de la preservación digital de los repositorios.

La mayor parte de la comunicación con los repositorios se realizó en forma de solicitudes de acceso a la información en la plataforma *Fala.BR*, lo que generó un cierto grado de formalidad y burocracia en la provisión de respuestas, que es lo contrario de lo que se pretendía en este estudio cuando las primeras comunicaciones se realizaron a través de mensajes de correo electrónico. No obstante, las respuestas fueron contestadas y permitieron realizar los análisis necesarios, siendo facilitadas por los responsables de los repositorios y miembros de sus equipos, añadiendo fiabilidad y conocimiento a las respuestas recibidas.

Con el fin de realizar una comparación objetiva entre los repositorios, se desarrolló una herramienta de puntuación para los criterios evaluados en el estudio, con pesos iguales entre las áreas y las preguntas dentro de las áreas. De esta forma, fue posible identificar puntos fuertes y aspectos a mejorar en los repositorios que permiten realizar un diagnóstico de la situación actual, pero que también pueden servir para monitorizar la evolución de los aspectos en el futuro. Además de servir para crear un ranking de repositorios, la herramienta puede ser utilizada para intercambiar soluciones

entre repositorios, fomentando el intercambio de experiencias que permitan madurar y evolucionar la preservación digital en la *Red Cariniana*.

Este apartado conclusivo presenta efectivamente las conclusiones extraídas del estudio. En el primer apartado se muestra el grado de cumplimiento de los objetivos propuestos, seguido de la confirmación o refutación de las hipótesis y las líneas de investigación futuras sugeridas.

6.1 Grado de cumplimiento de los objetivos

En este apartado se recapitulan los objetivos propuestos por este trabajo, identificando los apartados que indican su respectivo cumplimiento. El objetivo general OG1, "definir un modelo general de análisis y evaluación de repositorios digitales basado en fuentes normativas internacionales", se caracteriza por estar cumplido, ya que se han alcanzado todos los objetivos específicos que lo componen, tal y como se muestra a continuación. En la Tabla 1 se resumen los objetivos específicos de este objetivo general y los apartados del trabajo en los que se han cumplido.

El cumplimiento del objetivo específico OE1.1 "identificar y analizar las herramientas de evaluación de repositorios digitales utilizadas en el contexto internacional para obtener una visión general de los elementos comunes y las diferencias más significativas entre ellas" se demuestra en el apartado 4.2.2.1.

El cumplimiento del objetivo específico OE1.2 "identificar, seleccionar y clasificar los criterios de evaluación para la certificación de repositorios fiables identificados en las herramientas mencionadas en el punto anterior, con el fin de definir las áreas de gestión de repositorios digitales en las que debe centrarse un proceso de evaluación" se demuestra en el apartado 4.2.2.2.

El cumplimiento del objetivo específico OE1.3 "diseñar un cuestionario estructurado para la evaluación de repositorios digitales que pueda ser aplicado en cualquier entorno institucional, tecnológico y normativo" se demuestra en el apartado 4.2.3.

Tabla 1 - Cumplimiento de los objetivos específicos del Objetivo General1.

OG1: Definir un modelo general de análisis y evaluación de repositorios digitales basado en fuentes normativas internacionales	
	Sección del trabajo
OG1.1: Identificar y analizar las herramientas de evaluación de repositorios digitales utilizadas en el contexto internacional para obtener una visión general de los elementos comunes y las diferencias más significativas entre ellas.	4.2.2.1.
OE1.2: Identificar, seleccionar y clasificar los criterios de evaluación para la certificación de repositorios fiables identificados en las herramientas mencionadas en el punto anterior, con el fin de definir las áreas de la gestión de repositorios digitales en las que debe centrarse un proceso de evaluación.	4.2.2.2.
OE1.3: Desarrollar un cuestionario estructurado de evaluación de repositorios digitales que pueda aplicarse en cualquier entorno institucional, tecnológico y normativo.	4.2.3
OE1.4: Diseñar una herramienta de evaluación cuantitativa.	4.2.5.
OE1.5: probar el modelo de evaluación en un entorno limitado en el contexto brasileño, con el fin de probar la aplicación del modelo y validar su utilidad.	Capítulo 5.

Fuente: Elaboración propia

El cumplimiento del objetivo específico OE1.4 "diseñar una herramienta de evaluación cuantitativa, asociada al cuestionario indicado en el punto anterior, que permita realizar comparaciones objetivas entre las diferentes áreas de evaluación de un mismo repositorio y entre diferentes repositorios" se demuestra en el apartado 4.2.5.

Finalmente, el cumplimiento del objetivo específico OE1.5 "probar el modelo de evaluación en un ambiente limitado en el contexto brasileño, para comprobar la aplicación del modelo y validar su utilidad" se demuestra en el capítulo 5 de resultados, ya que el modelo fue aplicado a los repositorios de *la Red Cariniana*.

El objetivo general OG2 "analizar el estado de preservación digital de los repositorios digitales de la *Red Cariniana* mediante el modelo definido anteriormente" se considera cumplido porque se han alcanzado todos los objetivos específicos que lo componen, como se muestra a continuación. En la Tabla 2 se resumen los objetivos específicos de este objetivo general y los apartados del trabajo en los que se han cumplido.

Tabla 2 - Cumplimiento de los objetivos específicos del Objetivo General 2.

OG2: Analizar el estado de preservación digital de los repositorios digitales de la Red Cariniana utilizando el modelo definido anteriormente	
Apartado del trabajo	Sección del trabajo
OE2.1: Evaluar los aspectos relacionados con la planificación y gestión de los repositorios digitales de la <i>Red Cariniana</i> .	5.4.1, 5.4.2, 5.4.3 y 5.4.4
OE2.2: Evaluar los aspectos técnicos de los repositorios y el uso de metadatos de preservación	5.4.5 y 5.4.8.
OE2.3: evaluar las técnicas y mecanismos utilizados para garantizar la seguridad e integridad de los datos	5.4.6 y 5.4.7
OE2.4: evaluar cómo los repositorios difunden sus materiales y los hacen accesibles al público	5.4.9 y 5.4.10

Fuente: elaboración propia.

El cumplimiento del objetivo específico OE2.1 "evaluar aspectos relacionados con la planificación y gestión de los repositorios digitales de la Red Cariniana" se muestra en los apartados 5.4.1, 5.4.2, 5.4.3 y 5.4.4. El cumplimiento del objetivo específico OE2.2 "evaluar los aspectos técnicos de los repositorios y el uso de metadatos de preservación" se demuestra en las secciones 5.4.5 y 5.4.8. El cumplimiento del objetivo específico OE2.3 "evaluar las técnicas y mecanismos utilizados para garantizar la seguridad e integridad de los datos" se demuestra en las secciones 5.4.6 y 5.4.7. El cumplimiento del objetivo específico OE2.4 "evaluar cómo los repositorios difunden sus materiales y los hacen accesibles al público" se demuestra en las secciones 5.4.9 y 5.4.10.

El objetivo general OG3 "determinar la existencia de factores externos que pueden influir en la calidad de los repositorios digitales, estudiando el caso de la Red Cariniana" se considera alcanzado porque se han cumplido todos los objetivos específicos que lo componen, como se muestra a continuación. En la Tabla 75 se resumen los objetivos específicos de este objetivo general y los apartados del trabajo en los que se han cumplido.

Tabla 3- Cumplimiento de los objetivos específicos del Objetivo General 3.

OG3: Determinar la existencia de factores externos que pueden influir en la calidad de los repositorios digitales, estudiando el caso de la Red Cariniana	
	Sección del trabajo
OG3.1: Determinar la influencia del entorno geográfico en la calidad de los repositorios.	5.5.1
OE3.2: Determinar la influencia de los aspectos relacionados con el entorno institucional en la calidad de los repositorios	5.5.2, 5.5.3, 5.5.5.1, 5.5.5.2, 5.5.5.3 y 5.5.5.4
OE3.3: Determinar cómo influye el tiempo de creación (antigüedad) del repositorio en su calidad.	5.5.4.

Fuente: elaboración propia.

El cumplimiento del objetivo específico OE3.1 "determinar la influencia del entorno geográfico en la calidad de los repositorios" se muestra en el apartado 5.5.1. El cumplimiento del objetivo específico OE3.2 "determinar la influencia de aspectos relacionados con el entorno institucional en la calidad de los repositorios" se muestra en los apartados 5. 5.5.2, 5.5.3, 5.5.5.1, 5.5.5.2, 5.5.5.3 y 5.5.5.4. El cumplimiento del objetivo específico OE3.3 "determinar cómo influye el tiempo de creación del repositorio (antigüedad) en su calidad" se demuestra en el apartado 5.5.4.

6.2 Confirmación de las hipótesis

A partir de las hipótesis propuestas en este estudio, esta sección presenta su confirmación o refutación, como se muestra a continuación. Para facilitar la comprensión, la Tabla 4 resume las confirmaciones o refutaciones de las hipótesis propuestas en el estudio.

Tabla 4- Hipótesis y sus confirmaciones.

Hipótesis	Verificación
H1: los repositorios institucionales de la Red Cariniana poseen estructuras organizativas, físicas y tecnológicas para realizar preservación digital a largo plazo.	Confirmada con reservas
H2: los repositorios institucionales brasileños utilizan directrices establecidas para la preservación de sus materiales digitales	Parcialmente confirmada.
H3: Los repositorios institucionales brasileños no utilizan herramientas de evaluación de la fiabilidad.	Confirmada
H4: el desarrollo de un conjunto de criterios mínimos de fiabilidad, adaptados de las normas internacionales, es útil para evaluar los repositorios brasileños y formular planes de mejora.	Confirmada
H5: la calidad de los repositorios está influenciada por factores externos relacionados con su entorno.	Confirmada
H6: los repositorios más antiguos, creados más recientemente, deberían tener mayor calidad que los repositorios más recientes.	Refutada

Fuente: elaboración propia.

La hipótesis H1: "Los repositorios institucionales de la *Red Cariniana* cuentan con las estructuras organizativas, físicas y tecnológicas para llevar a cabo la preservación digital a largo plazo" se considera confirmada con reservas. Al observar los resultados de los criterios relativos a la infraestructura organizativa, disponibilidad presupuestaria y adecuación de los equipos, presentados en la sección 5.4.4, la hipótesis H1 podría refutarse, ya que los porcentajes no superan el 50% de aplicación práctica. Sin embargo, si se tiene en cuenta la aplicación parcial y la existencia teórica de los criterios, las cifras superan el 50%, de ahí la confirmación de la hipótesis con la condición de que los repositorios se comprometan a invertir esfuerzos para aplicar plenamente los criterios en la práctica. Los equipos de los repositorios se han propuesto desarrollar soluciones para la preservación de los materiales digitales, incluso con todas las dificultades que estas deficiencias imponen a esta actividad. Sin embargo, es fundamental que haya una evolución hacia la excelencia en la preservación digital, buscando crear o ampliar estructuras organizativas con puestos de trabajo, responsabilidades y atribuciones

definidas, así como presupuestos bien dimensionados que permitan el mantenimiento de las actividades e inversiones en equipos y capacitación.

La hipótesis H2 "Los repositorios institucionales brasileños utilizan directrices establecidas para preservar sus materiales digitales" se confirma parcialmente, a partir de los resultados mostrados en las secciones 5.4.2 y 5.4.7. Analizando las directrices utilizadas por los repositorios, la mayoría posee Políticas de Acceso y Planes de Respaldo que se aplican total o parcialmente en la práctica. Sin embargo, la mayoría de los repositorios no posee Políticas de Preservación y Planes de Continuidad, por lo que la hipótesis se confirma parcialmente. Cabe destacar que pocos repositorios han declarado tener una política de seguimiento de sus Comunidades Destinatarias, tratando de comprender e identificar sus necesidades y cómo reaccionar ante ellas, como se muestra en la sección 5.4.9.

El hecho de que casi todos los repositorios no tengan planes de continuidad es alarmante, ya que no son conscientes de las amenazas, porque no identifican ni evalúan los riesgos y esperan que no se produzcan eventos que les impidan funcionar. Esto pone de relieve el hecho de que no sólo los acontecimientos catastróficos pueden detener las actividades de los repositorios, sino que también pueden convertirse en el blanco de errores de gestión e injerencias políticas que pongan en peligro la autenticidad de sus materiales. Una posible razón que se aduce para explicar la falta de planes de continuidad es que el carácter público de la gran mayoría de las instituciones que albergan repositorios crea un escenario de seguridad, tanto por la estabilidad de la carrera profesional del personal como por la naturaleza de la financiación gubernamental, que hace innecesaria la elaboración de un plan de continuidad. Este estudio no permite identificar el grado en que esta situación sería similar en los repositorios privados, debido al bajo nivel de participación de estos repositorios, sin embargo, se considera que el establecimiento de contratos y acuerdos de nivel de servicio favorece el establecimiento de planes de continuidad sólidos y eficaces. Una vez implementado el plan de continuidad, se establece una vigilancia constante de los desencadenantes que puedan amenazar al repositorio y sus materiales, procurando que nunca cese la preservación digital confiable.

La falta de directrices de preservación nos lleva a la conclusión de que la preservación digital de los repositorios se efectúa de forma empírica, aunque la experiencia de los equipos de los repositorios les permite conocer el funcionamiento de sus instituciones y Comunidades Destinatarias. Las directrices de preservación constituyen un corpus teórico que orienta todo el funcionamiento de los repositorios, además de respaldar y delimitar sus responsabilidades y protegerlos de las situaciones legales y jurídicas que puedan surgir. Por lo tanto, recomendamos que los repositorios aborden estas cuestiones con urgencia para remediar esta debilidad que podría poner en peligro sus actividades.

La hipótesis H3 "Los repositorios institucionales brasileños no aplican herramientas de evaluación de confiabilidad" se confirma de acuerdo con lo dicho en la sección 5.4.3. Muestra que no existe una herramienta de evaluación vinculada a las prácticas, normas y auditorías internacionales, pero dirigida a la situación brasileña, lo que dificulta la aplicación de herramientas de esta naturaleza en los repositorios brasileños. También hay evidencias de un bajo grado de concienciación sobre las evaluaciones de confiabilidad y las certificaciones entre los equipos de los repositorios, lo que puede deberse a razones como la falta de conocimiento o de planificación de estudios en el área, o incluso a restricciones de personal y/o presupuestarias. Al conseguir certificaciones o aplicar métodos de auditoría, los repositorios contribuyen a establecer un sentimiento de confianza entre sus comunidades, ya que se considera que cumplen los requisitos mínimos necesarios para la preservación digital a largo plazo. La conclusión es que los repositorios dan poca prioridad a los conocimientos y a la obtención de certificaciones y auditorías en su conjunto, lo que indica una escasa base teórica sobre los criterios que evalúan la confiabilidad de la preservación digital. Se recomienda estudiar el ámbito de las certificaciones, los modelos de madurez y las auditorías para que la confianza pueda ser construida, establecida y realizada por sus comunidades de usuarios.

La hipótesis H4, "el desarrollo de un conjunto de criterios mínimos de fiabilidad, adaptados de las normas internacionales, es útil para evaluar los repositorios brasileños y formular planes de mejora", se confirma ya que la aplicación de la

herramienta de evaluación desarrollada, sus resultados y análisis se presentan en el capítulo 5 y en la sección 5.6.

Se confirma la hipótesis H5, "la calidad de los repositorios está influenciada por factores externos relacionados con su entorno", de acuerdo con los análisis realizados en la sección 5.5. Se ha podido comprobar la influencia directa que determinadas variables externas a los repositorios tienen sobre el rendimiento de la puntuación identificada. Estos factores permiten identificar una fuerte desigualdad entre los repositorios cuando se trata de temas ajenos a sus competencias. En el caso de los repositorios públicos, esta desigualdad podría reducirse estableciendo políticas públicas que, de alguna manera, prioricen a los repositorios con peor desempeño, con el objetivo de mejorar la preservación en su conjunto a nivel nacional.

La hipótesis H6, "los repositorios más antiguos deberían tener mayor calidad que los repositorios más recientes," queda refutada, según los análisis presentados en la sección 5.5.4. Los resultados mostraron repositorios creados más recientemente con puntuaciones globales mucho más altas que los más antiguos que, en teoría, deberían estar mejor organizados en términos de estructuras, directrices, etc. Por lo tanto, no ha sido posible establecer una relación directa entre la antigüedad de los repositorios y su puntuación.

6.3 Conclusiones y recomendaciones

Desde un enfoque más general, se buscó identificar la existencia de documentos que orienten la estrategia y la responsabilidad de los repositorios, tales como políticas y planes de preservación, políticas de acceso y planes de continuidad. A continuación, tratamos de identificar aspectos relacionados con la idoneidad de las estructuras organizativas, físicas y técnicas/tecnológicas para una preservación fiable a largo plazo. Desde un punto de vista más específico, tratamos de identificar cómo se relacionan los repositorios con las certificaciones de confianza y qué mecanismos y procesos se adoptan para garantizar que los materiales preservados sean fiables y no se corrompan o adulteren con el paso del tiempo. Por último, como forma de analizar la

forma en que los repositorios se promocionan a sí mismos y el uso de sus materiales almacenados, se verificaron aspectos relacionados con la visibilidad de los repositorios para sus Comunidades Destinatarias.

Una visión global de los resultados, presentada en la sección 5.3, mostró que incluso el repositorio con la puntuación más alta registrada, 66,86 puntos, tiene una puntuación relativamente baja en una escala con un máximo de 100 puntos, asimismo la media general calculada fue de 42,71 puntos, valor que se encuentra por debajo del 50% de la puntuación máxima posible. Por otra parte, el repositorio con la puntuación más baja ha obtenido 17,90 puntos, muy lejos de la media global y a una distancia muy preocupante del primer puesto. Por lo tanto, se puede concluir que, en general, los repositorios deberían invertir en iniciativas y proyectos destinados a mejorar las cuestiones relacionadas con la fiabilidad de la preservación digital. Otra conclusión es que existe un abismo de diferencias entre las realidades de los repositorios, lo que indica que la propia Red Cariniana podría desarrollar iniciativas que van desde la sensibilización hasta la promoción de medidas de financiación para la mejora de los repositorios.

La mayoría de los repositorios han declarado tener una Misión, Visión y Valores, tal y como se presenta en la sección 5.4.1, lo que puede considerarse un aspecto positivo, ya que esta información orienta el establecimiento de las políticas de los repositorios y su funcionamiento, además de transmitir a sus Comunidades Designadas los valores en los que se basan. Sin embargo, en el caso de los repositorios que no tienen declarada su misión, esta ausencia puede ser un impedimento para la realización de auditorías y certificaciones, ya que no se identifican sus fines principales y, por tanto, la evaluación de los criterios puede verse comprometida. Por ello, se recomienda que todos los repositorios elaboren y aprueben sus misiones, visiones y valores y los hagan públicos, permitiendo que sus comunidades conozcan los propósitos que persiguen.

Del mismo modo, la mayoría de los repositorios, 15 en total, declararon haber delimitado sus Comunidades Designadas, tal y como se presentó en la sección 5.4.1, lo que favorece un acercamiento dirigido y específico a las necesidades de sus usuarios, buscando satisfacer sus requerimientos. Se puede concluir que los repositorios tienen una idea de quiénes son sus usuarios y cuáles son sus necesidades, aunque falte una materialización y una cierta oficialización de esta definición, a partir de la cual se

pueden emprender otras acciones, como el seguimiento de las necesidades de la comunidad. Cabe señalar que los cuatro repositorios restantes deberían ocuparse de esta cuestión con cierto grado de importancia y urgencia, tratando de entender para quién preservan y cuáles son sus necesidades.

Un factor preocupante, identificado y presentado en la sección 5.4.2, fue la ausencia de Políticas de Preservación o Planes Estratégicos de Preservación en la gran mayoría de los repositorios. Ambos documentos pueden considerarse esenciales para el establecimiento de un repositorio, ya que a través de ellos se conoce su finalidad y se establece un punto de partida a partir del cual se emprenden todas las acciones, procesos y actividades encaminadas a realizar la preservación fiable del repositorio. La existencia de estos documentos puede considerarse obligatoria en los repositorios, saliendo a la luz el hecho alarmante de que los 17 repositorios que respondieron a las preguntas relacionadas con estos temas informaron no disponer de estos documentos. Se puede concluir que los repositorios participantes realizan la preservación de sus materiales sin una planificación específica, ya que carecen de políticas que orienten, dirijan y establezcan el alcance de la propia preservación.

La falta de un Plan de Continuidad en los repositorios, presentada en la sección 5.4.2, fue un hecho alarmante, identificado a partir de las respuestas de los participantes, en que apenas 1 repositorio, de 19 participantes, informó tener este documento. Como resultado de este escenario, no existen garantías de que los materiales depositados en los repositorios sean efectivamente preservados y accesibles en caso de eventos que imposibiliten su funcionamiento. Además, de acuerdo con la Ley nº 8.159, de 8 de enero de 1991, el Archivo Nacional es el órgano responsable en última instancia de la preservación y el acceso a los documentos del Poder Ejecutivo Federal (BRASIL, 1991), por lo que es necesario contar con un acuerdo y elaborar un plan para que los repositorios institucionales federales puedan transferir sus materiales al Archivo Nacional en caso de eventos relacionados con fallos.

Dado que el Archivo Nacional es el responsable último de la conservación y el acceso a sus propios documentos y a los de otras instituciones, la propia organización debe contar con un plan de continuidad elaborado y sólido que garantice que los materiales de investigación y los documentos de archivo e históricos se conservan

correctamente y siguen siendo accesibles, independientemente de los acontecimientos que puedan producirse. Cabe recordar que la discontinuidad de un repositorio puede ocurrir debido a eventos catastróficos naturales, dificultades financieras o incluso un cambio de alcance, propósito y cese de intereses, lo que hace que los repositorios públicos sean vulnerables a los intereses de los gobiernos de turno. Se puede concluir que los repositorios actualmente no comprenden la necesidad, la importancia y la urgencia de establecer un Plan de Continuidad y que sus gestores ven un horizonte completo para continuar sus operaciones, sin contar con que eventos imprevistos, catástrofes o cambios de políticas les afecten. Por ello, se recomienda que los repositorios aborden esta cuestión con la mayor urgencia posible.

Analizando los resultados del apartado 5.4.3 sobre la presencia e interés de los repositorios en las certificaciones de confianza, se observa un bajo porcentaje alcanzado, lo que puede deberse a diversas razones, como la falta de conocimiento o planificación de estudios en la materia o, incluso, restricciones de personal y/o presupuestarias. Al conseguir certificaciones o aplicar métodos de auditoría, los repositorios contribuyen a establecer un sentimiento de confianza con sus comunidades, ya que se considera que cumplen los requisitos mínimos necesarios para la preservación digital a largo plazo. La conclusión es que los repositorios dan baja prioridad al conocimiento y a la obtención de certificaciones, lo que indica una escasa base teórica sobre los criterios que evalúan la fiabilidad de la preservación digital. Se recomienda estudiar el ámbito de las certificaciones, los modelos de madurez y las auditorías para construir, establecer y hacer realidad la confianza ante sus comunidades de usuarios.

En general, teniendo en cuenta las prácticas existentes, las estructuras organizativas, físicas, técnicas y tecnológicas de los repositorios encuestados no son adecuadas para la preservación digital. Sin embargo, si las instituciones que han aplicado parcialmente determinados aspectos prácticos o aplicados solamente en teoría, los aplicaran plenamente, el escenario cambiaría, alcanzando la mayoría y, por tanto, estas estructuras serían adecuadas.

De acuerdo con los resultados de la sección 5.4.4, se identificaron graves problemas en cuanto a la falta de equipos formados que trabajan en los repositorios, así como la falta de recursos financieros que permitan el funcionamiento y las inversiones

en los repositorios. Con respecto a los presupuestos, de los 17 repositorios que dieron respuestas, sólo 2 informaron tener presupuestos adecuados que se aplican en la práctica, mientras que 6 informaron que el presupuesto existe sólo en teoría o no se aplica en su totalidad, mientras que en 9 repositorios, la mayoría, se informó que el presupuesto no es adecuado.

Dado que la mayoría de los repositorios pertenecen a instituciones públicas federales, dependen de los fondos de la administración federal, que los asigna en función de sus intereses y prioridades. Internamente, estas instituciones asignan los porcentajes adecuados de recursos para mantener sus actividades y realizar inversiones, y existe una baja priorización en la asignación de fondos adecuados a las actividades de preservación digital. Esta baja priorización, precariedad y escasez de inversiones se remonta a décadas atrás y ha sido señalada por Maher (1978), Beagrie (2003), Lavoie y Dempsey (2004) y Matlala, Ncube y Parbanath (2022). En este último trabajo se lleva a cabo una revisión del estado de la preservación digital en Sudáfrica, observando situaciones similares a las encontradas en el presente estudio (2022, p. 206):

La conservación de documentos digitales en el sector público sudafricano se ve obstaculizada principalmente por la falta o escasez de las herramientas y competencias necesarias. Estas herramientas incluyen políticas, directrices y normas apropiadas, infraestructura, apoyo de la alta dirección, un presupuesto específico y equipos adecuados, así como una mano de obra cualificada. La ausencia de políticas, directrices y normas afecta a la eficacia de la conservación de los documentos digitales. Además, el estudio reveló que la preservación de documentos digitales en Sudáfrica sigue enfrentándose a la escasez de equipos de almacenamiento; funciona en malas condiciones ambientales; tiene malas prácticas de manipulación de archivos tanto por parte del personal como de los usuarios; las instalaciones de almacenamiento son inadecuadas; trabaja con instalaciones poco desarrolladas, recibe una financiación inadecuada para la preservación y carece de personal formado y con experiencia.

La falta de presupuestos adecuados para los repositorios repercute directamente en la forma en que se lleva a cabo la preservación digital, limitando las inversiones que van desde la adquisición y el mantenimiento de equipos hasta la contratación de personal cualificado y culminando en la escasez de fondos para inversiones en formación, por ejemplo. La escasa prioridad concedida a la preservación digital en los repositorios puede ser un reflejo de la falta de inmediatez del rendimiento percibido de los fondos invertidos, ya que las inversiones realizadas construyen una base

para el futuro, cuando muchos gestores y responsables gubernamentales ya no estén en sus puestos de trabajo.

Es evidente la importancia y urgencia con la que los gestores de las instituciones de los repositorios deben abordar la preservación digital, para favorecer la creación y establecimiento de estructuras sólidas que permitan la conservación confiable y a largo plazo de los materiales para su uso futuro. Además, corresponde a los gestores y equipos de los repositorios concienciar y sensibilizar a los responsables de sus instituciones, a la comunidad y a la sociedad en su conjunto sobre la importancia de las tareas que llevan a cabo, destacando tanto los beneficios de una preservación digital correctamente realizada como los daños, impactos y sanciones que pueden producirse si la preservación no se lleva a cabo en su totalidad.

La conclusión es que, a pesar de las carencias presupuestarias, de estructuras técnicas y organizativas y de personal, los equipos de los repositorios llevan a cabo la preservación utilizando los recursos posibles y disponibles, adaptándose a las necesidades y exigencias que la preservación digital requiere, acumulando responsabilidades y apoyándose en el esfuerzo individual, señalando las áreas a las que se debe dar una alta prioridad en los repositorios y sus instituciones.

El análisis de las respuestas sobre las estructuras técnicas y tecnológicas de los repositorios arrojó resultados satisfactorios, en el sentido de que existen equipos y herramientas de preservación instalados en todos los repositorios, como se muestra en el apartado 5.4.5. Esto revela un alto grado de importancia de los aspectos tecnológicos implicados en la preservación digital, en detrimento de los demás aspectos existentes. Los resultados también nos permiten concluir que la existencia de una herramienta de preservación que se adhiere al modelo de referencia OAIS, basado en paquetes de información, lleva a pensar que el repositorio se convierte en un Archivo OAIS o que se establece un RDC-Arq. Si observamos la visión general de la norma ISO 14721 y de la Resolución 43, que define el RDC-Arq, vemos que ambas abarcan muchos más aspectos que los puramente tecnológicos, incluidos los requisitos organizativos y las políticas de preservación, por ejemplo. De ahí que sea errónea la afirmación de que los repositorios creen que se adhieren al OAIS o que son RDC-Arq.

Asimismo, la falta de planes de actualización de hardware y software fue señalada por los 14 repositorios que respondieron a la pregunta. Es esencial que haya un seguimiento para controlar la obsolescencia de la tecnología involucrada en la preservación digital, observando desde la vida útil de los medios de almacenamiento hasta las versiones de software utilizadas, por ejemplo, evitando la pérdida o corrupción de datos. Se concluye así que los materiales preservados en la mayoría de los repositorios corren un grave riesgo de volverse inútiles en algún momento en el futuro, ya que los equipos, los programas y los formatos de archivo pueden volverse obsoletos.

A partir de los análisis, se identificó que todas las soluciones de preservación, desde el nivel gerencial hasta el operativo, son realizadas por equipos y sectores dentro de las instituciones a las que pertenecen los repositorios, sin que se externalice la preservación de los materiales digitales. Esta situación puede deberse a diferentes factores, como el desconocimiento de la externalización de estos servicios o posibles restricciones presupuestarias, basadas en la creencia de que las instituciones no dispondrían de fondos para respaldar los contratos firmados. Dado que los informes señalan carencias en las estructuras organizativas, equipos y presupuestos de los repositorios, se sugiere, como alternativa, la realización de estudios de viabilidad sobre la contratación de servicios y empresas para la prestación de servicios de preservación digital, basándose en la norma ISO 17068 (INTERNATIONAL ORGANISATION FOR STANDARDIZATION, 2012), que trata sobre los Repositorios de Terceros de Confianza para Documentos Digitales. Al optar en esta dirección, los repositorios actuarían estableciendo sus políticas y acuerdos de nivel de servicio, por ejemplo, para ser utilizados en la supervisión de los servicios contratados, actuando más en las demandas de gestión y menos en las demandas operativas.

En lo que respecta a los metadatos, aspecto analizado en la sección 5.4.8, la mayoría de los repositorios declararon que disponían de una norma establecida y también declararon que existían metadatos que respaldaban la presunción de autenticidad de los materiales digitales conservados, ya fueran recogidos total o parcialmente en la práctica. Esta situación se ve corroborada por el hecho de que el área que evaluó los criterios relativos a los metadatos fue la que obtuvo el rendimiento medio más alto de todas las áreas. El análisis de la calidad de esta información recopilada para

demostrar la autenticidad requiere la aplicación de una metodología específica y orientada, que, según concluimos, exige un estudio exhaustivo y detallado, aunque escapa al alcance de este trabajo.

En lo que respecta a la salvaguarda de los materiales conservados, cabe destacar que la mayoría de los repositorios disponen de planes de copias de seguridad, que incluyen diferentes tecnologías entre las copias y diferentes ubicaciones geográficas, como se muestra en la sección 5.4.7. Puede concluirse que, en general, las soluciones de copia de seguridad de los repositorios están bien estructuradas y presentan cierto grado de resistencia a fallos y catástrofes.

Se puede observar que la preservación de los repositorios está fuertemente estructurada en torno al uso de tecnologías con herramientas de preservación y estrategias de copia de seguridad, aunque estas situaciones no se incluyan en los planes de preservación ni en los planes de copia de seguridad, por ejemplo. Se puede concluir que este estrecho enfoque en la tecnología se deriva del hecho de que la preservación digital sólo se asocia al aspecto de garantizar que los datos no se corrompan, en detrimento de los demás aspectos implicados. Se recomienda que los repositorios lleven a cabo estudios más amplios sobre la preservación digital, ya sea mediante proyectos internos o a través de colaboraciones entre instituciones, para que haya una comprensión más amplia de las características y responsabilidades que implica la preservación digital, con el fin de aprovechar y orientar correctamente las soluciones tecnológicas que ya se han desarrollado.

Dado que *DSpace* es la herramienta de preservación utilizada por la gran mayoría de los repositorios y que ya incorpora funcionalidades para acceder a los materiales y consultarlos, destaca positivamente el alto grado de disponibilidad de las colecciones en su conjunto. Sin embargo, los repositorios carecen de capacidad para establecer normas sobre qué materiales pueden ser consultados y en qué condiciones específicas, y una minoría de ellos supervisa las necesidades de sus comunidades designadas. La conclusión es que los repositorios difunden y dan acceso a sus materiales sin basarse en normas plenamente establecidas, corriendo el riesgo de infringir derechos o de poner a disposición formatos de archivo inadecuados. Los repositorios deberían hacer hincapié en sus esfuerzos por establecer políticas de acceso, incluyendo planes de

seguimiento de las Comunidades Destinatarias, tratando de comprender sus necesidades y deseos y reaccionando ante los cambios identificados.

En cuanto a la visibilidad de los repositorios, se encontró que, en casi la mitad de los casos, las páginas de inicio de sus instituciones disponen de formas de dirigir a los usuarios a los repositorios, como se muestra en la sección 5.4.10. Se cree que la situación de los repositorios que no tienen este acceso se puede resolver con un poco de esfuerzo por parte de los equipos de los repositorios y los equipos responsables de las páginas institucionales; Igualmente, esta mejora podría facilitar y aumentar el uso de los repositorios. En general, todos los repositorios ofrecen vías de contacto con la comunidad, facilitando la comunicación entre ésta y los equipos responsables. Se puede concluir que los repositorios disponen de herramientas de relación con las comunidades, y, además, que son necesarios algunos ajustes para que las páginas institucionales dispongan de formas de dirigir a los usuarios a los propios repositorios.

Finalmente, este estudio permitió identificar las herramientas, instrumentos, métodos y estándares existentes para la evaluación de repositorios digitales, a partir de los cuales se desarrolló una herramienta con amplia cobertura de los criterios de estas referencias existentes. Cuando se analiza un repositorio y se identifica el cumplimiento de un criterio, es posible inferir al mismo tiempo el cumplimiento de varios criterios de las herramientas utilizadas como referencia, favoreciendo y facilitando la aplicación de la herramienta desarrollada. De esta forma, los repositorios pueden trazar un mapa de sus puntos fuertes y de las debilidades que deben ser tratadas, sirviendo de base para la toma de decisiones por parte de los gestores de repositorios.

La aplicación de la herramienta en los repositorios brasileños posibilitó su propia verificación y, al mismo tiempo, la identificación del estado actual de la preservación en estos repositorios. A partir de los datos recogidos, se realizaron análisis y fue posible concluir que son muchos los desafíos que estos repositorios deben superar para alcanzar la excelencia en la preservación digital, que van desde la reformulación de sus estructuras organizativas, la priorización de las inversiones, la formación de los equipos y el seguimiento de las necesidades de las comunidades destinatarias. Sin embargo, se sabe que el camino hacia estos cambios pasa primero por sensibilizar a los

gestores para que se comprendan las prioridades, que es donde este estudio puede servir de apoyo.

6.4 Líneas futuras de investigación

Entre las posibilidades identificadas para futuros trabajos figura la diferenciación de ponderaciones y valores entre áreas y cuestiones según su grado de importancia para la preservación digital. Para llevar a cabo este trabajo, sería necesario identificar qué áreas y cuestiones son más críticas y esenciales para la preservación digital y cuáles podrían considerarse mejoras o perfeccionamientos del proceso, utilizando tanto las normas de certificación y auditoría disponibles como las opiniones de investigadores en la materia y realizando pruebas piloto.

Junto con la coordinación de la Red Cariniana, también es posible elaborar un calendario de evaluación de los repositorios miembros, realizado en el momento de la adhesión de los repositorios y, posteriormente, de forma periódica, en un intento de comprender cómo afecta la participación en la Red a la evolución de su preservación digital. De este modo, es posible diagnosticar áreas de deficiencia comunes a los repositorios que requerirían acciones de mejora o incluso establecer un nivel mínimo exigido para formar parte y permanecer en la Red.

La automatización de la recogida de información de los repositorios, en función de los criterios evaluados, podría realizarse mediante el uso de API gestionadas por la coordinación de la Red Cariniana, que tendría acceso a los datos de forma automática. En el caso de determinadas áreas y criterios, la API podría recoger información de las herramientas de preservación, enviándola a la coordinación y facilitando la tarea de dar respuesta al formulario de evaluación. En este caso, es necesario analizar la viabilidad técnica y las intenciones de la Red y de los repositorios participantes para aplicar esta solución.

Dado que la herramienta desarrollada permitió llevar a cabo la evaluación específica de los repositorios de la Red Cariniana, su aplicación podría extenderse a otros grupos de repositorios a nivel nacional o internacional, ya que la evaluación utiliza

criterios basados en normas internacionales de certificación y auditoría. Esta extensión contribuiría al avance de la preservación digital a nivel mundial, identificando los puntos fuertes existentes y las deficiencias a corregir, además de permitir la evaluación comparativa con otros países. La viabilidad de este estudio puede requerir la coordinación de un organismo central que lo normalice y se encargue de recoger y analizar los datos, sugiriendo o exigiendo los cambios identificados en los repositorios, de acuerdo con sus respectivos ámbitos administrativos. Esto requeriría la creación de futuras asociaciones para ampliar este trabajo.

