



**VNiVERSIDAD
D SALAMANCA**

CAMPUS DE EXCELENCIA INTERNACIONAL

FACULTAD DE TRADUCCIÓN Y DOCUMENTACIÓN
MÁSTER OFICIAL EN SISTEMAS DE INFORMACIÓN DIGITAL

Trabajo de Fin de Máster

**PROPUESTA DE DISEÑO DEL PROCESO DE GESTIÓN DE INFORMACIÓN
DEL ENTORNO LOCAL E INTERNACIONAL DE LA CONSULTORA
URBANÍSTICA ARNAIZ AND PARTNERS**

Francisco Javier Ramos Núñez
AUTOR

Dr^a. Yolanda Martín González
TUTORA

España, 2021



**UNIVERSIDAD
DE SALAMANCA**
CAMPUS DE EXCELENCIA INTERNACIONAL

UNIVERSIDAD DE SALAMANCA
FACULTAD DE TRADUCCIÓN Y DOCUMENTACIÓN
MÁSTER EN SISTEMAS DE INFORMACIÓN DIGITAL
TRABAJO FIN DE MÁSTER

Propuesta de diseño del proceso de gestión de información del entorno local e internacional de la consultora urbanística Arnaiz and Partners

Autor: Francisco Javier Ramos Núñez

Tutora: Dr^a. Yolanda Martín González

Salamanca, 2021

FICHA CATALOGRÁFICA

Título:

- **PROPUESTA DE DISEÑO DEL PROCESO DE GESTIÓN DE INFORMACIÓN DEL ENTORNO LOCAL E INTERNACIONAL DE LA CONSULTORA URBANÍSTICA ARNAIZ AND PARTNERS.**

Autor:

- **RAMOS NÚÑEZ, FRANCISCO JAVIER**

Directora:

- **MARTÍN GONZÁLEZ, YOLANDA**

Palabras Clave:

- **[ES] GESTIÓN DE INFORMACIÓN – MINERÍA DE TEXTO – ENTORNO EMPRESARIAL – BIG DATA – INTELIGENCIA DE NEGOCIOS – DATOS NO ESTRUCTURADOS – TECNOLOGÍAS DE LA INFORMACIÓN - DOCUMENTACIÓN**
- **[EN] INFORMATION MANAGEMENT – TEXT MINING - BUSINESS ENVIRONMENT – BIG DATA – BUSINESS INTELLIGENCE – UNSTRUCTURE DATA – INFORMATION TECHNOLOGY - DOCUMENTATION**

Fecha:

- **2021 – 03 - 12**

Descripción:

- **TRABAJO DE FIN DE MÁSTER EN SISTEMAS DE INFORMACIÓN DIGITAL, CURSO 2019 - 2020**

Resumen | Abstract

- [ES] Adaptar efectivamente los servicios empresariales a las demandas del mercado local e internacional, comprende una actividad elemental de inteligencia de negocios para la competitividad. Sin embargo, los rápidos cambios socioeconómicos y el ingente volumen de datos, conocido como *big data*, dificulta el desarrollo del conocimiento. Este trabajo proporciona un diagrama de gestión de información que involucra criterios metodológicos de las Ciencias de la Documentación y las Tecnologías de la Información para el monitoreo, selección, clasificación, análisis y disposición de la información de texto que produce el entorno de interés de la consultora urbanística Arnaiz & Partners.
- [EN] Effectively adapting the enterprise services to local and international market demands is an elementary business intelligence activity for competitiveness. However, the rapid socioeconomic changes and the huge volume of data, known as big data, make difficult the development of knowledge. This investigation provides a workflow of information management that involves methodological criteria of Documentation Sciences and Information Technology for the monitoring, selection, classification, analysis and availability of the text information produced by the environment of interest of the urban consulting firm Arnaiz & Partners.

Agradecimientos y dedicatorias

A Dios, mi fuente de sabiduría y optimismo.

A Luz Acosta, mi amor e inspiración.

A mis padres, M^a Yluminada y Juan Francisco,
por la incondicional confianza y apoyo.

A mis atentas hermanas: Elena, Nathalie.

A Da. Natividad Frías y Da. Mercedes Ramírez,
por su hospitalidad durante nuestra estadía en Europa.

A la profesora Yolanda Martín por su paciencia
entendiendo y encarrilando mis ideas.

A mis profesores y compañeras/os de la Universidad de Salamanca,
un espacio multicultural y pluridisciplinar de aprendizaje.

Glosario de Términos

Arnaiz & Partners: consultora de desarrollo territorial e inmobiliario con sede en España y múltiples delegaciones en Europa del Este, África y Latinoamérica.

Big data: conjunto de datos cuyos métodos de captura, almacenamiento, administración y análisis exceden las habilidades tradicionales, requiriendo para estos propósitos el uso de novedosos software (Manyika, 2011; cit. en Alcaíno, Arenas, & Gutiérrez, 2015).

Base de datos: colección de información organizada mediante un sistema informático para el fácil acceso y manejo.

Ciencia de datos: (*Data Science*) conjunto de conocimientos y prácticas para mejorar la toma de decisiones y la planificación empresarial mediante la selección, análisis y uso de la información en cualquier formato existente, proveyendo un mejor conocimiento de estas y de sus usuarios/clientes.

Clientes: se refiere a toda persona o institución beneficiada por el uso o recepción de un producto y servicio ofrecido por otra persona o institución.

Conocimiento: tiene como fuente la información misma, por lo que es un producto posterior y surge a partir de ella. El conocedor es capaz de descomponer, reorganizar e interpretar los datos que formaron la información influenciado por su experiencia, cultura y más información (Rojas & Ángel, 2005).

Contenido: se refiere a la información disponible en la web en cualquier tipo de formato multimedia.

Dato: entiéndase en el texto como una información completa que no ha sido sometida a un análisis o explotación acorde con los propósitos de una organización.

Entidad: hace referencia a instituciones públicas o privadas, organizaciones internacionales o sin fines de lucro, académicas o documentales.

Entorno de interés: (entorno empresarial) se refiere a toda entidad o factor que afecte los intereses de la entidad. Tales como medios de comunicación, instituciones públicas y competencia; así como los aspectos económicos, políticos, sociales, jurídicos, internacionales, entre otros.

Gestión de información: es el proceso que eficiencia y controla la creación, recepción, mantenimiento, uso y disposición de los archivos (ISO 15489, 2016). Los datos son desechados o utilizados para apoyar unos objetivos determinados (Ponjuan Dante, 2003). En ocasiones se utiliza gestión de datos o gestión de contenidos para aludir la misma idea.

Información: aplíquese también a documento o archivo. Es el conjunto de datos organizados de forma comprensible para comunicar un mensaje que incluye desde el contenido presentado en los más diversos formatos (texto, sonoro, audiovisual, etc.) hasta el conocimiento personal de los miembros de una organización (Prytherch, 2000).

Información no estructurada: son contenidos conformados de manera cambiante y de heterogénea procedencia, dificultando su organización en bases de datos. Aunque investigación hace énfasis en el texto, los archivos sonoros, imágenes y vídeos también son considerados como información no estructurada.

Inteligencia de negocios: (inteligencia competitiva, inteligencia económica, inteligencia de mercado, *business Intelligence*) son sistemas para mejorar la toma de decisiones y la competencia de la empresas ante su entorno comercial mediante el análisis de cualquier combinación de información y datos procedente de medios digitales y bases de datos.

Minería de texto: (*text analysis, text mining*) es una técnica de la ciencia de datos que haciendo uso de las tecnologías facilita la explotación de grandes volúmenes de información de textos, facilitando su búsqueda, selección, clasificación, organización, almacenamiento y uso.

Reportes de prensa: (*clipping press*) práctica tradicional de relaciones públicas. Reporta sobre la presencia de una organización, evento o personalidad en los medios de comunicación. La actividad puede ser continua o momentánea.

Producto: alude a cualquier bien material o servicio que una entidad ofrezca a sus clientes mediante su solicitud o transacción.

Texto: Conjunto de ideas presentadas en secuencia de manera escrita, sea digital o impresa.

Sumario

Glosario de Términos	7
1. Introducción	11
1.1. Justificación del tema	13
1.2 Problema de investigación	14
1.3 Objetivos	16
1.4 Proceso de investigación.....	16
1.5 Estructura de trabajo	17
1.6 Fuentes bibliográficas.....	18
2. Estado de la cuestión.....	19
2.1 Gestión de la Información.....	19
2.2. Gestión de la información en empresas	20
2.3. Gestión de información estratégica del entorno empresarial	23
2.4. Gestión de información de texto del entorno empresarial	25
2.5. Perfiles para la gestión de información del entorno	28
2.5.1. Perspectiva comunicacional para la gestión de información	30
2.5.2. Perspectiva informática para la gestión de información.....	31
2.5.3. Perspectiva documentalista para la gestión de información	32
2.6. Gestión de información en Arnaiz & Partners	35
3. Metodología	38
4. Resultados y discusión	42
5. Propuesta de diseño de un servicio de gestión de información.....	47
5.1. Objetivos para la gestión de información en Arnaiz & Partners	47
5.2. Selección de fuentes de información para Arnaiz & Partners	48
5.3. Diagrama de gestión de información digital	50
5.3.1. Descripción ampliada del diagrama de flujo	51
5.4. Personal responsable de la gestión de información en Arnaiz & Partners	55
6. Conclusiones	56
Bibliografía.....	58
Anexos	62
Anexo I: Guion de entrevista a informantes claves	63
Anexo II: Tabulación de respuestas en la entrevista y el cuestionario	64

Índice de Tablas

Tabla 1: Características de gestión estratégica de información según disciplinas académicas	34
Tabla 2: Sistema de categorías y subcategorías para el análisis de las entrevistas....	40
Tabla 3: Definiciones de subcategorías para el análisis de las entrevistas	41
Tabla 4: Criterios recomendados para seleccionar las fuentes de información	43
Tabla 6: Propuesta de criterios para la selección de fuentes externas	49
Tabla 5: Estructuración de datos de texto en bases de datos relacionales	52
Tabla 8: transcripción de respuestas de entrevista 1 - 30/07/2020.....	66
Tabla 9: transcripción de respuestas de entrevista 2 - 28/08/2020.....	69
Tabla 10: transcripción de respuestas de entrevista 3 - 28/07/2020.....	70
Tabla 11: transcripción de respuestas de entrevista 4 - 27/07/2020.....	72
Tabla 12: transcripción de respuestas de entrevista 5 - 29/07/2020.....	73
Tabla 13: transcripción de respuestas de entrevista 6 - 13/08/2020.....	74
Tabla 14: transcripción de entrevista a 7 - 28/07/2020.....	75
Tabla 15: transcripción de entrevista 8 - 27/07/2020.....	77
Tabla 16: transcripción de entrevista 9 - 30/07/2020.....	80
Tabla 17: transcripción de entrevista 10 - 09/09/2020.....	82
Tabla 18: transcripción de entrevista 11 - 30/08/2020.....	83
Tabla 19: transcripción de entrevista a 12- 31/08/2020.....	84
Tabla 20: transcripción de entrevista 13- 31/08/2020.....	85
Tabla 21: transcripción de entrevista14 - 30/08/2020.....	87
Tabla 22: transcripción de entrevista a entrevistado 15 - 02/09/2020.....	88

Índice de Ilustraciones

Ilustración 1: Potencial equipo de gestión de información del entorno	55
---	----

1. Introducción

“Infoxicación”, “sobrecarga informativa” o “infobesidad” son los términos más conocidos para referirse al estresante exceso de datos al que millones de personas son sometidas diariamente. Ante esta realidad las personas continuamente seleccionan las informaciones que les son necesarias para el desarrollo de sus actividades, valoran las fuentes más confiables y fundamentados en este conocimiento toman decisiones.

La abismal cantidad de datos producidos por el desarrollo de las tecnologías de la información también afecta a las organizaciones ya que los contenidos de interés para conocer las demandas de sus públicos se confunden en el inmenso mar de información web. También en el seno de las instituciones el volumen de información que se produce en un espacio corto tiempo es tanto que, de no existir métodos definidos para la administración de los datos que entran y salen, pudieran quedar ahogadas en sus propios contenidos.

La información puede ser perceptible como números, imágenes, vídeos, audios y textos, siendo esta última la que representa mayor interés en esta investigación. El texto es una de las modalidades más abundantes para describir la realidad, está en los documentos que conservan las bibliotecas, en los archivos médicos o los informes policiales, constituyéndose en una de las formas de la información más retadoras para ser analizadas.

El texto también se prolifera en la web como información digital en forma de artículos, *microblogs*, mensajería y comentarios. Las empresas también generan informes, memorias, correos, entre otros tipos de documentos.

La dinámica estructura de la información textual, su heterogénea procedencia y motivaciones de creación hacen que la gestión de los datos de texto sea compleja, clasificándose al momento de su análisis como un tipo de “dato no estructurado” (Feldman & Sanger, 2007). Que el texto no estructurado sea tan voluminoso, variado y veloz, le otorga las características necesarias para ser considerado como uno de los distintos tipos de información que componen la *big data*, definida como un conjunto de datos cuyos métodos de captura, almacenamiento, administración y análisis exceden las habilidades tradicionales, requiriendo para estos propósitos el uso de novedosos software (Manyika, 2011; cit. en Alcaíno, Arenas, & Gutiérrez, 2015).

Richard Benjamins, embajador de *big data* de Telefónica, señala la estructuración de los datos como la clave para su uso y lo contrapone con el inútil “lago de datos” cuando son totalmente disgregados siendo precisamente el procesamiento de los datos lo que genera mayor coste (Nogueira, 2019).

Los gestores de estos grandes datos (*big data*) son conocidos como científicos de datos, quienes se especializan para manejar la información según su forma (texto, sonoro, imagen, audiovisual) y la actividad profesional o institucional a la que está vinculada. Por tanto, la presente investigación también fue delimitada conforme al objetivo de proponer un diseño de gestión de las informaciones textuales producidas en el entorno externo nacional e internacional, de habla castellana, de la consultora urbanística Arnaiz & Partners.

Por lo general, la actividad de Arnaiz & Partners no proporciona servicios a personas, su público objetivo son instituciones públicas y privadas vinculadas a la gestión territorial y al desarrollo inmobiliario. Posee delegaciones en nueve países de Europa, América y África, con activa participación en múltiples instituciones internacionales como la Federación Latinoamericana de Ciudades, Municipios y Asociaciones Municipalistas (FLACMA), la Asociación de Municipios Mexicanos y Ciudades Inteligentes (AMMECI), entre otras.

El amplio alcance de la consultora Arnaiz & Partners hace que sean cientos las fuentes de información digital que sería necesario monitorear para atender a los constantes cambios en su entorno, conformado por gobiernos y actores políticos, la modificación de legislaciones; las innovaciones y tendencias del mercado privado y los competidores. Para desarrollar este trabajo de inteligencia de negocios proporcionando conocimiento y fortaleciendo la oferta de servicios de la consultora, deben obtenerse datos de los portales web gubernamentales, de empresas, medios de comunicación y otros canales informativos de interés, siendo los textos el más abundante contenido que producen estos espacios, permitiendo también hacer una importante aproximación a la realidad de su mercado.

Por eso se diseñó un diagrama de flujo con detalladas recomendaciones para efficientizar la gestión de información del entorno empresarial de Arnaiz & Partners, además de describir las aptitudes profesionales que debe exhibir el personal recomendado para esta tarea. Quedan excluidas las informaciones internas tales como informes, textos administrativos y correos electrónicos, entre otros documentos empresariales.

Tener métodos fundamentados en criterios objetivos permite priorizar las fuentes de datos, identificar los temas de mayor importancia, la forma de organización de las informaciones; su presentación, almacenamiento y disposición, es decir, una gestión de información eficiente.

Adicionalmente, se sugirieron tecnologías para la gestión de información textual (*text mining* o *text analytic*) que permitirán optimizar recursos humanos y tecnológicos Arnaiz & Partners, destacando software que integran arañas de navegación (*crawler*) para la extracción de información (*scraping*), la automatización de trabajos mediante programas de códigos y aplicaciones de aprendizaje artificial para encontrar patrones similares en los textos y clasificarlos. Aunque la automatización de estos procesos podrían restar algo de rigurosidad a los resultados por ciertas debilidades informáticas que prevalecen (Kantardzic, 2011), garantizan una exponencial mejora de eficiencia, trabajando en minutos significativas cantidades de datos que de forma tradicional podrían tomar horas o días.

Los procedimientos y herramientas informáticas propuestas facilitarán actividades como el análisis de los datos, la preservación de la información y su puesta en disposición de manera permanente y digital. Todo esto fue posible gracias a una exhaustiva actividad documental, las entrevistas de investigación realizadas y productivas conversaciones sobre la gestión y el análisis datos sostenidas con múltiples académicos, pudiendo destacar a Cesar Martínez, Félix Almonte, Raúl Hernández y Helen Hasbún de la Universidad Autónoma de Santo Domingo; Carlos Figuerola y José Frías, de la Universidad de Salamanca; y a Keneth Kwayu, de la *Western Michigan University*.

Al incorporar Arnaiz & Partners el diagrama recomendado con un equipo adecuado de gestión de información, este podrá apoyar los procesos de adaptación de los servicios empresariales a nivel nacional e internacional, así como la toma de decisiones en cada una de sus delegaciones. Puede constituirse en toda una estrategia de inteligencia territorial, conociendo los proyectos, soluciones y problemáticas de su mercado desde diversas perspectivas y produciendo colaboraciones entre instituciones con necesidades similares o complementarias.

Gestionar informaciones externas empleando métodos y tecnologías de análisis de *big data* es también de interés para la gestión de la reputación y la inteligencia competitiva de las organizaciones. Y es que los datos lo demuestran: en el 2018 en España el 50,4% de los especialistas en relaciones públicas afirmó que usar *big data* y algoritmos para la comunicación constituye el mayor reto de los comunicadores (Asociación de Directivos de Comunicación, 2018), en Latinoamérica es considerado como el segundo principal tema estratégico que produce preocupación con un 32.7% (Moreno, y otros, 2019); y en toda Europa se constituye en la quinta principal prioridad estratégica con un 29.2% (Zerfass, y otros, 2020).

Los profesionales interesados en la gestión de información de texto podrán encontrar en este trabajo un ejemplo integrador de la eficiencia metodológica de las ciencias de la documentación y la optimización de recursos que proporcionan las tecnologías.

1.1. Justificación del tema

Desde el año 2009, quien suscribe este estudio ha trabajado en los aspectos técnicos y profesionales de la información, dedicando, en especial, los últimos siete años a labores de comunicación, investigación y análisis de mercado, manejando los volúmenes de información que estas tareas involucran de manera casi análoga.

Tras experimentar la complicación y la pérdida de tiempo que significa no trabajar con los métodos y las tecnologías correctas, y estando percibido de que múltiples instituciones de distintos sectores económicos tienen la misma dificultad, el autor de esta investigación se propuso encontrar una solución para este problema, dando como resultado este Trabajo de Fin de Máster (TFM).

Es precisamente en la delegación dominicana de la consultora Arnaiz & Partners, donde implementando la técnica denominada “Análisis del Entorno Mediático”, quien suscribe se convenció de que es necesario contar con tecnologías y procedimientos innovadores para el correcto estudio de la competencia, así como del entorno socioeconómico y político de la empresa en temas tan cruciales para esta empresa como es la gestión urbana y el desarrollo territorial.

Aprovechando el alcance de la consultora Arnaiz & Partners y teniendo la autorización de sus directivos, se decidió investigar la problemática y proponer soluciones a escala Iberoamericana, complementando asimismo el Trabajo de Fin de Grado de comunicación realizado en el 2017, sobre el flujo internacional de comunicación con fines estratégicos que fue presentado en la Universidad Autónoma de Santo Domingo.

Con toda seguridad esta no es una dificultad solamente propia. En el estudio: “Estado de la Comunicación en España” (Asociación de Directivos de Comunicación, 2018), las empresas privadas consideran la evaluación y medición, junto con la comunicación internacional, como las principales actividades del área en la actualidad y en el futuro.

Los sistemas de información empresarial, según su tamaño y actividad, son altamente complejos y proponer recetas para la solución de todo tipo de problemas sería muy pretencioso. Incluso el establecimiento de principios de aplicación general requiere un esfuerzo inmenso de recursos y tiempo.

El tema abordado en la presente investigación responde a una necesidad concreta, pero cuya solución permitirá un conocimiento alineador de las empresas con las demandas y el clima de su negocio mediante una correcta selección, curación, presentación y almacenamiento de la información. A un nivel más personal, su elaboración nos permitirá completar el Máster en Sistemas de Información Digital de la Universidad de Salamanca.

1.2 Problema de investigación

La planificación, gestión y distribución de los productos o servicios ofrecidos por las instituciones públicas y privadas es efectiva en el tiempo cuando están respaldadas por un adecuado conocimiento de la demanda de sus *stakeholders*. Por esta razón los métodos de gestión y difusión de la información han evolucionado a través de los siglos, permitiendo el acceso al conocimiento.

Esta rápida transformación comenzó en el siglo XIX con el desarrollo de la Archivística, luego el siglo XX trajo avances tecnológicos, la globalización y el nacimiento de internet, entregando a las personas la posibilidad de producir y consumir contenidos sin restricciones editoriales (prosumir). Este fenómeno junto al desarrollo de la web 2.0 en el 2004, produjo la avalancha de datos hoy conocida como “*big data*”, siendo su gestión un reto para los profesionales de la información.

Los datos hoy adquieren una importancia inmensa para monitorear los mercados nacionales e internacionales, conocer los clientes e incidir en las decisiones gubernamentales, pero para aprovechar estas utilidades transversales se requiere combinar métodos adecuados a las necesidades de información de las instituciones con tecnologías innovadoras que se sobrepongan a los retos de sobreinformación y desinformación que cunde en internet.

La precisa recuperación de contenidos, el almacenamiento y el procesamiento de los datos para extraer información relevante, constituyen los principales desafíos de la gestión de información por el ingente y poco controlable volumen de contenidos en la web que dificulta la correcta selección de los datos fiables. En ese sentido, el profesor Carlos Figuerola (2019) señala la todavía incapacidad de la informática para identificar la mentira, el sarcasmo, la sátira y otros elementos que podrían ser estratégicamente valiosos durante el procesamiento de los datos.

La repercusión de la información en las ventas, las relaciones y la reputación empresarial queda evidenciada en el estudio “Estado de la Comunicación en España” (Asociación de Directivos de Comunicación, 2018), donde el 50,4% de los especialistas en Relaciones Públicas afirmaron que usar *big data* y algoritmos para la comunicación constituye el mayor reto de los comunicadores. Hacer coincidir los datos con los intereses de los diferentes públicos es la tercera mayor aspiración (40,1%). No es de sorprender que la monitorización, evaluación y medición de la información empresarial haya sido considerada como la cuarta principal actividad en la comunicación y en el futuro de la disciplina.

Todas estas consideraciones hacen que la inadecuada gestión documental llegue a ser una tarea enredada, cara y extensa, dificultando la toma de decisiones, los negocios y la evaluación de la realidad misma.

1.2.1 Preguntas de Investigación

Ante los retos descritos anteriormente, se nos plantean las siguientes preguntas:

1. ¿Qué instituciones nacionales e internacionales proporcionan información de interés estratégico para el devenir de la empresa?
2. ¿Qué volumen de información recibe la empresa actualmente? y ¿para qué cantidad de información debe prepararse?
3. ¿Qué métodos y tecnologías deben ser utilizadas para seleccionar, procesar y recuperar eficientemente una inmensa cantidad de datos de texto para adaptar la oferta comercial empresarial a las demandas del entorno empresarial?

Estas cuestiones delimitan la investigación a la gestión de información textual considerándola como la forma más natural de comunicación, pues constituye hasta un 80% de todos los datos extraíbles desde la web (Kantardzic, 2011). Cuando el conocimiento del entorno forma parte de la estrategia institucional gestionar la información es imprescindible.

La selección de las fuentes de información debe responder a criterios que permitan la valoración y medición de las mismas. Los datos textuales deben ser buscados, seleccionados, extraídos y organizados según la capacidad tecnológica que su gestión demande. Los métodos y tecnologías utilizados pueden marcar la diferencia entre la competitividad y la ineficiencia de la actividad.

1.3 Objetivos

Teniendo en cuenta todo lo expuesto, el propósito principal de este estudio será diseñar un proceso de gestión de información textual para el monitoreo del entorno de la empresa urbanística Arnaiz & Partners.

Este objetivo general se vertebrará entorno a otros objetivos específicos:

1. Identificar los temas de interés estratégico para la actividad de la empresa Arnaiz & Partners.
2. Establecer criterios cualitativos para la selección de las fuentes iberoamericanas de información textual útiles para la empresa Arnaiz & Partners.
3. Conocer los métodos y tecnologías utilizados por los departamentos de información y/o comunicación para monitorear el entorno mediante portales digitales oficiales y medios de comunicación digital.

1.4 Proceso de investigación

La investigación es de tipo cualitativa, involucra una intensa actividad de documentación y entrevistas a 15 informantes clave pertenecientes a la empresa Arnaiz & Partners, como escenario del trabajo, y a profesionales de la Documentación, las Ciencias de Datos, la Informática y la Comunicación, de esta forma se integran perspectivas metódicas, técnicas, tecnológicas y de negocios sobre la gestión de información, respectivamente. La identidad de todos estos participantes fue anonimizada.

Para estos encuentros se adaptó un guión a las áreas de estudios y experiencias de los entrevistados para identificar las tecnologías y técnicas de gestión de información que implementan y poder elaborar un diagrama de procesos adaptados a la consultora Arnaiz & Partners.

Los datos fueron analizados agrupándolos en tres categorías y siete subcategorías, acorde con los objetivos de investigación. La metodología es explicada con más detalle en el apartado 3 de esta memoria.

1.5 Estructura de trabajo

El trabajo se ha estructurado en varios capítulos que analizan teóricamente la gestión de información y desglosan los métodos de investigación utilizados, que a su vez, constituyen el fundamento de las propuestas realizadas.

El capítulo primero se ocupa de la gestión de la gran cantidad de datos que la tecnología pone a disposición de individuos y empresas, fundamentando estos conceptos en estadísticas y citas recogidas en el proceso de documentación. Introduce también a la empresa Arnaiz & Partners, en la cual se implementarán las propuestas de minería y análisis de textos, entre otros conceptos que tienen presencia en el resto del Trabajo de Fin de Máster, como *big data*, minería y análisis de texto, aplicaciones tecnológicas, entre otros.

También son presentadas las motivaciones que conducen a quien sustenta el trabajo a seleccionar este tema, las problemáticas que se abordarán, los objetivos a alcanzar y sus fuentes bibliográficas, todos estos delimitan el alcance del trabajo.

Los apartados incluidos en el estado de la cuestión, en el segundo capítulo, desarrollan ampliamente conceptos centrales de la investigación. Se inicia desglosando la gestión de información, luego la enfoca a las empresas hasta centrarse específicamente en la gestión de información como un elemento estratégico para conocer el entorno empresarial. Adicionalmente se destaca la importancia del texto como formato fundamental en el manejo de datos empresariales y son detallados perfiles profesionales cuya actividad requiere la implementación activa de técnica de gestión de datos para el desarrollo de sus actividades, como el documentalista, el informático y el comunicador.

Finalmente, el capítulo sintetiza la actividad de Arnaiz & Partners, su estructura tecnológica y actividad comunicacional para gestionar sus informaciones y valor dado por sus directivos otorgan a esta función.

De manera exclusiva el capítulo tercero detalla el proceso metodológico llevado a cabo desde octubre de 2019 hasta la presentación final del Trabajo de Fin de Máster. Son detalladas las técnicas de recopilación de información utilizadas, los criterios para estructurar los guiones de entrevistas, seleccionar los entrevistados y el análisis de los datos.

El capítulo cuatro analiza y compara los resultados de las entrevistas y las revisiones bibliográficas. Las respuestas son organizadas en categorías y subcategorías, permitiendo la fácil comparación de las opiniones de los entrevistados.

La cuidadosa exploración de los datos fue vital para presentar las conclusiones de la investigación, además de proponer tecnologías y técnicas para mejorar la gestión de información en la empresa Arnaiz & Partners, destacando un diagrama de flujo, software y procesos, criterios para la selección de fuentes y tipos de instituciones de interés para la empresa. También se sugiere un perfil profesional para gestionar las informaciones de la empresa y mecanismos estratégicos para utilizar la información a nivel interno y produciendo incidencia social.

1.6 Fuentes bibliográficas

Mediante los buscadores de fuentes científicas Google Académico y *Microsoft Academic* se accedió a múltiples repositorios, revistas, bases de datos y plataformas que permitieron construir un cuerpo teórico amplio y relevante para profundizar en la gestión profesional de la información, la ciencia de datos y el monitoreo del entorno empresarial.

Colecciones como *Eprints in Library and Information Science* (E-LIS), Gestión del Repositorio Documental de la Universidad de Salamanca (GREDOS), DIALNET proporcionan una perspectiva académica, mientras que los contenidos compartidos por los usuarios de *ResearchGate*, *ScienceDirect*, *Semantic-Scholar*, Google Libros y LinkedIn suelen estar muy involucrados con la práctica profesional.

El 99% de la bibliografía consultada ha sido publicada en los últimos 20 años y se encuentra en soporte digital.

2. Estado de la cuestión

2.1 Gestión de la Información

Un dato, idea, pensamiento o signo puede ser calificado como irrelevante o importante según el estímulo que produzca en sus receptores. La organización e interpretación de estas señales da lugar a la información. El proceso mediante el cual estos códigos son seleccionados o eliminados para apoyar unos objetivos determinados se contextualiza en la gestión documental (Ponjuan Dante, 2003). De forma amplia, la gestión de información administra los datos y la memoria, lo que permite el entendimiento mutuo para el desarrollo de actividades elementales como la educación y complejos procesos como la construcción de historias, la producción y la conservación de las experiencias, obras de valor y los documentos.

Aunque la gestión de información es una disciplina joven, su práctica es más antigua de lo que cabría pensar. Para Patricia Franks (2013), existe desde que los seres humanos pintaron símbolos en el muro de una cueva, e incluso hoy cuando colocamos publicaciones en el muro de una red social. Hasta la década de 1970, la “gestión de archivos” se refería a documentos físicos, siendo a partir de esta fecha que empieza a evolucionar la computación apareciendo así la figura del “gestor de información”. Actualmente más del 90% de la información en las instituciones es creada electrónicamente, siendo el 95% de esta no estructurada, esto es, archivos no organizados en bases de datos y cuyo formato es cambiante, como textos, audios, imágenes y vídeos, entre otros. La coexistencia de fondos documentales en papel y electrónicos trajo como resultado la disciplina denominada Gestión de archivos e información.

Para Read y Ginn (2016), como la gestión documental incluye funciones de la gestión de información, ambas deben ser integradas en el concepto “Gestión de archivos e información” (RIM, por sus siglas en inglés). Adicionalmente, la norma ISO 15489 sobre información, documentación y gestión de documentos, también asocia el término “archivo” con “documento”, definiendo la gestión de archivos como el “área de gestión responsable de la eficiencia y el control sistemático de la creación, recepción, mantenimiento, uso y disposición de los archivos, incluyendo procesos de captura y mantenimiento de evidencias en forma de archivos de información sobre las actividades de negocios y transacciones” (2016, p. 3).

Hoy, sea a nivel particular u organizacional, la debida administración de los datos constituye una actividad de poder e influencia. Este valor estratégico de la información motivó a la Unión Europea a tener como objetivo “convertirse en la economía basada en conocimiento más competitiva del mundo (...) y con mayor cohesión social” (*European Parliament*, 2000, p. 2) siendo esta una actitud defensiva y ofensiva ante la transformación gradual del mercado mundial.

Mientras la gestión de información alude a un proceso, el análisis de la información es la actividad instrumental en sí (Dulzaides & Molina, 2004) inicialmente se empleaban sencillos procesadores de texto en el caso de los individuos y complejos software de gestión de contenidos y programación para empresas, gobiernos y demás actores

socioeconómicos, teniendo como fin la autogestión y recuperación de sus informaciones de interés en la web (Osuna & De la Cruz, 2010). Simultáneamente, las instituciones educativas y de investigación consolidaron los repositorios digitales, la intranet se puso de moda en las organizaciones y la necesidad de acceder a una innumerable cantidad de servicios en la internet impulsó los motores de búsqueda en la web. El paulatino desarrollo de estos recursos permitió que fuera de estos círculos institucionales las personas también puedan intercambiar información a través de los medios sociales, blogs, wikis, foros y otros recursos que permiten producir todo tipo de contenidos multimedia, dando lugar al término “prosumidor” o “prosumer”.

La constante creación de información por todas estas organizaciones e individuos genera grandes cantidades de datos (*big data*) que dificultan su análisis, agravándose más por la constante aparición de perfiles y noticias falsas, portales inseguros y, software malignos. Estudios de la Universidad de Indiana, publicados en la revista *Computer In Human Behavior*, aseguran que solo el 16% de las personas son honestas en internet y el 98% de los usuarios consideran que sus contactos mienten (Sanz, 2016).

Los creadores de esta enorme y diversa cantidad de información en la web provienen de un entorno global con intereses heterogéneos, ante este reto, los gestores de información deben asumir métodos tecnológicos para acceder a los datos mediante desarrollos como como *Internet of Things (IoT)*, *Data Analytics*, *Machine Learning*, *Natural Language Processing (NLP)*, entre otras.

La periodista Andrea Nogueira al resaltar las luces de la gestión de la información, comenta que: “su uso y estudio continúa en una etapa incipiente” [sin embargo,] “indudablemente, (...) ya ha transformado los mercados, la forma de gestionar las ciudades e incluso la manera en la que las personas se relacionan entre ellas” (2019, p. 1).

2.2. Gestión de la información en empresas

Por sencilla que sea una empresa, incluso si carece de una estructura de trabajo formal, está obligada a administrar una cantidad esencial de información, como bases de datos de clientes, expedientes de trabajo, controles contables, entre otros.

La información es un aspecto intrínseco de los negocios. Sin discriminar tamaños, prácticamente todas las empresas han sido transformadas por la tecnología en la forma como gestionan sus datos, sea de manera espontánea, con el objetivo deliberado de ser eficientes o por imposiciones legislativas, como el manejo de los datos personales y los registros financieros.

Abordando los objetivos de la información digital en las empresas de hoy, Vargas, Ayala & Torres señalan que la meta es “crear un conjunto de normas técnicas y prácticas usadas para administrar el flujo de documentos de todo tipo en una organización, permitir la recuperación de estos, determinar el tiempo que los documentos deben guardarse, eliminar los que ya no sirven y asegurar la conservación indefinida de los más valiosos, aplicando principios de racionalización y economía.” (2013, p.38).

Pero en las empresas la información no es un fin en sí misma, su importancia radica en la posibilidad de tomar mejores decisiones haciendo uso de esta. Por eso la Asociación para la Gestión de la Información y la Imagen (2010) definió el término "gestión de contenidos para empresas" (ECM, por sus siglas en inglés), como las estrategias, métodos y herramientas usadas para capturar, gestionar, almacenar y entregar contenidos y documentos vinculados al proceso organizacional (Franks, 2013, p. 35). La ECM integra toda la información de la empresa, esté en formato papel, base de datos o correos electrónicos; incluye todos los flujos de trabajo, la creación de contenidos, los repositorios y las publicaciones.

La plataforma para la ejecución del ECM, apegada a las políticas institucionales y los objetivos de negocio, es la "gobernanza de información" (Franks, 2013) que involucra la unidad de asuntos legales, recursos humanos, tecnologías y negocios, obteniendo un conocimiento completo de la red de información empresarial y su administración. Así se determina:

1. Qué recursos informativos son más y menos importantes para la organización;
2. Qué acciones son necesarias para cumplir las normativas existentes;
3. Qué información debe ser destruida y cuándo.

La sobreabundante cantidad de información digital ha renovado el interés y el valor del gestor de información y archivos en la organización para priorizar la información, mantenerla, asegurarla, desecharla o preservarla (Franks, 2013), pues de toda la información producida en una empresa, solo califica entre el 10 al 15% (Read & Ginn, 2016), por eso se categorizan documentos según el valor de la información que posean.

Las especialistas Read & Ginn (2016) recomiendan la jerarquización siguiente:

- Archivos vitales: son documentos no reemplazables y cuya desaparición haría imposible el alcance de la misión y las operaciones empresariales. Requieren un alto grado de protección. Ejemplos: documentos de incorporación, títulos de propiedad, informes de accionistas, libros de beneficios y pérdidas (financieros), entre otros documentos en su mayoría jurídicos.
- Archivos de importancia alta: son informaciones reemplazables a un alto costo. Requieren de un alto nivel de protección. Ejemplos: archivos del personal; manual de políticas, memorias, datos y reportes de clientes.
- Archivos importantes: son documentos y/o informaciones reemplazables a un precio no tan alto. Requieren de un medio o bajo nivel de protección. Ejemplo: correos, cartas y ciertas memorias.
- Documentos no esenciales (no archivables): son documentos cuyo valor no es posible predecir más allá de su uso inicial. Se les ofrece un nivel mínimo de protección. Ejemplos: Anuncios y boletines a empleados, documentos rutinarios de correo electrónico y borradores de trabajos ya terminados.

Para clasificar la información también debe considerarse el origen de la información, que pueden ser:

- Archivos internos: contienen información necesaria para el funcionamiento de la empresa, que bien puede ser generada internamente o externamente. Ejemplo: renta de un servicio por correo electrónico, inventarios y reportes de funcionamiento.
- Archivos externos: son informaciones con incidencia en la institución, producidos de manera interna o externa. Ejemplo: regulaciones gubernamentales, facturas de servicios y comunicados de prensa.

El valor de los archivos no es estático, con el tiempo o circunstancialmente puede aumentar su importancia hasta ser vital o reducirse hasta considerarse no esencial, lo cual responde al ciclo de vida del archivo, conocerlo permite su correcta selección, preservación y expurgo.

Combinando las propuestas de Robek, Brown y Stephens (1995) con las de Vargas, Ayala, & Torres (2013) el ciclo de vida de la información podría resumirse de la forma siguiente:

- Modelado: se definen y clasifican los tipos de documentos (jurídicos, administrativos y científicos), los formatos de los archivos y los elementos que lo identifican con la institución. También se establece el personal responsable de su conservación y el periodo de vida de su formato físico y electrónico.
- Creación: es resultado de una solicitud, transacción, procedimiento o planificación. Debe estar debidamente identificado según las exigencias normativas y legales.
- Distribución: la información es suministrada y/o percibida por el responsable de la misma. Debe garantizarse que cada responsable tenga acceso a la información que le corresponde.
- Uso: los archivos son analizados y utilizados para la toma de decisiones, también pueden ser un producto en sí mismo (una investigación). Existen documentos de uso único, múltiple y de consulta constante.
- Conservación o expurgo: la información se mantiene asequible para su fácil disposición y preservación. Si un documento pasa a ser no esencial, si no existe ninguna obligación legal ni condiciones que puedan renovar su utilidad, debe ser desechado.
- Disposición: los archivos permanecen aptos para su recuperación hasta el momento de su destrucción.
- Evaluación: el modelo debe ser constantemente evaluado mediante encuestas y entrevistas.

2.3. Gestión de información estratégica del entorno empresarial

El entorno empresarial está compuesto por un conjunto de actores ajenos a la organización cuyas acciones e intereses afectan en mayor o menor medida la actividad de la organización. Estos públicos externos por lo regular son los clientes y la competencia; las instituciones gubernamentales, legislativas y medios de comunicación; en ocasiones también se incluyen agrupaciones sociales vinculadas a sus actividades. Estos agentes pueden tener un radio de acción internacional, nacional, regional o local.

Los también llamados “partes interesadas” o “*stakeholders*”, influyen en los aspectos socioeconómicos, legales y medioambientales que configuran el clima de negocios. Uno de los primeros términos utilizados para describir la gestión de estas informaciones fue “escaneo del ambiente” o “*environmental scanning*” (Mayeh, Scheepers, & Valos, 2012) mediante el monitoreo de los medios tradicionales de comunicación para ayudar a la alta gerencia de las organizaciones a trazar el curso de acciones de la empresa en el futuro (cit. en Aguilar, 1967). Hoy la mayor parte de las informaciones para desarrollar esta labor están disponibles en internet, recibiendo esta función el nombre de inteligencia de negocios (*business intelligence*), inteligencia económica, inteligencia de mercado (*marketing intelligence*) e inteligencia territorial, según el sector donde se aplique (Ortoll & García, 2015).

La supervisión del entorno es una actividad muy implementada como estrategia de inteligencia competitiva por instituciones gubernamentales y grandes empresas con exposición continua a los medios de comunicación, mientras la mayoría de empresas pequeñas y medianas subestiman monitorear su ambiente perdiendo capacidad para adaptar sus productos a las expectativas del mercado. Estas últimas exigen potencializar su cultura empresarial, capacitar su personal y fortalecerse tecnológicamente para realizar análisis del clima de negocios en las mejores condiciones posibles. Una exitosa gestión de información del entorno empresarial requiere establecer claros objetivos de índole cuantitativa y/o cualitativa:

Objetivos cuantitativos del análisis del entorno:

- Cantidad de apariciones y menciones de una empresa o un tema en los medios de comunicación.
- Monto de dinero ahorrado de publicidad mediante una táctica de difusión.
- Medir la eficiencia de una campaña observando la evolución de las ventas.
- Comparar los precios o el número de ventas de un producto de la competencia.
- Número de personas alcanzadas en una campaña de información.

Objetivos cualitativos del análisis del entorno:

- Análisis de sentimientos: aceptación de un producto y clima de negocios.
- Identificación de tendencias, tecnologías y prácticas.
- Prevención y respuesta a crisis.
- Desarrollo de conocimiento del mercado.
- Análisis de normativas, reglamentos y políticas públicas.

Los mencionados objetivos para gestionar la información del entorno deben estar vinculados a “lo que los directivos creen que necesitan, lo que realmente necesitan, y lo que es económicamente factible” (Kotler & Keller, 2012). También debe de considerarse la visión de la organización, la oferta de productos o servicios, la perspectiva de la dirección y los requerimientos legales existentes. Estos datos son la base de los criterios para identificar y descartar los actores del entorno y las fuentes para analizar el ambiente de negocios.

Las principales fuentes de información para escanear el ambiente de negocios son:

- Portales y medios digitales de instituciones gubernamentales y legislativas a nivel internacional, nacional, regional o local.
- Portales y medios digitales de institucionales internacionales y multilaterales.
- Portales públicos y privados de estudios y estadísticas.
- Portales web y medios digitales de la competencia.
- Repositorios académicos y centros de investigación.
- Portales y medios digitales de personalidades.
- Medios digitales y especializados.
- Seguimiento a tendencias en medios sociales en tiempo real.
- Seguimiento a palabras claves y etiquetas en medios sociales.
- Comentarios y recomendaciones en aplicaciones y herramientas digitales.
- Portales y medios digitales de grupos sociales.
- Medios de comunicación tradicional.
 - o Radio.
 - o Televisión.
 - o Periódicos.

Ciertamente estas fuentes no expresan necesariamente la opinión de la mayoría y podrían existir toda clase de intereses, perceptibles y disimulados. Sin embargo, para Wilcox, Cameron, & Xifra (2012): “los comentarios que ofrecen suelen ser una llamada de atención para las organizaciones en cuanto a problemas potenciales o asuntos importantes” (p. 204).

Los esfuerzos iniciales para investigar el clima de negocios eran muy rudimentarios, personas eran contratadas para tomar notas a medida que atendían los programas de televisión o radio, hacían recortes de prensa e informes en revistas, libros e investigaciones de interés. La llegada de internet y los potentes navegadores web trasladaron esta actividad a la computadora, sin embargo, los contenidos buscados recibían un trato tradicional: eran impresos o copiados y pegados de manera manual en procesadores de texto.

Actualmente, los avances tecnológicos permiten hacer un seguimiento en tiempo real de los contenidos que se producen en la red o extraer (*scrap*) la información de la web al instante de ser publicada (Figuerola, 2020), mientras que hace apenas ocho años la empresa *O'Dwyer's PR Report*, notificó que en un plazo de 10 a 15 minutos eran capaces de notificar dónde y cuándo se hablaba de un cliente (Wilcox, Cameron, & Xifra, 2012).

Los medios digitales son hoy un mecanismo idóneo para gestionar la información del entorno organizacional. Es un esfuerzo de inteligencia de negocios que para ser eficiente debe contar con criterios metodológicos y la implementación de tecnologías capaces de simular de manera ética y avanzada las acciones de análisis humano. Este análisis descubre patrones significativos de información que abrirá nuevas vías para alcanzar los objetivos organizacionales.

Prescindir del escaneo del entorno hace que se corran riesgos como la pérdida de la competitividad con respecto a otras instituciones, provocando una mala penetración en el mercado. También evita conocer las exigencias del ambiente de negocios causando miopía en la visión estratégica.

2.4. Gestión de información de texto del entorno empresarial

Aunque el estereotipo de un documento está vinculado a un conjunto de hojas impresas con información de texto, este concepto ha evolucionado para aludir a archivos que pueden estar en formato sonoro, de vídeo, imágenes, un fichero web y bases de datos, entre otros.

En la gestión de archivos y documentos empresariales el texto es el vehículo más común para el intercambio formal de información (Kantardzic, 2011) y la forma más natural para almacenar información, dándole un potencial comercial más alto que los datos fríos que pueden representar las bases de datos empresariales actuales. Algunos estudios indican que el 80% de la información gestionada por las organizaciones está contenida en documentos de texto (Kantardzic, 2011), presente en los correos electrónicos, los contenidos en la web, las presentaciones, los informes y memorias corporativas.

Cada vez es más común el análisis de las opiniones escritas de los clientes en portales web, aplicaciones, medios sociales y encuestas; las instituciones públicas producen información de índole jurídica o normativa de interés y se redactan artículos acerca de la institución o de su entorno en medios de comunicación, investigaciones universitarias, entre otras fuentes. Todos estos son contenidos de texto no estructurados del entorno y se desconoce qué porcentaje de información constituyen del total datos que producen las instituciones.

Los motores de búsqueda permiten encontrar de manera rápida las informaciones o páginas requeridas, reduciendo el tiempo de navegación. Sin embargo, para fines de investigación sería poca efectiva la tarea humana de: 1) entrar manualmente a decenas de portales; 2) identificar, leer y seleccionar los datos de interés; 3) copiar y pegar los textos uno por uno en un procesador de texto; 4) re-evaluar, resumir o extraer la información principal; 5) organizar en hojas de cálculo toda la información textual no estructurada; 6) hacer informes del análisis del entorno; 7) otras actividades

de preservación de datos. Todo el proceso descrito o cualquier otro parecido resulta obsoleto ante las exigencias competitivas del mercado y el gran volumen de datos en la web. La gran cantidad de contenidos involucrados y su compleja gestión hacen de este un problema de *big data*.

Uno de los sistemas en desarrollo que, haciendo uso de las tecnologías, facilita la explotación de estos volúmenes de información de textos, facilitando su búsqueda, selección, clasificación, organización, almacenamiento y uso, es conocido como análisis de texto (*text analysis*) o minería de texto (*text mining*)¹, esta actividad forma parte de la gestión de información desde mucho antes de la existencia del internet. Sus principios básicos son aplicados donde sea que existan grandes cantidades de contenidos que necesiten ser analizados (Kantardzic, 2011).

La minería de texto usa técnicas de la ciencia de datos para que “conociendo mejor la institución y sus usuarios/clientes, se pueda evaluar el pasado y el presente, planificar el futuro inmediato y a medio plazo, y mejorar la toma de decisiones” (Martínez-Uribe & Fernández, 2015, p. 6)

El uso de aplicaciones informáticas hace posible el monitoreo automatizado de las informaciones que producen las instituciones y actores que conforman el entorno empresarial mediante sus portales web, redes sociales, medios de comunicación digital y blogs (*web text mining*) adoptando técnicas de aprendizaje automático (machine Learning) para reconocer patrones y comprender nueva información, sea en formato *HTML* o software como PDF, *Word* y demás.

Deben darse instrucciones logarítmicas a un software araña (*crawler*) para que navegue en las bases de datos o la URL de los portales web predeterminados para extraer (*scraping*) grandes cantidades de contenidos relevantes imitando la forma de navegación de un humano hasta descargar los contenidos en un servidor con el formato que convenga.

Utilizar software y robots para el monitoreo del entorno empresarial, ha renovado importantes debates éticos vinculados a la violación de los derechos de autor por parte de los creadores y propietarios de la información; así como la protección de datos de los clientes y el secreto empresarial.

La efectividad de estas tecnologías está fundamentada en criterios metodológicos para la gestión, preservación, almacenamiento y disposición a largo plazo de la información textual no estructurada (Chen, Chiang, & Storey, 2012). Los criterios para determinar qué informaciones son de interés institucional y las fuentes para conseguirlas están basados tanto en técnicas documentales como en el pensamiento estratégico de la entidad interesada, sus objetivos y normativas. Estos principios son cardinales, dado que “rastrear todas las direcciones IP (...) es una tarea del todo inasumible para la mayoría de investigadores (...) aunque se dispongan de programas web crawler capaces de rastrear millones de sitios web (...) se necesita una infraestructura muy desarrollada con centenares de servidores y programas crawler recopilando

¹ Se enfatiza lo textual para diferenciar el análisis de texto (*text analysis*) de la minería de datos (*data mining*), actividad también propia de los científicos de datos que incluye tanto al texto como a la gestión de archivos sonoros, imágenes, videos, entre otros. Existen más términos como *text data mining*, *text knowledge mining*, *knowledge discovery in text*, *information mining*, *opinion mining*.

contenidos” (Ríos Ortega & Ramírez Velázquez, 2017, p. 177), pudiéndose agregar a esta declaración que además de inasumible es innecesario.

Existen aplicaciones de pago y código abierto que automatizan la extracción, clasificación, síntesis y análisis de los textos mediante el procesamiento del lenguaje natural, aunque algunas continúan siendo muy limitadas. Los software más recomendados para realizar análisis cualitativos y cuantitativos conllevan el uso de lenguajes de programación y funciones de estadística, algunos ejemplos son *R*, *Java* o *Python*.

Para el monitoreo de textos, mientras la intervención humana otorga menores niveles de error y mayor rigurosidad, las herramientas informáticas ahorran recursos y tiempo. Según los objetivos de gestión de información deben equilibrarse las tecnologías implementadas y el equipo de profesionales responsable.

Las principales aplicaciones de la gestión de información de texto del entorno empresarial de manera específica son:

- Análisis de sentimientos: es la medición de la percepción positiva, negativa o neutra de una comunidad sobre un servicio, producto, acontecimiento o proyecto. Este se encuentra muy focalizado en las redes sociales.

- *Knowledge discovery*: extrae unidades de información útil para la toma de decisiones y la producción de nuevas documentaciones fundamentadas en la gestión del conocimiento. Permite el monitoreo del mercado y el control de las instituciones públicas. Va más allá de parámetros binarios (bueno o malo).

- Análisis de clientes: retroalimenta detalladamente sobre las experiencias, expectativas y comportamiento de una comunidad a través del tiempo para fortalecer la fidelidad de estos con la institución.

- Publicaciones digitales: se recuperan informaciones de portales preestablecidos y se agrupan automáticamente acorde con categorías preestablecidas. Luego son evaluados para obtener artículos nuevos.

- Análisis de redes sociales: permite la formulación de negocios detectando organizaciones y actores que suelen colaborar, se etiquetan o tienen temas comunes.

2.5. Perfiles para la gestión de información del entorno

Tradicionalmente la gestión de archivos está vinculada a los documentalistas y a la Biblioteconomía. Sin embargo, la profanación de la disciplina por profesionales de todas las áreas se potencializó con los avances tecnológicos que pusieron a disposición de los usuarios ingentes cantidades de información y la posibilidad de producir contenidos, poderosos motores de búsqueda, amplio espacio de almacenamiento y herramientas para organizar el conocimiento fundamentadas en los principios de la gestión de la información, haciéndose adaptaciones a las técnicas acorde con la materia de estudio.

En el plano profesional, John Feather (2006) asegura que la gestión de información estaba siendo ya abordada como una herramienta básica por otros profesionales de diferentes tipos y clases restando protagonismo a los bibliotecarios (Tramullas (2016), como en el departamento de gestión de datos de la empresa española Telefónica, integrado mayoritariamente por informáticos; en Murcia profesionales del marketing monitorizan la satisfacción de los clientes de hoteles mediante los contenidos de texto generados en portales web institucionales, *Google Maps* o herramientas de referencia y compra como Trivago (Mellinas, 2015).

También en la ciudad estadounidense de Michigan, el equipo de ingenieros del Centro de Investigación de Transporte para Comunidades Vivibles analizó las narraciones de 12,671 reportes policiales para explorar los beneficios del aprendizaje automático en la clasificación de accidentes y recomienda este método para el análisis de los textos de los expedientes policiales (Kwayu, Kwigizile, Zhang, & Oh, 2020); y en República Dominicana, varios médicos, utilizan técnicas básicas de análisis de texto para caracterizar los pacientes con problemas cardiovasculares usando la información de los Archivos Electrónicos de Salud (Fabian, Ramos, & Fernández, 2018).

Como factor común en los recientes casos citados, la gestión de información de texto es asumida directamente por los profesionales involucrados en la actividad laboral o investigativa: ingenieros, mercadólogos y médicos. En vacantes vinculadas con el análisis de datos también se requieren profesionales de la economía, estadística y matemáticas.

La ausencia de profesionales de la información en estos proyectos y el uso intensivo de las tecnologías hacen que voces como Bekir-Kemal Ataman (cit. en Tramullas, 2016) prevean la extinción de la tradicional función de personas ayudando en el acceso a la información. Este autor considera que el modelo de educación está desfasado y coincide con Fernando Arnaiz (2020) en que todo profesional hoy está lo suficientemente equipado tecnológicamente para gestionarse su propia información.

Otros especialistas en comunicación consideran suficientes los servicios de monitoreo web como *Google Alerts* (Hernández, 2020), o el pago por este servicio (Almonte, 2020), los cuales remiten notificaciones vía correo electrónico acorde con las consultas registradas. Esta visión y herramientas serían útiles para instituciones poco diversificadas y con un entorno reducido, pero desaconsejadas cuando deben mantenerse en observación cientos de fuentes en la web.

Por demás, la gestión estratégica de los datos requiere tiempo, un recurso no abundante para quienes ocupan altas posiciones de responsabilidad en organizaciones con actividades y entornos muy activos. Estos requieren de profesionales capaces de procesar y poner a su disposición informaciones precisas sobre los temas de interés.

Aunque el servicio de gestión de información pudiera externalizarse, acorde con el tamaño, presupuesto y nivel de dependencia del entorno, lo recomendable es tener un personal con formación en la gestión documental y ciencia de datos, tecnologías de la información y un amplio conocimiento de la institución, pudiéndose involucrar profesionales formados o con experiencia en la actividad económica principal de la empresa.

Un equipo de gestión de información del entorno empresarial integrado a la organización tiene mayor conocimiento de la actividad empresarial y su proyección para identificar las informaciones de interés que podrían ser imperceptibles para un tercero. Estas ventajas hacen que incluso pequeñas empresas dispongan de personal que dedique segmentos de su tiempo a analizar el clima de negocios.

Los aportes de especialistas como Tramullas (2016), Mendez (2020) y Russom (2013), arrojan múltiples métodos para desarrollar equipos de gestión de información:

- *Domain knowledge*: propone que, en vez de importarse personal de áreas especializadas en gestión de datos, profesionales de la actividad principal de la institución y con conocimiento de esta sean capacitados para la actividad de investigación e inteligencia.
- Equipo multidisciplinarios: donde especialistas en gestión de información e informáticos controlarían los métodos y tecnologías a utilizar para el monitoreo, mientras capacitados profesionales de la actividad principal de la empresa y las áreas con mayor demanda de información participan en el análisis de los datos, la validación de las fuentes y la construcción de conocimiento organizacional.
- Profesionales de la información: el personal está especializado en los campos de la documentación y la gestión de información con el objetivo de curar los contenidos, desarrollo del conocimiento y creación de contenidos nuevos.
- Equipo de ciencias: involucra personal de áreas especializadas en manejo de códigos, bases de datos, algoritmos y lenguajes de programación como la informática, matemáticas, estadísticas, economía, entre otras.

Para evitar errores en cuanto al presente y futuro perfil del profesional de la información el profesor Jesús Tramullas (2016) no recomienda ir más allá del corto y medio plazo por el dinamismo producido por las nuevas tecnologías y las formas caricaturescas como se ha venido proyectando la sociedad de la información, recomendando no tener fijación en “lo que es” cada profesión, sino a través de las competencias y habilidades que manifiestan sus profesionales para manejar los datos. Bajo este principio, acorde con las características de este trabajo y el Máster en Sistemas de Información Digital, son evaluadas las perspectivas estratégicas y los retos de la comunicación, la informática y la documentación para gestionar la información del entorno.

2.5.1. Perspectiva comunicacional para la gestión de información

La gestión de la información del entorno cuando involucra el monitoreo de instituciones gubernamentales, grupos sociales y medios de comunicación suele ser realizada por especialistas en relaciones institucionales con el objetivo de prevenir contingencias y fortalecer la reputación empresarial (Wilcox, Cameron, & Xifra, 2012). Al integrarse públicos vinculados con las ventas y la competitividad, como los clientes y la competencia, también se requiere implementar estrategias de inteligencia de negocios vinculadas al marketing (Kotler & Keller, 2012). La correcta segmentación de estos públicos permite conocer sus heterogéneas necesidades, objetivos y actividades (la demanda), el análisis de estas informaciones permite el desarrollo de servicios rentables (la oferta), relaciones de cooperación e influencia.

Para esta actividad, los gabinetes y agencias de comunicación actúan como los ojos y los oídos de las instituciones públicas y privadas ante administraciones públicas, medios de comunicación y estableciendo redes de intereses, lo cual exige mucho tiempo reuniendo y procesando información, muchas veces con métodos rudimentarios como el *clipping* de prensa y desaprovechando las inmensas posibilidades de optimización de este trabajo que ofrecen las tecnologías.

En 1980 estudios de *Kodak* y otras organizaciones arrojaron que entre el 30% y el 40% de los artículos interesantes para las instituciones no podían ser evaluados mediante recortes de medios (CyberAlert, 2010). Por demás, los informes de prensa no sintetizaban la información, las principales noticias eran copiadas y pegadas en extensos informes que requerían bastos costos de personal.

A partir del 2004, el método de extraer la información manualmente desde los portales web hasta un procesador de texto fue rebasado por la cascada de información web. Los ejecutivos también habían sido alcanzados por la infoxicación y carecían de tiempo para leer las grandes memorias periodísticas colocadas en sus despachos.

Desde entonces los comunicadores tienen como reto el uso de las tecnologías para proveer conocimiento del entorno, consolidar la reputación y dirigir el diálogo con los *stakeholders*, a partir de las nuevas fuentes de información digital y las tradicionales: la *big data*. Esta actividad de valor puede garantizar el éxito de las organizaciones (Pereira, Portilla, & Rodríguez, 2018).

La transición tecnológica ha convertido la web en el escenario de intercambio de datos para escuchar y forjar vínculos, de ahí que agencias de relaciones públicas implementen herramientas de inteligencia artificial, como sistemas de lenguaje natural para mejorar la vigilancia de las redes sociales y los medios de comunicación digital (Cervantes Ramírez, 2017). Además de incorporar aplicaciones para la sumarización automática de textos, de manera que sean completos, incluir los datos de mayor relevancia de los textos, y legibles, la información debe ser coherente y gramaticalmente correcta (Modaresi, Gross, Sefidrodi, & Eckhof, 2017).

Sin duda alguna, prontamente los profesionales de la comunicación incluirán en su quehacer la gestión de análisis de datos, solo así mantendrá un flujo informativo bidireccional con el entorno. Así sugiere el informe *Latin America Communication Monitor* (2017) que atribuye bajas competencias a los comunicadores para el monitoreo y la interpretación de datos en medios sociales (37,7%) y para comprender

el uso de algoritmos (17,3%); los profesionales de esta disciplina siguen considerando como un reto alcanzar la evolución digital (2014), lidiar con la velocidad del flujo de la información (2016) y el uso de *big data* (2018). El uso de algoritmos con estos fines tiene una importancia de 32,7% (Moreno, y otros, 2019).

2.5.2. Perspectiva informática para la gestión de información

Desde la década de 1970 el desarrollo de la informática ha impulsado la evolución de la gestión documental mucho más rápido que en todo su pasado, permitiendo el casi ilimitado intercambio de datos para fines laborales, de investigación y actividades sociales. Este respaldo de las tecnologías de procesamiento del lenguaje natural y aplicaciones de inteligencia artificial permite automatizar la extracción clasificación y organización de los datos con aceptable precisión para la construcción del conocimiento y la toma de decisiones.

Con la correcta ejecución de estos software pueden monitorizarse en tiempo real los contenidos cualitativos y cuantitativos de una variada cantidad de fuentes de datos como los medios sociales, los medios de comunicación digital y páginas oficiales para analizar los sentimientos de los públicos, medir las ventas online, el alcance de las campañas publicitarias o mantener bases de datos cuyo análisis permita incidir en el comportamiento de los clientes y aumentar la competitividad.

Sin embargo, muchos datos no es sinónimo de calidad ni de un uso inteligente de la información. Tener una correcta visión institucional y documentalista sobre las necesidades informativas constituyen intangibles de altísimo valor para la gestión de información.

El analista de datos David Méndez (2020), considera que la mayoría de los informáticos no suelen prestar atención a los componentes sociales o humanos involucrados en la gestión de información porque el objetivo único de estos son los datos y los resultados que estos arrojen. Por eso recomienda equipos multidisciplinares: los técnicos de la información enfocados en el mantenimiento de las aplicaciones, bases de datos y el soporte en asuntos complejos de programación, mientras que los analistas deben enfocarse en la curación de los datos, su examen y la generación de contenidos.

Uno de los retos de la informática vinculada a la ciencia de datos es la adaptación y desarrollo de software acorde con las expectativas éticas y legales que aluden violaciones a los derechos de propiedad de la información y el manejo de datos personales, provocando que instituciones y empresas blinden sus portales con complejos sistemas de seguridad.

Otro desafío es el fortalecimiento de los software para contrarrestar informaciones infundadamente falsas (*fake news*). En la mesa redonda “La Verdad de las Mentiras” (Espada, Gascón, & Morel, 2018, p.7), de manera pintoresca, advierten lo siguiente: “Dos tipos en calzoncillos en un sótano de Macedonia pueden inventar que el papa apoya a Donald Trump y que esa noticia falsa circule”, y concluyen, “lo que vemos es un ecosistema donde la información llega de muchos lados: la información rigurosa y la mentira conviven como en un lago donde llega agua de distintos lugares. El objetivo

y el resultado es enfangar, dificultar las diferencias, introducir una especie de visión cínica”.

2.5.3. Perspectiva documentalista para la gestión de información

Las bibliotecas durante siglos han desarrollado metodologías para la conservación de los archivos, la clasificación de la información y el desarrollo del conocimiento. Al consolidarse la web empezaron a surgir colecciones digitales con importantes obras de índole histórica, literaria y filosófica; publicaciones científicas, archivos sonoros, creaciones artísticas, arquitectónicas y arqueológicas, entre otras ramas, con fines de conocimiento, competitividad y poder.

Este diverso conjunto de datos, precisa de expertos para su preservación, exposición y promoción, surgiendo así la figura del curador de contenidos (*content curation*).

El *content curator* es un campo de las ciencias de la documentación en plena evolución, cuya función consiste en proveer contenidos de calidad, accesibles, precisos y confiables acorde a su contexto, frente a la espesa nube de información al navegar en la web. Tiene como misión “encontrar el contenido más relevante y presentarlo” (Bhargava, 2009, p. 1) siendo especialmente útil en las organizaciones que ponen la creación de conocimiento como parte de su estrategia de desarrollo.

Para procurar la información pertinente a las comunidades de la organización, el curador genera contenidos de alta calidad y que responden a los requerimientos de los usuarios, aunque se vale de software para el análisis de los datos digitales en web, busca construir una cercanía social en su gestión, “personas, no meros algoritmos” (Urbano, 2016, p. 137). En vez de dejarse arrastrar por el continuo volumen informativo del *big data*, estos expertos en documentación revalorizan los archivos, adaptan la información al nuevo contexto y construyen el conocimiento para la toma de decisiones evitando la presión y la prisa.

Para los documentalistas la calidad del análisis es proporcional a la calidad de los datos, por eso la curación de contenidos se vale de criterios para confirmar la confiabilidad de las fuentes y su interoperabilidad, utilizan vocabularios controlados, validan los datos y la posibilidad de reutilizar la información (Cavanillas, Curry, & Wahlster, 2016).

En el libro “El *Content Curator*”, la bibliotecaria Natalia Arroyo (2013) resalta la metodología propia de esta área, que consta de las fases: *search*, *select*, *sense making*, *share* (búsqueda, selección, caracterización y difusión), y señala la intrusión que han hecho profesionales del periodismo y la mercadotecnia en este campo.

Mientras, Ríos Ortega & Ramírez Velázquez (2017) consideran que la gestión del *big data* está directamente vinculada con la figura del documentalista para lograr una inteligencia competitiva eficaz. El profesor José Antonio Frías (2020) asegura que aunque todo el mundo busca reivindicar una identidad profesional, los campos surgidos como fruto de los cambios tecnológicos no deben ser etiquetados por ninguna disciplina tradicional y excluyendo otras, coincidiendo parcialmente con

Guallar, Leiva-Aguilera (2013, cit. en Arroyo-Vázquez, 2014) cuando explican que el curador de contenidos surge en un nuevo contexto, requiriendo también habilidades normalmente atribuidas a la rama de los negocios.

Sin lugar a dudas el debate es producto de la evolución de la Documentación como consecuencia de las tecnologías. Una investigación de 173 ofertas de trabajo publicadas en Estados Unidos vinculadas a la gestión bibliotecaria y curación digital muestra un alto requerimiento de profesionales con habilidades en tecnologías de la información (50%) y gestión de proyectos (42%) para ocupar responsabilidades de curación (45%) y consultoría (55%) (Kim, Warga, & Moen, 2013). Por su parte, la bibliotecaria Melissa Horak (2020), recomienda también tener conocimientos básicos en lenguajes de programación.

Pero las herramientas que utilizan la mayoría de los documentalistas dedicados a la curación de contenidos continúan siendo las básicas: hojas de cálculo, editores de texto y correos electrónicos. Estas carecen de funcionalidades fundamentales para la captura y representación de las necesidades de los usuarios, el manejo eficiente de grandes cantidades de datos, la búsqueda de contenidos y la integración de diversos tipos de datos (Cavanillas, Curry, & Wahlster, 2016).

Según la empresa de investigación de las tecnologías de la información Gartner (2014) las empresas solo utilizan el 10% de los datos recogidos (Gartner, 2014; cit. en Singh, 2019). Buscando cambiar esta realidad, los documentalistas tienen como reto en lo relativo a la gestión de la información generar conocimientos aprovechando al máximo su colección documental con el uso de software, lo que disminuiría también los costos de curación (Cavanillas, Curry, & Wahlster, 2016).

Teniendo en cuenta lo expuesto, en la siguiente tabla-resumen se muestran cuáles son las características generales de la gestión estratégica de la información según las distintas disciplinas (Tabla 1).

CARACTERÍSTICAS GENERALES DE LA GESTIÓN ESTRATÉGICA DE LA INFORMACIÓN SEGÚN DISCIPLINAS ACADÉMICAS			
INDICADORES	DISCIPLINAS		
	INFORMÁTICO	DOCUMENTALISTA	COMUNICADOR
Perfil	<ul style="list-style-type: none"> ● Profesionales de tecnologías de la información, sistemas y ciencias de la computación. ● Alta capacidad de alcance a múltiples fuentes de información con varias herramientas. 	<ul style="list-style-type: none"> ● Profesionales de biblioteconomía y documentación. ● Especialistas en organización de información. ● Nivel de redacción alto. ● Altas aptitudes en la curación de contenidos. 	<ul style="list-style-type: none"> ● Profesionales de las relaciones públicas y marketing. ● Especialistas segmentando públicos. ● Nivel de redacción alto. ● Rápida integración institucional y aptitudes de negocios.
Retos	<ul style="list-style-type: none"> ● Desarrollo de software acorde con las expectativas éticas y contrarrestando las noticias falsas. ● Mejora de aptitudes para los negocios e integración en instituciones. ● Mejora en la selección de fuentes de información. 	<ul style="list-style-type: none"> ● Reducir la intrusión en su profesión. ● Eficientizar el desarrollo de reportes institucionales y con visión estratégica. ● Mejora del uso de las tecnologías y su relación con las personas para la administración de la información. 	<ul style="list-style-type: none"> ● Sustitución de métodos ineficientes de gestión de información con el uso de las tecnologías. ● Reducir la extensión de los informes sobre el análisis de datos. ● Establecer criterios específicos para la selección de la información de interés.
Competencias tecnológicas	<ul style="list-style-type: none"> ● Nivel Alto. ● Emplean procesadores de texto, hojas de cálculos, bases de datos y herramientas de inteligencia de negocios y ciencias de datos. 	<ul style="list-style-type: none"> ● Nivel Regulares. ● Emplean procesadores de texto, hojas de cálculos, bases de datos y gestores de información digital. 	<ul style="list-style-type: none"> ● Nivel Regular. ● Emplean navegadores web, procesadores de texto y herramientas de graficación.
Gestión de información	<ul style="list-style-type: none"> ● Utilizan plenamente las tecnologías con baja socialización y criterios. ● Uso de información para solucionar problemas. ● Generan diarios informes gráficos. 	<ul style="list-style-type: none"> ● Combinan métodos tradicionales y tecnológicos para la gestión de datos. ● Uso de la información para producir conocimiento. ● Generan informes periódicos. 	<ul style="list-style-type: none"> ● Utilizan métodos manuales de gestión de datos. ● Utilizan la información como un recurso estratégico. ● Generan informes diarios de lectura poco amigable.

Tabla 1: Características de gestión estratégica de información según disciplinas académicas

2.6. Gestión de información en Arnaiz & Partners

Desde el año 1975, el equipo que hoy lidera la empresa española Arnaiz & Partners colabora con diversos equipos vinculados al desarrollo sostenible del territorio hasta constituirse en una consultora con alcance internacional.

Por su naturaleza, las instituciones públicas son el principal cliente de Arnaiz & Partners, aunque sus servicios de asesoría técnica también son contratados por muchas entidades privadas y mantiene una activa relación con las universidades y organizaciones multilaterales. El director de comunicación y relaciones institucionales de la consultora, Fernando Arnaiz (2020), destaca las tres grandes áreas en las que trabaja:

- Planificación urbana: involucra actividades urbanismo, medio ambiente, ingeniería, asesoría jurídica y conservación urbana.
- Desarrollo inmobiliario: como la arquitectura, construcción, promoción, gestión de activos patrimoniales, mantenimiento de inmuebles.
- Desarrollo estratégico, que incluye:
 1. Gobernanza local: considerando que el desarrollo de las ciudades parte desde abajo;
 2. Tecnología aplicada al urbanismo: herramientas de inteligencia de ciudades como *UrbGIS* para la geolocalización de información.
 3. Formación académica y profesional: mediante el Instituto Arnaiz de Estudios e Investigación desde hace 15 años. Imparten cursos de proyectos estratégicos y urbanísticos; sobre estructura municipal; economía y tributación municipal; y tecnologías para datos territoriales.

Arnaiz & Partners tiene sede en Madrid y nueve delegaciones repartidas en Europa del Este, África y América, lo que impregna a sus intervenciones una amplia experiencia y visión global. Sin embargo, esta misma diversificación también dificulta la formación, la construcción del conocimiento e incluso el control de la información que debe manejarse con fines internos, pues la información se actualiza, crece y duplica con rapidez por causa de la digitalización (Arnaiz, 2020).

Para mantener los flujos de información la consultora utiliza software de gestión de contenidos, almacenamiento en nube, correos electrónicos y múltiples herramientas tecnológicas de comunicación, considerando la idiosincrasia y cultura de cada delegación; las coincidencias y diferencias jurídicas, políticas, sectores productivos que intervienen; así como las exigencias y velocidad del mercado.

A nivel interno, los departamentos que más intervienen en todo lo relativo a la gestión de información son: la Dirección de Comunicación y Relaciones Institucionales, responsable de la información y la producción de contenidos estratégicos; y el Departamento de Tecnología e Innovación (Arnaiz-Eguren, 2020).

El director de tecnología, Ignacio Arnaiz-Eguren (2020), explica que mientras los datos estructurados son protegidos en bases de datos internas, los documentos no estructurados son almacenados en la nube de *Dropbox* y *Box*. Sistemas como *VMware Horizon* permiten al personal intercambiar información de manera remota, manteniendo los documentos compartidos, ordenados y clasificados. El departamento está organizado en tres áreas:

1. Departamento de sistemas: asegura el correcto funcionamiento y disponibilidad de las plataformas de servidores y data center en la sede de Madrid. También ofrece soporte ante problemas de software o hardware.
2. Departamento de desarrollo: formado por programadores. Fundamentalmente dedicados a *UrbGIS*.
3. Departamento de arquitectura de base de datos: administra las bases de datos y garantiza su correcto rendimiento.

En cuando a la gestión estratégica de la información y la actividad productiva de Arnaiz & Partners la alta gerencia señala, de manera unánime, su amplia diversificación en el área de desarrollo territorial, utilizando expresiones como: “nuestro servicio no es como fabricar tornillos o vender camisetas” (Arnaiz-Ramos, 2020), “no vendemos productos enlatados” (Arnaiz-Rebollo, 2020), “nuestros proyectos no son estáticos, cambian acorde con la necesidad” (Tejeda, 2020).

Sin embargo, este mismo consenso no existe en cuanto a la administración de la información digital que produce su entorno de interés a nivel nacional y en sus delegaciones internacionales.

En España, el Departamento de Comunicación y Relaciones Institucionales hace un trabajo diario de localización de aquellas informaciones que puedan ser útiles para la consultora, las cuales son discutidas con los responsables a los que le pueda interesar (Arnaiz-Ramos, 2020). Se utilizan tecnologías de nueva generación para gestionar los clientes con los cuales mantienen una relación cercanía; tienen además una plantilla de consultores y *partners*, que proveen información y clientela a la empresa.

Fernando Arnaiz (2020) afirma que la empresa no está utilizando software ni estrategias de marketing especiales para obtener clientes porque la empresa hoy tiene capacidad de sobrevivir sin ellos. Aun así, investigan nuevas maneras para analizar las necesidades de los clientes.

Por otro lado, su delegación brasileña utiliza micro-datos para identificar informaciones relevantes sobre los públicos de interés, que son seleccionados en función del presupuesto que manejen y las posibilidades de acceder físicamente a estos (Rocha, 2020). Utilizan Software libres como: bases de datos *Postscript*, programas de inteligencia empresarial como *Pentaho*, entre otras *API*.

En República Dominicana, se sigue un método que prioriza medios digitales e instituciones en función de su actividad, presupuesto y localización para obtener informaciones relevantes del entorno, los clientes de interés y los concursos públicos. Con este propósito se optimizan los navegadores web implementado técnicas de búsqueda booleanas, los datos son organizados en hojas de cálculos, y analizan con la *API OpenRefine* y se presentan informes periódicos en PDF.

Cada delegación produce fuentes de información de forma independiente, sin parámetros formales que permitan a la alta gerencia evaluar y comparar de manera simple y sistemática la dinámica socioeconómica, técnica, política y legal de su entorno ni a nivel local ni internacional. En lugar de eso, sobre la marcha se interpreta la realidad y se adaptan las estrategias de negocio permanentemente a través de la internet. Por eso el presidente de la consultora, Leopoldo Arnaiz (2020) considera que las empresas deben gestionar la información relacionada con la actividad empresarial en temas que tengan utilidad práctica y debe ser ambiciosa en los cambios que puedan producirse con el propósito de generar conocimiento dirigido al interior de la empresa, incidir en la administración pública y la generación de un debate propositivo.

En cambio, el director ejecutivo en España, Luis Arnaiz (2020), está convencido de la suficiencia de los *partners* como suministradores de conocimiento empresarial y descarta la necesidad de incorporar acciones de gestión de información. Aunque Fernando Arnaiz (2020) e Ignacio Arnaiz (2020) también confían en estos métodos, no descartan la necesidad de implementar otras técnicas de inteligencia como el *bechnmarking* o el estudio de nuevas maneras de captar clientes.

Niwirka Tejeda, directora ejecutiva de Arnaiz & Partners en República Dominicana, fundamentada en los éxitos de su delegación al gestionar la información, recomienda aplicar técnicas similares a la alta gerencia para que los proyectos concebidos en Europa se apeguen mucho más a la realidad Latinoamericana. Adicionalmente, considera que podrán entenderse mejor las propuestas que se hacen desde la región y las necesidades del mercado Arnaiz & Partners, pues no siempre las perspectivas e intereses de los socios están alineados a los objetivos de la empresa. Empero, Fernando considera que acciones en materia de gestión de información podrían causar más infoxicación y confía en la capacidad de la dirección para manejar su propia información. Leopoldo Arnaiz (2020) ve imprescindible que las empresas definan áreas concretas para manejar informaciones con utilidad práctica. Agrega que la crisis del Covid19 ha provocado una serie de reflexiones en el mundo empresarial entorno a la correcta toma de decisiones a la hora de invertir tiempo y recursos en aspectos del negocio que permitan pronto resultados económicos. El propósito es conseguir “que la empresa sobreviva” (Arnaiz, 2020).

3. Metodología

Desde finales de octubre del 2019, quien suscribe este trabajo redactó un anteproyecto de investigación que vislumbraba la aplicación de los principios de la gestión de información digital para el manejo de la documentación empresarial generada por su entorno de interés, teniendo como resultado este Trabajo de Fin de Máster.

El tipo de investigación realizada es cualitativa, pues considera las opiniones de profesionales de la Documentación y de la Informática por ser estas las áreas fundamentales para la estructura metodológica de la gestión de la información y el desarrollo de software para el manejo de datos. También se incluyeron especialistas de la comunicación por ser el departamento encargado de desarrollar relaciones con las organizaciones que conforman el entorno de las empresas mediante estrategias de información.

El análisis se centra en la consultora urbanística Arnaiz & Partners por ser el escenario donde el autor de este trabajo desarrolló una labor de gestión de información con mecanismos obsoletos que a su vez motivaron al desarrollo de la investigación, como se explicó anteriormente, en el apartado de Justificación. Por eso se obtuvieron declaraciones de los integrantes de la alta gerencia en la empresa en España y directivos en Latinoamérica sobre la estructura empresarial para la gestión de la información.

La expectativa inicial era entrevistar un mínimo de 15 informantes clave y un máximo de 20, pudiéndose considerar esta meta como alcanzada. Las técnicas de recogida de información empleadas fueron las siguientes.

- Revisión bibliográfica: se consultaron más de 60 investigaciones científicas, libros, artículos periodísticos y blogs extraídos de fuentes digitales repositorios como el *Eprints in Library and Information Science* (E-LIS), la publicadora IGI Global, *ResearchGate*, *Dialnet* y *ScienceDirect*; además de redes sociales científicas *ResearchGate* y profesionales como LinkedIn. Para la redacción de la bibliografía se utilizó la norma APA, automatizada con el editor de texto Word. Este proceso se llevó a cabo desde el mes de octubre de 2019 hasta febrero 2021.
- Entrevista de investigación: fue de índole semiestructurada, aplicando un guión con preguntas abiertas en forma de embudo con el propósito de confirmar y ampliar los hallazgos documentales en cuanto a la gestión de información del entorno (Anexo I). En ocasiones, algunos elementos interrogatorios de control fueron agregados para encarrilar al entrevistado al tema en cuestión. Para ello se utilizaron plataformas de videoconferencias a distancia como Google Meet, Zoom y WhatsApp; esta tarea se realizó desde el 28 de julio hasta el 10 de septiembre de 2020. Quince personas fueron entrevistadas, a dos se les remitió por correo electrónico las preguntas por problemas de agenda; todos participaron de manera telemática desde España, Brasil, República Dominicana y Estados Unidos.

El mencionado guión fue adaptado a las áreas de estudios y experiencias de los encuestados sobre las herramientas y técnicas que utilizan para gestionar la información, a fin de captar sus reacciones y recomendaciones de manera abierta, por ejemplo, en el caso del profesional informático se le plantearon preguntas del tipo: “¿Cuáles son las tecnologías más recomendadas para gestionar la información del entorno en una empresa?”; al documentalista: “¿mediante qué actividades la documentación facilitará la gestión de información en una empresa?”; al comunicador: “¿Siguen siendo efectivos los recortes de prensa para gestionar la información del entorno?”. En otros casos se aplicaron preguntas diferentes conforme a la necesidad de información

Como informantes clave se consideró a los directivos de la empresa Arnaiz & Partners, cuya identidad fue anonimizada en la transcripción de sus respuestas así como los datos correspondientes al resto de entrevistados entre los que se incluyeron profesionales de la documentación por ser los conocedores de los métodos de gestión y preservación de los archivos; además de informáticos y científicos de datos por su amplio dominio de software para optimizar la extracción, clasificación y uso de lenguajes de programación, así mismo, comunicadores especializados en el monitoreo de medios de comunicación y asuntos públicos.

Para el procesamiento de toda esta información se utilizaron plantillas de recopilación de datos, parámetros de selección de entrevistados, grabación de audios de declaraciones de entrevistados, transcripción de textos a tabulación computacional, correo electrónico para el envío y recepción de cuestionarios, comparación e interpretación de datos.

Las preguntas y respuestas fueron tabuladas y se facilitan como anexo en este trabajo, los contenidos considerados irrelevantes para la investigación fueron descartados. La mayor parte de la información recabada fue analizada en el capítulo 4 de análisis de resultados, otras informaciones de índole teórica e institucional fueron incorporadas a secciones como el Estado de la Cuestión, un apartado dedicado a la institución estudiada en este trabajo y otros capítulos pertinentes.

Para el análisis de los datos cualitativos recogidos mediante la entrevista se escogieron tres categorías generales vinculadas a los objetivos de investigación y se establecieron una serie de subcategorías que permitirán analizar de manera detallada las opiniones de los informantes claves (Tabla 2). A las categorías preestablecidas por derivar de los objetivos se agregó una cuarta, cuyas subcategorías son consideradas serendipias, es decir, informaciones de alto interés para la investigación que no se tenían previstas. Es muy común en el desarrollo de actividades web o colaborativas donde se adquiere abundante información y puntos de vista.

Tabla 2: Sistema de categorías y subcategorías para el análisis de las entrevistas

Categorías	Subcategorías
Temas de interés estratégico para la actividad de la empresa Arnaiz & Partners.	Clasificación de la actividad de Arnaiz & Partners en temas.
Criterios para la selección de fuentes de información textual en Iberoamérica.	Cantidades de información que requieren técnicas y tecnologías especializadas para su gestión.
	Criterios de selección instituciones y medios como fuentes de información.
	Perfil profesional del gestor de información
Métodos y tecnologías para monitorear el entorno.	Importancia que las empresas otorgan a la gestión de información aplicando tecnologías.
	Facilidad que proporcionan las tecnologías para conocer el entorno.
	Tecnologías y técnicas para gestionar información textual.
	Entrega de resultados de análisis de información del entorno empresarial
Serendipia	Consecuencias de no implementar sistemas de información para conocer el entorno.
	Retos vinculados a la gestión de texto.

Cada subcategoría contiene una serie de datos elementales para el diseño de un proceso de gestión de información, cada una es desarrollada en la tabla 3.

Tabla 3: Definiciones de subcategorías para el análisis de las entrevistas

1. Temas de interés estratégico para la actividad de la empresa Arnaiz & Partners	
Subcategorías	Definición de subcategorías
Clasificación de la actividad de Arnaiz & Partners en temas.	Elaborar un listado de palabras que permita clasificar las actividades y temas de interés de la empresa Arnaiz & Partners para la búsqueda de información y la organización del conocimiento.
2. Criterios para la selección de fuentes de información textual en Iberoamérica	
Subcategorías	Definición de subcategorías
Cantidades de información que requieren técnicas y tecnologías especializadas para su gestión.	Conocer los niveles de datos digitales a partir de los cuales una institución debe plantearse establecer sistemas de información formales.
Criterios de selección de instituciones y medios como fuentes de información.	Consideraciones para seleccionar las instituciones y medios donde se originan los datos del entorno. Estando esto relacionado con la actividad y la clientela de empresa.
Perfil profesional del gestor de información.	Identifica la preparación profesional, habilidades y experiencias para la gestión de información o datos en empresas.
3. Métodos y tecnologías para monitorear el entorno.	
Subcategorías	Definición de subcategorías
Importancia que las empresas otorgan a la gestión de información aplicando tecnologías.	Describir la relevancia que las empresas dan a las tecnologías para la administración de los datos del entorno desde la perspectiva profesional y experiencial de los entrevistados.
Facilidad que proporcionan las tecnologías para conocer el entorno.	Ventajas de la aplicación de tecnologías en la gestión de información del entorno para las instituciones.
Tecnologías y técnicas para gestionar información textual.	Propuesta de herramientas tecnológicas y prácticas documentales para el manejo de la información textual.
Entrega de resultados de análisis de información del entorno empresarial.	Recomendaciones para la presentación de reportes vinculados al análisis del entorno empresarial.
4. Serendipia	
Subcategorías	Definición de subcategorías
Consecuencias de no implementar sistemas de información para conocer el entorno.	Pronostica lo que ocurrirá a las instituciones que no implementen técnicas y tecnologías adecuadas para conocer su ambiente de negocios.
Retos vinculados a la gestión de texto.	Avances pendientes para facilitar la eficiente gestión de información de texto para las instituciones.

Otras informaciones de índole empresarial y conceptual recogidas para fortalecer el cuerpo teórico de la presente investigación fueron excluidas de este capítulo. Los comentarios a continuación colocados son un extracto del colocado en los anexos, buscando responder de manera concreta las expectativas de información que se tenían.

4. Resultados y discusión

4.1. Categoría 1: Temas de interés estratégico para la actividad de la empresa Arnaiz & Partners.

Considerando las opiniones de la gerencia de la consultora Arnaiz & Partners, su estructura y las fuentes documentales consultadas se ha elaborado un listado con un conjunto de términos que, sin ánimo de que sean usados como tesaurus, han sido organizados de manera jerárquica. Los sistemas de información para el monitoreo del entorno requieren de descriptores para la búsqueda de información y la organización del conocimiento, y de ser requerido, para el entrenamiento de los programas y buscadores de información.

TG: Desarrollo territorial	TE: Centro América y El Caribe
TG: Arquitectura	TR: Cuba
TE: Tecnología urbanística	TR: República Dominicana
TE: Planificación urbana / urbanismo	TR: Haití
TE: Medio ambiente	TR: Puerto Rico
TG: Desarrollo inmobiliario	TR: Jamaica
TE: Promoción	TR: Belice
TE: Construcción	TR: Costa Rica
TG: Gestión estratégica territorial	TR: El Salvador
TE: Consultoría municipal	TR: Guatemala
TE: Consultoría a gobiernos nacionales	TR: Nicaragua
TE: Consultoría a gobiernos regionales	TR: Panamá
TG: Formación académica y profesional	TE: Comunidad Andina
TE: Proyectos urbanísticos	TR: Venezuela
TE: Información territorial	TR: Bolivia
TE: Estudios territoriales	TR: Colombia
TG: Regiones y países	TR: Ecuador
TE: Portugal	TE: Cono Sur
TE: México	TR: Paraguay
TE: España	TR: Uruguay
TE: Brasil	TR: Perú
	TR: Argentina
	TR: Chile

4.2. Categoría 2: Criterios para la selección de fuentes de información textual en Iberoamérica.

Los entrevistados reconocen que la cantidad de archivos que producen las instituciones, en el ejercicio de sus competencias, hace imprescindible el uso de tecnologías de la información. Para el monitoreo de las instituciones que componen el ámbito de la empresa, en este caso la consultora Arnaiz & Partners, deben tomarse en cuenta criterios asociados a la entidad para la que se gestiona la información, los clientes que componen su entorno y las fuentes que se utilicen a partir de una investigación diagnóstica. Los medios digitales han permitido que en ocasiones el mismo cliente también sea la fuente. Los criterios más comunes a considerar son los siguientes:

CRITERIOS RECOMENDADOS POR LOS ENTREVISTADOS PARA SELECCIONAR LAS FUENTES DE INFORMACIÓN		
Objeto	Criterio	Definición
Entidad	Interés de los usuarios	Los consumidores deben estar a gusto con el contenido y formato de los servicios informativos.
	Normativas y contratos	Evolución del marco legal y acuerdo de servicio, según las restricciones y libertades que brinda.
	Mercado y coyuntura	Cambios en la dinámica del negocio por razones de rentabilidad, contingencias internas o externas.
	Producto:	Tipo de servicio o producto que se busque promover.
	Pensamiento estratégico	Naturaleza de la institución, su misión y visión.
	Capacidad	Involucra los recursos humanos, la preparación profesional, arsenal tecnológico, posibilidades financieras y ética de la institución para dar el servicio.
Clientes	Calidad del cliente	Es el nivel de apertura, accesibilidad física o digital y la rentabilidad económica, de relación o reputación del potencial cliente para recibir los servicios.
	Relación	Fortaleza de vínculos con el potencial cliente y medio.
Fuentes	Credibilidad de la fuente	Nivel de confianza de los datos producidos por la fuente.
	Calidad de la fuente	Es proporcional al nivel de experiencia, los métodos utilizados para producir información, la importancia que el entorno otorga a sus datos, la permanencia de sus datos (URL) y formatos para su verificación; así como la accesibilidad a los recursos digitales.
	Producción de información	Cantidad de información que produzca la fuente.
	Difusión	Alcance de las informaciones publicadas, ya sea en función de su impacto o promoción.
	Estructura de datos	Las informaciones pueden estar en formatos multimedia, datos estructurados o tipo de texto como científico, periodístico, técnico, jurídico, entre otros.

Tabla 4: Criterios recomendados para seleccionar las fuentes de información

En cuanto al perfil del gestor de información, de manera unánime, los entrevistados consideran que el profesional de esta actividad debe ser multidisciplinar, con nociones de las disciplinas con las que trabaja, destacando además habilidades de documentación e informática. Algunos recomendaron, a quienes profundicen en la inteligencia de negocios y ciencia de datos, adquirir habilidades de estadística, matemáticas y programación.

4.3. Categoría 3: Métodos y tecnologías para monitorear el entorno.

Aunque la mayoría de los especialistas reconoce que no se están implementando métodos y tecnologías eficaces para la gestión de información, tienen una visión optimista en cuanto a esta de cara al futuro. Con el propósito de procesar y organizar gran cantidad de datos en tiempo real como soporte para la toma de decisiones, fueron recomendados los métodos y tecnologías siguientes:

Métodos

- Aunque exista un personal responsable de monitoreo, toda la organización debe estar orientada para reportar informaciones que considere de interés y sus impresiones sobre el servicio de información.
- El trabajo de clasificación es idóneo hacerlo de manera manual, las herramientas informáticas son potentes, pero aún tienen muchos errores.
- Integración de los medios sociales, periódicos digitales y otras fuentes en el *Clipping*. Mucha de esta información también es posible comprarla a consultoras.
- Deben consultarse las fuentes que estén en la región meta o la localidad del cliente, sea a nivel nacional e internacional.

Tecnologías

- Gestor de Relaciones con los Clientes (CRM): permite registrar las interacciones con los clientes y los datos que se generan para el análisis de las relaciones integrando todo el equipo y medios sociales.
- Herramientas de búsqueda y recuperación automática: versión comercial Google, Apache Solr (uso libre) y Elasticsearch. Las dos últimas son útiles también para automatizar la lectura de textos. Estas herramientas no clasifican la información.
- Lenguajes de programación: *R*, *Python*, *Java*, *SQL*. Permiten realizar *scraping*, crear recolectores, aplicar técnicas de recuperación y clasificación de información, además de detección de *topics*.
- Herramienta de visualización de datos: *Power BI*, *Tableau* y *Pentaho*. Permite acceder a múltiples fuentes de datos en línea. Admiten y exportan datos en hojas de cálculo, incluyendo el formato CSV, estos funcionan mejor con datos estructurados.

- Hojas de cálculo (preferible formatos CSV, ODS y *Excel*) y *OpenRefine*: las tablas siguen siendo las herramientas de estructuración de datos y migración de datos por excelencia. La última permite limpiar los datos, identificarlos y cruzarlos con agilidad y sencillez.
- API's y productos de Google - *Cloud, Forms, Chrome, Maps*; Hojas de cálculo, *Drive, My Bussines, Analytics*; Alertas, *Meet*: el uso combinado de estas herramientas de manera gratuita ofrece un potencial enorme para la gestión de información. Más aún cuando se combina con herramientas de pago que integran funciones de inteligencia artificial para el Lenguaje Natural para análisis de opinión y clasificar textos no estructurados; convertidor de voz a texto y de texto a voz con alta calidad; detectores de sentimientos mediante textos, entre otros.
- Redes sociales y aplicaciones de monitoreo: para dar seguimiento a comentarios de la comunidad Web de la empresa.
- Sistemas de gestión y alojamientos de archivos: *Box, Dropbox* y *OneDrive* son plataformas en línea (nube) para la administración de documentos.

Una vez analizados los resultados, es indispensable ponerlos al alcance de una serie de receptores o usuarios. Los entrevistados sugieren, desde la no entrega de informes confiando en la capacidad del personal para gestionar su propia información (incluyendo la no gestión de información del entorno), hasta actividades más complejas. Consideran que el Departamento de información debe tener libertad para compartir los informes con los directivos en igualdad de condiciones, así como para recibir y hacer sugerencias:

- Entrega periódica de informes: semanal o previo acuerdo entre usuarios y gestor de información. En el caso de las instituciones públicas, este puede ser diario.
- Difusión selectiva de información: para la gerencia un informe sencillo y breve, mientras que el personal técnico puede tener datos más detallados.
- Tiempo de discusión: establecer en la agenda de algunas reuniones momentos de discusión de las informaciones recibidas vinculadas al accionar de la empresa.
- Forma de entrega: desde el envío de documentos tradicionales por correo electrónico, hasta la creación de un gestor de contenidos con resúmenes ampliables mediante un enlace o la impresión de un informe con datos graficados. El usuario o receptor debe sentirse conforme con los informes, acorde con la capacidad técnica de los gestores de la información.
- Accesibilidad: Los usuarios deben tener acceso permanente a los datos e informes mediante sistemas de información en línea.
- Retroalimentación: Los gestores de información deben responder las necesidades de información concreta de los usuarios, de los cuales también se tomarán sugerencias para mejorar la calidad del servicio.

4.4. Categoría 4: Serendipias

Los entrevistados pronostican que las instituciones que no implementan métodos y técnicas de gestión de información pueden incluso dejar de funcionar en algún momento o perderán competitividad. También los profesionales que se dediquen a esta tarea deben adaptar la gestión de información a la actividad e intereses de la empresa y utilizar tecnologías de inteligencia artificial. Las plataformas digitales suelen variar con frecuencia, siendo necesario actualizarlas, sobre todo, si la información es recogida de manera automática. Tomar en cuenta las redes sociales en esta dinámica es esencial.

5. Propuesta de diseño de un servicio de gestión de información

Los sectores económicos locales e internacionales en los que Arnaiz & Partners desarrolla su actividad evolucionan constantemente. La administración pública es importante para la empresa, porque además de impactar su actividad administrativa, es su principal cliente. La discusión de políticas, la promulgación de normativas y el desarrollo de proyectos por las instituciones del poder ejecutivo y legislativo influyen en la adaptación de los productos y servicios que se ofrecen.

Adicionalmente, Arnaiz & Partners pone interés en las instituciones educativas, multilaterales y privadas; los medios masivos y especializados, entre otras. Ante la inmensa cantidad de información que producen las citadas entidades, las cuales constituyen una parte importante del entorno empresarial, solo implementando tecnologías y técnicas de gestión de información con sentido estratégico será posible utilizar los datos para añadir valor a la oferta empresarial y su marca.

Profundizar en la importancia de la gestión información mediante entrevistas de investigación a académicos y especialistas de las Ciencias de Datos, Profesionales de la Información, Ciencias de la Computación y de la Comunicación, permitió cumplir el objetivo de elaborar un diagrama de flujo fundamentado en objetivos específicos, medibles y alcanzables. Además, analizando las prácticas presentes y futuras del gestor de información, se propusieron criterios para la selección los datos de texto y el uso estratégico la información.

5.1. Objetivos para la gestión de información en Arnaiz & Partners

Los propósitos de la gestión de información son determinantes para el desarrollo de los procedimientos de trabajo, la planificación de los equipos de trabajo y tecnologías que sostengan la actividad en el tiempo. Algunas recomendaciones son:

- Apoyar la toma de decisiones vinculada a la gestión de relaciones, proyectos, administración y comercialización de Arnaiz & Partners y sus delegaciones internacionales mediante recursos informativos sobre el entorno empresarial.
- Proveer acceso al conocimiento de manera precisa, dinámica y acorde con las necesidades informativas de Arnaiz & Partners en sus distintos niveles de estructura.
- Implementar estrategias y software para la efectiva monitorización de entidades que puedan generar oportunidades de negocio, analizar la evolución de los mercados y sus tendencias, prevenir y manejar crisis.
- Añadir valor a los proyectos, servicios e incidencia de la empresa acorde con los requerimientos de su entorno.

5.2. Selección de fuentes de información para Arnaiz & Partners

La calidad de los datos está vinculada a la idoneidad de las fuentes de información. Considerando la naturaleza de Arnaiz & Partners y su actividad económica, las fuentes de datos textuales iniciales serían las siguientes:

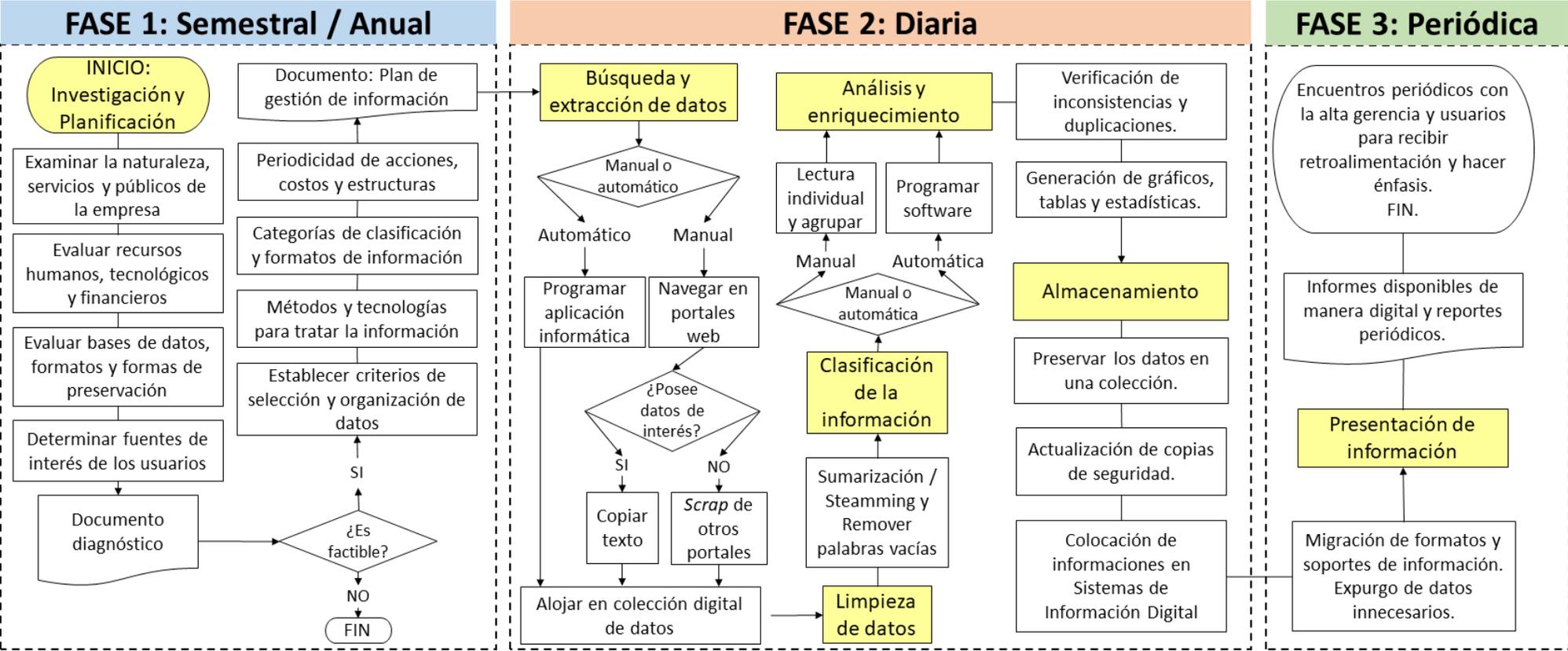
- Instituciones públicas.
 - o Gobiernos nacionales, regionales y locales: además de sus principales ejecutivos, incluye la actividad y concursos públicos de ministerios, direcciones y funcionarios en asuntos de interés empresarial.
 - o Órganos legislativos: monitoreo de alcance nacional, regional y local de las normativas que incidan en los aspectos de negocio o administrativos de la empresa.
 - o Asociaciones municipales nacionales, regionales e internacionales.
- Medios de Digitales.
 - o Prensa digital.
 - o Medios especializados y revistas científicas.
- Universidades y ONG: instituciones que suministran innovación e investigaciones.
- Alusión a Arnaiz & Partners: directa o indirecta, así como a sus integrantes.
- Competencia.

Con excepción de los órganos legislativos, únicos en cada país, existe una prolífica cantidad de instituciones cuya selección precisa de una serie de criterios que harán la gestión de información más específica. (Tabla 6).

Tabla 5: Propuesta de criterios para la selección de fuentes externas

CRITERIOS PARA LA SELECCIÓN DE FUENTES DE INFORMACIÓN		
Objeto	Criterio	Definición
Características de la fuente	Tipo	Las fuentes de interés son las instituciones públicas, privadas y actores destacados en los temas de interés empresarial, incluyendo sus cuentas o perfiles de medios sociales. También incluye entidades educativas, de investigación o revistas científicas, así como relevantes medios especializados y de prensa con presencia digital.
	Navegabilidad	Amabilidad del portal para ser explorado, proveyendo fácil acceso a sus recursos informativos. Baja cantidad de enlaces rotos y ventanas emergentes. Volúmenes aceptables de datos en formato accesibles.
	Credibilidad	Confianza en las publicaciones por los métodos de gestión de datos utilizados, la valoración que reciba de su audiencia meta y los públicos empresariales; su oficialidad institucional y su fundamento en datos.
	Continuidad de publicaciones	Las fuentes deben publicar un mínimo de información de interés: los medios de comunicación digital 15 contenidos mensual, las fuentes científicas y especializadas 05 anual. Mientras que las instituciones oficiales y privadas se monitorean en función del debate público.
Aspectos institucionales	Recursos	Contar con los medios técnicos, financieros y humanos suficientes para el monitoreo y análisis de las fuentes.
	Relevancia	Preferencia de la fuente por su valor técnico, científico, periodístico, jurídico y acorde con la orientación comercial de la empresa.
Legalidad y ética	Scraping autorizado	No desautorización expresa para extraer contenidos de la fuente utilizando herramientas robot.
	Normativa	Considerar las restricciones de las legislaciones locales sobre la extracción de contenidos.

5.3. Diagrama de gestión de información digital



5.3.1. Descripción ampliada del diagrama de flujo

Se describen a continuación de manera detallada aquellas tareas principales del diagrama de flujo: 1) Investigación y Planificación, 2) Búsqueda y extracción de Datos, 3) Limpieza de datos, 4) Clasificación de la información, 5) Análisis y enriquecimiento, 6) Almacenamiento, y 7) Presentación de información.

1. Investigación y Planificación

El propósito es evaluar todos los factores que permitirán una gestión de información de calidad. Se estudia la naturaleza empresarial, los intereses y aptitudes de los usuarios, las normativas internas y externas.

Se establecen los objetivos, procedimientos y tecnologías para la búsqueda, tratamiento, conservación y suministro de información; también se definen las fuentes de información y su forma de clasificación; la periodicidad de estas acciones, los formatos de almacenamiento y los metadatos.

Son definidos los perfiles laborales, sus responsabilidades y su asignación presupuestaria. Las iniciativas para involucrar el personal de la institución y concienciarle sobre la importancia de atender, retroalimentar y aportar información también es pertinente.

2. Búsqueda y extracción de datos

La identificación de los datos textuales puede realizarse de manera automatizada o manual, acorde con las restricciones legales y de las fuentes. El personal empresarial también debe ser motivado a remitir a los gestores de información los contenidos cuyo análisis considere relevante para la institución. La periodicidad de esta actividad se establece en la fase de planificación.

Pueden utilizarse técnicas tradicionales de búsqueda (métodos booleanos) y, aunque con poca eficiencia, la extracción manual de información usando navegadores cuando la descarga con aplicaciones informáticas produzca una pérdida considerable de datos o si las normativas prohíben el uso de las tecnologías de *scrapping*. Sobre estas últimas, se recomiendan las siguientes:

- *Google Cloud Natural Language API*: extrae información de personas, lugares, eventos y mucho más, que sean mencionadas en documentos de texto, artículos de noticias o publicaciones web utilizando procesadores de lenguaje natural. Impone una estructura que facilita el análisis de los datos.
- *WordStat*: es un software avanzado de análisis de contenido y minería de texto. Debería ser usada por alguien que necesita rápida extracción y análisis de información de una gran cantidad de documentos. Utiliza sistemas de clasificación *N-bayes*, *K-nearest neighbour*, *steamming*. Útil en 18 idiomas.

- Alertas de google: avisos de actividad en la web sobre términos preestablecidos y con periodicidad controlada. Son enviados por correo electrónico.

Adicionalmente, pueden emplearse lenguajes de programación, capaces de leer la estructura de códigos HTML de los portales web mediante un robot (*crawler*), siendo los más recomendados:

- *Rstudio*: lenguaje de programación de código abierto ampliamente aplicado para ciencia de datos, estadística computacional y gráficos.
- *Python*: lenguaje de programación de código abierto caracterizado por ser intuitivo en el desarrollo de aplicaciones y ciencia de datos.

Los datos de texto pueden ser recopilados en software y sistemas como:

- Procesadores de texto plano: los más populares son el Bloc de notas y *Notepad++*.
- Sistemas de Gestión de contenidos: *Joomla*, *WordPress*, *Blogger*, entre otros.
- Colecciones estructuradas de información: como *MongoDB* o *MySQL* (Tabla 5); en un sistema de relaciones con los clientes (CRM, por sus siglas en inglés); y hojas de cálculo formato *csv*, *ods*, *xlsx* y *html*. En este caso, sería indispensable que los datos sean organizados según:
 - o categoría, fecha, texto, nombre de la fuente, metadatos, países aludidos, entidades involucradas, URL, autor y espacio de notas.

ESTRUCTURACIÓN DE DATOS DE TEXTO EN BASES DE DATOS RELACIONALES						
Fuente	Información textual	Categoría	Etiquetas	Fecha	comentario	URL

Tabla 6: Estructuración de datos de texto en bases de datos relacionales

3. Limpieza de datos:

Para facilitar la futura clasificación y análisis de la información según los métodos que se vayan a utilizar, como lenguajes de programación, inteligencia artificial o procedimientos manuales, son aplicadas las técnicas siguientes:

- *Stopwords* (palabras vacías): son removidas las palabras, caracteres, mayúsculas, hipervínculos y signos que puedan representar una dificultad para que algunos sistemas informáticos. También son eliminadas las terminaciones en plural, tiempos verbales y género (*stemming*). Según el sistema que colecciona la información las herramientas para purgar estos elementos pueden variar: expresiones regulares, lenguajes de programación y procesadores de lenguaje natural.

- Sumarización: es la extracción de la información de mayor interés en un texto, eliminando también las informaciones innecesarias. Con este propósito, suelen utilizarse herramientas de Procesamiento de Lenguaje Natural.
- Otros: deben evaluarse los problemas de sinonimia, polisemia, homonimia, etcétera.

4. Clasificación de la información

Las informaciones son agrupadas en las categorías predefinidas y se asignan etiquetas para organizar, recuperar, interpretar y preservar los datos. Esta actividad puede realizarse implementando las técnicas siguientes:

- Clasificación manual: común cuando la información es coleccionada en hojas de cálculo o archivos de texto. Los datos de texto son leídos de manera individual y agrupados según los criterios del gestor de información.
- Software de clasificación automática: los documentos se reúnen usando algoritmos, cuando están previamente entrenados con palabras clave se considera que la clasificación es supervisada, mientras que cuando los contenidos son analizados sin entrenamiento previo la clasificación es no supervisada.
- Algunos software son:
 - o *IBM SPSS Text Analytics*: para transformar los datos textuales de las encuestas en información cuantitativa útil para el análisis de sentimientos. Automatiza el proceso de categorización para eliminar el tiempo de una codificación manual. Usa procesadores de lenguajes naturales para reducir la ambigüedad del lenguaje humano, ayudando a descubrir patrones de actitudes, creencias y opiniones de las personas.
 - o *Microsoft Azure Text Analytics API*: detecta sentimientos, palabras clave, tópicos y lenguajes de un texto mediante técnicas de categorización. Incluye funciones de clasificación como *N-grams*. Funciona con idiomas como inglés, francés, español y portugués.

5. Análisis y enriquecimiento

Los datos, procedentes de una heterogénea cantidad de organizaciones, son homogenizados en una estructura que pueda estar completa o no, según los métodos y tecnologías que precedan a esta fase.

Herramientas como *OpenRefine* facilitan la integración de los textos ya estructurados para evitar inconsistencias y duplicaciones. Otras aplicaciones como *Tableau* o *Power BI*, también permiten crear gráficos, tablas y estadísticas para la presentación de los datos.

Las modificaciones de los datos y su estructura en esta fase no son automatizadas. El gestor de información debe tener un amplio dominio del tema para descubrir datos que no están presente de manera explícita en ningún documento y que no busca expresamente. Se extrae fruto de la relación y comparación adecuada de documentos

(Justicia de la Torre, 2017) para mejorar servicios, aprovechar oportunidades, identificar o pronosticar comportamientos en el mercado y prevenir crisis. Se localizan patrones, hechos, tendencias, lugares, nombres.

6. Almacenamiento.

La colección de información debe ser preservada en una base de datos, sea esta relacional (*MySQL*), no relacional (*MongoDB*) o en hojas de cálculo. Con fines de seguridad, se recomienda hacer copias de los datos en un servidor local o en la nube y crear copias periódicas de seguridad en un sistema de almacenamiento en la nube. La migración de formatos y soportes, así como el tiempo de conservación de los datos y su forma de expurgo, deben ser ejecutadas acorde con el plan de gestión de información.

Finalmente, los informes de texto y gráficos elaborados mediante el análisis de los datos, son colocados de manera permanente en un sistema de gestión de contenidos como *Joomla* o *Wordpress*, los cuales deben clasificar la información según los departamentos y sus niveles de responsabilidad exigiendo una identificación. El gestor de contenidos debe permitir recuperar las informaciones por etiqueta y descargar ciertos datos de manera editable.

Con un perfil de ingreso, los usuarios podrán utilizar manera permanente la información del entorno empresarial con valor añadido, de manera simple y abierta para ser reutilizada en proyectos, planes y en la toma de decisiones.

7. Presentación de la información:

Los usuarios deben recibir informes periódicos de las actualizaciones disponibles en el sistema de información de contenidos (semanal preferiblemente), estos deben ser breves, simples y tener un respaldo gráfico de alta calidad y simplificación.

El gestor de información también debe solicitar espacios en las reuniones directivas para ser retroalimentado por los responsables principales, enfatiza cambios en el entorno de interés para la empresa y recomienda usos de los contenidos para fines de formación, difusión o desarrollo de recursos.

5.4. Personal responsable de la gestión de información en Arnaiz & Partners

Se recomienda a Arnaiz & Partners contar con un especialista en gestión de información, responsable de monitorizar y analizar el entorno empresarial a nivel nacional e internacional. Este deberá mantenerse en constante interacción con la dirección empresarial, en sus diferentes niveles, para atender de manera precisa sus necesidades de información.

Trabjará en permanente asociación con el personal de las áreas de comunicación e informática. Preferiblemente, debe encabezar un equipo integrado por un especialista en ciencias de la computación y otra/o de urbanismos (Ilustración 1). El perfil de este profesional de la información debe ser el siguiente:

Formación y experiencia

- Diplomado a nivel superior en una de las áreas siguientes: ciencias de la información, ciencias de la computación o economía.
- Alguna certificación en temas vinculados al territorio es deseable.
- Tres años de experiencia en de gestión inmobiliaria y/o desarrollo territorial.
- Haber ejecutado tareas de documentación, gestión de información, inteligencia de negocios o ciencias de datos por un mínimo de dos años.

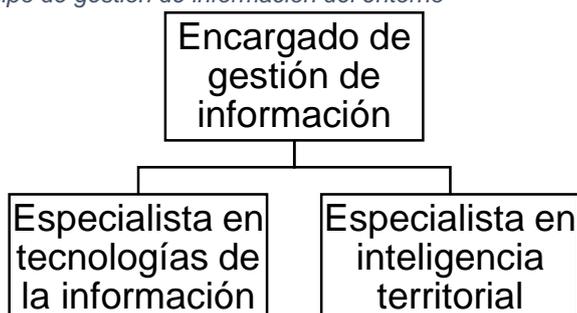
Aptitudes

- Alto nivel en investigación, análisis y organización de información.
- Conocimiento social, político y económico iberoamericano.
- Pensamiento estratégico en negociación y servicios.
- Estadística.
- Comunicación oral, contenidos de texto y gráficos estadísticos.
- Idiomas: español (avanzado), inglés y portugués (intermedio).

Tecnologías

- Manejo avanzado de navegadores web y gestión de redes sociales.
- Manejo intermedio de gestores de contenidos de información (CMS).
- Análisis de datos intermedio (*Python o R; Power BI o Tableau*; bases de datos *SQL o MongoDB*).
- Nivel intermedio en ofimática: *Excel, Word y Power Point*.

Ilustración 1: Potencial equipo de gestión de información del entorno



6. Conclusiones

El objetivo del presente trabajo era realizar una propuesta de diseño para la gestión eficaz y eficiente de información del entorno de la empresa Arnaiz & Partners. Para ello, se ha reunido la opinión, así como la experiencia en el uso de tecnologías y la aplicación de métodos para la gestión de información de múltiples profesionales de la información, ciencias de datos y especialistas en información estratégica de negocios recogida mediante una serie de entrevistas y una importante revisión bibliográfica.

Como resultado, se ha identificado que la consultora urbanística Arnaiz & Partners no tiene un sistema para gestionar los datos textuales de su mercado nacional e internacional. Las técnicas y tecnologías implementadas son informales, insuficientes y no uniformes para analizar de manera dinámica la gran cantidad de datos que produce su entorno. Aunque la Dirección de la empresa no considera de manera unánime la pertinencia de una estrategia de inteligencia de negocios los hallazgos durante el proceso de investigación y las opiniones de los expertos sugiere que el monitoreo de los *stakeholders* es una necesidad para las empresas del presente y del futuro si no quieren perder competitividad o desaparecer del mercado.

Se desarrolló un diagrama de flujo flexible y apegado a la actividad de la empresa para la selección, la clasificación, el análisis, la presentación y preservación de la información del entorno, tomando en cuenta las características de las fuentes, los aspectos institucionales, la legalidad y ética de la gestión de información según criterios técnicos detalladamente definidos.

El propósito es poner a disposición información digital, usable, accesible, segmentada, visual y acorde con los intereses de los usuarios. Para la búsqueda y organización de todo el conocimiento Iberoamericano afín con la actividad de Arnaiz & Partners, se recomendaron 44 términos, útiles también para la programación de las aplicaciones informáticas que se empleen.

La propiedad de los datos y su adecuada explotación mediante las tecnologías es prioritaria a un insustancial lago de datos. Por eso, los principios fundamentales para proveer inteligencia competitiva a la empresa son:

- La eficiencia en la gestión de los datos aplicando software informáticos que permitirán el manejo de los grandes volúmenes de información de texto utilizando programas de lenguajes de programación, procesadores de lenguaje natural, bases de datos y sistemas de gestión de contenidos.

Ciertamente las aplicaciones informáticas han demostrado su potencial para facilitar la gestión de grandes volúmenes de datos, aun así, existen condiciones éticas e informáticas cuando el uso de ciertos software no está sugerido, en estos casos, según la importancia de la fuente, se recomienda utilizar o combinar métodos tradicionales para adquirir los datos de interés.

- La calidad de los datos, proporcionales al nivel de relación de estos con la naturaleza y servicios de la consultora, así como las necesidades de información de los usuarios según su nivel en la organización y los criterios para identificar las mejores fuentes.

También se logró presentar el perfil del gestor de información cuando incide directamente en la toma de decisiones, el desarrollo de proyectos, la gestión de crisis y la construcción de relaciones, el cual debe guardar cercanía con la dirección empresarial y sus múltiples delegaciones internacionales para proveer información acorde con los objetivos y la filosofía empresarial. Adicionalmente, debe tener habilidades de documentación, estar abierto a trabajar con profesionales de bibliotecología, informática, comunicación, mercadología e incluso ciencias, así como aptitudes en tecnologías de la información, bases de datos relacionales y no relacionales; sistemas de gestión de contenidos, entre otras herramientas.

La implementación del ya mencionado diagrama de gestión de información y el alcance de los objetivos, proporcionará a la consultora de una vasta cantidad de información digital, accesible y abierta para adaptar los servicios a las necesidades de cada mercado y en el desarrollo integral del negocio. Esta estrategia de inteligencia competitiva permitirá identificar patrones y prever condiciones que complementarán significativamente las herramientas de gestión de información geográfica que actualmente realiza Arnaiz & Partners, para que instituciones generadoras de mapas compartan datos vinculados al territorio; pues como afirma el Director de Arnaiz & Partners en España, Luís Arnaiz (2020), las empresas y gobiernos no van a reconocer sus necesidades, por lo que es necesario ser lo suficiente hábiles para detectarlas.

Por tanto, la gestión de datos de texto añade valor en la obtención de conocimiento de la consultora, pudiendo considerarse como una estrategia integral de inteligencia territorial, implementando efectivas técnicas y tecnologías para monitorear los contenidos de interés que producen los organismos gubernamentales, multilaterales y de investigación; empresas y medios de comunicación. La mejor descripción de inteligencia territorial es provista por Ana Guzmán (2013): “Dedicado a crear y fortalecer las capacidades de innovación en el territorio, incluye el proceso de vigilancia e inteligencia, que consiste en involucrar información estratégica del entorno y de todas las variables que tienen algún impacto en la competitividad; que luego se transforman en ideas, proyectos y recursos (Guzmán Peña, 2013, página 6).

La combinación de esta dinámica estructura de datos sumada a la vocación académica del Instituto Arnaiz de Estudios e Investigación y el interés de incidencia social manifestado por algunos directivos, abre paso al establecimiento de un Observatorio de Urbanístico y Territorial Iberoamericano, para difundir artículos fundamentados en datos que reconozca las mejores prácticas de desarrollo, promueva la marca de la empresa y fortalezca sus relaciones con el entorno.

Las finalidades de este trabajo de fin de máster pueden considerarse altamente satisfechas. A su vez, abre la puerta a futuras investigaciones para determinar qué cantidad de las informaciones textuales almacenadas son objetivamente útiles para mejorar la toma de decisiones en las empresas, pudiendo incluir también la documentación interna como parte de la gobernación de datos.

Bibliografía

- Abels, E., Jones, R., Latham, J., Magnoni, D., & Gard-Marshall, J. (2003). *Competencias para profesionales de la información del siglo XXI*. Asociación de Bibliotecas Especializadas. Recuperado de <http://dbiosla.org/Competencias%20for%20Information%20Professionals%20of%20the%2021st%20Century.pdf>
- Alcaíno, M., Arenas, V., & Gutiérrez, F. (2015). *Modelos de negocios Basados en Datos: Desafíos del Big Data en Latinoamérica*. Santiago, Chile.
- Arroyo-Vázquez, N. (2014). El content curator. Guía básica para el nuevo profesional de internet. *Revista Española de Documentación Científica*, 2.
- Asociación de Directivos de Comunicación. (2018). *El Estado de la Comunicación en España*. Madrid.
- Bhargava, R. (30 de septiembre de 2009). *Manifesto For The Content Curator: The Next Big Social Media Job Of The Future ?* Recuperado de Rohit Bhargava: <https://www.rohitbhargava.com/2009/09/manifesto-for-the-content-curator-the-next-big-social-media-job-of-the-future.html>
- Cavanillas, J., Curry, E., & Wahlster, W. (2016). *New Horizons for a Data-Driven Economy*. Madrid, Spain: Springer Open. Recuperado de <https://library.oapen.org/bitstream/handle/20.500.12657/27764/1002241.pdf?sequence=1#page=103>
- Cervantes Ramírez, Á. (2017). DERECHOS DE AUTOR DE LAS OBRAS PERIODÍSTICAS EN LA ERA DIGITAL Y EL PRESS CLIPPING. *Revista Republicana*(23), Págs. 39-66. Recuperado de <http://ojs.urepublicana.edu.co/index.php/revistarepublicana/article/view/405/358>
- Chen, H., Chiang, R., & Storey, V. (4 de diciembre de 2012). Business Intelligence and Analytics: From Big Data to Big Impact. *Management Information Systems Research Center, Vol. 36(N. 4)*, pp. 1165-1188. Recuperado de <https://www.jstor.org/stable/pdf/41703503.pdf>
- CyberAlert. (2010). *Media Monitoring: The Complete Guide*. Stratford.
- Debons, A., Horne, E., & Cronenweth, S. (1988). *Information Science: An Integrated View*. Boston: G. K. Hall.
- Dulzaides, M., & Molina, A. (2004). Análisis documental y de información: dos componentes de un mismo proceso. *Scielo*. Recuperado de http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1024-94352004000200011
- Espada, A., Gascón, D., & Morel, S. (18 de junio de 2018). *10 apuntes sobre la posverdad*. Obtenido de Letras Libres: <https://www.letraslibres.com/espana-mexico/politica/10-apuntes-sobre-la-posverdad>
- European Parliament. (2000). *Lisbon European Council 23 and 24 march 2000*. Lisbon.
- Fabian, M., Ramos, L., & Fernández, P. (2018). *Caracterización de pacientes con eventos cerebrovasculares*. Hospital Dr. Francisco Eugenio Moscoso Puello. Santo Domingo.

- Feldman, R., & Sanger, J. (2007). *The Text Mining Handbook*. Cambridge, United Kingdom: Cambridge University Press.
- Franks, P. (2013). *Record and Information Management*. Estados Unidos: Editorial Neal-Schuman. Recuperado de <https://books.google.es/books?id=CNGcAwAAQBAJ&lpg=PP1&hl=es&pg=PR2#v=onepage&q&f=false>
- Gómez de Silva, G. (1988). *Breve diccionario etimológico de la lengua española*. México: México: El Colegio de México, Fondo de Cultura Económica.
- Guzmán Peña, A. R. (2013). *Propuesta de un modelo de inteligencia territorial*. México: Journal of Technology Management & Innovation.
- Higgins, S. (2008). The DCC Curation Lifecycle Model. *International Journal of Digital Curation*, vol 3(Nº 1). Recuperado de <http://www.dcc.ac.uk/sites/default/files/documents/publications/DCCLifecycle.pdf>
- Justicia de la Torre, M. (2017). *Nuevas Técnicas de Minería de Textos: Aplicaciones* (Universidad de Granada. Tesis Doctorales ed.). Granada: Universidad de Granada.
- Kantardzic, M. (2011). WEB MINING AND TEXT MINING. En *Data Mining: Concepts, Models, Methods, and Algorithms* (págs. 300-327). Institute of Electrical and Electronics Engineers.
- Kim, J., Warga, E., & Moen, W. (2013). *Competencies Required for Digital Curation: An Analysis of Job Advertisements*. University of North Texas, College of Information. Texas: International Journal of Digital Curation. Recuperado de https://untlis-ir.tdl.org/bitstream/handle/123456789/6/2013_kim_Competencies%20Required%20for%20Digital%20Curation.pdf?sequence=1&isAllowed=y
- Kotler, P., & Keller, K. (2012). *Dirección de Marketing* (14 ed.). México: Pearson.
- Kwayu, K., Kwigizile, V., Zhang, J., & Oh, J.-S. (2020). Semantic N-Gram Feature Analysis and Machine Learning–Based Classification of Drivers’ Hazardous Actions at Signal-Controlled Intersections. *American Society of Civil Engineers*.
- Malik, K. (05 de febrero de 2017). *Pandemonium*. Recuperado de not post-truth as too many ‘truths’: <https://kenanmalik.wordpress.com/2017/02/05/not-post-truth-as-too-many-truths/>
- Martínez-Uribe, L., & Fernández, P. (2015). Servicios de datos: función estratégica de las bibliotecas del siglo XXI. *El profesional de la información*, n. 2, 184-190.
- Mayeh, M., Scheepers, R., & Valos, M. (2012). Understanding the Role of Social Media Monitoring in Generating External Intelligence. *23rd Australasian Conference on Information Systems*. Geelong.
- Mellinas, J. (2015). *Análisis y aplicaciones de los sistemas de medición de la reputación online de los hoteles*. Recuperado de Research Gate: https://www.researchgate.net/profile/Juan_Mellinas/publication/280095623_Analisis_y_aplicaciones_de_los_sistemas_de_medicion_de_la_reputacion_online_de_los_hoteles/links/55a8c5f808ae815a04217fb1/Analisis-y-aplicaciones-de-los-sistemas-de-medicion-de-la-re

Modaresi, P., Gross, P., Sefidrodi, S., & Eckhof, M. (2017). *On (Commercial) Benefits of Automatic Text Summarization Systems in the News Domain: A Case of Media Monitoring and Media Response Analysis*.

Moreno, Á., Molleda, J., Álvarez-Nobell, A., Herrera, M., Athaydes, A., & Suárez, A. (2019). *Comunicación estratégica y sus retos: fake news, confianza, información para la toma de decisiones, liderazgo y compromiso laboral. Resultados de una encuesta en 19 países*. Bruselas & Madrid: EUPRERA/Dircom.

Nogueira, A. (04 de enero de 2019). Obtenido de Sobre pasados por el 'big data'. *Periódico El País*. Recuperado de https://elpais.com/cultura/2019/01/04/actualidad/1546601759_505183.html

Ortoll, E., & García, M. (2015). *La inteligencia competitiva*. Barcelona: UOC. Recuperado de https://books.google.es/books?hl=es&lr=&id=R47LDAAAQBAJ&oi=fnd&pg=PT3&dq=inteligencia+competitiva&ots=HCESCr5B74&sig=TsdnO0ktnQJEbQMvQvzf4tfP_C0&redir_esc=y#v=onepage&q=documenta&f=false

Osuna, M., & De la Cruz, E. (2010). Los Sistemas de Gestión de Contenidos en Información y Documentación. *Revista General de Información y Documentación*, Vol. 20, 67-100.

Pereira, T., Portilla, I., & Rodríguez, N. (2018). Big data y Relaciones Públicas. Una revisión bibliográfica del estado de la cuestión. *Revista de Comunicación*, 151-166.

Ponjuan Dante, G. (2003). Gestión documental, de información y del conocimiento... puntos de contacto y diferencias. *Revista Ciencias de la Información*, 34. Recuperado de <https://biblat.unam.mx/hevila/Cienciasdelainformacion/2003/vol34/no3/7.pdf>

Prytherch, R. (2000). *Harrod's Librarian's Glossary and Reference Book* (9ª ed.). Aldershot: Gower.

Read, J., & Ginn, M. L. (2016). *Records Management* (Décima edición ed.). Estados Unidos: Editorial Cengage Learning. Recuperado de https://books.google.es/books?id=ui5-BAAAQBAJ&lpg=PR3&ots=-UuJS_Smzu&dq=records%20and%20information%20management&lr&hl=es&pg=PA51#v=onepage&q&f=false

Ríos Ortega, J., & Ramírez Velázquez, C. (2017). *Uso Ético de la Información: Implicaciones y Desafíos*. México, México: Universidad Nacional Autónoma de México. Recuperado de http://ru.iibi.unam.mx/jspui/bitstream/IIBI_UNAM/L134/2/uso_etico_informacion2.pdf#page=176

Robek, B. y. (1995). *Citado por Patricia C. Franks en Record and Information Management*. Estados Unidos.

Rojas, J. L. (2015). *Gestión de las Relaciones Públicas: Antecedentes, Naturaleza y Casos* (4ª ed.). República Dominicana: Editora Búho.

Rojas, R., & Ángel, M. (2005). Relación entre los conceptos: información, conocimiento y valor. Semejanzas y diferencias. *Ciência da Informação*, 52-61. doi:<https://doi.org/10.1590/S0100-19652005000200006>

- Russom, P. (2013). *Managing Big Data*. Washington: Data Warehousing Institute.
- Sanz, E. (14 de noviembre de 2016). El 84% de las personas miente en internet. *El Mundo*. Recuperado de <https://www.elmundo.es/papel/todologia/2016/11/14/5825b011e2704efe038b46a1.html>
- Singh, D. (28 de abril de 2019). *The Role of Data Curation in Big Data*. Recuperado de <https://www.datasciencecentral.com/profiles/blogs/the-role-of-data-curation-in-big-data>
- Tramullas, J. (2016). HANNIBAL AD PORTAS, O LOS FUTUROS PERFILES PROFESIONALES DE LA INFORMACIÓN. *OBSERVATORIO*, 1-7.
- Urbano, C. (2016). Control y evaluación de recursos de información en clave de 'continent' curation. *Anuario ThinkEPI*, pp. 137-141. doi:<http://dx.doi.org/10.3145/thinkepi.2016.29>
- Vargas, D., Ayala, E., & Torres, M. (2013). Diseño de un repositorio para la gestión documental y administración automática de la información digital en las empresas. *Cuadernillo de investigaciones, semilleros de investigación Facultad de Ingeniería*, 38-43.
- Vicente Stenhouse, N. (2018). *HabScraper: herramienta automatizada para la extracción de datos con web scraping*. Universitat de les Illes Balears, Escola Politècnica Superior. Recuperado de https://dspace.uib.es/xmlui/bitstream/handle/11201/151095/Memoria_EPSU1195.pdf?sequence=1&isAllowed=y
- Wilcox, D., Cameron, G., & Xifra, J. (2012). *Relaciones Públicas Estrategias y tácticas* (10 ed.). Madrid, España: PEARSON.
- Zerfass, A., Verhoeven, P., Moreno, Á., Tench, R., & Verčič, D. (2020). *European Communication Monitor 2020*. Brussels: EUPRERA.

Anexos

Anexo I: Guion de entrevista a informantes claves

Plantilla de entrevista sobre cómo eficientizar la gestión de información de texto sobre el entorno en las empresas. Las preguntas son abiertas, y fueron adaptadas al área profesional de los entrevistados.

Informaciones generales

1. Nombre de entrevistado/a.
2. Posición:
3. Nombre de la entidad que representa o está vinculada/o.
4. Naturaleza de la entidad que representa o está vinculada/o
 - a. Institución pública.
 - b. Institución de educación superior.
 - c. Institución documental o bibliotecaria.
 - d. Empresa de servicios de consultoría.
 - e. Empresa o posición de ciencia de datos o gestión de información.
 - f. Otras, especifique: _____

La gestión de información de texto

5. A nivel interno, ¿Cómo determinan las instituciones las informaciones del entorno que son de su interés? (ponga ejemplos).
6. ¿Qué nivel de importancia otorgan las instituciones a las demandas y las novedades de su entorno de interés?
7. ¿Cuáles tecnologías son las más recomendadas para gestionar la información del entorno de interés para las empresas? * **(Especificar)**
8. Al seleccionar la información del entorno o escoger las fuentes para extraerlas, ¿Cuáles criterios considera más relevantes? **(Justificar respuesta)**
9. Al trabajar como gestores de información, ¿Qué cantidad de archivos del entorno debemos esperar semanalmente? ¿Cuánto es mucho para considerar el uso de tecnologías? ¿Cuánto es poco para no requerir de aplicaciones informáticas?
10. Si debiéramos entregar informes periódicos sobre las demandas y novedades de los entornos de interés de la institución, ¿Cuáles características deberían tener?
11. ¿Cuál es el perfil del gestor de información?

Anexo II: Tabulación de respuestas en la entrevista y el cuestionario

RESPUESTAS DE ENTREVISTA 1	
Perfil: Personal de la dirección de Arnaiz & Partners	Fecha: 30/07/2020
Modalidad de entrevista: vídeo conferencia	
<p>1. Arnaiz tiene unos servicios que se adaptan a las necesidades particulares de sus clientes, ya sean inmobiliarios, municipales o de cualquier otro proyecto. ¿Cuáles son los mecanismos para conocer el amplio mercado de la empresa?</p> <p>Tenemos varios problemas relacionados con la pregunta. La primera: somos una empresa que se dedica a muchas actividades distintas puesto que engloba todo lo que afecta el desarrollo urbanístico y la transformación del territorio; la transformación completa urbana afecta a los proyectos de edificación de las ejecuciones de las obras de cualquier tipo de usos, actividades, logística, etcétera. [La empresa] tiene una labor muy concreta también en aspectos de tecnología y desarrolla trabajos de gobernanza y trabajo institucional. Por lo tanto, al tener unas actividades tan amplias, prácticamente no es posible nunca hacer correctamente lo que correspondería a una formación en la representación de todos estos campos, puesto que permanentemente se están actualizando.</p> <p>Por otro lado, tenemos dos problemas añadidos, como la extensión geográfica: trabajamos en distintos países. También la multiplicación de información que llega a las empresas a través de todas las tecnologías de redes, es tan enorme y se está multiplicando de forma tan exagerada que es muy difícil tener un control, ni de la información que recibes y ni parte de la información que tú mismo deberías sintetizar para ser más efectivo; y por último, hay una complicación añadida, la crisis del Covid19 conlleva la pérdida de una crisis económica muy importante y eso nos induce a tener una preocupación mayor, a tener una cierta incidencia en trabajos y en actuaciones concretas que permitan tener un resultado económico.</p> <p>Es decir, aquellas empresas o tantas empresas que tienen mucho equipo de personal, que tienen muchos gastos y que tienen que afrontar el equilibrio para la supervivencia hoy tienen una necesidad nueva que consiste en matizar todos los campos en los que se actúa para concretarlos y a aunar esfuerzos en temas en los que tengan un resultado económico. Hay sectores como notaremos que van a tener y han tenido una afectación absoluta, como por ejemplo el turismo, que afecta tanto en este caso a la República Dominicana, como el propio España; y hay quienes tienen menos afectación, pero al mismo tiempo, las actuaciones que realizan las administraciones también están o van a obligar a hacer una reconsideración de los puntos de inversión y por lo tanto, las empresas tendremos que hacer una reconversión derivada, no solo un número de personas y gastos que se asumen que es muy importantes, sino en ocasiones de los temas en los que uno actúa para tener un resultado de interés que permita conseguir que la empresa sobreviva.</p>	
<p>2. Si yo te preguntara cuáles son los principales temas para la actuación de ARNAIZ tanto en España como en Latinoamérica, o que los clasifiques, ¿Cuáles serían estos temas?</p> <p>Es necesario hacer la consideración previa, porque en estos cuarenta y seis años de experiencia y perfeccionamiento ha ido cambiando muchas veces el enfoque como consecuencia no solamente de un mercado, sino de una actitud con acciones en las que uno se encuentra más cómodo y no siempre, además han sido basadas exclusivamente en la rentabilidad, sino encontrarse satisfecho.</p> <p>Yo creo que no solo debemos considerar tres o cuatro aspectos, sino todo lo que afecta a los cuerpos de transformación del territorio. Es decir, no sólo atender lo que la gente demanda, que no es solo los problemas de planificación o gestión, sino que lo que la gente valora es que el producto que hagas lo entregues terminado, por lo tanto, aquí hay una implicación enorme entre el campo del urbanismo, el campo de la arquitectura y el de la construcción.</p> <p>La segunda es que hoy como consecuencia de la reubicación del dinero, es imprescindible que las construcciones que se hagan más rápida, de mayor calidad al mismo tiempo y a precios muy competitivo. Por lo tanto, como política en nuestra empresa, este es tema importantísimo donde estamos entrando en construcción, puesto que la única manera de poder soportar el gasto, las tramitaciones y precios que pudiesen reducir costes en cada latitud geográfica. Las características</p>	

son distintas, pero estamos ante un esfuerzo importante en la concentración de algo que para el futuro puede tener muchas oportunidades para hacer realidad de estos conocimientos y con esa capacidad empresarial.

El tercer grupo corresponde a la tecnología. Hoy la tecnología se ha convertido en la revolución, nosotros tenemos 30 años trabajando con tecnología relacionada con el territorio y con las ciudades. Y esos 30 años han sido en muchas ocasiones muy poco satisfactorios, porque la tecnología relacionada con el territorio conlleva transparencia, y hay muchos países donde no hay interés. Pero llevamos unos cinco años reflexionando en la realidad que supone que la tecnología tenga acceso a todos los ciudadanos, cosa que no ocurría antes, o porque la tecnología sea prácticamente gratuita situación que tampoco ocurría antes. Son estos cambios tan sustanciales de concepto tecnológico lo que va a permitir impulsar muchos proyectos que afectan al territorio.

En el caso concreto, además de la situación provocada por el Covid19 se ha dado un salto espectacular. Anteriormente había terminologías que la gente no conocía, ahora como consecuencia de los confinamientos la tecnología es un elemento clave para relacionarse y para trabajar. Por lo tanto, todo el mundo ha tenido que hacer un esfuerzo en todos los tipos de edades y en todos países para adaptarse a la tecnología y por lo tanto nosotros tenemos mucha fe en que hagamos un trabajo que sea útil y que cumpla precisamente las pautas que hoy se marcan para esa tecnología gratuita y generalizada con la colaboración de los ciudadanos.

El cuarto punto que podríamos considerar que es interesante es que previamente nuestros conocimientos globales en materia de urbanismo, arquitectura, organización de las ciudades, tecnología y, en definitiva, el territorio, pongamos a disposición estos conocimientos en los campos institucionales. Es decir, en aquellos órganos que los gobiernos que tienen que tomar decisiones importantes en materia de inversión o en materia de prioridad de intervenciones y en este caso afecta no solamente a los municipios, en primera instancia ellos son los que tienen que llevar a cabo una transformación directa de muchos de sus territorios, sino también a los gobiernos nacionales que tienen que propiciar que haya un conocimiento más generalizado y más efectivo de las inversiones y por lo tanto del reparto justo del dinero que hay para que la gente sobreviva en primera instancia y mejore en segunda instancia a su propia calidad de vida.

3. ¿Qué criterios consideras para identificar los temas de interés vinculados a la empresa?

Estos criterios no son estáticos, en mis 46 años de experiencia han cambiado bastante. Las oportunidades estarán vinculadas a las relaciones y las relaciones con gente. La autonomía que pueda mantener la empresa manteniendo antiguos clientes, aunque esto signifique diversificarse haciendo construcción, promoción y urbanismo.

La actitud del servicio, útil para cualquier campo. Un profesional que presta servicios debe hacerlo bien y manejar niveles de conocimiento para que tenga más o menos.

El segundo aspecto, en el que llevo más de 25 años dedicado una parte de mi actividad puntual, es la formación, promoviendo cursos en materia territorial, con la creación de un instituto que hicimos hace ya más de casi 15 años para aglutinar el conocimiento y propiciar la formación de la gente joven y por lo tanto que las empresas hubiera mecanismo permanente de cambio de gente, pero sobre todo de cambio de gente para que los jóvenes tengan un mayor conjunto de oportunidades del conocimiento en cualquiera de los campos en los que estemos trabajando. Por tanto, nuestro futuro también va a estar vinculado precisamente a esa formación que deben tener los jóvenes en materias complejas, como son prácticamente casi todas las que te he comentado en cuanto a la transformación del territorio, como de construcción que es un tema que no ha tenido el desarrollo tecnológico que habría tenido otras actividades en empresas. Pero seguimos construyendo casi en parte como construían los romanos, porque ha habido poca transformación tecnológica en materia de construcción. La tecnología es un tema que cambia día a día y por lo tanto exige una atención permanente.

Y la forma de gobernar es un tema apasionante, independientemente los sistemas políticos que se tengan de la izquierda a la derecha, de los centralismos, etcétera. Todo ello mezclado con los aspectos legales, las competencias y de las capacidades de los responsables políticos y de la solución a los problemas humanos; pueden ser la falta de coordinación contra la corrupción permanente. Lógicamente, a medida que vamos metiendo en temas más complejos, se hacen imposible de solucionar, pero que tienen un componente de actuación. Por tanto, a las preguntas tuyas en concreto: son la prestación del servicio, hoy también la formación de los jóvenes para transformar las empresas...

4. Si pudieras recibir informes que te actualicen sobre tu mercado ¿Cuáles deberían ser sus características?

Pues mira, la información que se puede dar a los directivos de manera que sea útil para tomar decisiones es imprescindible. Creo que es fundamental que las informaciones que se den sean sobre temas muy concretos en los que la empresa haya depositado la definición de un eje o de un proyecto estratégico en ejecución. O sea, hay mucha información que puede ser muy difusa y es difícil a veces completar informaciones que respondan a las necesidades de las que la empresa tiene que sacar conclusiones para actuar, es decir cuando tu vendes tornillos la información que necesitas sobre quien compra tornillos, cual es la competencia, hasta qué punto pueden atender la evolución futura del mercado y es fácil, o por lo menos es más fácil, son cosas muy concretas.

Cuando tienes una actividad mucho más dispersa, entonces es imprescindible que el gestor de esa información, que consigas que la empresa defina cuál es el trabajo concreto que pretenden realizar para que la información que se refiera precisamente a que tenga utilidad práctica.

Al mismo tiempo hoy, es de manera especial importante por los cambios de todo tipo económicos, sociales, administrativos, políticos, todos los aspectos que afectan el mecanismo de funcionamiento de la sociedad que hoy van a estar en entredicho. Lo cual conduce u obliga a un cambio de paradigmas, de modelo de cómo se debería actuar y del que debemos estar atentos.

En ese sentido, yo creo que debería ser una información que recogiese aspectos ambiciosos en el tiempo, que no se circunscribe solo a temas muy puntuales y muy concretos porque los puntuales y los concretos son útiles para trabajar, pero hoy también hay que tener un marco de ambición en cambios que tengan relación con nuestras propias, actividades: sea de construcción, de tecnología, de transformación del territorio o de organización de la gobernanza.

Creo que aquellas empresas que tengamos un conocimiento extenso y una ambición deberíamos estar teniendo información, de cosas en las que tendríamos que tener presencia para generar discusión social, un debate en el que también seamos propositivos de cosas que se puedan implementar y que tratemos de convencer. Por lo tanto, no es una actividad receptiva de lo que quieren hacer unos señores, sino que con el espíritu del servicio en se entienda que las proposiciones tienen que ser un servicio general a la sociedad en su conjunto. Porque todos hoy necesitamos un componente importante no solo de ilusión, sino además necesitamos una realidad de actividad. Pretendiendo acometer las cosas de forma ambiciosas.

La información que tenemos que recibir de nuestros campos de intervendría tiene que ser también aspectos concretos, en algún caso puntual; aspectos generales con respecto al modo como se están enfocando el desarrollo; y una información precisa de aquellas propuestas que están encima de la mesa para que podamos discutir y debatir. También podemos usar experiencia para encontrar como espíritu de trabajo y como capacidad de interrelación, podáis ser agentes efectivos dentro de la sociedad Latinoamericana. Igual que tenemos que seguir siendo agentes efectivos en el proceso de formación de tanto joven dominicano, trabajador, inteligente y dispuesto a hacer cosas por pertenecer, incluso de cambiar una parte muy importante de la percepción que se quiera ver desde el exterior, de la falta de interés

Tabla 7: transcripción de respuestas de entrevista 1 - 30/07/2020

RESPUESTAS DE ENTREVISTA 2	
Perfil: Personal de la dirección de Arnaiz & Partners	Fecha: 28/08/2020
Lugar: Edificio Arnaiz. Madrid, España.	
<p>1. ¿Cómo defines a Arnaiz & Partners y su actividad?</p> <p>Es una consultora privada y de servicios técnicos. Tiene equipos nacionales e internacionales. Se dedica fundamentalmente al desarrollo sostenible en aspectos relacionados con el territorio y ciudad. Tiene una experiencia superior a los 37 años. Sus profesionales tienen diversas disciplinas como arquitectos, ingenieros, diseñadores, urbanistas, medioambientalistas, geógrafos, geólogos, inmobiliarios, financieros o informáticos. De ahí ARNAIZ & Partners cuenta con colaboraciones de otras empresas para el desarrollo de los proyectos u obras y negocios.</p> <p>El objetivo principal es “crear ciudad, crear territorio”. Es decir, el desarrollo sostenible, territorio y ciudad, lo hacemos en cumplimiento de los principios de sostenibilidad y los principios de colaboración entre todos los agentes que intervienen en el modelo.</p> <p>Si hubiera que definir el modelo de negocio de ARNAIZ podríamos decir que ha evolucionado, se ha internacionalizado. Trabajamos en un sector muy complejo, hay un conjunto amplio de intereses entre todos los agentes que intervienen. El CEO, Leopoldo Arnaiz considera como aspecto fundamental no basta aportar al cliente soluciones técnicas parciales, sino que esas soluciones tienen que ser integrales y con una visión amplia de la rentabilidad en el escenario de renovación de las ciudades y los territorios en el futuro. Por eso la experiencia y el trabajo acumulado que tiene ARNAIZ con oficinas en distintos países permite a sus profesionales permite tener una amplia visión con una información mucho más global en el desarrollo de sus trabajos. Las tres grandes áreas del trabajo de ARNAIZ son:</p> <ul style="list-style-type: none"> • Planificación urbana: urbanismo, medio ambiente, ingeniería, asesoría jurídica, conservación urbana. • Desarrollo inmobiliario: arquitectura, construcción, promoción, gestión de activos patrimoniales, mantenimiento de inmuebles. • Desarrollo estratégico: 1. Gobernanza local (gobiernos locales) porque todo el desarrollo de las ciudades parte desde abajo; 2. Somos pioneros en tecnología aplicada al urbanismo (UrbGIS: Territorializar la información municipal.) 3. Formación académica y profesional: cursos en materia de proyectos estratégicos y urbanísticos; de la estructura de los municipios; economía y tributación municipal; tecnología para tratar los datos. <p>Es una gran consultora con los ojos abiertos, con expectativas de negocio no solo de interno, sino de negocio con los colaboradores. No se puede trabajar en una ciudad si no es con colaboración. Para Latinoamérica existen diferencias importantes en negocio del desarrollo de la ciudad o el propio negocio inmobiliario no tiene las mismas características en España que en Latinoamérica. Los trabajos se realizan en función de la idiosincrasia y la cultura de cada país.</p>	
<p>2. ¿Qué importancia da Arnaiz & Partners a las demandas y necesidades de su entorno y cuáles técnicas o actividades implementan para conocer su entorno?</p> <p>La empresa y sus representantes, encabezadas por Leopoldo Arnaiz siempre ha sido muy cercano al cliente. Porque su mayor esfuerzo interno es solucionar el problema y necesidades del cliente más allá de conseguir el negocio o de practicar una política de captación del cliente. Es decir. Absoluta, el 100%. Lo más importante son las necesidades del cliente, teniendo una cercanía muy importante logrando fidelizar los clientes.</p> <p>El Perfil de nuestros clientes es: Administración pública: administraciones locales, regionales (sub-nacionales) o nacionales. Clientes privados. Agentes que intervienen en el desarrollo de las ciudades. Arnaiz tiene unos criterios definidos, pero en función de una previsión de trabajo. Es decir, nunca hemos tenido una inflexibilidad a la hora de realizar los trabajos, e incluso los hemos hecho sin cobrar solamente para poder mantener al cliente. Todos los trabajos que se desarrollan tienen dos características fundamentales: 1) están en nuestro ámbito de actuación; 2) son trabajos que nos interesan no solo por nuestro propio desarrollo de actividad laboral sino por el interés en las renovaciones de las actividades que se hacen. Desde un punto de vista textual no existen unos parámetros escritos objetivos que delimitan cual es el tipo de cliente. Sino la linealidad, esta es una empresa muy horizontal, nos permite evaluar cualquier negocio antes de emprendernos en su trabajo o actividad.</p>	

Consideramos que hoy en día no es posible que una consultora que se dedique a temas relacionados con el desarrollo de la ciudad no esté en primera línea con las tecnologías. Por eso nuestro esfuerzo en aplicar tiempo para I+D, además venimos aplicando tecnologías desde hace 30 años.

Trabajamos con un montón de consultores, nuestros propios *partners* son casi stakeholders prescriptores de nuestra propia actividad y por lo tanto muchos de los trabajos que recibimos llega de nuestros clientes habituales por ese sistema de fidelización al que me referí. Por tanto, hay volumen importante de negocio que entra en Arnaiz como consecuencia del trato a lo largo del tiempo con los clientes. Casi todas las personas que trabajan en Arnaiz se han convertido en verdaderos *partners* o especialistas en su propia materia con capacidad de captar clientela.

No es una empresa que anda buscando maneras de captación de clientes porque los propios clientes ya llegan a Arnaiz por la experiencia de saber hacer. No obstante, si estamos investigando nuevas maneras de obtener clientes intentando analizar cuáles son sus necesidades. No estamos trabajando con ningún software básico o complicado, pero si con tecnologías de nueva generación para la gestión del cliente. No podríamos decir que no lo necesitamos, sino simplemente que hoy tenemos una capacidad de sobrevivir sin necesidad de tener que hacer un trabajo permanente de captar clientes.

3. ¿Qué cantidad de noticias o documentos concerniente al entorno de interés la empresa gestiona?

Ya tenemos una implementación prácticamente en su totalidad de los sistemas de trabajo en redes en la nube como Microsoft 365. Esta es una empresa que genera multitud de documentos digitales. Pero a su vez, tenemos sistemas y software que nos permiten tener perfectamente ordenados, clasificados, con criterio y con necesidad todos los documentos que tenemos.

Respecto a las informaciones de nuestra área de comunicación todos los días hace un trabajo intenso de localización de aquellas informaciones que puedan ser útiles para la consultora, por un lado, desde el punto de vista de la actualidad, y útiles, desde la mejora de nuestras técnicas a la hora de trabajar. Es decir, nosotros tenemos en cada uno de los ámbitos hay informaciones útiles en el ámbito jurídico las novedades respecto de la legislación; en el ámbito técnico, las informaciones relativas a las mejoras que se realizan en ese sector desde el punto de vista inmobiliarios, los cambios permanentes que se producen en aquellos indicadores desde el punto de vista inmobiliario.

Por tanto, el área de información o comunicación, desarrolla de forma casi permanente una actualización y un estar al día o en línea con lo que está ocurriendo para trasladarlo a cada uno de los responsables.

7. ¿Proveen informes periódicos a la alta gerencia?

Hemos llegado a la conclusión de que las áreas y departamentos de comunicación han cambiado mucho. Mi propia opinión es que hacer dossier de prensa para los directivos de la empresa muchas veces es inocuo, porque no tienen tiempo para leerlos. Por lo tanto, el avance digital en internet ya les permite de alguna manera, porque cada uno de los profesionales de las empresas en los últimos años se han convertido en gestores propios de la información, no necesitan de un tercero que les dé información.

Lo que importante en el área de comunicación de las empresas en el siglo XXI, es la capacidad del responsable de comunicación de filtrar la información y por lo tanto poder ofrecer a esos responsables una información filtrada sobre aquellas cuestiones que le puedan resultar de interés sin necesidad de desarrollar la información, simplemente lanzándoles lo que sería un “flash” informativo para que ellos posteriormente pudieran investigar si es de su agrado hacerlo o no. Nunca se ha hecho nada en Arnaiz basado en la imposición, y, por lo tanto, no obligamos a la gente que trabaja en esta empresa a que se lea cantidades ingentes de documentos, simplemente se produce una serie de actualizaciones e informaciones y se dice: os invito a que podáis investigar o trabajar en estos tipos de informaciones.

Se envía por correos electrónicos. Un Newsletter o vídeo es más útil para la comunicación externa, pero no a lo interno. Para la comunicación interna lo que funciona es la relación persona a persona, acercando a los directivos las informaciones de interés de forma diaria, pero de manera directa. Lo que no quiere decir que en otras empresas no funcione, pero aquí no.

7.1 ¿Has pensado en estructurar un sistema de gestión de información? Si, lo que es cierto es que en muchas ocasiones hay un problema de infoxicación, de generación de información que está muy por encima de la capacidad de las personas administrarla o gestionarla.

8. ¿Con qué periodicidad la empresa repiensa, ajusta, actualiza su oferta y estrategia para actuar en España y Latinoamérica?

Se modifican en función de las líneas generales del mercado. No se replantea la actividad que desarrolla Arnaiz, pero si la estrategia de negocios para mantenerla o ampliarlas con nuevos negocios.

Estamos permanentemente de forma online trabajando en las modificaciones de las ofertas en función de las necesidades del cliente. No somos una fábrica que fabrica tornillos ni una tienda que vende camisetas. Somos una consultora que está permanentemente viva.

La discusión del presente, la evolución y el futuro de la empresa se realiza permanentemente. Así como la forma como evoluciona el negocio. Una cosa es la interpretación de lo que ocurre y otra es la adaptación de lo que está ocurriendo. Nuestra adaptación es online, nuestra interpretación es semanal, mensual, bimensual, permanente.

Tabla 8: transcripción de respuestas de entrevista 2 - 28/08/2020

RESPUESTAS DE ENTREVISTA 3	
Perfil: Personal de la dirección de Arnaiz & Partners	Fecha: 28/07/2020
Modalidad de entrevista: vídeo conferencia	
1. ¿Qué esfuerzos realiza la empresa para monitorear sus audiencias del sector público y privado y así estar al tanto de sus acciones?	
<p>No. No se hace monitoreo de las informaciones que producen las instituciones públicas o privadas, el posicionamiento en el mercado y la diversidad de consultores hace que estemos al día de lo que sucede. No necesitamos tener nadie buscando información, nuestro día a día, al estar en tantísimos trabajos y en tantísimos departamentos conseguimos toda la información. Pero es un proceso de día a día por estar en el mercado, no tenemos que estar atentos a lo que publiquen las instituciones públicas, sabremos lo que se publica porque te llega por correo, por una noticia o de alguna reunión. Por eso no es lo mismo nuestra empresa con esa diversidad de funciones que una empresa decida implantarse y entrar en un mercado donde sí tendría que estar más atento. Nuestra participación (en el mercado) y a través de los colaboradores nos permite saber que está pasando sin tener que hacer un esfuerzo específico en la búsqueda de esa información.</p> <p>Como no tenemos un producto enlatado sino de consultoría cada proyecto requiere una actividad específica. No podemos determinar una estrategia de fuentes de información común para todos los proyectos, <i>stakeholders</i> y <i>partners</i>. Cuando vamos a gestionar la información debemos saber que muchas veces está en la calle. No hacemos [entrevistas, encuestas <i>benchmarking</i>]. Porque no se forzará al cliente una necesidad. Tengo que ser lo suficientemente hábil para identificar el problema sin preguntárselo, porque el cliente no me dirá su problema. La administración pública no te va a decir su necesidad, porque si te dice que está haciendo algo mal está reconociendo su fallo. Por tanto, no tendría sentido su puesto en la administración pública. A los clientes no puedes ir a decirle lo que necesita, sino a darles una solución de sus problemas.</p>	
2. ¿Cómo se gestiona la información con los <i>partners</i> y clientes?	
<p>Los <i>partners</i>: son aquellos que realmente aportan valor añadido. Siempre proporcionan consultorías porque los <i>Partners</i> son de conocimiento. Para mí las instituciones no suelen tener el conocimiento suelen ser órganos de transmisión y tramitación de documentos.</p> <p>Sobre los clientes: ellos tienen un problema y nuestra entrada a los clientes es gracias a saber la solución de sus problemas porque los clientes no vienen a comprar un coche porque necesitan un coche, tenemos que convérseles de cuál es su problema y ofrecerles que confíen en nosotros su necesidad. Entonces son cosas distintas, por eso te digo que la consultoría y las ventas no es tan directa. No salimos al mercado a través de LinkedIn para buscar clientes a preguntarles si tienen un problema, porque no lo van a decir, pero tienen problemas que hay que identificarlos. Sin la confianza no tienes el cliente.</p>	

3. En la estructura de Arnaiz & Partners, ¿Qué función desempeñan los departamentos de sistemas de información y de relaciones institucionales en la gestión de información empresarial?

Ignacio dirige el departamento de desarrollo e innovación. Está enfocado en aportar al desarrollo territorial a través de una plataforma web, y maneja un tema de cultura muy duro. Las relaciones institucionales son encabezadas directamente por Leopoldo Arnaiz y mi persona y los otros directivos. No se hace una comunicación activa de marketing, ni queremos hacerla porque no vendemos un producto enlatado.

Tabla 9: transcripción de respuestas de entrevista 3 - 28/07/2020

RESPUESTAS DE ENTREVISTA 4	
Perfil: Personal de la dirección de Arnaiz & Partners.	Fecha: 27/07/2020
Modalidad de entrevista: vídeo conferencia	
<p>1. - ¿Cómo está estructurado el departamento de sistemas de información de Arnaiz & Partners?</p> <p>El departamento está organizado en tres áreas:</p> <ol style="list-style-type: none"> 1. Departamento de sistemas: asegura que toda la plataforma del servidor funcione, nosotros tenemos un Data Center en nuestra sede de Madrid y este departamento se encarga de que esté siempre disponible. Se encarga también de que si en algún momento algún usuario tiene un problema con el software o hardware se pueda sustituir cualquier complicación con el sistema. [Data center y problemas de los usuarios]. 2. Departamento de desarrollo: está formado por programadores prácticamente fundamentalmente dedicados a Urbigis. [Programación] 3. Departamento de arquitectura de base de datos: se ocupa de que las bases de datos que nosotros utilizamos estén bien administradas, tengan el mejor rendimiento posible y que todas las consultas que se realizan a la base de datos se resuelvan en el mínimo tiempo posible. [Administrar la base de datos] <p>En el departamento de sistemas en este momento hay dos personas, en desarrollo en estos momentos hay tres personas y en administración de base de datos somos dos. Un total de 7 personas.</p>	
<p>2. ¿Cómo se administran los datos en la empresa, especialmente los no estructurados?</p> <p>La base de datos solo contiene información estructurada toda la información no estructurada, que antes estaba toda en la data center, ahora se está trasladando en gran parte a data center externo sobre todo a Dropbox y a Box. Lo que está fuera es sobre todo información no estructurada.</p>	
<p>3. ¿Cuáles son las fuentes que se exploran para obtener las informaciones del entorno y cómo las administras?</p> <p>El DBA, el administrador de la base de datos, maneja dos bases de datos fundamentales la de Navision (herramienta empresarial de Microsoft y que sirve para la gestión Administrativa y de impuestos del grupo de empresas) para manejo con SQL Server, toda esa es información exclusivamente interna. Bien separada está una base de datos que fundamentalmente maneja datos externos. Como tercera base de datos tenemos donde se aloja Urbigis, que proporciona una plataforma de base de datos espaciales a cualquier usuario en cualquier parte del mundo, la cual está bien diferenciada y están en servidores separados. Incluso toda la información que nosotros producimos y guardamos en Urbigis la consideramos externa nosotros somos un usuario más de Urbigis como cualquier otro usuario. La información empresarial está bien protegida. Pero son dos bases de datos muy diferentes una totalmente externa y otra totalmente interna.</p>	

Toda la información que maneja el departamento de comunicación es el responsable en el diseño y contenido de la información web, nosotros manejamos decenas de sitios web que maneja el departamento de comunicación. Nosotros le damos el soporte para el alojamiento en los servidores, normalmente son servidores WordPress y nosotros le damos el mantenimiento, es decir, si aparece un sitio nuevo nosotros generamos la base de datos de WordPress y se pasa comunicación para que determine el contenido y el diseño. Si en alguna ocasión el departamento de comunicación tiene algún problema con alguna plantilla de WordPress o la inserción de algún plugin de pago, normalmente nosotros le apoyamos.

4. ¿Cuál es el valor que la empresa da a las demandas del entorno externo?

La importancia es máxima, puesto que ya llevamos 4 años de desarrollo y un dineral invertido para confeccionar servicios como Urbigis, porque adicionalmente nos enfrentamos a la creación de una cultura. Aunque la gente utiliza mapas en sus móviles y en sus computadoras, el uso generalizado de los mapas tiene relativamente poco tiempo.

La gente, los técnicos municipales, las empresas de *facility*... siempre han pensado que fabricar información Geoespacial era algo complejo que solo estaba al alcance de los geómetras, topógrafos y algunos técnicos. Hasta que como decía un brasileño: "hay un momento mágico en el que una persona sin ningún conocimiento comprueba que puedes fabricar mapa", que no es una esfera del conocimiento especializado y que está fuera de su alcance. Hoy en día, no solamente son Urbigis o Google Maps, existen muchas herramientas con las que las personas pueden producir mapas y eso es un momento mágico y que todavía tienen muy poca difusión. Nuestro problema en Urbigis es precisamente convencer a las personas que pueden fabricar mapas su territorio y que no es complicado y que no dependen de un geómetra para hacerlo.

5. ¿Qué mecanismos se utilizan para identificar las necesidades del mercado?

No se ha hecho, y hay que hacerlo. Todo nuestro esfuerzo en este momento y durante estos años ha estado en construir un producto. Bien, es verdad que construir un producto sin saber cuáles son las necesidades del mercado, pues es un poco absurdo, pero nosotros llevamos muchos años trabajando con los ayuntamientos y con los técnicos, conocemos bien cuáles son sus problemas. Hemos creado un producto que desde nuestro punto de vista resuelve los problemas reales que tiene los técnicos municipales, los políticos y en su relación con los ciudadanos para conseguir la información bidireccional, tanto se produzca por el Gobierno cómo los ciudadanos. Hacer un benchmarking de hasta qué punto hemos logrado eso, lo estamos empezando a hacer, aunque no podemos gastar mucho dinero en hacerlo; y segundo, a medida que vamos ganando unos pocos clientes pues vamos contrastando cuál es la brecha entre lo que quieren los clientes y lo que nosotros le ofrecemos. Lo que sí hemos llegado a contrastar es que normalmente los clientes quieren menos complejidad, porque está si es una buena herramienta compleja.

Pero por otra parte nosotros estamos convencidos de que el valor añadido de Urbigis es su enorme funcionalidad, incrementar la funcionalidad sin aumentar la complejidad no es sencillo. Esta es una discusión eterna tanto dentro de nuestro grupo de trabajo y de la empresa como con las personas que colaboran fuera de la empresa. Si me preguntas a mí, pues quiero dar el máximo de funcionalidad, aunque eso signifique aumentar la complejidad porque si no fuese así pues la gente usaría *Google Maps* o compraría una licencia de *ArcGis* online. Pero entonces tendríamos que competir en igualdad de funcionalidades y complejidad con gigantes de la información geoespacial y no podemos competir con gigantes si tenemos que competir con gigantes de la información geoespacial es porque hacemos cosas que ellos no hacen, aunque eso signifique el aumento de la complejidad.

6. Si tuvieras que clasificar las principales actividades de la empresa en temas, ¿Cuáles serían?

No es sencillo. Porque hay gran variedad. Excepto Urbigis, la parte empresarial y todas las demás oscilan alrededor de proyectos como un encargo, una obra, un plan, pero son proyectos que regularmente están cerrados en equipos donde solo interviene el cliente con ese equipo que realiza el proyecto. Además, son proyectos que suelen ser confidenciales es decir que son grupos cerrados de información a los que solo tienen acceso los técnicos que están involucrados en ese trabajo. Son islas de información.

Tienen que ser islas información porque lo que está haciendo un grupo de trabajo a veces no puede ser visto por los otros grupos de trabajo, eso en algunas ocasiones es desventajoso y puede provocar problemas de comunicación. Pero desde nuestro punto de vista los posibles conflictos que se pueden dar son menos importantes que los conflictos que se generan cuando se hace de otra forma.

A nivel de dirección lo que hay es un seguimiento al estado de ejecución. La herramienta empresarial de gestión de proyectos es informada periódicamente por los grupos de trabajo sobre el estado de la ejecución desde que se carga el trabajo, se ejecuta, se entrega, se factura y se cobra. Pero el departamento de gestión empresarial no conoce el interior del proyecto, no tiene ni acceso a la información del proyecto, solo tiene acceso a la información del seguimiento del proyecto Leopoldo o Luis, la dirección normalmente tienen una estadística de seguimiento de proyectos que lleva el Departamento de Gestión y normalmente conocen con distinto nivel de profundidad el proyecto, y en ocasiones la dirección nunca accede a la interioridad del proyecto.

Cada proyecto tiene un número de referencia para su identificación. Toda la información de los proyectos se guarda en una carpeta con Box en Dropbox con ese número de referencia, y solo están invitados a compartir esa carpeta aquellas personas que el jefe del proyecto autorice.

7. ¿Qué nivel de complejidad tiene gestionar la información de un sector donde hay tantos actores e intereses?

El nivel de complejidad en basto precisamente porque la empresa está muy compartimentada (no en varios niveles de información, ni hay una relación jerárquica entre proyectos), sino que como no se comparte información (como te he dicho) a la hora de trabajar con los clientes no se optimiza el espacio que ocupa la información. Me explico: yo tengo un proyecto donde tengo que utilizar una fotografía aérea que pesa 1GB y hay otro proyecto que tiene que utilizar la misma fotografía, lo más probable es que la fotografía esté duplicada en las carpetas de cada proyecto porque como los técnicos de cada proyecto tienden a aislarse dentro del proyecto van cogiendo todo lo que encuentran en las carpetas compartidas y van copiando dentro de su proyecto para que su proyecto sea lo más autónomo posible. Si en algún momento se tiene que hacer una copia del proyecto o si se debe entregar al cliente no debe tener componentes externos o que dependan de carpetas compartidas de manera general.

Tabla 10: transcripción de respuestas de entrevista 4 - 27/07/2020

RESPUESTAS DE ENTREVISTA 5	
Perfil: Personal de la dirección de Arnaiz & Partners.	Fecha: 29/07/2020
Modalidad de entrevista: vídeo conferencia	
<p>1. En la delegación de RD implementamos un sistema de información del entorno durante casi cuatro años, ¿Cómo valoras las iniciativas de conocimiento que fueron implementadas?</p> <p>Para mí fue muy importante la implementación de esas herramientas porque nos dieron las bases para dirigirnos a un público específico y ofrecerles lo que ellos quieren sin necesidad de que nos digan. Es decir, nosotros llevamos un proyecto puntual porque ya sabíamos cuál era su necesidad. Antes de eso, pues íbamos pensando que nuestro proyecto era el mejor del mundo que nos que nos lo debían comprar sí o sí.</p> <p>Como teníamos la información vamos a algo seguro y hablábamos con más propiedad, la gente tenía que admitir su dificultad y reconocer que la solución propuesta era la correcta. No solo se vendía un producto, sino una solución que responde a lo que el cliente necesita.</p>	
<p>2. ¿Qué nivel de sintonía tenían estos informes con la realidad y con los intereses que tenías como directora?</p> <p>Todos los proyectos que fuimos creando son adaptables. En base a ese análisis moldeamos el servicio que estamos ofreciendo. Los proyectos de nosotros no son estáticos, van cambiando acorde con la necesidad.</p>	

3. El reporte de la gestión de información evolucionó con el tiempo, de ser diario a semanal y semestral; así como su formato, de mucho texto a cuadros e imágenes. ¿Qué formato sugeriría para hacerlo atractivo, legible y deseado?

Me gustaba más cuando era entregado todos los viernes, porque era más digerible y no tenía tanta información. Entiendo que el nivel de complejidad del documento tiene que ir por nivel. El que llegue la gerencia tiene que ser un documento limpio, sencillo y en pocos párrafos debería decir lo esencial; para el responsable de un área que tiene un componente técnico, en el caso de la inmobiliaria o de consultoría, esa persona si debe tener un documento detallado con todas las tablas y formas de lugar.

La gerencia no siempre tiene tiempo de leer el documento completo, desde que vemos que el documento tiene más de 10 páginas, ya nos sentimos exhaustos. También el informe puede poner una referencia de la página donde se encuentra la información completa, y si quieres más, vas al documento.

Sería un éxito si fuera un informe dinámico: gráficas, estadísticas, textos y demás elementos. Deben haber aplicaciones que puedan ayudar a hacer eso.

4. Arnaiz & Partners trabaja simultáneamente con múltiples áreas. ¿Cómo las clasificarías?

Por un tiempo hacer eso fue un problema incluso para los empleados porque en realidad hacemos muchas cosas, pero luego hicimos un taller muy bueno y logramos sistematizar todo lo que hacemos en inmobiliaria, consultoría a nivel municipal y en levantamientos, y con ALIADE en sus buenas prácticas sociales para el desarrollo de las comunidades producción agropecuaria y desarrollo de familias.

5. Un componente importante de Arnaiz son sus socios empresariales, que le acompañan en el desarrollo de los proyectos y le mantienen al día con el acontecer en el mercado. ¿Es suficiente la información que entregan para considerar la empresa actualizada o se debería la empresa gestionarse su propia información?

Lo ideal es que la misma empresa haga el reconocimiento del entorno porque siempre los *partners* tienen una perspectiva e intereses específicos. Entiendo que es válida la información que ellos nos pueden suministrar, pero si nosotros tenemos un estudio previo considerando nuestro interés y cómo encaja en el entorno sería lo mejor.

6. ¿Cuáles son las instituciones del entorno que es más necesario poner atención?

Estamos muy enfocados en los temas municipales en República Dominicana. Por eso señalaría:

- Los Ayuntamientos.
- La asociación de ayuntamientos y la de Distritos Municipales.
- Ministerios y direcciones vinculadas a la actividad territorial
- Organizaciones sin fines de lucro que participan activamente en el desarrollo territorial.

7. Si se pidiera tu perspectiva, ¿recomendarías un sistema de conocimiento similar enfocado en Latinoamérica?

Sí. Me parece que es de suma importancia porque muchas veces los proyectos nacen con una perspectiva europea que al aterrizarlos en Latinoamérica sufren muchos cambios. Un instrumento como ese ayudaría a adaptar los proyectos a la realidad sociopolítica de manera más efectiva.

Tabla 11: transcripción de respuestas de entrevista 5 - 29/07/2020

RESPUESTAS DE ENTREVISTA 6	
Perfil: Personal de la dirección de Arnaiz & Partners	Fecha: 13/08/2020
Modalidad de entrevista: vídeo conferencia	
<p>1. ¿Cómo se gestionan las informaciones en Brasil?</p> <p>Aproximadamente el 90% de mi trabajo tiene un modelo de almacenamiento de datos, porque la publicación de datos públicos de Brasil está entre los mejores del planeta pues tenemos datos de catastro de personas. La salud y la educación son públicas, si los municipios desean tener presupuestos deben estar al día con las informaciones.</p> <p>Para determinar cuáles alcaldías tienen mayor probabilidad de hacer negocios con nosotros organizo los datos para manejarlos y traducirlos en modelos que puedan ser utilizados. Con este protocolo puedo analizar los datos y generar conocimientos.</p>	
<p>2. ¿Cómo determinar cuándo una institución es más importante que otra para hacer negocios o gestionar sus informaciones?</p> <p>En Amazonía existen 45 alcaldías, de estas solo es posible llegar a 8 en coche, para las demás habría que tomar un barco y otras no tienen conexión a internet. Si llegar es complicado y no tengo informaciones de ellos, ya esos son datos que me permiten ir haciendo el filtro. Además, otras alcaldías tienen menos de 15000 ciudadanos y no ven como una necesidad las tecnologías, ni tienen fondos propios; otras gracias a su petróleo tienen más fondos para invertir. Entonces voy persiguiendo esos caminos, no existe una fórmula concreta. Accesibilidad</p>	
<p>3. ¿Qué valor das a las informaciones publicadas por los portales web institucionales y en los medios de comunicación?</p> <p>Las informaciones publicadas por la prensa para mí no tienen ningún valor. No me fío de los datos que se publican en la prensa de Brasil, conozco como estos datos son generados. Solo presentan la realidad de un momento, entonces no sirven para nada.</p> <p>Yo me fío de micro-datos, de datos brutos. Están en bases de datos públicas que son accesibles a todos. Yo tengo toda la información de Brasil de todas las personas que nacieron en Brasil: día, sitio, hora, edad de la madre, incluso de quienes nacieron hace 15 días. Pero para analizar estos datos una tabla me es insuficiente, uso 12 tablas y las sintetizo en una.</p>	
<p>4. ¿Cuáles tecnologías utilizas para manejar la información?</p> <p>Tecnologías libres. Las principales son:</p> <ul style="list-style-type: none"> • Bases de datos PostScript. • Pentaho. • Entre otras API. 	
<p>5. La empresa tiene un amplio nivel de acción simultáneamente, ¿Cómo clasificarías estos servicios?</p> <ul style="list-style-type: none"> • Identidad digital: Tecnología para la creación de catastros de personas. • Anales de datos públicos. • Dirección de datos territoriales. • Consultoría de gestión de las alcaldías. 	
<p>6. ¿Cómo intercambias informaciones con la sede de la empresa?</p> <p>Mi relación con la empresa es diferente que con las demás delegaciones. Soy un socio de Arnaiz. Nosotros hablamos muy informalmente y como decidimos y hacemos.</p>	

Tabla 12: transcripción de respuestas de entrevista 6 - 13/08/2020

RESPUESTAS DE ENTREVISTA 7	
Perfil: Director de Comunicación del Ministerio de Educación Superior, Ciencia y Tecnología RD.	Fecha: 28/07/2020
Modalidad de entrevista: vídeo conferencia	
<p>1. ¿Cómo identificar las fuentes de información que nos permitirán monitorear el entorno externo de las instituciones?</p> <p>Aquí hay 2 cuestiones transversales que debemos trabajar, tienen que ser la eficacia y la eficiencia, porque si no trabajamos estos dos componentes es posible que se dé un tercero, que se llama la dispersión. Tenemos que ver la pertinencia del tipo de información, si es importante o no es importante, o si es superflua, sin ningún tipo de información o de importancia.</p> <p>Los entornos deben estar codificados por tipos de importancia. La información del entorno que es importante para la gerencia, no necesariamente es la competente para los demás componentes de la misma o quienes no pertenecen a la institución. Eso lo identificamos por el orden de importancia y pertinencia. Aquí entra en función los esquemas que hay en las instituciones, que sirven para unificar y para dividir, para determinar los espacios donde la información juega un papel más importante.</p> <p>Otra cosa muy importante tiene que ver con la cultura, que trae consigo la disciplina y a su vez está vinculada con la eficiencia y la eficacia. En cualquier trabajo humano que se dé socialmente tenemos que tener bien claro estos ejes transversales para tener resultados. Cuando no existen o no se han identificados estos roles ocurren accidentes de comunicación e información, como que se filtre información de un producto nuevo antes del lanzamiento o la noticia de reducción de personal, causando incertidumbre entre los empleados. Por eso la información y la comunicación deben cumplir un rol de racionalidad en su distribución.</p>	
<p>2. ¿Cuáles criterios o características debemos tomar en consideración para seleccionar las informaciones que son relevantes para la institución sobre sus clientes y entornos de interés?</p> <p>No podemos perder de vista el basamento jurídico recuerde que las instituciones se rigen por leyes y por estatutos. Lo primero que usted tiene que ver es cuáles son sus alcances jurídicos y el perfil de las exigencias que le están haciendo. En ocasiones hay que ver también cuáles son los parámetros nacionales e internacionales para saber cuáles son los alcances. Luego tengo que ver cuál es el alcance financiero y humano para ejecutar, para vigilar o para evaluar. Vivimos en una cultura que constantemente está violando los límites y para poder implementar el seguimiento a esas instituciones en estar yo conocer los límites legales que existen y éticos.</p>	
<p>3. ¿Podría mencionar de manera concreta métodos y tecnologías para gestionar adecuadamente las informaciones de su entorno externo?</p> <p>La calidad del trabajo más que en los medios, se aprecia en los resultados. Deben implementarse técnicas funcionales para cada institución, aunque definitivamente el dominio de la tecnología nos va ayudar a aglutinar información para proceder a su análisis. Ahora, ¿Qué haremos con tanta información una vez la tengamos y nos sea útil para algo específico? ¿Desecharla? ¿Y cuándo tengamos otro problema mañana? La información debe protegerse, ¿Pero ¿dónde la almacenamos, porque es mucha? A estas cuestiones deben responder los métodos.</p> <p>Para evitarse todos estos dolores de cabeza, es posible acceder a plataformas que monitorean las informaciones y que es posible comprarlas. Aun así, hay que tener capacidad de análisis una vez uno recibe estos informes.</p>	

Tabla 13: transcripción de entrevista a 7 - 28/07/2020

RESPUESTAS DE ENTREVISTA 8	
Perfil: Profesor en el Departamento de Informática y Automática, Universidad de Salamanca	Fecha: 27/07/2020
Modalidad de entrevista: vídeo conferencia	
<p>1. - ¿Cómo la informática facilita la gestión de información a las empresas que desean monitorear sus públicos externos?</p> <p>La informática y la tecnología en general lo facilita de alguna manera. 1. Permite transmitir y recibir información en tiempo real y grandes cantidades cada vez con mayor calidad. 2. Permite procesar y organizar grandes cantidades de información, que es un importante en este tiempo porque cada vez es más fácil producir información y comunicarla, lo que produce la necesidad de organizar y hacer esa información utilizable. Ya en muchos casos, no es que la tecnología ayude, sino que es la única forma de enfrentarse a esa gran cantidad de información. No es concebible manejar esa cantidad de información sin ayuda tecnológica.</p> <p>Es probable que todavía hoy aparezcan instituciones que subestimen las tecnologías, pero creo que son empresas u organizaciones de pequeñas dimensiones y muy poco ambiciosas. A la hora de emprender nuevos proyectos, crecer o de entrar en otros ámbitos, si se quedan circunscrito a lo que llevan haciendo, puede que llegue un momento cuando no puedan funcionar. Pero si quieren crecer sin hacer un trabajo sistemático de gestión de información, lo veo muy difícil.</p>	
<p>2. - ¿Considera que las instituciones están dando suficiente importancia a la implementación de las tecnologías para conocer su entorno?</p> <p>Se le da mucha importancia, pero no sé si se valora adecuadamente. Me explico: hay una serie de elementos tecnológicos que forman parte de la visualización, de la imagen de una determinada empresa. Por ejemplo: si tú vas a visitar a un despacho a un directivo de una empresa u organización, es inconcebible que no tenga una buena pantalla o un buen ordenador en su despacho, preferiblemente de una marca prestigiosa, aunque luego no se utilice. Simplemente por la apariencia o imagen pública de la empresa. Es inconcebible que una empresa no tenga una página web e incluso aplicaciones móviles, aunque luego no se utilice ni le sirva mucho. Forma parte de la identidad visual de la empresa.</p> <p>Hay factores en la tecnología que importan y se valoran mucho, e incluso más importancia que la ayuda real que brinda. Pero no sé si valoran adecuadamente todos los aspectos que involucra.</p>	
<p>3. - Usted ha participado en proyectos de recuperación y clasificación automática de información de texto. Al día de hoy, ¿Cuáles son las tecnologías más recomendadas para esta actividad?</p> <p>Hay dos grandes bloques de tecnologías.</p> <p>1. Destinadas a la búsqueda y recuperación automática: de compra la versión comercial de Google; y uso no comercial: Solr, seguro es el más utilizado para identificar información interna de las organizaciones; otra más sería Elasticsearch. Estas últimas no son fáciles de aplicar, pero son potentes.</p> <p>2. Extracción de información de estos textos: es más avanzada. Lee, examina y extrae a una velocidad sumamente útil a diferencia con la actividad humana. No hay una aplicación o programa que haga de todo. Hay herramientas que cada una hace parte del trabajo y luego se juntan.</p>	
<p>4. - ¿Qué criterios de selección de la información deben de considerarse al utilizar las herramientas informáticas?</p> <p>La gracia de estas herramientas es que la selección la pueden hacer ellas. Algunas son muy simples: le das una serie de palabras claves y ellas seleccionan los documentos que tienen estas palabras. Es simple, pero funcional en algunos campos. Hay otras herramientas que es posible entrenarlas a reconocer las informaciones que son interesantes, ellas aprenden y las seleccionan. Este último grupo ha sido probado con resultados muy buenos, con errores inferiores al 5%.</p>	

5. - ¿A partir de que volumen de información las empresas deberían considerar el uso de aplicaciones informáticas especializadas en extracción y agrupación de documentos? (puede expresarse en cantidades de documentos [miles, cientos] o en espacio de almacenamiento [GB, MB, KB])

Yo creo que depende. Una organización puede manejar 500 documentos diarios y pensar que contratando más personal en lugar de tecnología sería la mejor opción, es posible. Pero con las cantidades que nos movemos hoy en día, como poco de miles, es imprescindible. Yo no creo que salvo organizaciones muy pequeñas que se muevan en el orden de las decenas. Pero cuando hablamos de centenares a miles nos obligan a utilizar herramientas tecnológicas.

6. - Acorde con su experiencia gestionando las informaciones de texto, ¿Qué recomendaciones da a la hora de entregar los informes de resultado?

Depende a quien vaya dirigido. Normalmente hacen dos versiones, una ejecutiva: un resumen muy breve de los resultados y conclusiones; y una versión completa con todo lo que se ha encontrado.

Periodicidad: si es sobre los medios: informes diarios; informes trimestrales de tendencias.

7. - Imaginemos que deseamos extraer, agrupar y analizar las informaciones de texto de 4 periódicos producidas en un día sobre los temas de municipalidad, construcción y desarrollo socioeconómico, ¿Qué tiempo nos tomaría aproximadamente?

Salvo en análisis, las otras pueden ser instantáneas. Y que muestren incluso resúmenes de las informaciones. Es prácticamente instantáneo. Pocos minutos.

8. - ¿Es muy difícil implementar las tecnologías en la gestión de estos tipos de información?

Según donde se quiera llegar. Un ejemplo: casi todas las empresas tienen tecnologías web y los servidores registran las cantidades de visitas, pero recientemente es que [nombre de una institución] utilizan estas informaciones para relacionar los clics y cantidades de visitas con los servicios más consumidos.

9- Perfil profesional del gestor de información.

Cada vez más vamos a perfiles multidisciplinarios. Se acabó el tiempo en que el químico solo sabe de química, hoy ese profesional debe entenderse con profesionales de otras disciplinas, de las que no tienes que ser experto, pero debes tener nociones de las múltiples áreas con las que trabajas.

Tabla 14: transcripción de entrevista 8 - 27/07/2020

RESPUESTAS DE ENTREVISTA 9	
Perfil: Director de Comunicación y profesor en la Universidad Autónoma de Santo Domingo	Fecha: 30/07/2020
Modalidad de entrevista: vídeo conferencia	
<p>1. - La comunicación tiene un rol vinculador de las instituciones con su entorno interno y externo, concierne a este último ¿Cómo identificar las fuentes de información que nos permitirán monitorear el entorno externo de las instituciones?</p> <p>Cualquier relación de comunicación de una organización con sus públicos es intencional, no casual. Primero se definen cuáles son los públicos que nos interesan, aunque se escapen algunos. Además de los públicos obvios vale la pena preguntar, ¿A quién le importa lo que pase en la organización? Pues a las personas que vive en las comunidades, quienes se benefician directa o indirectamente de su actividad, entre otros que puedes identificar si están agrupados o el nivel de trascendencia de los actores. Luego tengo que incluir en mi plan de trabajo, ¿Cómo mantenerme al tanto de lo que está pasando con los sectores externos y cómo relacionarme con ellos? Es prever o construir mecanismos de retroalimentación que permitan a la organización mi estar al tanto con las reacciones que esos públicos están teniendo con lo que yo hago, como también me serviría a mí organización conocer lo que ellos entienden que deberíamos saber de ellos. Deben identificarse cuáles son esos mecanismos de retroalimentación que voy a tener: todos los medios caben, pero hay algunos que son presenciales, ejemplo: lograr que participen en una actividad anual de planificación, informes o actividades concretas de la empresa donde ellos participen en la organización o yo pueda meterme entre ellos, es decir toda una dinámica de relación. Si les facilito que cuando tengan una asamblea, ellos me invitan, es una gran ventaja. Otro medio es si la gente te llama, escribe o accede a los medios digitales, todos esos son medios de relacionamiento con esos públicos externos. Obviamente, no todos tendrán el mismo nivel de participación.</p>	
<p>2- ¿Cuáles criterios o características debemos tomar en consideración para para seleccionar las informaciones que son relevantes para la institución sobre sus clientes y entornos de interés?</p> <p>Investigación. No se puede seleccionar por mera impresión porque los amigos míos están en tal o cual lugar. Pero adelantamos: lo planes deben estar fundamentados en una investigación y eso establece cuales públicos están más cercanos, tomando en cuenta los objetivos de la organización, la misión y visión; habrá muchos públicos a quienes ponga más caso o trate diferente acorde con su necesidad. Hay algunos públicos cuya atención a sus representantes es suficiente, mientras otros más fragmentados requieren un esfuerzo adicional para alcanzar más sus integrantes.</p>	
<p>3. - ¿Qué importancia las instituciones deben otorgar a las demandas y los cambios de su entorno externo?</p> <p>Las empresas deben tener un sistema de monitoreo o vigilancia constante sobre todo lo que tiene que ver con ella. Por eso las actividades de monitoreo por pequeño o grande que sea un proyecto, utiliza las herramientas que considera necesaria. Si es en televisión quizás sea algo que deba ser contratado. Pero en todo momento debe haber dentro de la organización un poco de colaboración de todos los componentes para que, si algo se ve o escucha sea reportado al área competente. Por eso decimos en el caso de las relaciones públicas, que es algo ejecutado no por una persona, sino por la organización.</p>	
<p>4. - ¿Sigue siendo efectivo el recorte de prensa (Clipping) y las síntesis periodísticas? ¿Coexistirán con otras técnicas modernas o debemos evolucionar?</p> <p>Lamentablemente en nuestro medio ni siquiera el Clipping se usan bien. Tenemos un atraso debido a la diversidad de medios y la inmediatez de la información que no hay tiempo para ponerse a elaborarlos y enviarlos, si lo haces, cuando lo vengas a entregar los recortes a las 8 AM (trabajados desde las 6 AM), ya la información es vieja. Además, los periódicos impresos están muriendo, no son las fuentes principales para ser ágiles. Las versiones digitales son versiones rápidas y que la gente da por hecho de forma inmediata, en este sentido la fuente es importante: si viene del Diario Libre o del Listín Diario tienen mayor credibilidad, pero hay otros medios susceptibles de ser manipulados como el Twitter, donde incluso los mensajes del presidente han sido falsificados en cuanto a los nombramientos. Ahí también hay que tener mucho cuidado.</p>	

En las redes sociales, donde también circula mucha información que puede ser de interés, no es posible llevar control de lo que pasa ahí. La televisión todavía la considero es más fácil de trabajar enfocándonos en los programas especializados y dando seguimiento a los temas que nos interesan, ya los otros son las noticias del día que en ocasiones son extraídas de los mismos periódicos.

Si hablamos de las regiones, provincias o foráneos, hay que tener en cuenta los medios locales, regionales o internacionales que más afinidad tienen con la actividad empresarial. El Clipping como lo conocíamos antes tiene una función muy limitada. Probablemente los directivos o miembros de la organización que lean en la computadora o algún medio que reciban a su domicilio, ya están enterados de las principales noticias. Es indispensable enfocar los esfuerzos en medios (como los digitales) o informales que también son importantes. También deben seguirse a las figuras en los medios sociales. Síntesis informativa del congreso de México: <https://congresocdmx.gob.mx/sintesis-informativa-503-1.html>

5.- Es posible observar que para el monitoreo de redes sociales y medios se involucran cada vez más economistas, informáticos, documentalistas, entre otros profesionales no involucrados a la comunicación, ¿Qué impresión tiene?

Considero eso como una función de la comunicación. Si un departamento no incorpora esa habilidad se va a retrasar. Un profesional de la documentación o la biblioteconomía debe tener una base informática porque incluso las bases de datos hoy están informatizadas. Como eso las estrategias de búsqueda de Data mining tienen una base muy importante para eso. Las expresiones booleanas tienen que ser manejadas también por el área de comunicación, si puede capacitarse un comunicador para eso es excelente, pero si no tienen esa capacidad tienen que dotarse de esa habilidad o traer a la persona que pueda hacer eso.

Incluso existen programas, que, utilizando expresiones booleanas, hacen que te arroje de manera inmediata las informaciones vinculadas a las palabras con que has automatizado. No existe una profesión específica para hacer esto [gestión de información de prensa], muchas veces profesionales de otras áreas han demostrado ser más efectivos. Aun así, no debe pensarse que la profesión perderá preponderancia o algo así.

6. Aproximadamente, ¿Qué cantidad de información de texto [documentos digitales y físicos] concernientes al entorno externo de una institución gestiona un departamento de comunicación semanalmente?

Dependerá de la organización. Hay fuentes que son dinámicas, muy activas, como la policía; a diferencia de otras fuentes como el congreso. En la universidad próximo y durante la inscripción, las redes sociales y el centro de información tenía mayor demanda de información. Las instituciones en vez de establecer máximos de información deberían establecer mínimos que le permita tener una idea de qué está pasando, sin crear desencanto porque un día no se generó información suficiente.

Dependiendo el nivel la institución debe haber un mínimo de información, aunque este sea cero. Solo no salió nada. Incluso el periodo del año puede influir en la cantidad de información, debe pensarse también que esa información después hay que darle tratamiento para determinar si es favorable o no para la institución. Observar sencillamente si da un tratamiento neutro (como noticia) o si esta inclinada de manera positiva o negativa.

7.- ¿Cuáles son sus recomendaciones para la entrega de los informes?

Comenzando por detrás: hay que garantizar que la función de comunicación sea vista como una actividad gerencial. Debe participar en igualdad de condiciones con planificación, RRHH, administración y área técnica. ¿Qué ventajas tiene? El área podrá plantear sus orientaciones a todo el grupo, y todos a su vez demandarán a comunicaciones que esperan del departamento.

Tiene que haber información periódica que esté dentro del plan de la empresa. Debe establecerse un tiempo dentro de los encuentros con la dirección junto a la entrega de un dossier físico o electrónico que diga a toda la gerencia los éxitos, novedades, entre otros detalles de importancia. Adicionalmente, hay informaciones que, así como la gerencia debe conocerlas, deben compartirse con toda la organización, aunque sea en un mural, un correo electrónico o un portal que permita que sea de conocimiento público. Todo esto es a nivel general. A nivel particular deberían existir reuniones periódicas con la gerencia para recibir orientación, discutir planes e informar. Hay cosas que no se dicen frente a todos, esto puede involucrar una reunión con un determinado departamento llevando una propuesta para que se apruebe. Planificar las actividades y recursos que se implementarán para conocer los públicos (gestión de información, presentación o eventos de intercambio) no pueden ser

espontáneas, debe tener una presentación previa y rendición de cuentas posterior, aunque no se pida. Este informe deben ser breves líneas, nada agotador.

8. ¿Qué nivel de complejidad tiene la gestión de información para su departamento?

El nivel de complejidad está definido por los objetivos, misión y visión de la organización. Los objetivos de manejo de la información deben ser claros, no puede ser “mejorar la calidad de vida de una población determinada”, pero si “Proveer los conocimientos elementales para que una población determinada desarrolle una actividad”.

Tabla 15: transcripción de entrevista 9 - 30/07/2020

RESPUESTAS DE ENTREVISTA 10	
Perfil: Profesor de Comunicación en la Universidad Autónoma de Santo Domingo	Fecha: 09/09/2020
Modalidad: cuestionario.	
<p>1. La comunicación tiene un rol vinculador de las instituciones con su entorno interno y externo, concierne a este último ¿Cómo identificar las fuentes de información que nos permitirán monitorear el entorno externo de las instituciones?</p> <p>La pregunta está enfocada en cómo se hace un proceso determinado que, en el caso que nos ocupa, tiene que ver con la identificación de las fuentes de información necesarias para monitorear el entorno externo de las instituciones. En tal sentido, lo primero que se debe hacer en aras de lograr ese cometido es realizar un diagnóstico respecto a las necesidades de información requeridas que pueden ser detectadas mediante la aplicación de una auditoría de comunicación, ejercicio clave para estos casos. Por otro lado, una vez concluida esa fase del proceso, entonces estamos en condiciones de planificar las acciones destinadas a clasificar las fuentes de información identificadas previamente, sin pasar por alto su finalidad que consiste en monitorear el entorno externo.</p> <p>En efecto, cada fuente de información debe guardar una relación directa con el propósito planteado de cara a la gestión oportuna, eficiente y eficaz de la Comunicación Externa como parte de los esfuerzos tendentes a monitorear el entorno externo. Un ejemplo de ello es la aplicación de las encuestas para conocer la opinión, valoración y percepción de los grupos de interés, entre los cuales están los proveedores, así como los consumidores o clientes, de quienes interesa saber cuál es el nivel de satisfacción o no de los productos o servicios provistos por la institución.</p>	
<p>2. ¿Qué nivel de importancia las instituciones deben otorgar a las demandas y los cambios de su entorno externo?</p> <p>Una herramienta clave que nos ayuda a entender el nivel de importancia que las instituciones deben otorgarle a las demandas y los cambios que tienen lugar en el entorno externo lo constituye la Comunicación Estratégica cuyo objetivo principal consiste, precisamente, en procurar una relación armónica, positiva entre las organizaciones y los grupos de interés que forman parte de su ecosistema como algunos prefieren denominar ahora al entorno externo de las organizaciones.</p> <p>La Comunicación Estratégica ofrece un esquema metodológico idóneo para gestionar las demandas y los cambios que tienen lugar en el ecosistema o entorno externo (social, económico, político, cultural, ambiental, entre otros) que inciden o gravitan de manera directa o indirecta en la sostenibilidad de las organizaciones. Este esfuerzo implica entender la importancia que tienen esos cambios, más aún en el contexto de crisis permanente en que vivimos, acentuado por los riesgos, dificultades, barreras y problemas que caracterizan el momento actual. Muchas organizaciones y proyectos que fracasan por no prestarles la debida atención a las demandas y los cambios del entorno. Por ejemplo, marcas emblemáticas como Kodak, Blackberry y Nokia desaparecieron por ese motivo.</p>	
<p>3. ¿Siguen siendo efectivo los recortes de prensa (clipping) y las síntesis periodísticas? ¿Coexistirán con otras técnicas modernas o debemos evolucionar?</p> <p>El autor español Alfons Cornella, quien acuñó el término “infoxicación”, describe con mucha razón el mundo que vivimos actualmente, caracterizado por la sobreabundancia de información. De ahí que estamos infoxicados. El mejor ejemplo de esto lo representa la pandemia del Coronavirus o COVID-19, pues hemos sido sometidos a un bombardeo interminable de informaciones de toda índole (noticias, reportajes, artículos, editoriales, entrevistas, videos, noticias falsas o fake news),</p>	

situación que incide notablemente en la actitud de muchas personas a no prestar la debida atención al tema que ha saturado la atención de la opinión pública mundial en la era de la posverdad.

Ahora bien, el cansancio que provoca la sobreabundancia de información respecto a la pandemia obedece a que se ha convertido en un tema que ocupa la principalía en los medios de comunicación, redes sociales e intercambio de información por otros medios tradicionales y no tradicionales.

De lo anterior se desprende la importancia y efectividad de los servicios de Clipping y las síntesis de periodísticas para las instituciones porque permite superar el carácter monotemático que imponen los medios de comunicación y otros espacios donde compartimos informaciones. Entiendo que, a pesar de cualquier vaticinio pesimista, la importancia y efectividad de este tipo de servicio adquiere mayor importancia a raíz de lo que ha sucedido en el mundo con la pandemia y coexistirán con otras técnicas modernas debido a que los avances tecnológicos lo que han hecho es optimizar sus resultados, pues de recortes en papel como se hacía años atrás, ahora se hace este tipo de servicios apoyándose en los recursos digitales disponibles.

4. ¿Podría mencionar de manera concreta métodos y tecnologías para gestionar adecuadamente las informaciones de su entorno externo?

- Suscribirse a las alertas de Google.
- Utilizar las encuestas en líneas a través de plataformas como los formularios de Google.
- Hacer uso de las RRSS para difundir informaciones a interesados y recabar opiniones.
- Utilizar páginas web, blogs, cuentas en redes sociales para mantener una interacción en líneas con los públicos.
- Realizar videoconferencias, a través de las diferentes plataformas populares existentes como Zoom, Teams, Google Meet, entre otras.

5. Aproximadamente, ¿Qué cantidad de información de texto [documentos digitales y físicos] concernientes al entorno externo de una institución gestiona un departamento de comunicación semanalmente?

La respuesta a esta pregunta va de la mano con la naturaleza de cada institución. No me parece correcto establecer una cantidad fija. Eso se puede inferir en función de los hallazgos que se deriven de una Auditoría de Comunicación, cuyos resultados cuantitativos podrán evidenciar cuál sería la dimensión o tamaño de la información que deberá ser gestionada semanalmente por un departamento de comunicación.

6. ¿Qué dinámica desarrollan los departamentos de comunicación para orientar la dirección de la institución sobre el estado de su entorno externo y recomendar la toma decisiones a partir de estas?

Reitero la importancia de realizar una Auditoría de Comunicación para determinar la realidad existente en términos de la comunicación interna y externa. A renglón seguido, adquiere relevancia diseñar y ejecutar un Plan de Comunicación Institucional, a los fines de guiar el proceso de gestión de la comunicación de manera efectiva, incluyendo los productos necesarios para hacer los análisis y presentar las recomendaciones de lugar.

El análisis de coyuntura constituye un elemento singular en este tipo de proceso, considerando no solo su utilidad, sino la periodicidad y a quien o quienes les toca la responsabilidad de elaborarlos atendiendo las necesidades de la institución.

7. Aproximadamente, ¿Qué tiempo toma la extracción y organización de la información necesaria para conocer el entorno?

Todo dependerá de la naturaleza de la institución, así como del equipo que trabaje en el área de comunicación. Mi recomendación es seguir al pie de la letra los planteamientos del español Joan Costa, quien ha trabajado de manera profusa el tema de la figura del Dircom debido al papel estratégico que le corresponde desempeñar al Director de Comunicaciones en la estructura organizacional.

En resumidas cuentas, el factor tiempo dependerá de otras variables. Ahora bien, la redacción de un discurso tiene que ver con las competencias de la persona o el equipo que tenga esa responsabilidad sobre sus hombros. Por otro lado, conocer el comportamiento de los grupos de interés requiere de la aplicación de modelos metodológicos diseñados para tales fines y el conocimiento, sumado a la experiencia, constituyen elementos clave para la consecución de ese objetivo. Finalmente, las actividades de capacitación dependen de la planificación que se realice, junto al equipo humano y los recursos disponibles.

8. ¿Qué nivel de complejidad tiene la gestión de información para su departamento?

La gestión de información acusa un papel crítico en cualquier departamento de comunicación. Las tareas y actividades son múltiples, empezando con la necesidad de comprobar si el mensaje de la institución es recibido y percibido de manera correcta por parte de los públicos. Prevenir el surgimiento de cualquier crisis, anticipación que es posible si se aplican los mecanismos adecuados de una comunicación de crisis efectiva que debe anticipar cualquier escenario negativo y la manera correcta de enfrentarlo.

Hay que valorar el tipo de información que afecta positiva o negativamente a la institución. Se debe interpretar, analizar y gestionar de manera efectiva el impacto del mensaje que se pretende transmitir en los públicos. Procurar el posicionamiento deseado y lograr una percepción positiva de la imagen institucional.

En resumidas cuentas, el nivel de complejidad que implica llevar a cabo una gestión efectiva de la información en el Departamento de Comunicación constituye un desafío permanente. El éxito dependerá de muchos factores, pero la actitud hacia el cambio representa una herramienta valiosa si se quiere tener resultados en el corto, mediano y largo plazo.

Una visión estratégica, sin olvidar el accionar táctico que demandan las circunstancias, representa un recurso crítico cuando se pretende alcanzar las metas de la organización.

Tabla 16: transcripción de entrevista 10 - 09/09/2020

RESPUESTAS DE ENTREVISTA 11	
Perfil: Profesor en el Departamento de Informática y Automática, Universidad de Salamanca	Fecha: 30/08/2020
Modalidad: cuestionario.	
1. - ¿Considera que las instituciones están dando suficiente importancia a la implementación de las tecnologías para conocer su entorno?	
No. Las instituciones tienen una visión bastante miope de la utilización de las nuevas tecnologías. No están preparadas más allá de indicar que utilizan nuevas tecnologías, pero en muchos casos sin hacer estudios previos de las necesidades reales.	
2. De manera específica, ¿Cuáles son los software más recomendados para la búsqueda y recuperación de informaciones y/o documentos de textos en la web? (Pueden ser de compra o no comerciales)	
Todo dependerá del entorno en el que se quiere implementar el sistema de búsqueda. De forma global los buscadores ofrecen una buena forma de acceder a información global. Si se quisiera implementar un sistema de RI y su posterior búsqueda en una empresa o en un entorno de intranet existen productos como Elasticsearch, Apache Solr que pueden dar soluciones a coste cero.	
3. Una vez extraídos los textos de la web, ¿Cuáles son los software más recomendados para automatizar su lectura y extraer información? (Si precisa nombrar más de uno, puede hacerlo. Pueden ser de compra o no comerciales)	
Elasticsearch, apache Solr	
4. Durante su materia, observé a "R" como una herramienta muy completa para el análisis de información digital. A través de esta ¿Sería posible hacer Scraping o crawler, extraer información y hacer el agrupamiento de las informaciones por tema?	
En R se puede realizar scraping, se pueden crear recolectores, se pueden aplicar técnicas de recuperación de información y se pueden aplicar técnicas de clasificación, detección de topics.	

5. A la hora de determinar las fuentes para extraer la información de la web ¿Qué criterios de selección se deberían obedecer?

Siempre hay que establecer un conjunto de URLs que actúen como semillas en el proceso de recogida de información. Deben ser URLs de alto nivel de importancia, que determinarán que los resultados sean más satisfactorios.

6. - Imaginemos que deseamos extraer, agrupar y analizar las informaciones de texto de 4 periódicos producidas en un día sobre los temas de municipalidad, construcción y de políticas socioeconómicas, ¿Qué tiempo nos tomaría aproximadamente esta actividad?

Todo depende de las características de los periódicos y de si imponen restricciones a la descarga. El sistema que tengan implementado a efectos de URL también tiene su influencia.

Recoger la información de un día concreto, en secciones bien determinadas, con URLs que no requieren filtros o análisis podría llevar mucho menos de 1 hora. Pero cuidado, porque en la recogida web las trampas y problemas en la recogida pueden ser muchas.

7. - ¿Qué orientaciones adicionales considera apropiadas a la hora de estudiar la gestión de información de texto que podrían ser de utilidad para esta investigación?

Para trabajar con información textual es preciso utilizar fuentes de recogida automáticas y la utilización de técnicas de recuperación de información avanzadas. Además hay que complementarlos con sistemas de acceso a la información, yo diría que mediante plataformas web.

Tabla 17: transcripción de entrevista 11 - 30/08/2020

RESPUESTAS DE ENTREVISTA 12	
Perfil: Especialista en inteligencia de negocios y administración	Fecha: 31/08/2020
Modalidad: vídeo conferencia	
<p>1. - ¿Existe un perfil para el / la gestión de <i>Big Data</i> y la inteligencia de negocios? ¿Cuál sería y cuál es el tuyo?</p> <p>Realmente sí. Existen diversas actividades que nacen del <i>Big Data</i>, como el científico de datos, gestor o analista de datos y el especialista en la visualización de los datos. Estos son los principales. Eso produce un nexo entre el área de negocios y el área de informática. Una persona que pueda manejar el conocimiento técnico y del negocio sería idónea para llenar ambos requerimientos.</p> <p>Actualmente existe mucho solapamiento de funciones, pues quienes tienen un perfil de negocios, de por sí, están obligados a conocer el uso de herramientas técnicas y lenguajes de programación más allá que el simple análisis. Lenguajes de programación: <i>Python, R, Java, SQL</i>. ¿Cuál es recomendable para texto? <i>Python</i> y <i>R</i> están enfocados a programación con fines de visualización para datos estadísticos.</p>	
<p>2. - ¿Cómo las tecnologías de la información facilitan la gestión de datos a las empresas que desean monitorear sus públicos externos?</p> <p>El propósito es analizar las industrias para tomar decisiones, promover iniciativas y conocer en qué dirección enfocar sus esfuerzos o donde están sus nichos de mercado. Todo esto y más es posible hacerlo con las informaciones de los públicos, pero también de los <i>partners</i>. En esta materia seguimos en pañales en RD, pero la preocupación es creciente. En ocasiones forma parte del departamento de marketing y en otras ocasiones es una posición transversal de la empresa.</p>	
<p>3. - Para determinar qué informaciones son útiles para las empresas y cuáles no. ¿Qué criterios de selección de la información deben considerarse?</p> <p>Eso se adapta a la naturaleza de negocio para la que se trabaje. Si es una inmobiliaria, quizás necesites un CRM para manejar tus clientes, un programa de contabilidad y un programa de geolocalización para saber dónde están tus ingenieros. Con toda la información recopilada con todos estos sistemas, hará falta un analista de TI o base de datos que pueda unificar esa información para acceder a esa información con un reporte.</p> <p><i>Power BI</i>: herramienta de visualización. Conecta la información desde diversas fuentes y permite ver la información en un tablero. En las elecciones presidenciales pasadas de RD se utilizó este programa. Pertenece a Microsoft. En Europa se utiliza más <i>Tableau</i>.</p>	
<p>4. - Una vez trabajados los datos estos deben ser útil para el conocimiento del negocio y el mercado, así como para la toma de decisiones. Por eso se formulan informes periódicos. ¿Qué recomendaciones darías a la hora de entregar estos informes de resultado?</p> <p>No debe ser largo, sino claro y conciso. Debe resaltar las informaciones más importantes. Debe ser gráfico. Debe sobreentenderse el mensaje: sea un cuadro o un mapa.</p> <p>Debe hacerse énfasis en que se está dando algo que la empresa no tenía antes.</p>	

Tabla 18: transcripción de entrevista a 12- 31/08/2020

RESPUESTAS DE ENTREVISTA 13	
Perfil: Profesor en el Departamento de Biblioteconomía y Documentación USAL	Fecha: 31/08/2020
Modalidad: vídeo conferencia	
<p>1. Cuándo aplicamos los principios de la documentación y biblioteconomía a una empresa, ¿Cómo se identifican las fuentes externas de información que son de interés para de la empresa? (ponga ejemplos)?</p> <p>Si existen criterios, lo que ocurre es que estos criterios cambian según el tipo de institución al que se aplique. Información hay mucha, al punto que se ha convertido en un problema, el criterio de conseguir información de mayor calidad, poder aportar un valor añadido y eligiendo las mejores fuentes. La calidad se puede deber por distintos motivos: porque su calidad es más fiable, por su nivel de proliferación, la metodología utilizada para adquirir los datos, la reputación de la institución responsable, entre otras.</p> <p>Puede seguir variando según el tipo de empresa y sus fuentes. También es bueno considerar los departamentos a los que a nivel interno el gestor de información sirva.</p>	
<p>2. Si las instituciones contratan personal especializado y conocedor de las necesidades del mercado y como desarrollarlo, ¿Para qué necesitan una persona o departamento que gestione la información de sus mercados y públicos de interés? ¿Existe un perfil profesional para el gestor de información?</p> <p>No te diré que sí, o que no. Yo creo que probablemente se necesite un perfil profesional que tenga algunas características de los documentalistas, pero necesita tener otras competencias en las que no está formado un documentalista. Me refiero por ejemplo al perfil <i>Community Manager</i>, entre otras denominaciones que surgen. Yo creo se necesitan una serie de competencias documentales: 1. Saber trabajar con la información, organizar y procesar la información, entre otras.</p> <p>El perfil para gestionar la información digital es nuevo. Las tareas que están apareciendo nunca han sido propias de los documentalistas. Lo que sí creo es que también hay una invisibilidad social muy grande de las personas que estudian documentación. Muchas veces no se menciona sencillamente porque no se conocen, tampoco saben cuáles son las capacidades del personal egresado de esta área. Cada cierto tiempo esta invisibilidad de los perfiles profesionales resurge en tono de queja en las revistas profesionales. A nivel personal, también conozco personas que han empezado a trabajar por casualidad en una empresa y se han hecho imprescindibles. Pero no les estaban buscando inicialmente.</p> <p>En cuanto al curador de contenidos, esa es el tipo de pregunta que cuando la haces un estudiante de documentación te va a decir que sí, porque todo el mundo trata de reivindicar una identidad profesional. Yo, como soy mayor, no le veo mucho sentido a estar poniendo etiquetas, sobre todo cuando se hacen de manera excluyente. Tiene que tener habilidades documentalistas. Lo de curador, habría que definirlo, no siempre significa lo mismo.</p> <p>Sobre los perfiles multidisciplinares: Si, sin duda. Sobre todo, quienes trabajan en empresas privadas. Los titulados de documentación también tienen perfiles muy diferentes. No es lo mismo trabajar en una empresa que tiene un fondo antiguo que en una empresa.</p>	
<p>¿Imaginemos que te solicitan un trabajo? ¿Cómo debería ser este informe?</p> <p>La mayor actividad del documentalista está en la recopilación de la información y luego en la organización de la misma para la realización del informe.</p>	

Tabla 19: transcripción de entrevista 13- 31/08/2020

RESPUESTAS DE ENTREVISTA 14	
Perfil: Científico de Datos, Biblioteca de Fundación Juan March	Fecha: 30/08/2020
Modalidad: vídeo conferencia	
<p>1. - ¿Considera que las instituciones están dando suficiente importancia a la implementación de las tecnologías para conocer su entorno?</p> <p>No sé si todas las instituciones lo están haciendo. No conozco de estudios o encuestas para determinar eso. Considero que, aunque de una forma no tan metódica, más o menos todos lo hacen. En mi caso trabajamos toda la actividad, museística y de conciertos para complementar toda la cultural que se desarrolla en esta ciudad [Madrid], haciendo cosas que no hacen los demás respondiendo a una necesidad. El entorno es Fundamental para desarrollar la actividad de la Institución, siguiendo parámetros más específicos en redes sociales: número de seguidores y como aumentan, lo comparamos con nosotros mismos para ver cómo vamos nosotros.</p>	
<p>2. Software como Sorl, Elasticsearch o la versión comercial de google aparecen como los más recomendados para la búsqueda y recuperación de información de texto en la web. Pero una vez con los textos identificados ¿Cuáles programas deben ser utilizados para clasificarlos por materias?</p> <p>Sorl y Elasticsearch son para buscar información, pero no te clasifican. Una de las cosas que nosotros hacemos es capturar los clips que mencionan a la Fundación, esos los almacenamos en una especie de bases de datos no estructuras [lo podríamos almacenar en los programas almacenados, para extraerla para acceder a ellas, pero no las clasificaría en política, cultura y así] para hacer eso tienes que usar herramientas de clasificación como IBM o Google que te ayuda a clasificarlos.</p> <p>Nosotros no contamos con un equipo de catalogadores que lo hagan manual. Por eso hacemos clasificaciones automáticas que hacen que los resultados no sean del todo buenos. Aunque hemos montado nuestra propia taxonomía con más de 4,000 términos. El automático te ayuda mucho, pero realmente después necesitas esa clasificación manual.</p> <p>2.1. Determinar fuentes de interés. ¿Qué criterios?</p> <p>Se depende mucho de los temas que se estén buscando: impresión de los clientes, perspectiva periodística, perspectiva académica (publicaciones científicas). Las fuentes deben ser de calidad, conocidas y que te den confianza.</p>	
<p>4. ¿Se necesita ser matemático o informático para la ciencia de datos?</p> <p>Yo creo que es un área nueva, en la que no hay una educación reglada. No tienes por qué ser matemático o informático, pero seguramente necesites tener nociones avanzadas de matemáticas o estadísticas; e informáticas de programación o base de datos. La ciencia de datos yo suelo decir que es una herramienta, que depende el entorno en que esté, es un poco distinta. Tiene elementos comunes como son el tema del análisis, la inteligencia artificial, el aprendizaje de máquinas, etc.</p> <p>Pero en bibliotecas yo digo que va más allá, porque se une con toda la parte de la curación de datos que no se suele hacer en las otras disciplinas no está. La ciencia de datos se adapta a los temas médicos, ciencias, etc; lo importante es el tema de la limpieza de datos, y es una dinámica fundamental en las bibliotecas. Es una posibilidad para que en el mundo de documentación surja un perfil nuevo, ya un especialista en curación de datos que facilita todas una técnicas de ciencias de datos.</p>	
<p>5. - ¿Qué tiempo nos tomaría aproximadamente extraer, agrupar y analizar las informaciones de texto provenientes de medios digitales utilizando adecuadas herramientas tecnológicas?</p> <p>Depende mucho de la capacidad que se tenga para trabajar con información de este tipo y los mismos medios de comunicación. Montar el sistema puede tomar varias semanas a tiempo completo. Sin el uso de herramientas de inteligencia artificial puede tomar meses.</p> <p>Sobre la presentación de los resultados, no puedo dar una regla general. Nosotros montamos un sistema que permite acceder a los directivos cuando quieran, sin depender de nadie para hacer el análisis. Algunas veces también respondemos preguntas concretas que nos hacen. Se presentan gráficos de redes de manera principal.</p>	

6. - ¿Qué retos se deben afrontar para eficientizar la gestión de información de textos?

El análisis de texto. El reto es hacerlo de manera automática contextualizando, no como google que lo pones y lo organiza según su taxonomía, sino acorde con las necesidades de cada uno y arroje los resultados según las necesidades que se tengan. Incluso estableciendo nuestros propios algoritmos.

Tabla 20: transcripción de entrevista 14 - 30/08/2020

RESPUESTAS DE ENTREVISTA 15	
Perfil: Especialista en análisis de datos	Fecha: 02/09/2020
Modalidad: vídeo conferencia	
1. - ¿Existe un perfil para el / la gestión de <i>Big Data</i> y la inteligencia de negocios?	
<p>El <i>data science</i> siempre se ha usado, pero es relativamente nuevo su resurgir en importancia. Tanto que se han hecho cursos académicos vinculados con este. Es multidisciplinario, combina matemática, estadísticas, entre otras áreas. Hoy hay tantos datos, que no se sabe qué hacer con tantos, por eso surge el <i>Big Data</i> como disciplina que para la evaluación de los datos usa la <i>data mining</i>. Existen muchos tipos de <i>Data Mining</i>, para mí el <i>Text data mining</i> es uno de los más complicados porque hay que llevar al texto a una condición de hacerlo digerible para su análisis.</p> <p>El analista de datos se encarga de minar datos, hacer regresiones lineales, Análisis de varianza (Anova); se utilizan herramientas como programación en R, programación en <i>Python</i>, <i>Power BI</i>, <i>Excel</i>, entre otros programas. Hay mucha similitud entra esta función y el <i>Big Data</i> [solapamientos]. Se utiliza mucho en publicidad, pero también tiene aplicaciones en ingeniería, deportes, sociología (<i>sentiment analysis</i>). <i>Data Science</i> es un campo más complejo de matemáticas y estadística avanzada para cuestiones no tan comunes. Se utiliza por ejemplo en los grandes equipos deportivos.</p>	
2. ¿Pueden los profesionales de las matemáticas entender el análisis de datos mejor que los documentalistas o incidir en un negocio mejor que los mercadólogos? ¿Existe un perfil?	
<p>Es un área multidisciplinaria. Si eres matemático podrás entender la materia a profundidad, pero el perfil se adapta a lo que sea. Un economista o un administrados, involucrados ambos en los negocios, que se especialice, puede ser gestor de datos. Un ingeniero industrial [de calidad] es un analista perfecto.</p> <p>Pero no solo la profesión. La experiencia que se tenga es muy importante para poder incidir e innovar en un área. Como te dije, hay economistas que evalúan estadísticamente el rendimiento de los peloteros para determinar si su contratación es viable, aunque no sepan en sí de Baseball. Econometría keynesiana.</p>	
2. - ¿Cómo las tecnologías de la información facilitan la gestión de datos a las empresas que desean monitorear sus públicos externos?	
<p>Para este tipo de actividad no se necesita un científico de datos, es demasiado, un analista de datos sería suficiente o un especialista de <i>big data</i>. Sería una persona que pueda hacer una programación en <i>Machine Learning</i> para hacer un scraping de las distintas páginas web para llevarlos a <i>Excel</i>, <i>Access</i> u otra página. Se pueden usar distintas herramientas de programación.</p>	
2.1. ¿Qué software serían los más eficaces para realizar Scraping de texto?	
<p>Con lenguajes de programación como <i>Python</i>. Sería ideal y muy versátil con los programas de <i>Big Data</i>.</p>	
4. - Para determinar qué informaciones son útiles para las empresas y cuáles no. ¿Qué criterios de selección de la información deben considerarse?	
<p>Desde el libro: hay que investigar los factores que motivan a las personas a utilizar el servicio que promuevo y utilizaría una data representativa para tener los criterios que debemos de tomar en cuenta. Se adaptarían las condiciones del país también.</p>	

4.1. Selección de las fuentes ¿Qué criterios?

Determinar si los datos estadísticos que usa el país son confiables. También se podrían hacer muestreos (encuestas) sobre lo que ganan, ahorran, proyecciones a futuro de mi población meta, áreas del país con mayor proyección para el desarrollo del negocio. Usar herramientas digitales también sería útil. También hay empresas que se encargan de ofrecer datos.

5. - ¿Considera que las instituciones están dando suficiente importancia a la implementación de las tecnologías para conocer su entorno?

Quien no se adapte morirá. Hay muchos que no lo han hecho, pero lo harán.

8. - Una vez trabajados los datos estos deben ser útil para el conocimiento del negocio y el mercado, así como para la toma decisiones. Por eso se formulan informes periódicos. ¿Qué recomendaciones darías a la hora de entregar estos informes de resultado?

El Data visualizing se encarga de la estética y ser convencibles de los datos para su presentación. No se necesitan presentar muchos números o letras, pero tampoco gráficos cargados. Se presenta lo más importante de manera comprensible.

Tabla 21: transcripción de entrevista a entrevistado 15 - 02/09/2020