

LA ALFABETIZACIÓN EN DATOS EN EL CONTEXTO UNIVERSITARIO DE LAS CIENCIAS SOCIALES (ALFADATAUniv_CCSS)

Ana Iglesias Rodríguez
Yolanda Martín González
Críspulo Travieso Rodríguez
Azucena Hernández Martín
Josefa Gallego Lorenzo
José Carlos Toro Pascua



Proyecto PID2020-116233RB-I00 financiado por:



Publicaciones GIR INFORAL

GRUPO DE INVESTIGACIÓN EN INCLUSIÓN, INNOVACIÓN, FORMACIÓN DOCENTE Y ALFABETIZACIÓN INFORMACIONAL DE LA UNIVERSIDAD DE SALAMANCA

Iglesias Rodríguez, A., Martín-González, Y., Travieso Rodríguez, C., Hernández Martín, A., Gallego Lorenzo, J., & Toro Pascua, J. C. (2024). *La alfabetización en datos en el contexto universitario de las ciencias sociales (ALFADATAUniv_CCSS)*. Zenodo. <https://doi.org/10.5281/zenodo.14503056>

La alfabetización en datos en el contexto universitario de las ciencias sociales

(ALFADATAUniv_CCSS)

(versión 2)

DOI: 10.5281/zenodo.14503056

Ana Iglesias Rodríguez; Yolanda Martín González; Crispulo Travieso Rodríguez; Azucena Hernández Martín; Josefa Gallego Lorenzo y José Carlos Toro Pascua

2024

Diseño y maquetación: cincocuatro comunicación



Reconocimiento

La presente publicación ha sido posible gracias a la financiación concedida por la Agencia Estatal de Investigación, Convocatoria 2020 de «Proyectos de I+D+i» de los Programas Estatales de Generación de Conocimiento y Fortalecimiento Científico y Tecnológico del Sistema de I+D+i y de I+D+i Orientada a los Retos de la Sociedad. Ministerio de Ciencia e Innovación del Gobierno de España. A través de la ayuda de referencia PID2020-116233RB-I00, denominada: «Alfabetización en datos en el contexto universitario: detección de necesidades, diseño de escenarios formativos en abierto y elaboración de un referencial de competencias (ALFADATAUniv)».

Coordinadoras del Proyecto: Ana Iglesias Rodríguez y Yolanda Martín González

Equipo de investigación: Crispulo Travieso Rodríguez; Azucena Hernández Martín, Josefa Gallego Lorenzo y José Carlos Toro Pascua



LA ALFABETIZACIÓN EN DATOS EN EL CONTEXTO UNIVERSITARIO DE LAS CIENCIAS SOCIALES (ALFADATAUniv_CCSS)

GUÍA DEL CURSO

1. CONTEXTO EN EL QUE SURGE EL CURSO

Este curso de formación es el resultado de la participación del Grupo de Investigación Inforal de la Universidad de Salamanca, integrado por docentes de Educación y de Biblioteconomía y Documentación, en el Proyecto I+D titulado: *Alfabetización en datos en el contexto universitario: detección de necesidades, diseño de escenarios formativos en abierto y elaboración de un referencial de competencias* (ALFADATAUniv) (REFERENCIA: PID2020-116233RB-100). Se trata de un proyecto financiado por el Ministerio de Ciencia e Innovación.

Los objetivos planteados en este proyecto están relacionados con el diseño e implementación de este curso:

1. Identificar las fortalezas y debilidades que los profesores universitarios poseen en el uso de los datos y su incorporación a su práctica docente.
2. Diseñar escenarios formativos online que favorezcan el desarrollo de competencias y habilidades para el acceso, análisis, reutilización y difusión de datos abiertos y de investigación para profesores universitarios del ámbito de las ciencias sociales.

Para la consecución del primer objetivo se procedió al diseño e implementación de un cuestionario dirigido a todos los docentes de las Universidades españolas en el ámbito de las ciencias sociales. Los resultados obtenidos a partir de este instrumento han permitido detectar las principales necesidades formativas que el profesorado universitario de este área de conocimiento posee en relación con la alfabetización en datos para poder incorporarlos en su práctica docente.

Como resultado de estas carencias /debilidades se decidió, por parte del grupo de investigación, el diseño del siguiente curso cuyo contenido se centra en fortalecer aquellas temáticas en las que los docentes expresaron poseer unas menores competencias.

2. OBJETIVOS

Objetivo general:

Ofrecer al profesorado universitario del área de conocimiento de ciencias sociales una formación específica en diferentes aspectos vinculados a la alfabetización en datos.

Objetivos específicos:

- Conocer e identificar los principales aspectos relacionados con la protección de datos personales, académicos y de investigación.
- Analizar el concepto de confiabilidad y los criterios para la evaluación y selección de datos abiertos.
- Conocer en qué consiste la gestión de datos de investigación y diseñar un proceso de elaboración de un plan de gestión de datos.
- Analizar los principales requerimientos en la reutilización y citación de los datos.

- Identificar y poner en práctica los procesos implicados en la visualización y representación gráfica de datos.

3. DURACIÓN Y MODALIDAD

La realización del curso tiene una duración estimada de cinco semanas, que se distribuyen en cinco módulos con sus correspondientes actividades prácticas.

La modalidad será virtual. No existe un horario fijo de participación, sino que cada usuario/a puede elegir, de un modo flexible, el momento más adecuado para realizar las tareas de cada semana.

4. CONTENIDOS Y TEMPORALIZACIÓN

La adquisición de los objetivos propuestos se llevará a cabo a través de los contenidos que conforman el curso.

Estos contenidos se organizan en torno a cinco módulos que son los que se detallan seguidamente:

Módulo 1: PROTECCIÓN DE DATOS PERSONALES, ACADÉMICOS Y DE INVESTIGACIÓN

Críspulo Travieso Rodríguez y Azucena Hernández Martín

1. ¿Qué es un dato personal?
2. ¿Cuáles son estos principios y qué implican cada uno de ellos?
3. ¿Cuáles son los derechos del ciudadano en lo que respecta a la protección de datos y cómo se pueden ejercitar?
4. ¿Y qué ocurre con las Garantías de datos digitales?
5. ¿Cuáles son las autoridades de protección de datos personales en España?
6. ¿Cómo se aplica el actual régimen jurídico de protección de datos de carácter personal en el ámbito docente universitario?
7. ¿Qué entendemos por datos de investigación o datos científicos en abierto?
8. ¿Qué leyes o normativas contemplan la regulación de los datos abiertos de investigación?
9. ¿Qué consideraciones éticas debe contemplar la preservación de datos de investigación?
10. ¿Por qué es importante llevar a cabo una descripción de los datos de investigación?
11. ¿En qué tipos de formatos documentales se pueden materializar los datos de investigación?
12. ¿Qué acciones se pueden llevar a cabo para mejorar la descripción y la visibilidad de los datos?
13. ¿Existen formatos o esquemas de descripción de datos específicos para datos de investigación?

Referencias

Módulo 2: CONFIABILIDAD EN LA EVALUACIÓN Y SELECCIÓN DE LOS DATOS ABIERTOS

Josefa Gallego Lorenzo

1. Conceptos de interés: Dato, información, fake news, confiabilidad y alfabetización en datos
2. Búsqueda y selección de datos
3. Criterios para seleccionar y evaluar la calidad de los datos
4. Recomendaciones y directrices para evaluar la calidad de los datos abiertos
 - 4.1 Recomendaciones generales
 - 4.2 Recomendaciones específicas
5. Acceso a los datos abiertos y datos de investigación
 - 5.1 Directorios y recolectores de datos
 - 5.2 Portales de datos abiertos
 - 5.3 Repositorios de datos multidisciplinares
 - 5.4 Repositorios institucionales

Referencias

Módulo 3: VISUALIZACIÓN Y REPRESENTACIÓN GRÁFICA DE DATOS

José Carlos Toro Pascua

1. ¿En qué consiste la representación gráfica o visualización de datos?
2. Importancia y beneficios de la visualización gráfica de datos
3. Necesidades y características en un proceso de visualización
4. Tipos de visualizaciones
5. Tipo de gráficos y utilidad
6. Tipo de software o plataformas de servicios generadores de gráficos
7. Fases de creación de una visualización de datos
8. Software o plataformas de servicios generadores de gráficos
9. Errores a evitar
10. Bases de datos de imágenes y utilidades gráficas

Referencias

Módulo 4: LA GESTIÓN DE DATOS DE INVESTIGACIÓN

Yolanda Martín González y Ana Iglesias Rodríguez

1. La gestión de datos de investigación
2. El proceso de gestión de datos de investigación (fases, principios FAIR, ...)
3. El plan de gestión de datos
 - 3.1 Qué es
 - 3.2 Por qué es necesario elaborar un plan de gestión de datos
 - 3.3 Requerimientos de las agencias de financiación relativos a la gestión de datos

Referencias

Módulo 5: LA REUTILIZACIÓN Y CITACIÓN DE DATOS

Ana Iglesias Rodríguez y Yolanda Martín González

1. Qué es la reutilización de información
2. Normativa europea y estatal sobre reutilización
3. Datos abiertos y reutilización de la información
 - 3.1 Datos abiertos y reutilización en el sector público
 - 3.2 Datos abiertos de investigación y su reutilización
 - 3.3 Citación de los datos

Referencias

5. METODOLOGÍA Y MATERIALES

La metodología será online. Se trabajará en la Plataforma Moodle de la Universidad de Salamanca (denominada Studium), en la que se incluirá, para cada uno de los módulos que conforman el curso, distintos materiales elaborados por los miembros del grupo de investigación que participan en el proyecto. Dichos materiales han sido diseñados específicamente para este curso y son los siguientes:

- a) Un documento en el que se encontrarán explicados los principales contenidos de cada módulo, así como las referencias bibliográficas necesarias para la ampliación y/o consulta.
- b) Una o varias píldoras formativas de corta duración para cada módulo en las que se presentarán los aspectos esenciales de la temática tratada.
- c) Distintos tipos de actividades en función del contenido tratado para reforzar su aprendizaje.

La inscripción es libre y puede realizarse desde cualquier lugar y en cualquier momento.

6. ACTIVIDADES PROPUESTAS

Módulo 1. Protección de datos

Actividad 1. Reflexión sobre la protección de datos personales, académicos y de investigación. Tiempo previsto 2h / 2h30min.

Píldora formativa: Protección de datos personales, académicos y de investigación.

Módulo 2. Confiabilidad de los datos

Actividad 1. Búsqueda y selección de datos confiables. Tiempo previsto: 2h.

Píldora formativa: Confiabilidad y selección de los datos abiertos.

Módulo 3. Visualización de datos

Actividad 1. Evaluación de visualizaciones de datos: ejemplos de diseño correctos e incorrectos. Tiempo previsto: 20min.

Actividad 2. Diseño de una infografía con Piktochart. Tiempo previsto: 1h / 1h30min.

Píldora formativa: Visualización y representación gráfica de datos.

Módulo 4. Gestión de datos

Actividad 1. Plan de gestión de datos. Tiempo previsto: 1h.

Píldora formativa: La gestión de los datos de investigación.

Módulo 5. Reutilización de datos

Actividad 1. Reutilización de datos de investigación. Tiempo previsto: 1h.

Actividad 2. Reutilización de datos del sector público. Tiempo previsto: 1h.

Píldora formativa: La reutilización y citación de datos.

7. EVALUACIÓN

Todos los módulos se abrirán de forma conjunta, sin que haya que superar cada uno, para seguir avanzando en el curso.

Cada uno de los módulos constará de una última parte de autoevaluación que permitirá conocer el grado de adquisición de los conocimientos adquiridos en relación con los datos.

8. EQUIPO DOCENTE

- **Críspulo Travieso Rodríguez.** Departamento de Biblioteconomía y Documentación. Universidad de Salamanca. E-mail: ctravieso@usal.es
- **José Carlos Toro Pascua.** Departamento de Biblioteconomía y Documentación. Universidad de Salamanca. E-mail: jctoro@usal.es
- **Azucena Hernández Martín.** Departamento de Didáctica, Organización y Métodos de Investigación. Universidad de Salamanca. E-mail: azuher@usal.es
- **Yolanda Martín González.** Departamento de Biblioteconomía y Documentación. Universidad de Salamanca. E-mail: ymargon@usal.es
- **Josefa Gallego Lorenzo.** Departamento Patrimonio Artístico y Documental. Universidad de León. E-mail: josefa.gallego@unileon.es
- **Ana Iglesias Rodríguez.** Departamento de Didáctica, Organización y Métodos de Investigación. Universidad de Salamanca. E-mail: anaiglesias@usal.es

LA ALFABETIZACIÓN EN DATOS EN EL CONTEXTO
UNIVERSITARIO DE LAS CIENCIAS SOCIALES
(ALFADATAUniv_CCSS)

PROTECCIÓN DE DATOS PERSONALES, ACADÉMICOS Y DE INVESTIGACIÓN

Crispulo Travieso Rodríguez
Azucena Hernández Martín



PROTECCIÓN DE DATOS PERSONALES, ACADÉMICOS Y DE INVESTIGACIÓN

ÍNDICE

1. ¿Qué es un dato personal?
2. ¿Cuáles son estos principios y qué implican cada uno de ellos?
3. ¿Cuáles son los derechos del ciudadano en lo que respecta a la protección de datos y cómo se pueden ejercitar?
4. ¿Y qué ocurre con las Garantías de datos digitales?
5. ¿Cuáles son las autoridades de protección de datos personales en España?
6. ¿Cómo se aplica el actual régimen jurídico de protección de datos de carácter personal en el ámbito docente universitario?
7. ¿Qué entendemos por datos de investigación o datos científicos en abierto?
8. ¿Qué leyes o normativas contemplan la regulación de los datos abiertos de investigación?
9. ¿Qué consideraciones éticas debe contemplar la preservación de datos de investigación?
10. ¿Por qué es importante llevar a cabo una descripción de los datos de investigación?
11. ¿En qué tipos de formatos documentales se pueden materializar los datos de investigación?
12. ¿Qué acciones se pueden llevar a cabo para mejorar la descripción y la visibilidad de los datos?
13. ¿Existen formatos o esquemas de descripción de datos específicos para datos de investigación?

Referencias

1. ¿Qué es un dato personal?

Los datos personales son cualquier información relativa a personas físicas que se puedan identificar. Dichos datos pueden ser identificativos (nombre, apellidos, número del documento nacional de identidad), referidos al origen étnico o racial de la persona; a su situación laboral, financiera, de salud, opiniones políticas, convicciones religiosas, así como otros que puedan identificar su vida y orientación sexual.

Actualmente, la protección de datos personales se rige por:

- El reglamento (UE) 2016/679 del parlamento europeo y del consejo de 27 de abril de 2016 relativo a la Protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos.
- La Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y Garantía de los Derechos Digitales.

La protección de las personas físicas en lo que concierne al tratamiento de los datos personales es un derecho fundamental protegido por el artículo 18.4 de la Constitución Española.

La legislación considerada ha de ser cumplida por los responsables de tratamientos de datos personales (Empresas, Administraciones Públicas, Centros educativos...) a los que se han facilitado los datos de ca-

rácter personal, de acuerdo con el Reglamento General de Protección de Datos (RGPD). En este sentido, el consentimiento para utilizar los datos personales debe darse mediante un acto que refleje una voluntad libre, informada e inequívoca del interesado, bien por escrito, verbalmente, por medios electrónicos, o mediante la señalización de una casilla en un sitio web de Internet. Dicho consentimiento se puede revocar en cualquier momento y de forma sencilla.

La normativa anteriormente citada configura una serie de principios que deben ser respetados cuando se produzca este tratamiento.

2. ¿Cuáles son estos principios y qué implican cada uno de ellos?

1. **El principio de licitud, lealtad y transparencia** implica que cuando se recojan datos de carácter personal han de ser tratados de manera lícita, leal y transparente, de acuerdo con las Bases Jurídicas del RGPD.
2. **El principio de limitación de la finalidad** supone que los datos personales sólo serán recogidos para fines concretos, explícitos y legítimos.
3. **El principio de minimización o limitación de los datos** a aquellos que se precisan en relación con los fines para los que son recogidos.
4. **El principio de exactitud de los datos**, y, por tanto, tienen que estar actualizados. Si no es el caso, se deberán adoptar las medidas necesarias para suprimir o rectificar los datos inexactos.
5. **El principio del plazo de conservación de los datos**, que se mantendrán por un periodo de tiempo no superior al necesario para cumplir con los fines del tratamiento. La conservación de datos debe limitarse a las finalidades para las cuales se han recabado dichos datos. Una vez cumplidas estas finalidades deben ser borrados o, al menos, desprovistos de todo elemento que permita identificar a los interesados. Ello se recoge en el RGPD, al igual que en el Título V de La Ley de Protección de datos Personales y Garantía de los derechos digitales (LOPDG); al igual que la necesidad de otorgar una responsabilidad activa al responsable del tratamiento de los datos, obligándole a bloquear estos (impedir su tratamiento y visualización) cuando se vayan a rectificar o suprimir.
6. **El principio de integridad y seguridad** garantiza la adecuada seguridad de los datos personales, aspecto que incluye la protección contra el tratamiento ilícito o no autorizado, su pérdida, daño accidental, destrucción. Para ello, se aplicarán las medidas técnicas y organizativas que se consideren apropiadas.
7. **El principio de responsabilidad proactiva** supone que los responsables y encargados del tratamiento de datos deben cumplir estos principios y ser capaces de demostrar dicho cumplimiento. El RGPD establece un catálogo de medidas que ambos deben aplicar para garantizar que los tratamientos de los datos se ajustan a todos los principios anteriores: *el análisis de riesgo* con la finalidad de adoptar las correspondientes medidas de seguridad, *la notificación de brechas de seguridad* o la realización de *evaluaciones de impacto de protección* de datos. Si una entidad va a recabar datos de menores de 14 años, se exigirá el consentimiento de los padres o tutores. Para ayudar a profesores, tutores y padres en el respeto a la privacidad de los menores, la web de la Agencia Estatal de Protección de Datos (AEPD) contiene una sección dedicada específicamente a los mismos, denominada Canal Joven.

3. ¿Cuáles son los derechos del ciudadano en lo que respecta a la protección de datos y cómo se pueden ejercitar?

La normativa en materia de protección de datos permite que se pueda ejercitar ante el responsable correspondiente los derechos de acceso, rectificación, oposición, supresión, limitación del tratamiento de los datos, así como su portabilidad.

El derecho de acceso permite que cualquier ciudadano conozca a través del responsable del tratamiento de datos, si se están tratando o no sus datos personales, y, en caso de que sea así, poder obtener copia de los datos que están siendo objeto de tratamiento, ser informado sobre los fines del tratamiento, y a qué destinatario/s se comunicaron.

El derecho de rectificación supone la corrección de datos personales que sean inexactos. Para ello será necesario cumplimentar la solicitud correspondiente a objeto de completarlos o corregirlos, pudiendo añadir, si es necesario, la documentación justificativa de la inexactitud o el carácter incompleto de ciertos datos.

El derecho de oposición implica que cualquier ciudadano puede oponerse a que el responsable del tratamiento de los datos pueda utilizar estos, a excepción de que que el responsable acredite motivos imperiosos de interés público que prevalezcan sobre los intereses, derechos y libertades personales.

El derecho de supresión, también denominado “derecho al olvido”, se podrá ejercitar cuando los datos personales ya no sean necesarios, o si se retira el consentimiento prestado al responsable del tratamiento de los datos. Este derecho, como se ha indicado, se relaciona con el derecho al olvido.

En el caso de Internet y sus motores de búsqueda, el derecho al olvido tiene en cuenta los siguientes supuestos:

- La eliminación de fotos y vídeos que han sido publicados en internet sin consentimiento, tratándose en ambos casos de datos personales, de acuerdo con el derecho de supresión de los mismos. Para ello, la persona debe dirigirse, acreditando su identidad e indicando los enlaces donde aparecen los vídeos y fotos, ante quien los haya subido a la red, solicitando, por tanto, el borrado de estos. Si no responden o la respuesta es insatisfactoria, se puede interponer una reclamación ante AEPD.
- La usurpación de identidad, el acoso, los vídeos ofensivos y otros supuestos en los que se haya publicado contenido sin consentimiento. En todos estos supuestos es necesario retirar el contenido y denunciar el abuso o acoso, la vulneración de la privacidad, los contenidos sexuales y/o violentos.

El derecho a la portabilidad de los datos con la finalidad de reforzar todavía más el control de los datos personales, de forma que cuando el tratamiento se efectúe por medios automatizados, se reciban los datos personales en un formato estructurado, de uso común, de lectura mecánica e interoperable, y se puedan transmitir a otro responsable del tratamiento, en el marco, por ejemplo, de la ejecución de un contrato.

Todos estos derechos son (1) gratuitos, (2) el responsable debe informar sobre los medios para ejercitar esos derechos, (3) serán medios accesibles.

4. ¿Y qué ocurre con las Garantías de datos digitales?

En el Título X de la LOPDPGDD se incorporan una serie de normas para garantizar los derechos digitales de la ciudadanía. En este sentido son objeto de regulación en esta ley los derechos y libertades propios de la era digital, como es el caso de la neutralidad de Internet, proporcionando una oferta transparente sin discriminación por motivos técnicos o económicos, el derecho de acceso universal a Internet, con independencia de la condición personal (género, brecha generacional, necesidades especiales), social, económica o geográfica; el derecho a la seguridad digital. El derecho a la Educación Digital, garantizando la plena inserción del alumnado en la sociedad digital por parte del sistema educativo, así como el aprendizaje de un consumo responsable y un uso crítico y seguro de los medios digitales. También se ha de garantizar la intimidad personal y familiar y la protección de los datos personales. Así, se explicita que las administraciones educativas deberán incluir en el currículo la competencia digital y evitar las situaciones de riesgo, derivadas de la inadecuada utilización de las TIC. En lo concerniente al alumnado universitario se establece la necesidad de que los planes de estudio de las distintas Titulaciones garanticen una formación en el uso y seguridad de los medios digitales.

De igual forma, se asegura el derecho a la intimidad y el uso de dispositivos en el ámbito laboral por parte de los trabajadores y empleados públicos. Así, por ejemplo, sólo podrán acceder a los contenidos derivados del uso de medios digitales, a efectos de controlar el cumplimiento de las obligaciones laborales. Esta intimidad se hace extensible al uso de dispositivos de videovigilancia y de grabación de sonidos en el lugar de trabajo, informando con carácter previo y de forma expresa, clara y concisa; y nunca en zonas de esparcimiento o de descanso.

5. ¿Cuáles son las autoridades de protección de datos personales en España?

En el título VII de La LOPDGDD, se explicita que serán (a) La Agencia Española de Protección de Datos, que ejercerá sus funciones de acuerdo a la normativa estatal; (b) Las Autoridades autonómicas de protección de datos, ejerciendo sus funciones y potestades de acuerdo con la normativa autonómica.

6. ¿Cómo se aplica el actual régimen jurídico de protección de datos de carácter personal en el ámbito docente universitario?¹

Partiendo del hecho de que en la acción docente se maneja un elevado número de datos personales, tanto de los estudiantes como de los docentes, las herramientas que actualmente se utilizan para la docencia virtual deben cumplir con todos los requisitos que se plantean en la normativa aplicable. En este sentido, han de emplearse las herramientas oficiales que haya provisto cada institución universitaria para garantizar la legalidad y seguridad adecuadas.

- Las prácticas y documentos que incluyan datos personales del alumnado se podrán conservar un máximo de un curso académico, salvo que las mismas formen parte de las evidencias de evaluación del alumnado. Hay que tener en cuenta la posibilidad de conservarlos durante más tiempo, pero, a través de una copia debidamente bloqueada, es decir, a la que no se podrá tener acceso.
- Si un profesor deseara reutilizar estos materiales para otras actividades académicas deberá eliminar previamente cualquier dato personal de los estudiantes o de terceras personas.
- La publicación de calificaciones se efectuará en los espacios virtuales que la institución universitaria tiene habilitados para tal fin, de modo que sólo pueda acceder el alumnado implicado en el proceso de evaluación. No se podrán publicar nunca en espacios abiertos. En el caso de que no fuera posible, podrá realizarse en los tablones de anuncios del centro, siempre que no se encuentren en las zonas comunes, se garantice que el acceso a los mismos queda restringido a dichas personas, por parte de quienes carecen de interés en el mismo.
- Dicha publicación debe contener los datos mínimos necesarios para que el estudiante en cuestión pueda identificarse, de acuerdo con lo indicado en el artículo 5.1.c de Reglamento General de Protección de Datos, que hace alusión a la necesidad de seguir el criterio de minimización de los datos. Y seguir también las Orientaciones de la Disposición Adicional 7ª de la LOPDGDD, de acuerdo con la cual se explicita que se identificará al afectado mediante su nombre y apellidos, añadiendo cuatro cifras numéricas aleatorias del documento nacional de identidad. En este sentido, las calificaciones se publicarán con nombre y apellidos, así como la nota obtenida, y en el caso de que en un mismo grupo dos o más estudiantes puedan coincidir en nombre y apellidos, se podrá utilizar el llamado DNI ofuscado. Ej. ***3212**.
- En cuanto al acceso de los progenitores a las calificaciones de sus hijos, económicamente dependientes, este acceso se podrá realizar cuando la finalidad sea la de utilizarlas en un proceso judicial para la solicitud de modificación, por ejemplo, de la pensión de alimentos. En definitiva, en el presente caso concurre un interés legítimo del solicitante para poder obtener las calificaciones académicas de un hijo mayor de edad.

1 Contenido recogido a partir de la consulta del documento elaborado por el Servicio Jurídico de la Agencia Estatal de Protección de Datos.

- Otra cuestión es la relativa a la grabación y divulgación de imágenes en actos públicos organizados por la Universidad. Al tratarse de actos organizados por las propias Universidades, de acuerdo con sus estatutos o normas de funcionamiento para el cumplimiento de una misión de interés público, la grabación y divulgación podrá realizarse siempre que se informe previamente de ello y de los fines a los que se destinan los datos personales.
- La grabación de los exámenes orales puede ser necesaria como medio de prueba para el ejercicio de los derechos del alumnado, así como para que el profesor pueda justificar la evaluación realizada, sin perjuicio de que puedan admitirse otros medios probatorios (por ejemplo, exigir al estudiante un esquema de lo que va a exponer). Por consiguiente, con la finalidad señalada, y siempre que las normas internas de la Universidad prevean la grabación de los exámenes orales, el tratamiento se encontrará fundamentado, de acuerdo con una misión realizada en interés público.
- En cuanto a la grabación de las sesiones docentes, es conveniente e incluso necesaria, en el caso, por ejemplo, de las Universidades a Distancia, para el ejercicio de la función educativa, debiendo limitarse su acceso al personal docente y a los alumnos a los que vaya dirigida, sin que pueda ser utilizada ulteriormente para otros fines, como su divulgación pública, que requeriría del consentimiento expreso de los afectados. Estas grabaciones pueden realizarse por parte de la propia Universidad con fines de control laboral, para la ejecución de un contrato en el que el interesado es parte o para la aplicación a petición de este de medidas precontractuales. En cuanto a la grabación de la sesión docente por parte de los alumnos debería identificarse la finalidad concreta de dicha grabación, que requerirá, en todo caso, del consentimiento del docente, así como del resto de los alumnos si su imagen o su voz pudieran ser objeto de grabación.

7. ¿Qué entendemos por datos de investigación o datos científicos en abierto?

Los datos abiertos de investigación comparten las características de los *open data* en lo que tiene que ver con la transparencia de su proceso de creación y también en cuanto a la rendición de cuentas y la reutilización (Peset y Fernández-López, 2014), especialmente al hablar de investigaciones llevadas a cabo mediante fondos públicos. Pero, además, se definen por el contexto en el que surgen y al que sirven. Así, pueden considerarse datos de investigación “todo aquel material que ha sido registrado durante la investigación, reconocido por la comunidad científica y que sirve para certificar los resultados de la investigación que se realiza” (Torres-Salinas y otros, 2012). Si a ello sumamos su carácter abierto, obtendremos que han de poder ser utilizados, reutilizados y redistribuidos libremente por cualquier persona, siempre que se reconozca su autoría y que se vuelvan a compartir una vez se hayan transformado durante un uso posterior.

8. ¿Qué leyes o normativas contemplan la regulación de los datos abiertos de investigación?

Hay varios documentos legislativos que abordan la cuestión de la protección de datos resultantes de investigaciones. En el ámbito europeo debemos señalar el Reglamento (UE) 2016/679 relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos establece en su artículo 98.1:

(...) El tratamiento con fines de archivo en interés público, fines de investigación científica o histórica o fines estadísticos estará sujeto a las garantías adecuadas, con arreglo al presente Reglamento, para los derechos y las libertades de los interesados. Dichas garantías harán que se disponga de medidas técnicas y organizativas, en particular para garantizar el respeto del principio de minimización de los datos personales. Tales medidas podrán incluir la seudonimización, siempre que de esa forma puedan alcanzarse dichos fines. Siempre que esos fines pueden alcanzarse mediante un tratamiento ulterior que no permita o ya no permita la identificación de los interesados, esos fines se alcanzarán de ese modo.

Este reglamento comunitario (RGPD) se complementa con la legislación estatal, en este caso, a través de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y Garantía de los Derechos Digitales.

Se ha de considerar que el tratamiento de datos con finalidades de investigación científica siempre estará condicionado a asegurar y garantizar los derechos y deberes de las personas. De nuevo debe prestarse especial atención al principio de minimización, es decir, al uso exclusivo de los datos que sean realmente necesarios. Así, las investigaciones que requieran de datos de salud, datos sensibles de colectivos en situación de vulnerabilidad, datos genéticos, biométricos, etc., deberán incluir el registro de actividades de tratamiento, así como un informe de evaluación de impacto elaborado con la ayuda y asesoramiento del delegado de Protección de Datos.

9. ¿Qué consideraciones éticas debe contemplar la preservación de datos de investigación?

Habitualmente, se ha de contemplar una serie de requisitos éticos en cualquier política de preservación de datos de investigación. Cada institución puede tener sus propia reglamentación o recomendaciones en las que se desarrolla, pero a grandes rasgos, coinciden en que cualquier investigación que implique la participación de humanos o de seres vivos ha de estar sujeta a revisión ética. Del mismo modo, también lo estarán aquellas investigaciones referidas a personas o que conlleven el almacenamiento de la identificación de los individuos, a menos que los datos sean obtenidos a través de estudios puramente observacionales de comportamiento público.

Con estudios de observación se alude a aquellos que:

- Se producen en un foro abierto al público general.
- Son no invasivos.
- No da lugar a identificar a los participantes.

Además, las asociaciones profesionales que representan campos académicos concretos tienden a publicar sus propias directrices éticas, que deben estar en constante actualización y consenso.

10. ¿Por qué es importante llevar a cabo una descripción de los datos de investigación?

Para que los fines que persigue el paradigma de la ciencia abierta se alcancen es fundamental establecer una política y una planificación de la preservación de los datos científicos. En ese sentido, para realizar una adecuada previsión de esas acciones habrá que tener en cuenta, entre otros aspectos, el volumen de datos y su crecimiento potencial, la variedad en cuanto a la tipología y formato de los datos de investigación, que pueden estar derivados de metodologías y fuentes de recogida muy distintas, la probable dependencia para su uso y manipulación a determinadas tecnologías de *hardware* y *software* y la futura degradación a la que probablemente deban hacer frente.

Junto con esos planes, será la descripción que se asocia y acompaña a los conjuntos de datos abiertos la que permitirá que estos puedan ser recuperados y, a la postre, reutilizados (Travieso Rodríguez y Ferreira Araújo, 2019); por lo tanto, ese resultado de la indización puede ser considerado en sí mismo un indicio de la calidad de un dato de investigación abierto (Peset y Fernández-López, 2014; Peset Mancebo y otros, 2017).

11. ¿En qué tipos de formatos documentales se pueden materializar los datos de investigación?

- Documentos, hojas de cálculo.

- Cuadernos de laboratorio, cuadernos de campo, diarios de investigación.
- Cuestionarios, transcripciones, libros de códigos.
- Cintas de audio o vídeo.
- Fotografías, películas.
- Respuestas de prueba.
- Diapositivas, artefactos, especímenes, muestras.
- Modelos, algoritmos, scripts.
- Contenido de una aplicación (entrada, salida, archivos de registro para *software* de análisis, *software* de simulación, esquemas).
- Metodologías y flujos de trabajo.
- Procedimientos y protocolos operativos estándar.

12. ¿Qué acciones se pueden llevar a cabo para mejorar la descripción y visibilidad de los datos?

Una de las estrategias más eficaces para fomentar la comunicación y reutilización posterior de los datos es asociar los mismos a un esquema de metadatos enriquecidos: los metadatos proporcionan la contextualización y descripción de un objeto de datos, que permite que ese objeto de datos sea correctamente interpretado y convertido en información. Se trata, en definitiva, de desarrollar un sistema para crear metadatos (o aplicar uno ya existente) que describa con precisión sus datos, incluyendo, como mínimo, información como tipo de datos, fecha de creación y creador.

Por otro lado, es esencial documentar todo el proceso. En ese sentido, se ha de mantener una documentación detallada del proceso de creación y manipulación de los datos, utilizando herramientas adecuadas que garanticen su integridad y usabilidad. En la práctica, documentar los datos significa incluir información sobre cómo fueron recopilados, procesados y analizados.

13. ¿Existen formatos o esquemas de descripción de datos específicos para datos de investigación?

Podemos encontrar varios esquemas de metadatos que pueden ser empleados para la descripción de datos de investigación. Debido a la enorme variedad de datos derivados de la actividad científica, sería impensable que un mismo sistema fuese válido para todos los tipos y para todas las disciplinas científicas y metodologías de investigación.

Pueden servir formatos de descripción generales, como el estándar *Dublin Core* u otros específicos según el tipo documental, como EAD (Encoding for Archival Data) CDWA (Categories for the Description of Works of Art), *VRA Core* o la disciplina científica. También existen formatos específicos para descripción de conjuntos de datos, como *DataCite Metadata Schema* o *Data Documentation Initiative* (DDI).

Independientemente de que se opte por un sistema ya establecido o se adapte la descripción a un determinado entorno *ad hoc*, hay una serie de campos de descripción comunes que debieran aparecer en la representación resultante de este proceso. Entre ellos destacan: creador, título, año de creación, editor, tema, palabras clave, fechas de recogida de los datos, idioma, tipo de recurso, identificador, tamaño, formato, versión de los datos, derechos, descripción, geolocalización, referencia de financiación.

En adición, es conveniente emplear lenguajes controlados para designar la materia y el análisis de contenido de los datos, para facilitar la integración con otras colecciones de datos y situarlos en las disciplinas científicas correspondientes.

Referencias

Agencia Española de Protección de Datos. *Guía de Protección de Datos. Derechos del ciudadano*. <https://www.ismsforum.es/ficheros/descargas/guiaciudadanoaepd1496923753.pdf>

Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y Garantía de los Derechos Digitales. <https://www.boe.es/buscar/act.php?id=BOE-A-2018-16673>

Reglamento (UE) 2016/679 del parlamento europeo y del consejo de 27 de abril de 2016 relativo a la Protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos. <https://www.boe.es/buscar/doc.php?id=DOUE-L-2016-80807>

Peset, F.; Fernández-López, A. (2014). Carencias informativas de los datos abiertos en España. *Anuario ThinkEPI*, 8, 318-321.

Peset Mancebo, F.; Aleixandre Benavent, R.; Blasco Gil, Y; Ferrer Sapena, A. (2017). Datos abiertos de investigación. Camino recorrido y cuestiones pendientes. *Anales de Documentación*, 20(1). <https://doi.org/10.6018/analesdoc.20.1.272101>

Torres-Salinas, D.; Robinson-García, N.; Cabezas-Clavijo, Á. (2012). Compartir los datos de investigación: introducción al *data sharing*. *El profesional de la información*, 21(2), 173-184. <https://doi.org/10.3145/epi.2012.mar.08>

Travieso Rodríguez, C., & Ferreira Araújo, R. (2019). Aspectos metodológicos de los datos abiertos de investigación: análisis de los conjuntos de datos de la colección SciELO incluidos en Figshare. *Revista Española De Documentación Científica*, 42(3), e242. <https://doi.org/10.3989/redc.2019.3.1597>

LA ALFABETIZACIÓN EN DATOS EN EL CONTEXTO
UNIVERSITARIO DE LAS CIENCIAS SOCIALES
(ALFADATAUniv_CCSS)

PROTECCIÓN DE DATOS PERSONALES, ACADÉMICOS Y DE INVESTIGACIÓN

Actividad

Críspulo Travieso Rodríguez
Azucena Hernández Martín



PROTECCIÓN DE DATOS PERSONALES, ACADÉMICOS Y DE INVESTIGACIÓN

[Actividad propuesta]

Tiempo aproximado de realización de la actividad:
2 horas / 2,5 horas.

1. Objetivo

El objetivo de esta actividad es reflexionar sobre la protección de datos en distintos ámbitos, para asimilar y profundizar en los contenidos teóricos y normativos.

2. Enunciado explicativo de la actividad

Tras la lectura de los textos facilitados a continuación, responda brevemente a las correspondientes preguntas relacionadas con la protección de datos.

- A. *El examen de apps: los datos.*
- B. *Guía sobre privacidad y seguridad en Internet.*
- C. *Los riesgos de investigar con datos personales.*

3. Desarrollo e instrucciones para realizar la actividad

Preguntas para la reflexión para lecturas A y B.

- ¿Cuáles pueden ser las principales causas de un posible ciberataque que atente contra los datos personales, bien sea en su dispositivo móvil o en su ordenador? Le sugerimos que enumere y describa tres causas.
- ¿Qué consecuencias puede traer consigo un ciberataque de este tipo? Le animamos a que reflexione y explique dos de ellas.
- Considerando las recomendaciones realizadas en la primera de las lecturas, así como en la guía sobre privacidad y seguridad en Internet,
 - ¿Qué consejos y recomendaciones proporcionaría a sus estudiantes para proteger sus datos personales?
 - ¿Qué medidas adoptaría para proteger los principales datos académicos y de investigación?

Preguntas para la reflexión para lectura C.

- ¿Ha tenido situaciones de conflicto relacionadas con la protección de datos científicos en el marco de una investigación? Puede hacer un somero relato de estas y de las soluciones que encontró para resolverlas.
- En su opinión, ¿sobre quién debe caer la responsabilidad de gestionar y mantenerse al día de la normativa sobre protección de datos de investigación?

- En función de su experiencia, ¿a qué instancias ha recurrido para plasmar el plan de protección de datos al elaborar un proyecto de investigación?

4. Recursos requeridos para realizar la actividad

- A. *El examen de apps: los datos*. Recuperado el día 18 de abril de 2024 de <https://www.aepd.es/prensa-y-comunicacion/blog/el-examen-de-apps-y-iv-los-datos>
- B. *Guía sobre privacidad y seguridad en Internet*. Recuperado el día 25 de abril de 2024 de [guia-privacidad-y-seguridad-en-internet.pdf](https://www.aepd.es/guia-privacidad-y-seguridad-en-internet.pdf) (aepd.es)
- C. Martínez Martínez, R. (2019, 26 de noviembre). Los riesgos de investigar con datos personales. *The Conversation*. Recuperado el día 13 de abril de 2024 de <https://theconversation.com/los-riesgos-de-investigar-con-datos-personales-127388>

LA ALFABETIZACIÓN EN DATOS EN EL CONTEXTO
UNIVERSITARIO DE LAS CIENCIAS SOCIALES
(ALFADATAUniv_CCSS)

CONFIABILIDAD EN LA EVALUACIÓN Y SELECCIÓN DE LOS DATOS ABIERTOS

Josefa Gallego Lorenzo



CONFIABILIDAD EN LA EVALUACIÓN Y SELECCIÓN DE LOS DATOS ABIERTOS

ÍNDICE:

1. Conceptos de interés: Dato, información, fake news, confiabilidad y alfabetización en datos.
2. Búsqueda y selección de datos.
3. Criterios para seleccionar y evaluar la calidad de los datos.
4. Recomendaciones y directrices para evaluar la calidad de los datos abiertos.
 - 4.1. Recomendaciones generales.
 - 4.2. Recomendaciones específicas.
5. Acceso a los datos abiertos y datos de investigación.
 - 5.1. Directorios y recolectores de datos.
 - 5.2. Portales de datos abiertos.
 - 5.3. Repositorios de datos multidisciplinares.
 - 5.4. Repositorios institucionales.

Referencias

1. CONCEPTOS DE INTERÉS: DATO, INFORMACIÓN, FAKE NEWS, CONFIABILIDAD Y ALFABETIZACIÓN EN DATOS

Vivimos en una sociedad “datificada” donde los datos digitalizados forman parte de nuestras vidas y cualquier operación que realizamos en internet deja un rastro de datos e información que, una vez analizados, pueden convertirse en conocimiento. Es por ello que, debido a la gran cantidad de datos generados y disponibles a partir del uso de aplicaciones web, redes sociales y diferentes dispositivos tecnológicos, la sociedad necesita ser instruida en el uso de datos, con el fin de fomentar el conocimiento y la inquietud por procesar información de una forma autónoma.

Para comprender mejor este tema, se revisan los siguientes conceptos: dato, información, *fakes news*, confiabilidad y alfabetización en datos.

Dato es la unidad básica de información, no procesada, es decir, la materia prima con la que se trabaja para extraer información valiosa (conocimiento) y tomar decisiones. Los datos pueden ser de diferente naturaleza y procedencia: datos de investigación (ciencia abierta), datos públicos (open data), datos gubernamentales; por el origen de la información, pueden ser datos primarios o secundarios, cuantitativos o cualitativos; por su formato, pueden ser textuales, numéricos, estructurados, multimedia; por su contenido, educativos, científicos, estadísticos, financieros, tecnológicos, geográficos, culturales, etc. Como podemos observar, los datos pueden proceder de diversas fuentes.

Información es el conjunto de datos procesados que tienen significado. La información, además de ser un conjunto de datos, necesita contar con ciertas propiedades, como relevancia (en función del propósito o tema considerado), precisión o exactitud (para que obtenga un cierto nivel de confiabilidad), completa para tomar decisiones; adecuada, oportuna y comprensible. Dichos aspectos hacen que la información procesada por un sistema de información se convierta en un recurso útil, valioso y relevante para la toma de decisiones y necesaria para una gestión inteligente.

Otro factor importante que hay que tener en cuenta, ligado a la sobresaturación de información que recibimos,

mos a través de diferentes canales de difusión, es lo que se conoce como las *fake news* (noticias falsas), que ocupan un papel protagonista a nivel político, mediático, y también, científico, que ante la imparcialidad de la cobertura informativa se aprovechan de la situación creando demanda informativa a través de noticias basadas en datos sensacionalistas y contenidos no contrastados. Son contenidos de intención pernicioso, que apelan a noticias llamativas, impactantes o alarmantes que se propagan con gran virulencia y rapidez, incrementando contenidos falsos y produciendo desinformación, como ocurrió con la crisis sanitaria provocada por la Covid-19.

En esta línea, la red de verificación *Internacional Fact-Checking Network* (IFCN) puso en marcha un proyecto para combatir la desinformación y los bulos que circulaban por las redes, a través de plataformas que ayudan a certificar la credibilidad de los datos. En España, el *fact-checking* se compone de cuatro proyectos, Maldita.es¹, Newtral², Verificat³ y EFE Verifica⁴. La IFLA⁵ también ha publicado una infografía sobre “¿cómo detectar noticias falsas?”. Se trata de dotar a la ciudadanía de conocimientos y destrezas que fomenten la calidad de los datos que reciben a través de una alfabetización digital y mediática.

Ligado a estos conceptos está la **confiabilidad** que representa la seguridad que se otorga a los datos y que tiene que ver con la veracidad y franqueza del procedimiento que se utilice para medir la credibilidad de los datos (Almada Martínez, 2019). Un criterio que puede dar confianza y credibilidad es la procedencia de los datos que siempre deben estar avalados bien por una entidad profesional o bien por un proceso de investigación, ambos buenos indicadores de fiabilidad.

En este sentido, las bibliotecas, archivos y museos son excelentes proveedores de datos confiables, al igual que los profesionales de la información cuya función y responsabilidad es la de proveer datos e información de calidad, ya que son considerados agentes facilitadores de confianza, para ayudar a la ciudadanía a tomar decisiones informadas; y en el ámbito académico denominados infomediarios esenciales por su labor de ayuda para que el docente pueda seleccionar y evaluar datos y obtener recursos de calidad.

En este ámbito académico, la inteligencia artificial está transformando la forma en que accedemos, procesamos y compartimos información, lo que plantea desafíos y oportunidades en términos de confiabilidad. Uno de los principales desafíos es la propagación de la desinformación y las noticias falsas, por ello, tanto los bibliotecarios como los docentes y profesionales de la información en diversos campos tienen la oportunidad de integrar y aplicar servicios de inteligencia artificial (IA) en su trabajo y adaptarse a estos avances tecnológicos (Alonso Arévalo, 2024).

El crecimiento exponencial de los datos en todos los ámbitos de nuestra sociedad nos obliga a estar alfabetizados en datos para saber localizar y comprender la información que contienen. El denominador común de estos datos, para que puedan ser útiles a cualquier profesional, ciudadano, estudiante o particular, es que los datos deben prepararse cuidadosamente antes de su publicación, tienen que estar bien organizados, bien presentados y disponibles para su posterior consulta en acceso público como datos abiertos. Por ello, es recomendable que los datos cumplan con los principios (FAIR) de estar localizables, accesibles, interoperables y reutilizables (Wilkinson et al., 2016). Y, una forma de preservar los datos de forma segura es depositarlos en repositorios digitales, ya que disponen de marcos organizativos de gobernanza sostenible, infraestructura confiable y políticas que respaldan las prácticas acordadas por cada comunidad (Li et al., 2020).

En relación con los conceptos anteriores, la **alfabetización en datos** nos puede ayudar a identificar, analizar y evaluar los datos que estamos buscando, a la vez que nos capacita en habilidades competenciales, necesarias para responder al por qué y para qué los estamos buscando, cómo encontrarlos, cómo comunicarlos de manera crítica y cómo interpretarlos en la toma de decisiones.

En el contexto educativo no existen unas normas o directrices concretas para la alfabetización en datos, pero sí debemos destacar las competencias recogidas, tanto en el *Marco Común de Competencia Digital*

1 <https://maldita.es/>

2 <https://www.newtral.es/>

3 <https://www.verificat.cat/>

4 <https://verifica.efe.com/>

5 <https://www.ifla.org/es/news/como-las-bibliotecas-pueden-ayudar-a-conseguir-soluciones-reales-para-las-noticias-falsas/>

Docente (INTEF 2022) como en la reciente actualización, DigComp 2.2, *Marco de Competencias Digitales de la Ciudadanía* (2022) que sirven de marco referencial. En ambos marcos, se desarrolla la competencia de alfabetización en información y en datos, que a su vez se desdobra en el área competencial de búsqueda y gestión de información y datos. Esta competencia digital es clave para el profesorado, puesto que ofrece conocimientos sobre la búsqueda de datos y localización de contenidos en entornos digitales. El objetivo es que el docente sea capaz de:

- Localizar datos confiables y relevantes, en función de sus necesidades, navegando a través de los datos, seleccionando recursos de forma eficaz, y creando estrategias de búsqueda.
- Analizar, interpretar y evaluar de forma crítica los datos presentados en cualquier formato para comunicar y tomar las decisiones adecuadas.

Estas competencias confluyen, a su vez, en la alfabetización mediática que se define como la capacidad de adquirir competencias y habilidades técnicas, cognitivas, sociales, cívicas, éticas y creativas para desenvolverse en los medios actuales, producir contenido y comprender de manera crítica la información que se recibe (*Alfabetización Mediática e Informativa*. AMI)⁶. En dicho documento, también se plantean cuestiones emergentes sobre inteligencia artificial, educación para la ciudadanía, educación para el desarrollo sostenible y la alfabetización cultural, y el aumento exponencial de información falsa y la desinformación. Este marco combina las tres áreas de alfabetización mediática, alfabetización informativa y alfabetización digital.

2. BÚSQUEDA Y SELECCIÓN DE DATOS CONFIABLES

El gran volumen de contenidos disponibles en la red significa que no siempre se puede acceder fácilmente a lo que se busca. Internet se ha convertido en el principal medio de difusión de la información científica y técnica, tanto para la información que ha pasado por ciertos controles de calidad y procesos de selección que se encuentra en bibliotecas y centros de documentación: bases de datos, repositorios, plataformas de revistas electrónicas; como para la publicada por cualquier persona u organización que sube libremente en la web, ya sea páginas personales, redes sociales, blogs, plataformas de contenidos, sedes web de empresas, organismos o asociaciones.

La explosión de información en la que estamos inmersos, tanto a nivel profesional como en la vida diaria, hace imprescindible el uso de sistemas de recuperación de información (SRI) que nos permitan acceder y gestionar la gran cantidad de datos que nos rodea.

¿Cómo planteamos una búsqueda?

En todo proceso de búsqueda y selección hay que tener en cuenta varias etapas:

- Identificar bien la necesidad de información.
- Seleccionar la herramienta o sistema de recuperación más adecuado (buscadores de datos, catálogos, bases de datos, repositorios de datos).
- Trasladar nuestra necesidad de información al lenguaje documental.
- Aplicar el lenguaje de consulta más adecuado, lenguaje natural, lenguaje documental o palabras clave (tesauro, listas de materias), o lenguaje de interrogación.
- Formular la estrategia de búsqueda, utilizando operadores booleanos.
- Seleccionar y evaluar los resultados obtenidos con el fin de saber si son relevantes o no para nuestra necesidad.
- La búsqueda puede ser sencilla o avanzada.

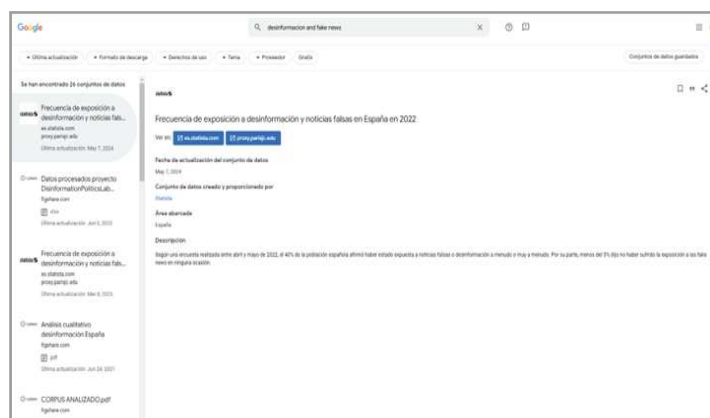
6 UNESCO. <https://www.unesco.org/es/media-information-literacy>

A la hora de realizar una búsqueda podemos utilizar los operadores lógicos. Estos comandos ayudan a recuperar resultados relevantes de las bases de datos, incluidas las búsquedas generales en Internet, ya sea en un motor de búsqueda especial o en redes sociales o en servicios de comercio electrónico, todos ellos utilizan operadores booleanos: AND, OR, NOT (Y, O, NO). El uso de estos operadores en la búsqueda eliminará resultados irrelevantes y ahorrará tiempo y esfuerzo (AND localiza documentos que contengan los dos términos, reduce resultados; OR localiza documentos que contenga sólo uno de los términos y NOT que contenga un término, pero no el otro).

Un motor de búsqueda interesante por su fiabilidad de resultados es *Dataset Search* <https://datasetsearch.research.google.com>

Google Dataset Search es un motor de búsqueda especializado que permite encontrar conjuntos de datos almacenados en la web proveniente de bases de datos gubernamentales, recursos públicos, librerías digitales y websites. Los datos están alojados en miles de repositorios en la web, haciendo que sean universalmente accesibles y útiles. Encuentra conjuntos de datos que se pueden explorar, refinar y filtrar por última actualización, derechos de uso, tema, formato.

Planteada la búsqueda proporciona los datos estructurados en una lista que se encuentra a la izquierda y en la parte derecha los datasets con información sobre: el título, el formato del documento para descargar, el identificador DOI, fecha de actualización, el conjunto de datos, los autores, la licencia de atribución *Creative Commons* que tienen y la descripción del documento. Los datos obtenidos se pueden guardar, compartir y citar.



Fuente: Dataset search. Búsqueda: desinformación and *fake news*

Este buscador tiene una guía de búsqueda que es importante consultar para conocer mejor su funcionamiento: <https://newsinitiative.withgoogle.com/es-mx/resources/trainings/dataset-search-quickstart-guide/>

3. CRITERIOS PARA SELECCIONAR Y EVALUAR LA CALIDAD DE LOS DATOS

La convivencia de recursos de calidad con otros que no han pasado ningún tipo de filtro profesional, que no están estructurados ni controlados, y que se encuentran dispersos en diferentes servicios, con diferentes formatos y contenidos, sin normalizar y sin filtrar, crea dificultades a la hora de identificar su calidad y fiabilidad. Por eso es necesario disponer de una serie de conocimientos y habilidades para evaluar recursos y contenidos web y saber discernir la información fiable y de calidad de la que no lo es (UNESCO, 2024).

Es necesario que desde las bibliotecas se fomente el uso de herramientas que capaciten en la evaluación y selección de la alfabetización en datos, tanto para el docente como para el discente, puesto que son los profesionales que mejor preparados están para localizar información en portales de datos abiertos generales y especializados, sitios web gubernamentales, portales de instituciones de estadísticas, repositorios de datos de investigación, repositorios institucionales, bases de datos...

Para llevar a cabo cualquier investigación científica es fundamental contar con la calidad de los datos disponibles y que estos dispongan de fiabilidad y validez y estén exentos de errores técnicos y metodológicos.

La revisión bibliográfica efectuada apunta que existe una abundante literatura científica sobre los diferentes criterios y herramientas que se pueden utilizar para evaluar la credibilidad de las fuentes de datos en Internet, checklists, formularios, cuestionarios como el *CRAAP Test*⁷ de California State University (*currency, relevance, authority, accuracy, purpose*), etc. Hay que señalar que las bibliotecas en el campo de selección y evaluación fueron pioneras en la elaboración de los criterios para evaluar los contenidos de la Web, destacando principalmente los parámetros de autoridad, cobertura, actualización y precisión.

A continuación, destacamos algunas propuestas sobre evaluación de recursos web que llevaron a cabo algunas bibliotecas para apoyar las necesidades de información, tanto de docentes como de estudiantes.

A nivel de bibliotecas internacionales, la *Berkeley Library*⁸ propone criterios de autoridad, propósito, publicación y formato, relevancia, fecha de publicación y documentación, mientras que la *University Libraries Virginia Tech*⁹, propone tener en cuenta la autoría del documento (si es un experto en el tema), la cobertura (si el tema se trata de forma amplia y desarrollada), la objetividad con que se trata el tema, la exactitud (si es un artículo académico o divulgativo), la actualidad de la información, el tipo de audiencia a quien va dirigido.

A nivel nacional, destacamos los criterios de evaluación que propone la Biblioteca Complutense y que podemos ver en la siguiente infografía¹⁰ (Figura 1): relevancia, autoridad, actualidad, exactitud y fiabilidad.

Figura 1.

Criterios para evaluar la información



Fuente: Biblioteca Universidad Complutense. <https://biblioteca.ucm.es/file/criterios-de-evaluacio%CC%81n/?ver>

Ante la variada oferta de datos en fuentes y recursos que circulan por las redes es importante que los usuarios sepamos filtrar aplicando algunos criterios que proporcionen veracidad, credibilidad, fiabilidad y calidad de los datos; seamos críticos con los datos que encontramos, en relación a su cobertura y actualización; y sepamos buscar evidencias para verificar si los datos tienen credibilidad y chequear con otras fuentes.

A continuación, proponemos los siguientes criterios de calidad de forma general:

Autoridad. Se refiere al responsable del sitio, ya sea una persona, un grupo, una institución pública o privada, educativa, cultural, etc. Se evalúa la identificación y solvencia del autor responsable, se valora su conocimiento, experiencia, credenciales, contacto, dominio, reputación, credibilidad.

7 CRAAP Test. <https://library.csuchico.edu/sites/default/files/craap-test.pdf>

8 Evaluating resources. <https://guides.lib.berkeley.edu/evaluating-resources>

9 Evaluating Information. <https://guides.lib.vt.edu/c.php?g=212535&p=1402527>

10 Biblioteca Complutense. <https://biblioguias.ucm.es/como-buscar/como-evaluar>

Fiabilidad de la información/datos. Es importante saber seleccionar la procedencia de las fuentes de datos que nos interesan en función de nuestras necesidades de información.

Consideramos fuentes solventes:

- Instituciones educativas y centros de investigación.
- Organismos públicos oficiales, nacionales e internacionales.
- Instituciones culturales, académicas y científicas.
- Editoriales científicas y académicas de prestigio.
- Entidades y asociaciones científicas, profesionales o empresariales relevantes.

Son fuentes poco solventes las que proporcionan:

- Datos anónimos cuya procedencia no se puede comprobar.
- Datos que no contienen bibliografía para contrastar.
- Datos que no tienen fecha.
- Sitios de alojamiento libre de contenidos pero que no tienen control selectivo.
- Blogs, páginas personales, espacios de entretenimiento.

Acceso a la información: navegación y recuperación. La navegación se refiere a la facilidad con que se puede desplazar para buscar información. Si la web del sitio es clara sencilla, comprensible. Para que la información pueda ser recuperada, la navegación debe ser ordenada. La web debe tener un mapa web de los contenidos que facilite su recuperación, con posibilidad de sistemas de búsqueda de información. También debe incluir índices temáticos de contenido. Diseño funcional y compatible con diferentes versiones del navegador.

Publicación y formato. Dónde fue publicado, difundido. Si está publicado en una publicación académica, quién fue el editor, si fue revisado por pares, en qué medio, si es una publicación impresa o en línea, si es un blog, o un vídeo.

Tipos de licencias. No todo lo que está en Internet se puede utilizar, por tanto, debe estar sometido a una serie de licencias de uso.

4. RECOMENDACIONES Y DIRECTRICES PARA EVALUAR LA CALIDAD DE LOS DATOS ABIERTOS

La calidad de los datos es un tema de actualidad y en creciente demanda, especialmente los datos que están disponibles públicamente y que pueden reutilizarse. Un factor importante que afecta a la calidad es el origen de la diversidad de fuentes de donde proceden los datos. Es importante que los datos estén bien definidos, sin faltas, que sean exactos, que no contengan discrepancias y que dispongan de la documentación pertinente sobre la estructura y disponibilidad del formato (HTML, GIF o PDF). Los editores de datos deben familiarizarse con los principios FAIR antes de publicarlos. Este marco proporciona calidad de los datos a través de cuatro dimensiones: localización, accesibilidad, interoperabilidad y reutilización (FAIR, 2020). Estas dimensiones es lo que genera confianza y validez en el conjunto de datos. Y, para que los datos sean seguros, deben estar en repositorios confiables que cumplan con los requerimientos FAIR y los principios TRUST, de transparencia, responsabilidad, foco en el usuario, sostenibilidad y tecnología, componentes esenciales para definir repositorios de datos confiables (De Giusti, 2021). Los repositorios deben ganarse la confianza de las comunidades a las que pretenden servir y demostrar que son confiables y capaces de gestionar adecuadamente los datos que poseen. Los repositorios deben proporcionar evidencia transparente, honesta y verificable de su práctica y garantizar la integridad, autenticidad, precisión, confiabilidad y accesibilidad a los datos durante mucho tiempo (Lin et al., 2020).

No existen muchos estudios que analicen las pautas de cómo evaluar la calidad de los datos abiertos, quizás porque es un campo de estudio demasiado reciente y todavía faltan estudios rigurosos. Por ello, para abordar este tema, tomaremos en consideración las recomendaciones europeas que propone Data.europa.eu y Datos.gob.es:

- *Data.europa.eu. Data Quality Guidelines* (2021). <https://dataeuropa.gitlab.io/data-provider-manual/data-quality/>
- https://op.europa.eu/webpub/op/data-quality-guidelines/en/#_idsec112 (Trad. castellano)
- Datos.gob.es. (2017). *Manual práctico para mejorar la calidad de los datos abiertos*. <https://datos.gob.es/es/documentacion/manual-practico-paramejorar-la-calidad-de-los-datos-abiertos>
- *Guía para la evaluación de repositorios institucionales de investigación*. (2021). FECYT. https://calidadrevistas.fecyt.es/sites/default/files/informes/2021guiaevaluacionrecolecta_vf.pdf

Las *Directrices de calidad Data.europa* están basadas en las dimensiones de los principios FAIR que propone unas recomendaciones generales y otras específicas.

4.1. Recomendaciones generales

(a) Localización de los datos. Describir los datos con metadatos para mejorar el descubrimiento de datos. La calidad de los datos depende de su disponibilidad. Los datos deben estar disponibles públicamente para que puedan reutilizarse con diferentes propósitos. Los datos deben ser válidos y limpios. Estar bien definidos y ser exactos; eliminar duplicados y contener la documentación sobre la estructura o disponibilidad del formato (HTML, GIF, PDF).

(b) Accesibilidad. Publicar datos sin restricciones, proporcionar una URL de descarga accesible.

(c) Interoperabilidad:

- Formato de fecha y hora.
- Formato de números decimales y números de miles.
- Utilizar codificación de caracteres estandarizado.
- Utilizar identificadores uniformes (URI) para identificar entidades.

(d) Reutilización:

- Proporcionar una cantidad adecuada de datos.
- Considerar los estándares comunitarios.
- Eliminar duplicados.
- Aumentar la precisión de los datos.
- Proporcionar información sobre el tamaño de bytes.

4.2 Recomendaciones específicas

(a) Formatos:

- Reutilizar conceptos inequívocos de vocabularios controlados (Encontrabilidad).
- Armonizar etiquetas, usando identificadores únicos (Accesibilidad).
- Desreferenciar la traducción de una etiqueta (Reusabilidad).
- Vincular y aumentar los datos (Interoperabilidad).

(b) Documentación de los datos:

- Publica tu documentación (Encontrabilidad).
- Usar esquemas para especificar la estructura de datos (Accesibilidad).
- Documentar la semántica de los datos (Reusabilidad).
- Cambios en los datos del documento (Interoperabilidad).
- Desaprobar versiones antiguas (Encontrabilidad).
- Vincular versiones de un conjunto de datos (Accesibilidad).

(c) Recomendaciones para mejorar el nivel de apertura:

- Utilice datos estructurados (Encontrabilidad).
- Utilizar un formato no propietario (Accesibilidad).
- Utilice URI para indicar cosas (Reutilizabilidad).
- Utilizar datos vinculados (Interoperabilidad).

(d) Y, por último, una lista de verificación para publicar datos de alta calidad. Consta de seis fases:

- Definir el plan de gestión de los datos: características, estructura, cómo compartir datos.
- Enriquecer los datos.
- Documentarlos.
- Validarlos.
- Difundirlos / Publicación: preservación de los datos y publicación en diferentes formatos.

El informe, *Manual práctico para mejorar la calidad de los datos abiertos de Datos.gob*, redactado por un equipo de expertos, plantea también una serie de recomendaciones prácticas sobre la calidad de los datos abiertos, proponiendo un listado de indicadores:

- Exactitud de los datos.
- Consistencia.
- Disponibilidad.
- Completitud.
- Conformidad.
- Credibilidad.
- Precisión.
- Actualidad.
- Comprensibilidad.

5. ACCESO A LOS DATOS ABIERTOS Y ACCESO A LOS DATOS DE INVESTIGACIÓN

La apertura de los datos es fundamental para el avance de la ciencia y es imprescindible para la reproducibilidad de los resultados científicos, facilita la cooperación interdisciplinar, estimula el crecimiento económico a través de oportunidades para la innovación, permite la reutilización de datos, aumenta la eficiencia de los recursos, mejora la transparencia, la rendición de cuentas y la confianza en los resultados de la in-

investigación científica tal y como recomienda el consejo de la OCDE, sobre el acceso a datos de investigación procedentes de financiación pública (OCDE, 2021).

Los datos abiertos responden a una iniciativa internacional de poner a disposición de la sociedad parte de los datos creados por las administraciones públicas que pueden ser de interés para el dominio público. Y el sector público es uno de los principales productores y poseedores de información. Estos datos pueden ser accesibles desde portales de datos abiertos, repositorios de datos institucionales y plataformas. Son herramientas que almacenan y permiten la consulta de los datos, a la vez que se ocupan de su preservación a largo plazo, ya que pueden presentarse en múltiples formatos y soportes.

Por otro lado, los repositorios de datos almacenan los datos que provienen de investigaciones, instituciones o administraciones y pueden ser utilizados libremente como base para otras investigaciones, artículos, etc. Sirven al propósito de transparencia y ayudan al conocimiento colaborativo. Algunos de estos datos pueden ser datos de gestión, datos académicos de universidades, datos de investigación. La publicación en abierto tiene una doble finalidad: por un lado, la validación de la investigación y ser reutilizada por otros proyectos, aunando esfuerzos.

Mientras que los datos de investigación son todos aquellos materiales que han sido generados o recolectados durante el proceso de una investigación. Dependiendo de la disciplina científica, los datos se presentan en diferentes formatos y en diferentes soportes y estos pueden variar en función de la disciplina o área de humanidades, ciencias sociales o ciencias de la salud. La clasificación más común diferencia entre datos cuantitativos y datos cualitativos. Pero, otros autores hacen alusión a datos numéricos, descriptivos o visuales, en estado bruto o analizado, experimentales u observacionales, etc.

A continuación, mostramos una selección de portales de datos abiertos y repositorios de datos internacionales y nacionales que garantizan la recuperación y el acceso a los datos públicos y a los datos de investigación.

5.1 Buscadores, Directorios, recolectores de datos

- **COAR** (*Confederation of Open Access Repositories*). <https://www.coar-repositories.org/>
Confederación de Repositorios en Acceso abierto que representa a más de 150 instituciones de todo el mundo. COAR trata de mejorar la visibilidad y posicionar los resultados de investigación a través de la red de repositorios global.
- **Google Dataset Search**. <https://datasetsearch.research.google.com/>
Buscador de conjunto de datos. Localiza conjuntos de datos para ayudar a investigadores y periodistas a descubrir datos disponibles en línea que pueden ser de ayuda para los proyectos.
- **DataCite**. <https://www.datacite.org/>
Organización global que proporciona identificadores persistentes (DOI) para datos de investigación. Ayuda a la comunidad de investigación a localizar, identificar y citar datos de investigación, participando activamente en las comunidades de investigación y realizando labores de divulgación. Tiene como prioridad hacer que los datos de investigación sean visibles y accesibles.
- **OpenAire**. <https://explore.openaire.eu/search/find>
Plataforma europea de información que promueve la comunicación académica abierta y la mejora de la accesibilidad, reutilización, reproducibilidad y monitoreo de resultados de investigación, incluyendo los datos, de cualquier disciplina científica y área temática, a través de Europa y en el exterior.
- **OpenDoar**. <https://www.jisc.ac.uk/opendoar>
Directorio mundial de repositorios científicos de acceso abierto.
- **Re3data.org**. <https://www.re3data.org/>
Registro internacional de repositorios de datos de investigación. Recopila los metadatos de los repositorios especializados que contienen datos procedentes de investigación.
- **Recolecta**. <https://www.recolecta.fecyt.es/>
Recolector de ciencia abierta o agregador nacional de repositorios institucionales y temáticos de acceso abierto.

5.2 Portales de datos abiertos

Un portal de datos abiertos facilita la búsqueda a los usuarios, garantiza la calidad de los datos publicados y su disponibilidad. Estos portales pueden ser internacionales, nacionales, autonómicos, locales, universitarios, culturales, etc.

- **European data.** <https://data.europa.eu/en>
El portal de datos abiertos de la UE es el punto de acceso a los datos publicados por las instituciones, agencias, y otros organismos de la Unión Europea.
- **European Open Science Cloud.** <https://open-science-cloud.ec.europa.eu/>
Plataforma que apoya la investigación multidisciplinar y promueve el uso de los datos FAIR en un entorno multidisciplinar abierto en el que pueden participar el personal investigador, sectores innovadores, empresas y ciudadanos.
- **DataUsa.** <https://datausa.io/>
Portal de acceso a datos públicos del gobierno de EE.UU.
- **Open data CERN.** <https://opendata.cern.ch/>
Portal de datos abiertos de la investigación que se hace a través del CERN.
- **Datos.gob.es.** <https://datos.gob.es/es/catalogo>
Plataforma que organiza y gestiona el Catálogo Nacional de Datos Abiertos de la información del sector público. Portal de datos abiertos gubernamental que permite recuperar datos del sector público. Carácter multidisciplinar.



- **Datos.bne.es.**
Portal de datos bibliográficos de la Biblioteca Nacional de España, publicados como *Linked Open Data*.
- **Catálogo de datos abiertos de la Comunidad de Madrid.** <https://www.comunidad.madrid/gobierno/datos-abiertos>
El catálogo de datos permite el acceso centralizado a los conjuntos de datos, o datasets en formatos abiertos para su reutilización por parte de la ciudadanía. El catálogo dispone de un buscador libre y un buscador facetado.
- **Catálogo de datos abiertos de Castilla y León.** <https://datosabiertos.jcyl.es/web/es/datos-abiertos-castilla-leon.html>
El objetivo del portal de datos abiertos es aumentar la transparencia y conseguir la participación de los ciudadanos y empresas en el intercambio de conocimientos. A través del catálogo se accede al listado de todos los conjuntos de datos que se encuentran en el portal.
- **CORA (Catalan Open Research Area).** <https://www.csuc.cat/es/servicios/cora-catalan-open-research-area>
Portal web que centraliza y promueve la ciencia y el acceso abierto en el ámbito académico y de investigación de Cataluña.
- **InvestigaM.** <https://www.consorcioamadrono.es/investigam/>
Portal de ciencia abierta del Consorcio Madroño. Su principal objetivo es la difusión de los resulta-

dos de la actividad investigadora que se lleva a cabo en sus instituciones miembro y asociadas, lo que incluye tanto publicaciones como datos de investigación.

5.3 Repositorios de datos multidisciplinares

- **E_cienciaDatos.** Consorcio Madroño. <https://edatos.consorciomadrono.es/>
Repositorio de datos de carácter multidisciplinar que aloja conjuntos de datos científicos de las investigaciones de las universidades públicas de la Comunidad de Madrid (Universidad de Alcalá, Universidad Autónoma, Universidad Carlos III, UNED, Universidad Politécnica de Madrid, Universidad Rey Juan Carlos y Fundación ara el conocimiento madri+d), que forman el Consorcio Madroño.
- **FigShare.**
Repositorio de ciencia abierta en línea donde los investigadores pueden preservar y compartir los datos y resultados de un proceso de una investigación (figuras, imágenes, vídeos...). Proporciona datos estadísticos.
- **Mendeley Data.** <https://data.mendeley.com/>
Plataforma gratuita que permite a los investigadores cargar datos en bruto de su investigación. Les proporciona un identificador DOI para vincular los artículos de las webs de las revistas con el conjunto de datos de investigación.
- **RDR.** Repositorio de datos de investigación (RDR) de las universidades catalanas, centros de investigación CERCA y otras entidades. Repositorio de datos federado y multidisciplinar para la publicación de conjuntos de datos de investigación con principios FAIR y siguiendo las directrices de la EOSC.
- **Zenodo.** <https://zenodo.org/>
Repositorio de datos de investigación en acceso abierto desarrollado bajo el programa europeo OpenAire y operado por CERN. Recoge todo tipo de formatos y documentos publicados. Permite a los investigadores depositar artículos de investigación, conjuntos de datos, *software* de investigación, informes y cualquier otro artefacto digital relacionado con la investigación.
- **Digital Curation Center (DCC).** <https://www.dcc.ac.uk/search>
Centro de curación de datos de investigación que apoya a las instituciones de educación superior del Reino Unido.

5.4 Repositorios institucionales

Responden al compromiso de una institución de hacer visible, preservar y difundir su producción científica y académica. En este grupo se incluyen los repositorios de datos de Universidades.

La implementación de repositorios de conjuntos de datos de investigación en las universidades públicas está en fase de desarrollo. Del total de las 49 universidades públicas (47 presenciales y 2 a distancia) destacan sólo 10 universidades con el porcentaje mayor de conjuntos de datos publicados en sus repositorios institucionales, según el estudio realizado por Martínez Méndez et al. (2023) (Tabla 1).

Tabla 1.

Universidades con porcentaje mayor de conjuntos de datos publicados en sus repositorios institucionales.

Universidad	#datasets
Alcalá de Henares	375
Carlos III	259
Zaragoza	222
Granada	112
Barcelona	105
Politécnica de Catalunya	75
Lleida	59
Pompeu Fabra	44
Politécnica de Valencia	43
Politécnica de Madrid	41

- **Digital CSIC.** <https://digital.csic.es>
Repositorio institucional del Consejo Superior de Investigaciones Científicas que **organiza**, preserva y difunde en acceso abierto los resultados de investigación del CSIC. El mandato institucional de publicar en acceso abierto rige a partir del 1 de abril de 2019 y dispone que las referencias bibliográficas de todas las publicaciones revisadas por pares (artículos, capítulos de libros, libros y comunicaciones de congresos), los textos y los datasets, realizados por la comunidad investigadora estén disponibles en acceso abierto en el repositorio.

Repositorios institucionales de datos en universidades

- **Docta Complutense. Universidad Complutense de Madrid.** <https://docta.ucm.es/home>
Contiene datos de investigación científica, documentos de trabajo, materiales docentes, trabajos académicos.
- **e-Archivo. Repositorio institucional de la Universidad Carlos III de Madrid.**
<https://e-archivo.uc3m.es/home>
- **Repositorio de datos UC3M. E-ciencia datos. Consorcio Madroño.**
<https://edatos.consorciomadrono.es/dataverse/UC3M>
- **Repositorio de datos UAH. E-ciencia datos. Consorcio Madroño.**
<https://edatos.consorciomadrono.es/dataverse/UAH>
- **Repositorio de datos UAM. E-ciencia datos. Consorcio Madroño.**
<https://edatos.consorciomadrono.es/dataverse/UAM>
- **Repositorio de datos UNED. E-ciencia datos. Consorcio Madroño.**
<https://edatos.consorciomadrono.es/dataverse/UNED>
- **Repositorio de datos UPM. E-ciencia datos. Consorcio Madroño.**
<https://edatos.consorciomadrono.es/dataverse/UPM>
- **Repositorio de datos URJC. E-ciencia datos. Consorcio Madroño.**
<https://edatos.consorciomadrono.es/dataverse/URJC>
- **Open data UGR.** <https://opendata.ugr.es/dataset/>
Portal de datos abiertos de la Universidad de Granada sobre su actividad académica, investigadora, servicios, etc.) en formato accesible, reutilizable y compartible que pone a disposición del público en general y miembros de la universidad.
- **Repositorio de Gredos. Universidad de Salamanca.** <https://gredos.usal.es/>
Contiene datos académicos, colecciones patrimoniales, producción científica, materiales docentes.
- **Repositorio institucional de documentos. Zagan. Universidad de Zaragoza.**
<https://zagan.unizar.es/collection/opendata?ln=es>
Contiene datos académicos, datos de recursos humanos y datos abiertos de investigación medioambiental.
- **Repositorio institucional Digibug. Universidad de Granada.** <https://digibug.ugr.es/>
Contiene documentos digitales de carácter científico, docente e institucional.

Portales de información estadística

- **Instituto Nacional de Estadística (INE).** <https://www.ine.es/>
- **Eurostat. Oficina Estadística de la Unión Europea.** <https://ec.europa.eu/eurostat/data>
Publica estadísticas e indicadores de alta calidad a nivel europeo, elabora informes y planes estratégicos.

Referencias

- Almada Martínez, S. A. (2019). *Utilización de los métodos de validación y confiabilidad de los instrumentos de recolección de datos en lo trabajos de tesis de postgrado*. [Tesis doctoral]. Universidad Tecnológica Intercontinental.
- Alonso Arévalo, J. (23 abril 2024). Confiabilidad de la información en un mundo de IA. En *I Seminario Internacional de Bibliotecas en Ciencias de la Salud*. AMBAC. México.
- Data.europa.eu data quality guidelines. (2021). Oficina de Publicaciones de la Unión Europea. <https://op.europa.eu/webpub/op/data-quality-guidelines/en/> o <https://data.europa.eu/doi/10.2830/333095>
- Datos.gob.es. (2017). *Manual práctico para mejorar la calidad de los datos abiertos*. <https://datos.gob.es/es/documentacion/manual-practico-paramejorar-la-calidad-de-los-datos-abiertos>
- De Giusti, (2021). Calidad en los repositorios digitales. Los principios TRUST para repositorio de datos. *Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología*, 29, 55-59. [https://doi.org/10.24215/18509959.29.e6\(2021\)](https://doi.org/10.24215/18509959.29.e6(2021)).
- FAIR Data Maturity Model Working Group (2020). *FAIR Data Maturity Model. Specification and Guidelines (1.0)*. <https://doi.org/10.15497/rda00050>
- FECYT (2021). *Guía para la evaluación de repositorios institucionales de Investigación*. <https://www.fecyt.es/es/publicacion/guia-para-la-evaluacion-de-repositorios-institucionales-de-investigacion>
- INTEF (enero 2022). *Marco de referencia de la Competencia Digital Docente*. Recuperado el 15 de mayo de 2024, de https://intef.es/wp-content/uploads/2022/03/MRCDD_V06B_GTTA.pdf
- Lin, Dawei, Crabtree, J., Dillo, I., Downs, R., Edmunds, R., Giaretta, D., DeGiusti, M., L'Hours, H., Hugo, W., Jenkyns, R., Khodi-yar, V., Martone, M.E., Mokrane, M., Navale, V., Petters, J., Sierman, B., Sokolova, D.V., Stochhause, M., y Westbrook, J. (2020). The TRUST Principles for digital repositories. // *Scientific Data* 7:18, 1-5. Recuperado el 15 de mayo de 2024, de <http://digital.cic.gba.gob.ar/handle/11746/10566>.
- DigComp 2.2 Marco de Competencias Digitales de la Ciudadanía (2022). https://somos-digital.org/wp-content/uploads/2022/04/digcomp2.2_castellano.pdf
- Martín González, Y., Gallego Lorenzo, J., e Iglesias Rodríguez, A. (2023). Alfabetización en datos para los docentes universitarios: necesidades formativas, escenarios didácticos y referencial de competencias. En Cuevas Cerveró, A., Prieto Gutiérrez, J.J. y Simeño E., (Coords.). *Discurso de odio, desinformación e inclusión digital* (pp. 155-161). Trea.
- Martínez Méndez, F.; Baptista, A. A.; López Carreño, R.; Delgado Vázquez, A. M. (2023). Implementación de los repositorios de datos de investigación en las universidades públicas españolas: estado de la cuestión. *Scire*, 29, 39-49.
- OECD/LEGAL/0347 (2021). *Recommendation of the Council concerning Access to Research Data from Public Funding*. <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0347>
- Raffaghelli, J. (2020). "Datificación" y Educación Superior: Hacia la construcción de un marco para la alfabetización en datos del profesorado universitario. *Revista Interamericana de investigación, Educación y Pedagogía*, 13(1). <https://doi.org/10.15332/25005421>
- Royo-Montañés, S., y Benítez-Gómez, A. (2019). Portales de datos abiertos. Metodología de análisis y aplicación a municipios españoles. *El Profesional de la Información*, 28(6), 1-13. <https://doi.org/10.3145/epi.2019.nov.09>
- UNESCO (2024). *Plataforma electrónica internacional de alfabetización mediática e informacional*. Módulo 3: Investigación, ciclo de la información, procesamiento digital de la información y propiedad intelectual. <https://www.unesco.org/mil4teachers/es/module3>
- Wilkinson, M. D.; Dumontier, M.; Aalbersberg, I.; Brandt-Jan, A.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J. W.; Da-Silva-Santos, L. B.; Bourne, P. E.; Bouwman, J.; Brookes, A. J.; Clark, T.; Crosas, M.; Dillo, I.; Dumon, O.; Edmunds, S.; Evelo, Chris T.; Finkers, R.; ... Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, 3, 160018. <https://doi.org/10.1038/sdata.2016.18>

LA ALFABETIZACIÓN EN DATOS EN EL CONTEXTO
UNIVERSITARIO DE LAS CIENCIAS SOCIALES
(ALFADATAUniv_CCSS)

CONFIABILIDAD EN LA EVALUACIÓN Y SELECCIÓN DE LOS DATOS ABIERTOS

Actividad

Josefa Gallego Lorenzo



BÚSQUEDA Y SELECCIÓN DE DATOS CONFIABLES

[Actividad propuesta]

Tiempo aproximado de realización de la actividad:
2 horas.

1. Objetivos

- Valorar la calidad, procedencia y fiabilidad de los datos.
- Aprender a ser crítico y contrastar la información.
- Conocer los datos de investigación y conjuntos de datos.

2. Enunciado explicativo de la actividad

La sociedad está cada vez más datificada y cada vez resulta más difícil encontrar datos de calidad. Esta actividad permite adquirir habilidades en la búsqueda y selección de los datos abiertos y conocer los diferentes tipos de repositorios sostenibles y confiables generalistas, temáticos, institucionales para encontrar resultados confiables.

3. Desarrollo e instrucciones para realizar la actividad

Una vez leído el módulo 2, correspondiente al tema *Confianza en la evaluación y selección de los datos abiertos*, podrás realizar las diferentes actividades.

4. Actividad práctica

Lee atentamente los diferentes enunciados y responder las siguientes preguntas.

Debido a la gran cantidad de noticias que circulan por internet, marca los criterios de calidad que consideres que se deben tener en cuenta a la hora de realizar una búsqueda.

- Autoría.
- Actualización.
- Contenido.
- Objetivo.
- Pertinencia.
- Mucho ruido.
- Bien presentado.

¿Cuáles de los siguientes recursos consideras que proporcionan datos fiables sobre la LOSU? Justifica la respuesta.

- a. Una noticia en la Radio.
- b. Una página web recomendada por el profesor.
- c. Una página web que comience por <https://datos.gob.es>.
- d. Una página web de la universidad.
- e. Una noticia en Instagram.
- f. Documento publicado en el BOE.

¿Qué recursos de datos abiertos consideras más seguros para localizar conjuntos de datos sobre la Desinformación y las Fake news? Formula la estrategia de búsqueda más adecuada utilizando los operadores booleanos y justifica la respuesta.

- Un buscador genérico de datos abiertos.
- Un portal de datos abiertos.
- Un repositorio institucional.
- Un post de una red social.

Cómo documentar un proceso de investigación (TFG, Tesis doctoral) a través de los datos FAIR (Encontrables, Accesibles, Interoperables y Reutilizables). Reflexionar a partir del artículo.

Wilkinson, MD, Dumontier, M., Jan Aalbersberg, I. *et al.* (2019). Anexo: Principios rectores de FAIR para la gestión y administración de datos científicos. *Datos de ciencia* 6. <https://doi.org/10.1038/s41597-019-0009-6>

¿Qué tipos de conjuntos de datos encontramos en el catálogo de datos sobre “Universidades”: <https://datos.gob.es/es/catalogo> ?

Marca con una cruz los repositorios institucionales que consideras que tienen datos de investigación.

- Universidad de Zaragoza. <https://zaguan.unizar.es/collection/opendata?ln=es>
- Universidad de Salamanca. Gredos. <https://gredos.usal.es/>
- Universidad Complutense. <https://docta.ucm.es/home>
- Alcalá de Henares. <https://edatos.consorciomadrono.es/dataverse/UAH>
- Universidad de Granada. <https://digibug.ugr.es/>

LA ALFABETIZACIÓN EN DATOS EN EL CONTEXTO
UNIVERSITARIO DE LAS CIENCIAS SOCIALES
(ALFADATAUniv_CCSS)

VISUALIZACIÓN Y REPRESENTACIÓN GRÁFICA DE DATOS

José Carlos Toro Pascua



VISUALIZACIÓN Y REPRESENTACIÓN GRÁFICA DE DATOS

ÍNDICE:

1. ¿En qué consiste la representación gráfica o visualización de datos?
2. Importancia y beneficios de la visualización gráfica de datos
3. Necesidades y características en un proceso de visualización
4. Tipos de visualizaciones
5. Tipo de gráficos y utilidad
6. Tipo de *software* o plataformas de servicios generadores de gráficos
7. Fases de creación de una visualización de datos
8. *Software* o plataformas de servicios generadores de gráficos
9. Errores a evitar
10. Bases de datos de imágenes y utilidades gráficas

Referencias

1. ¿EN QUÉ CONSISTE LA REPRESENTACIÓN GRÁFICA O VISUALIZACIÓN DE DATOS?

La visualización y representación gráfica de los datos es uno de los aspectos más importantes y esenciales del Big Data. Actualmente recibimos una gran cantidad de información por múltiples canales, datos que en numerosas ocasiones nos llegan en forma de cientos de impactos visuales, que no podemos procesar, pero la visualización nos ayuda a seleccionarlos (Alcalde Perea, 2015). Es tal el volumen de información recibido, que, en 1996, el consultor Alfons Cornellá acuñó el término de infoxación para referirnos a una intoxicación de información, a un escenario de recepción de cientos de informaciones diarias en el que vamos saltando de una a otra, sin ser capaces de dedicarles el tiempo adecuado. Además, muchas de ellas, acaban siendo olvidadas (Cornellá, 2013).

Los estudios han demostrado que resulta más fácil captar y recordar información a través de imágenes que a través de texto. Esta característica neurológica es fundamental cuando tratamos de comunicar grandes cantidades de datos.

El cerebro humano procesa 60.000 veces más rápido la información visual que la textual. Según un estudio de Potter et al. (2014), se necesitan unos 6 segundos para leer 20 palabras, pero sólo un cuarto de segundo para procesar y dar significado a un símbolo o a una información visual. Además, es capaz de recordar el 80% de las imágenes, el 20% del texto y el 10% del sonido. En el mismo sentido, en un experimento neurológico con estudiantes recogido en el libro *Las reglas del cerebro* de Jhon Medina (2008), se describe cómo únicamente un 10% de la muestra recordaba la información oral que se les comunicaba previamente, y que ese porcentaje ascendía al 65% cuando a esa misma información se le añadía una imagen.

Los datos se componen, a simple vista, de hojas de cálculo interminables o tablas llenas de cifras y textos con filas y columnas, sin interrelaciones ni posibilidad de lecturas útiles para proceder a distintos análisis. De esta manera, serán difícilmente legibles y accesibles, además de tener una enorme complejidad de visualización y comprensión.

La visualización de datos es una técnica que nos facilita la presentación visual de los datos que, en ocasiones, forman un gran volumen de informaciones, muy complejas y difíciles de recordar. Mediante su repre-

sentación visual facilitamos el procesamiento de estas grandes cantidades de datos (*¿Qué es la visualización de datos?*, 2023).

Representando los datos mediante gráficos, infografías o mapas, las visualizaciones ofrecen y aportan una manera mucho más sencilla de identificar aspectos, patrones, relaciones, tendencias, características o comportamientos atípicos que no son evidentes o perceptibles a simple vista. Se convierte en un proceso fundamental para la comprensión, la interpretación y la reutilización de los datos (*Guía de visualización de datos*, 2023).

Un proyecto de visualización tiene que cumplir una serie de requisitos (Esteban Gracia, 2017):

- Que los datos que utilicemos no sean comprensibles a primera vista, que no podamos interpretarlos hasta que no los visualicemos.
- Que los datos sean susceptibles de convertirse en imágenes, en una visualización gráfica.
- Que el resultado del proyecto sea una visualización que se pueda entender e interpretar.
- Un proyecto de visualización traduce datos en imágenes.

Actualmente contamos con una gran variedad de posibilidades para visualizar conjuntos de datos debido a la evolución y multiplicación del *software*, de lenguajes y entornos de programación y de la disponibilidad de numerosas plataformas digitales, tanto en línea como de instalación local, para llevar a cabo visualizaciones muy complejas y variadas. Además de las mejoras producidas en el *hardware* y las tecnologías de la comunicación.

2. IMPORTANCIA Y BENEFICIOS DE LA VISUALIZACIÓN GRÁFICA DE DATOS

De manera general, la visualización de datos nos permite convertir grandes volúmenes de datos en información útil y comprensible. Asimismo:

- Permite que se pueda comunicar de forma efectiva grandes volúmenes de información, en ocasiones con un alto grado de complejidad, de forma rápida y sencilla, simplificando el mensaje.
- Evitamos el exceso de información, ya que podemos seleccionar de manera más eficaz y con mayor rapidez la información visual que la textual.
- La información que queremos comunicar se puede adaptar a todo tipo de consumidores (lectores, oyentes...), incluso a individuos de diferentes lenguas.
- Se facilita el análisis de manera que se puede realizar en un menor tiempo y de manera más sencilla la obtención de conclusiones y los valores de los datos.
- Se detectan fácilmente tendencias, valores atípicos, relaciones entre variables o errores con mucha mayor eficacia y rapidez.

3. NECESIDADES Y CARACTERÍSTICAS EN UN PROCESO DE VISUALIZACIÓN

Para llevar a cabo un proyecto de visualización de datos tenemos que planificar y cumplir una serie de aspectos:

- Recuperar un conjunto de datos confiables sobre la materia que vamos a trabajar.
- Definir claramente qué información queremos comunicar.
- Seleccionar el tipo de visualización más correcta o conveniente para los datos que queremos mostrar.
- Adaptar la representación gráfica al tipo de público-usuarios-consumidores, con preguntas como: ¿quiénes son?, ¿en qué están especialmente interesados?, o ¿qué queremos comunicar?

- Transmitir mensajes adecuados. Simplificando y seleccionando lo que realmente queremos transmitir.
- Potenciar la visualización de lo realmente importante y no lo interesante o anecdótico.
- Realizar un producto visualmente atractivo, claro y con buena legibilidad, mediante la correcta utilización de los colores, las tramas, las fuentes o etiquetas.
- Incluir una cantidad de datos proporcionada, priorizando la simplicidad y evitando la sobrecarga informativa.
- Adecuar la representación de la información cuantitativa, mediante ponderaciones o datos porcentuales.
- Complementar con explicaciones claras de los procesos, los resultados y las conclusiones o el valor de los datos.
- Citar las fuentes y toda aquella información referente a las licencias de uso del material que utilicemos y del propio material resultante. En ocasiones, algunas aplicaciones de generación de visualizaciones exigen unas condiciones de uso determinadas, desde no permitir cualquier utilización hasta firmar la visualización resultante de una manera concreta.

4. TIPOS DE VISUALIZACIONES

En el entorno digital, las visualizaciones pueden ser estáticas, dinámicas y/o interactivas.

- **Gráficos estadísticos.** Representación o visualización sencilla de datos mediante una serie de símbolos y codificaciones: barras, líneas, puntos, superficies, etc.
- **Infografías.** Las infografías combinan texto, imágenes y gráficos para comunicar información específica de manera atractiva y sencilla.
- **Sistemas de información geográfica.** Sistema gráfico que nos permite visualizar cualquier tipo de datos sobre un lugar geográfico determinado o que sean informaciones que contengan una variable geográfica asociada.
- **Dashboards o cuadros de mando o tableros de mando.** Se trata de una visualización compleja que presenta una gran cantidad de información de una manera concisa y fácil de entender mediante varios gráficos relacionados entre sí para monitorizar un aspecto determinado. Estos gráficos se pueden actualizar en tiempo real, además de contar con la posibilidad de filtrar determinadas variables.

5. TIPOS DE GRÁFICOS Y UTILIDAD

Existen diversos tipos de gráficos, con distintas configuraciones. Cada uno de ellos puede ser conveniente para presentar un determinado tipo de datos. Vamos a presentar los más habituales:

- **Gráficos de barras o columnas.** Suelen ser los más habituales. Mediante barras verticales u horizontales se expresan cantidades de una o varias variables. Son especialmente valiosos para analizar tendencias durante un rango determinado de tiempo o comparar variables. También podemos agrupar datos o apilarlos para series más complejas.
- **Gráficos de líneas.** Similares a los de barras, pero facilitan la representación de un conjunto mayor de variables, ya que las cantidades se reflejan únicamente con líneas que facilitan el estudio y evolución de un mayor número de datos y sus tendencias.

- **Gráficos de áreas.** Semejante a los gráficos de líneas, pero, en este caso, el espacio inferior a la línea se rellena con una trama. Podemos apilar cantidades o porcentajes sobre una cantidad total, al igual que en los gráficos de columnas.
- **Gráficos circulares o de sectores.** Sirven para mostrar proporciones sobre una totalidad. No son muy aconsejables si tenemos que mostrar una gran cantidad de categorías, porque pierde legibilidad. La totalidad puede estar representada por un círculo o un sector con diferentes formas geométricas como el rectángulo.
- **Gráficos de dispersión o de puntos.** Sirven para analizar distribuciones, relaciones, dispersión de datos y comportamientos entre dos variables, representándolos en un eje mediante puntos.
- **Gráfico de burbujas.** Igual que el anterior, pero, en este caso, se codifica el tamaño de los puntos para indicar mayor o menor cantidad.
- **Gráficos de combinación.** Pueden combinar un gráfico de barras, con líneas y puntos en una misma representación. Muy útiles para relacionar varios tipos de datos.

6. TIPO DE SOFTWARE O PLATAFORMAS DE SERVICIOS GENERADORES DE GRÁFICOS

Los usuarios de las aplicaciones para generar gráficos son muy diversos: desde grandes compañías y multinacionales con unas necesidades muy amplias, hasta usuarios individuales para los que puede ser suficiente alguna herramienta muy sencilla. Por tanto, el universo de *software* que nos podemos encontrar es muy amplio: desde complejas aplicaciones, con costes económicos muy elevados y necesidades de grandes conocimientos para su uso; a otras herramientas mucho más sencillas y gratuitas o de bajo coste y muy intuitivas.

Por consiguiente, en la selección que llevemos a cabo del *software* que utilicemos habrá que tener en cuenta las necesidades que tenemos para mostrar nuestros datos, el presupuesto, y el conocimiento de las herramientas.

En cuanto a los costes económicos, podemos distinguir herramientas comerciales y herramientas gratuitas, además de una serie de fórmulas intermedias o mixtas. Cada una de ellas tendrá una serie de características y, normalmente, el coste económico influye en la complejidad y multiplicación de opciones, soporte y posibilidades de la aplicación.

- **Software de pago.** El usuario adquiere una licencia o paga una suscripción. Normalmente posee funciones más avanzadas, soporte técnico, actualizaciones y mejoras continuas. También suelen ser más complejas en su utilización y se necesitan mayores conocimientos para manejarlas. Habitualmente incorporan posibilidades de trabajo en equipo y multitarea. En ocasiones, se ofrece una prueba gratuita durante un periodo de tiempo limitado. Al finalizar este periodo, el usuario debe adquirir una licencia para seguir utilizando la aplicación.
- **Software gratuito.** No es necesario abonar cantidad económica o pagar una licencia. Normalmente tienen funciones más básicas y sencillas, con menor soporte técnico o carece de él y poca madurez en las aplicaciones: sin actualizaciones e incluso con cancelaciones de este tipo de *software*. En ocasiones, pueden ser de código abierto para que los usuarios puedan hacer modificaciones, mejoras o personalizar los programas.
- **Software freemium.** Ofrece una serie de opciones gratuitas que son las más básicas y versiones de pago con funciones añadidas y más avanzadas. Normalmente, existen varias posibilidades de pago con diferente nivel de permisos en la utilización de las aplicaciones, en función de las necesidades de los usuarios.

7. FASES DE CREACIÓN DE UNA VISUALIZACIÓN DE DATOS

1. Seleccionar el tipo de visualización que mejor pueda representar nuestros datos y con el que sea más adecuado para lo que deseamos comunicar.
2. Elegir las variables que queremos mostrar en función del sistema de visualización seleccionado.
3. Preparar los datos. Limpieza de los archivos con datos, eliminando aquella información no necesaria, normalizando el contenido de los datos, comprobando si hay errores o valores atípicos.
4. Seleccionar el *software* teniendo en cuenta nuestras necesidades para mostrar los datos, además de las posibilidades económicas y de conocimiento.
5. Realización del gráfico con los datos previamente preparados, teniendo en cuenta las distintas codificaciones, el diseño en cuanto a colores, tramas, etiquetas, datos descriptivos, etc.
6. Análisis de los datos a través de la visualización llevada a cabo. Establecimiento de conclusiones con la obtención del valor determinado.

8. SOFTWARE O PLATAFORMAS DE SERVICIOS GENERADORES DE GRÁFICOS

Plataformas avanzadas y de pago:

- **Microsoft Office:** Excel y Power Point fundamentalmente. Suite de pago de la empresa Microsoft con una gran variedad de realización de gráficos con datos.
- **IBM SPSS Statistics.** *Software* de pago para el análisis estadístico avanzado que permite realizar una gran variedad de gráficos.
- **Microsoft Power BI.** Herramienta *freemium*, con pocas opciones gratuitas. Especializada en análisis de datos y en la creación de visualizaciones más complejas: gráficos, infogramas, *dashboard*, o información geográfica. Muy enfocado al big data.
- **Tableau.** Una de las principales aplicaciones para crear visualizaciones de grandes volúmenes de datos de innumerables tipos. Es una herramienta de pago con licencia de prueba gratuita durante un año para estudiantes.
- **Qlik Sense.** Una herramienta de pago de visualización de datos que le permite analizar datos fácilmente.
- **ArcGIS.** Principal aplicación comercial para el desarrollo de sistemas de información geográfica.
- **Otras aplicaciones comerciales para mapas:** MapBox o Carto.

Plataformas gratuitas y freemium:

- **Looker Studio** (anteriormente Google Data Studio). Plataforma de Google gratuita para la generación de visualizaciones complejas de datos, infográficos o *dashboard*.
- **Canva.** Plataforma *freemium* de creación de todo tipo de gráficos e infografías, con miles de prediseños gratuitos.
- **Infogram.** Es una potente herramienta en línea *freemium* que permite generar infografías, gráficos y mapas interactivos de manera sencilla.
- **Genially.** Plataforma virtual española con un enorme catálogo de presentaciones prediseñadas e interactivas. Muy sencilla y utilizada.
- **Piktochart.** Plataforma virtual para crear presentaciones, infografías, informes, etc. Muy sencilla y utilizada. Nos posibilita crear gráficos en mapas de forma gratuita.

- **Datawrapper.** Plataforma para realizar infografías en mapas muy sencilla y utilizada en medios.

Plataformas gratuitas de mapas:

- **gvSIGs** proyecto para desarrollar sistemas de información geográfica en lenguaje libre de la asociación gvSIG.
- **Proyecto QGIS** Es una aplicación profesional de SIG en *software* libre y abierto.
- **Google My Maps y Google Earth:** Plataformas gratuitas de Google para generar mapas.
- **Instamaps.** Plataforma española gratuita y abierta.
- **Open Street Map** Es una plataforma colaborativa para crear mapas libres y editables, con los que podemos realizar cualquier producto o mapa e incluir nuestros propios datos.
- **Lenguajes o entornos de programación:** Lenguajes gratuitos orientados al big data y a la visualización de datos. Son sistemas que necesitan un alto grado de conocimiento y manejo.
- **Python.** Lenguaje de programación libre de código abierto.
- **R:** Lenguaje de programación, y entorno para al análisis interactivo de datos.

9. ERRORES A EVITAR

Cuando realizamos un proyecto de visualización de datos, existen una serie de errores que se cometen y que hay que intentar evitar. Estos problemas de diseño nos permiten diferenciar los proyectos bien ejecutados o los que tienen problemas de visualización que, en ocasiones, se cometen a propósito para presentar la información de una manera determinada.

- **Tipo de visualización incorrecta.** El tipo de visualización que elegimos se tiene que adaptar a la información que queremos transmitir. Por ejemplo, cada gráfico está dirigido a un tipo de información determinada o a una forma de comunicación precisa. Si los datos contienen un elemento geográfico asociado, lo más adecuado puede ser mostrar la información sobre un mapa, o si los datos explican algún tipo de evolución temporal, la mejor visualización puede ser una línea de tiempo.
- **Cantidad de variables o datos.** Si utilizamos demasiadas variables o datos, la visualización puede llegar a ser ininteligible e incluso caótica. Debemos utilizar una cantidad de datos coherentes con la información que queremos transmitir.
- **Escalado y ponderación adecuados.** Es importante utilizar un escalado proporcional, sobre todo si vamos a mostrar diferencias entre cifras. También ponderar las variables que vamos a mostrar, cuando sea preciso.
- **Elecciones de colores, fuentes o tramas.** Existen una serie de recomendaciones básicas para mejorar la legibilidad, por ejemplo, la utilización de letras minúsculas que se leen mejor que las mayúsculas porque son más diferentes entre sí. Fuentes en formato redondo, mejor que cursivas o negritas. Evitar textos largos y con poco interlineado. En cuanto a los colores o las tramas, cuidar las incompatibilidades de tonos, o paletas de colores mal utilizadas, al igual que evitar utilizar demasiados colores. En gráficos o mapas, con codificación por intensidad del color, los colores más oscuros serán los de las variables más altas o para las mayores cantidades. El empleo de unos colores u otros, pueden ayudarnos a comunicar mejor el mensaje previsto o hacerlo más atractivo. Tampoco se recomienda huir de convencionalismos interiorizados por la sociedad, como que el rojo marque peligro o prohibición y el verde lo contrario; los tonos oscuros mayor cantidad de cosas que los claros, etc.

10. BASES DE DATOS DE IMÁGENES Y UTILIDADES GRÁFICAS

En internet podemos encontrar cualquier tipo de imágenes, videos, o material de cualquier formato en páginas web o portales de todo tipo para utilizar en nuestros proyectos de visualización de datos. Pero, debemos tener muy claro que no todo lo que podemos visualizar en internet se puede utilizar para cualquier fin. Asimismo, que puede haber material que sea válido para unas determinadas utilizaciones o productos, pero no para otras. No contamos con licencias de uso por estar viendo un determinado material.

Por tanto, antes de utilizar una fotografía, algún tipo de símbolos o vectores e incluso fuentes tipográficas, habrá que consultar las posibilidades de uso. En primer lugar, hay que asegurarse de que el material recuperado se puede utilizar; y, en segundo lugar, bajo qué condiciones: podemos encontrar imágenes válidas para cualquier tipo de uso u otras que únicamente nos permiten una utilización no comercial. Igualmente, debemos tener clara la forma de firmar las imágenes o símbolos o el material que recuperemos en internet, ya que, en muchas ocasiones, el autor que cede el uso del material exige que se le acredite la autoría de una manera determinada.

Para facilitarnos la tarea de recuperar material con posibilidades de reutilización, existen una serie de bases de datos o buscadores especializados en documentos con licencias *copyleft*. Son muy numerosas. Veamos algunos ejemplos:

- **Buscador de Creative Commons:** <https://search.creativecommons.org/>
Metabuscaador en el que podemos interrogar a diversas bases de datos según los distintos tipos de documentos que contienen.
Creative Commons es una organización sin ánimo de lucro que surgió para promover el intercambio y el acceso gratuito a la cultura. Puso en marcha sus licencias Creative Commons que permiten a un autor gestionar sus derechos de autor para ceder las licencias de uso a terceros bajo una serie de condiciones especificadas mediante atributos.
- **Wikimedia Commons:** <https://commons.wikimedia.org/>
Metabuscaador que da acceso a más de 100 millones de imágenes, vídeos y sonidos de la Fundación Wikimedia con licencias de uso gratuitas.
- **Pixabay:** <https://pixabay.com/es/>
Se trata de una fototeca con más de un millón de imágenes gratuitas que permiten utilizaciones de todo tipo. También contiene vídeos, efectos de sonido, música y gifs. Otras bases de datos similares pueden ser Pexels o Morguefile.
- **Videvo:** <https://www.videvo.net/>
Base de datos de más de un millón de vídeos con licencias de todo tipo, aunque mayormente gratuitas. Es necesario consultarlas antes de utilizar el material.
- **Flaticon:** <https://www.flaticon.es/>
Base de datos con más de 15 millones de iconos y *stickers* que se pueden utilizar de manera gratuita.
- **Undraw:** <https://undraw.co/>
Base de datos de ilustraciones para diseñadores.
- **Motores de búsqueda** como Google o Bing.
Nos ofrecen la posibilidad de buscar imágenes y videos, e incluso filtrar por licencias que permiten la reutilización, sin embargo, no nos informan de manera precisa de las condiciones de uso. Para consultar fehacientemente estas licencias tendremos que ir a la fuente primaria de los documentos y asegurarnos de sus posibilidades de uso.

En todos los casos que tengamos necesidad de utilizar cualquier tipo de material recuperado de internet, es fundamental asegurarse de las licencias de uso de ese archivo en particular en la propia fuente de información. Y, en caso de duda, consultar de manera personal las condiciones de uso. Existen bases de datos en las que se almacenan documentos con diferentes licencias y no todas gratuitas, ni tampoco con las mismas posibilidades de utilización.

Referencias

Alcalde Perea, I. (2015). *Visualización de la información*. Editorial UOC, S.L.

Cornellá, A. (2 de octubre de 2013). Infoxicación.... *Alfons Cornella Semper proguedi*. Recuperado el 16 de mayo de 2024, de <https://alfonscornella.com/2013/10/02/infoxicacion/>

Esteban Gracia, C. (2017). *La visualización de datos: Evolución de la infografía en el siglo XXI* [Tesis Doctoral]. Universidad Complutense de Madrid. <https://hdl.handle.net/20.500.14352/21163>

Guía de visualización de datos: Definición, ejemplos y recursos de aprendizaje. (2023). Recuperado el 16 de mayo de 2024 de <https://www.tableau.com/es-es/learn/articles/data-visualization>

Medina, J. (2008). *Brain Rules: 12 Principles for Surviving and Thriving at Work, Home, and School*.

Potter, M.C., Wyble, B., Hagmann, C.E., & McCourt, E.S. (2014). Detecting meaning in RSVP at 13 ms per picture. *Attention, Perception, and Psychophysics*, 76(2), 270–279. <https://doi.org/10.3758/s13414-013-0605-z>

¿Qué es la visualización de datos? - Explicación de la visualización de datos - AWS. (2023). Amazon Web Services, Inc. Recuperado el 16 de mayo de 2024 de <https://aws.amazon.com/es/what-is/data-visualization/>

LA ALFABETIZACIÓN EN DATOS EN EL CONTEXTO
UNIVERSITARIO DE LAS CIENCIAS SOCIALES
(ALFADATAUniv_CCSS)

VISUALIZACIÓN Y REPRESENTACIÓN GRÁFICA DE DATOS

Actividad

José Carlos Toro Pascua



EVALUACIÓN DE VISUALIZACIONES DE DATOS: EJEMPLOS DE DISEÑO CORRECTOS E INCORRECTOS

[Actividad propuesta 1]

Tiempo aproximado de realización de la actividad:
20 minutos.

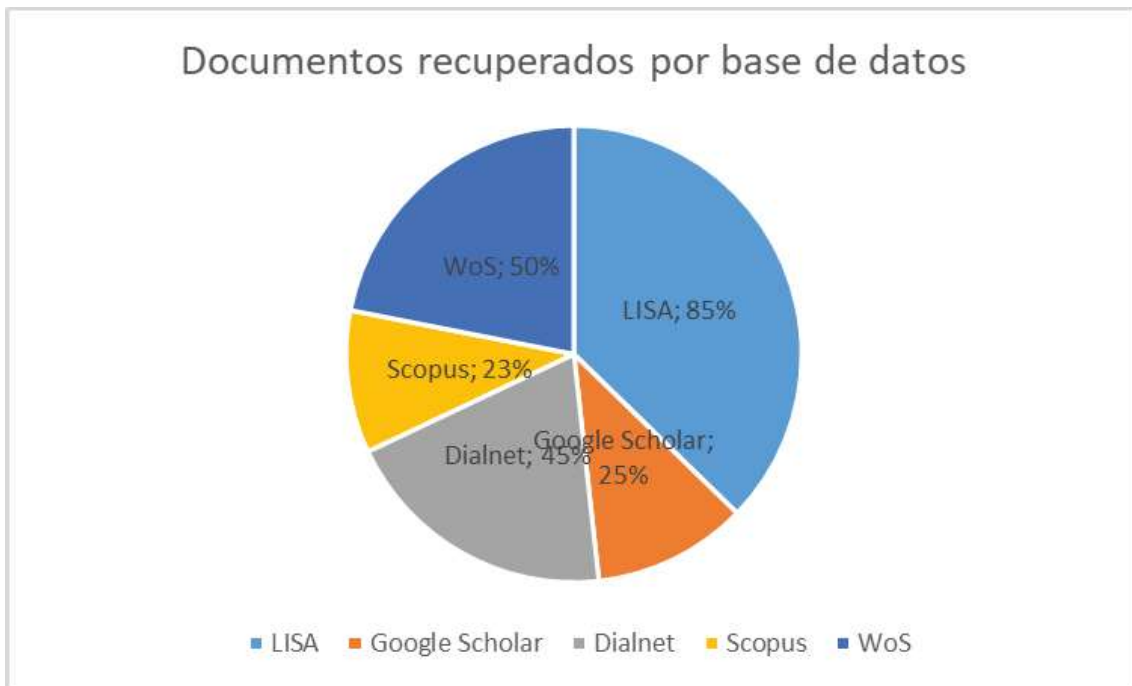
1. Objetivo

El objetivo de esta actividad es valorar y seleccionar aquellas visualizaciones de datos correctas, sabiendo elegir las representaciones gráficas de datos más adecuadas en función de la información que deseemos transmitir.

2. Enunciado explicativo de la actividad

Tras la lectura de la píldora formativa vamos a presentar una serie de visualizaciones gráficas de datos que se tendrán que valorar como correctas o incorrectas. Tras la respuesta, obtendremos la explicación pertinente.

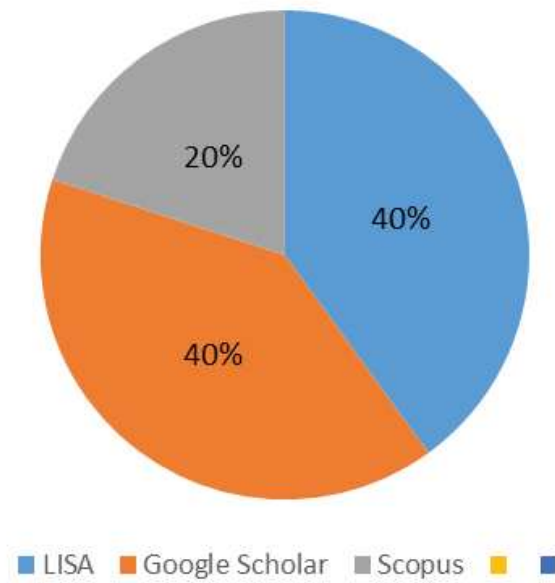
Desarrollo de la actividad



Incorrecto

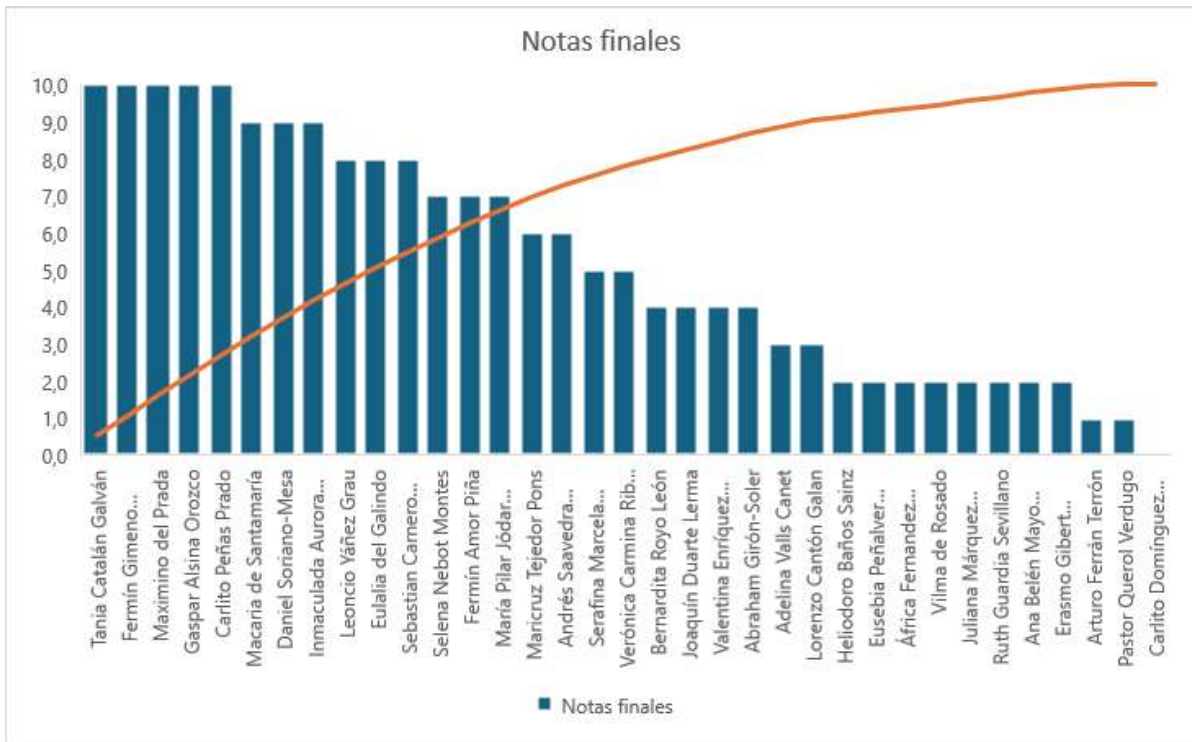
Retroalimentación: En un gráfico circular se visualizan porcentajes ponderados al 100%, se muestra una variable y su relación de las partes con el todo.

Documentos recuperados por base de datos



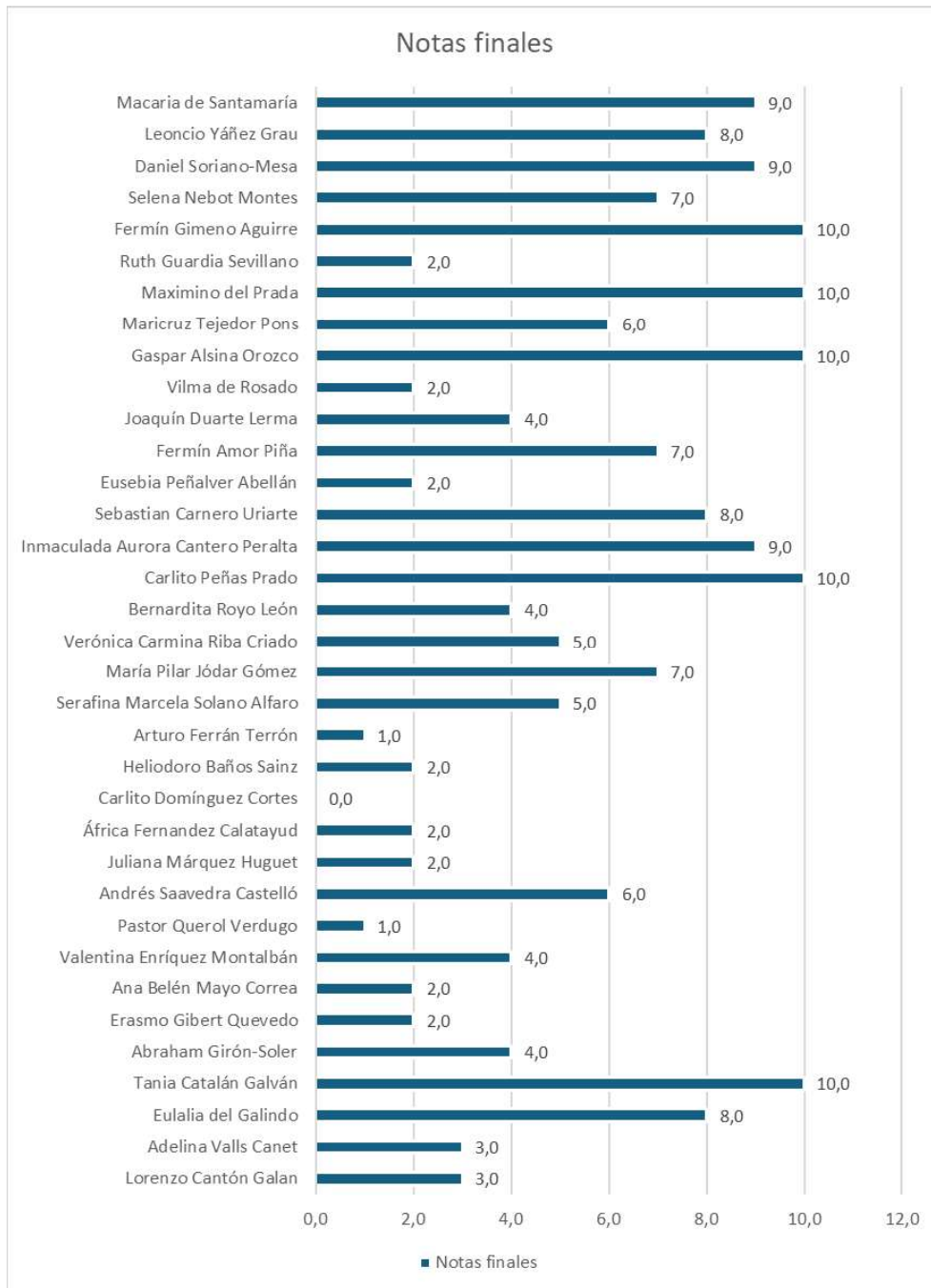
Correcto

Retroalimentación: En un gráfico circular se visualizan porcentajes ponderados al 100%. Se recomiendan de 3 a 5 porciones.



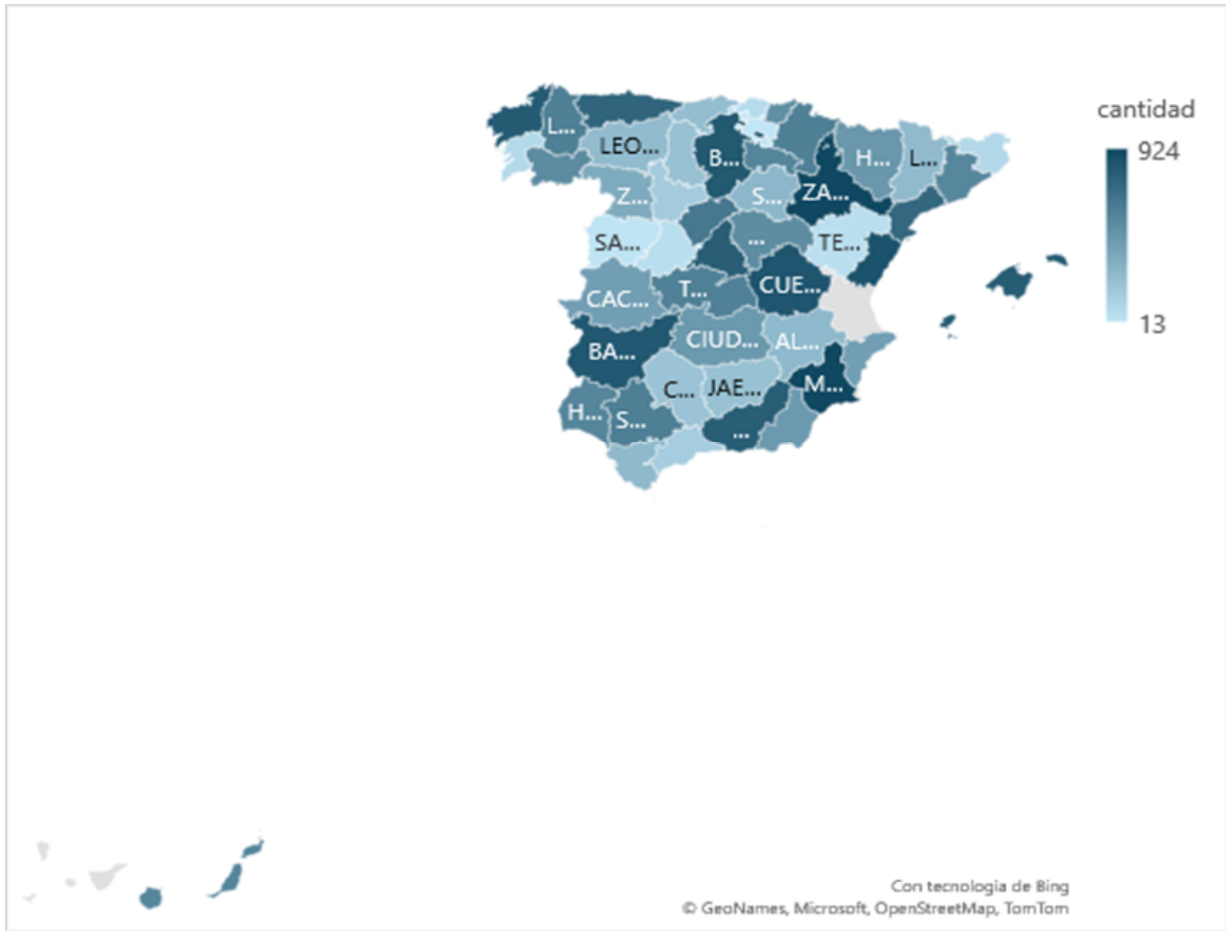
Incorrecto

Retroalimentación: Exceso de datos y no legibles. Nada amigable.



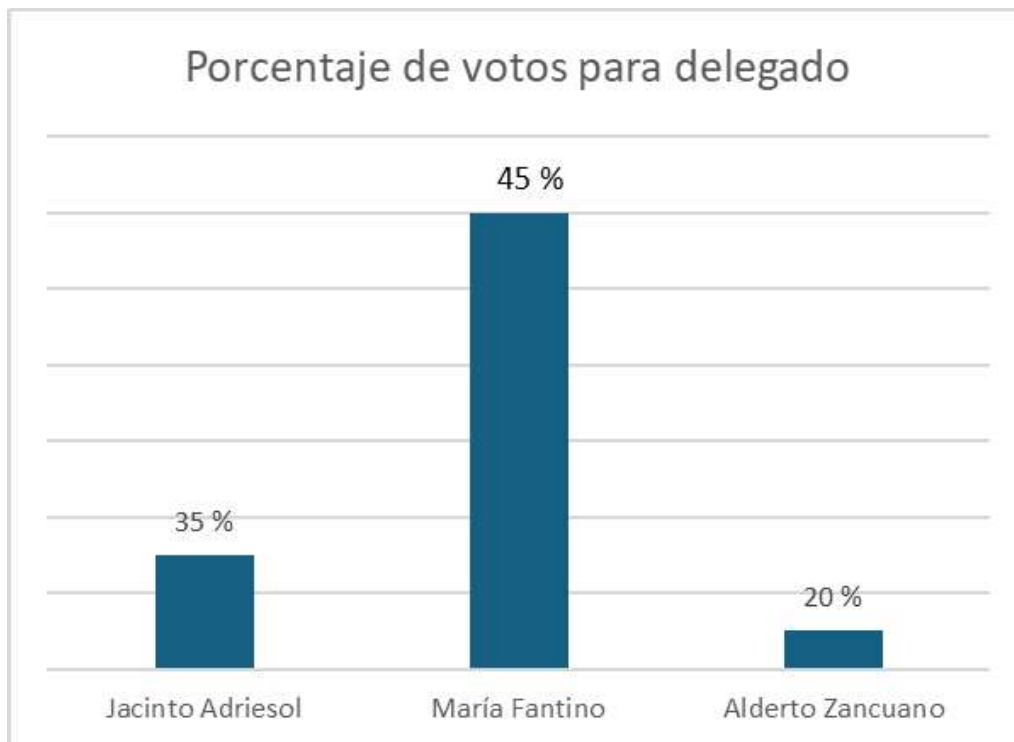
Correcto

Retroalimentación: Opción correcta para mostrar números globales y muchas variables. Podemos ordenar por cantidades o alfabéticamente por la variable.



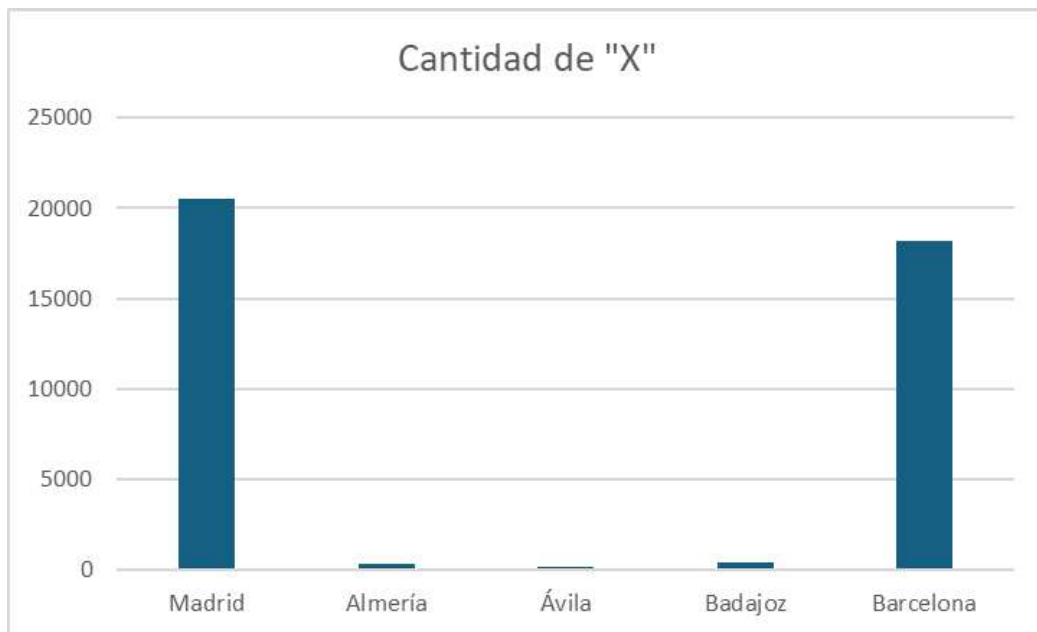
Incorrecto

Retroalimentación: sobrecarga de información al intentar mostrar el nombre de las categorías. Pero son datos conocidos y el mapa de calor sería correcto.



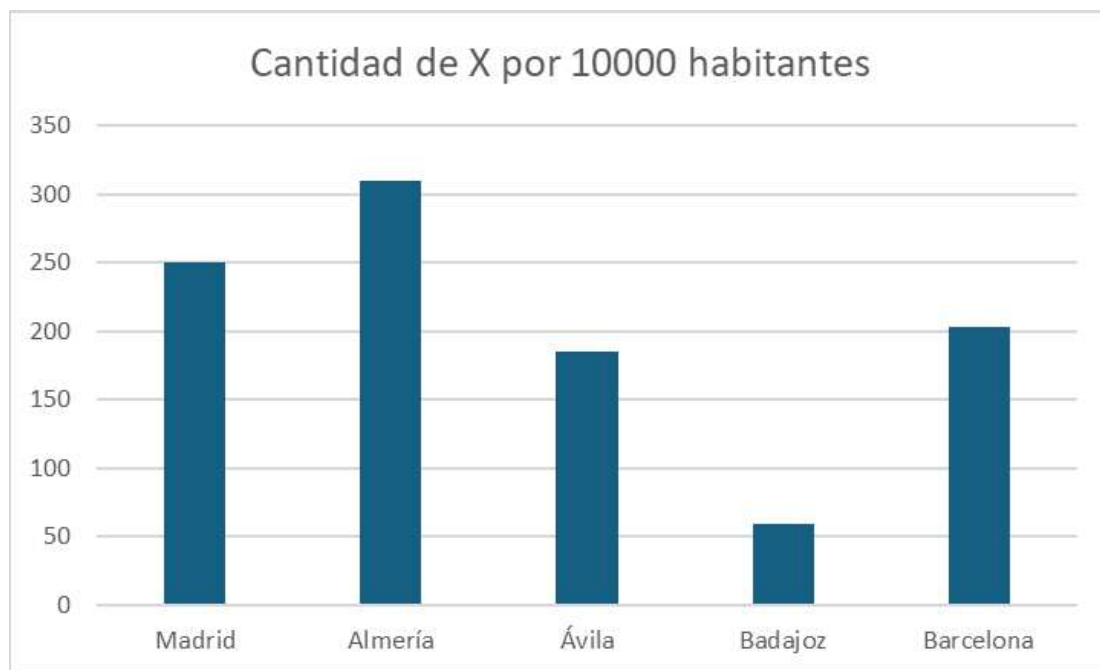
Incorrecto

Retroalimentación: La proporcionalidad del escalado no es correcta.



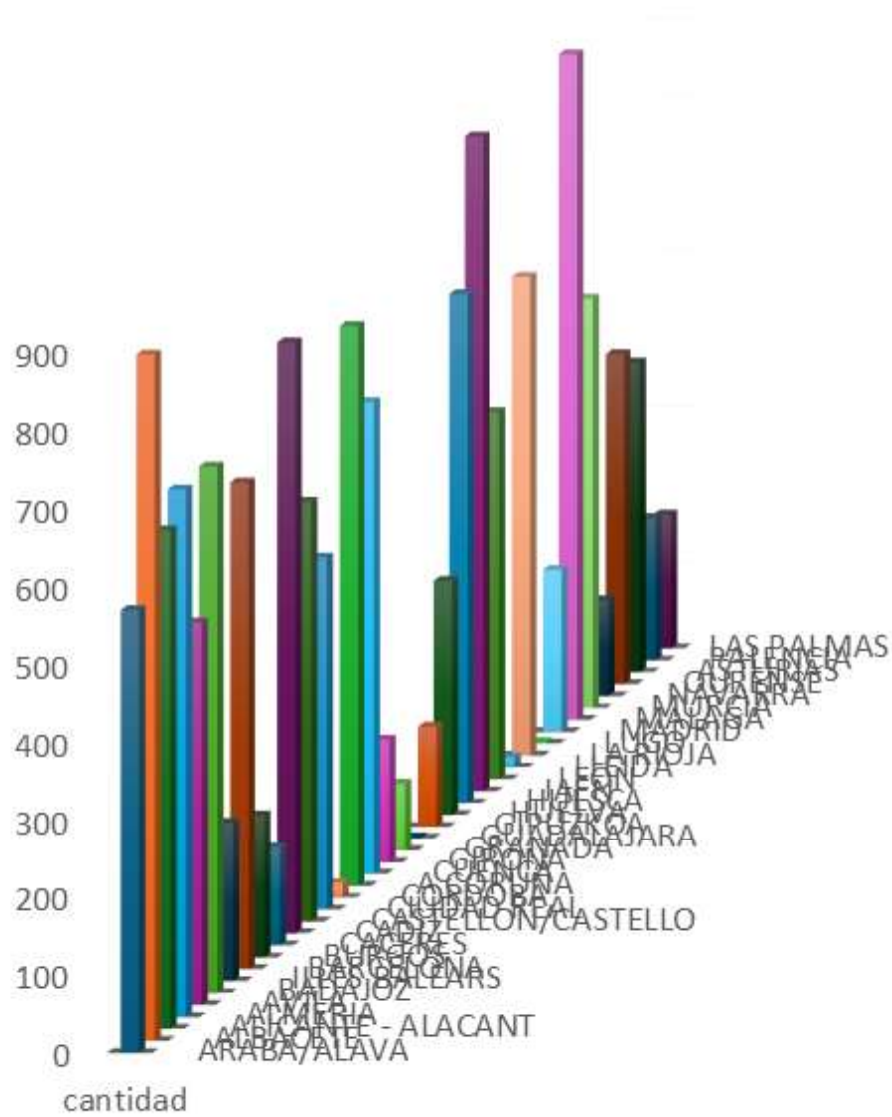
Incorrecto

Retroalimentación: Sería necesaria la ponderación de la cantidad de x por habitante.



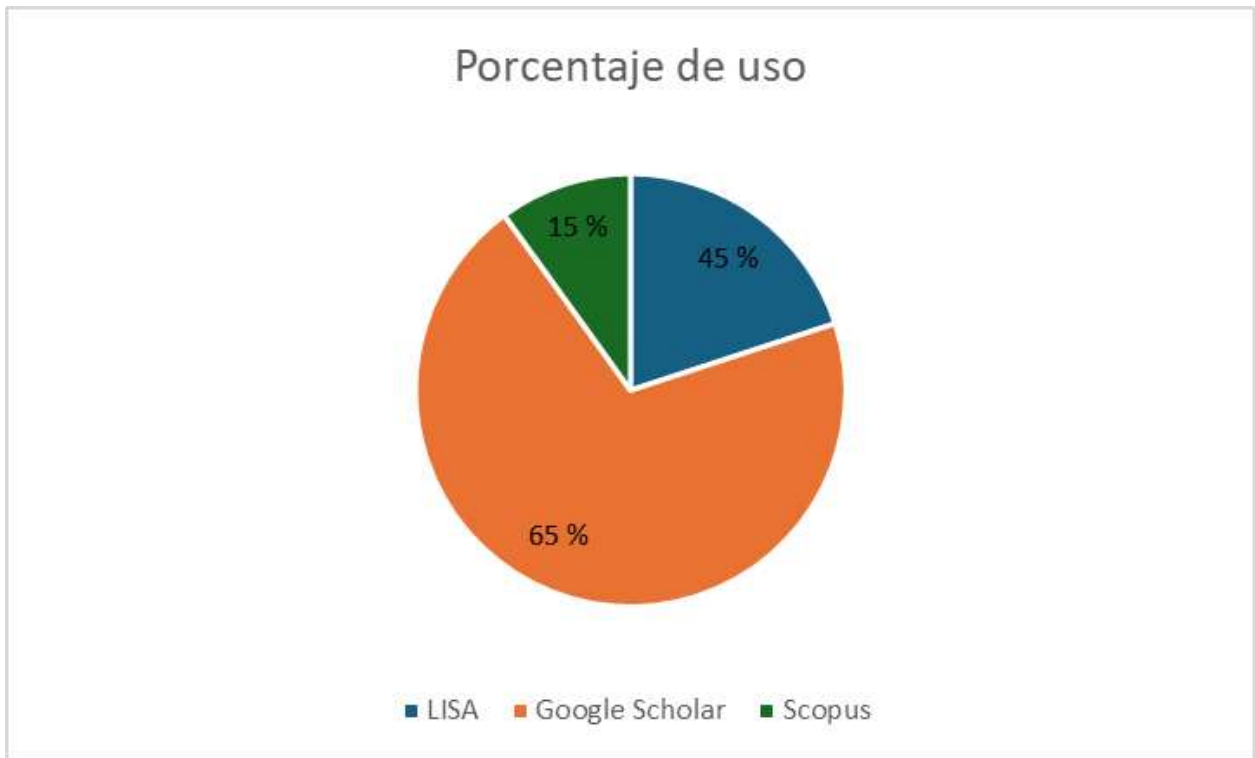
Correcto

Retroalimentación: Ponderación necesaria. Correcto.



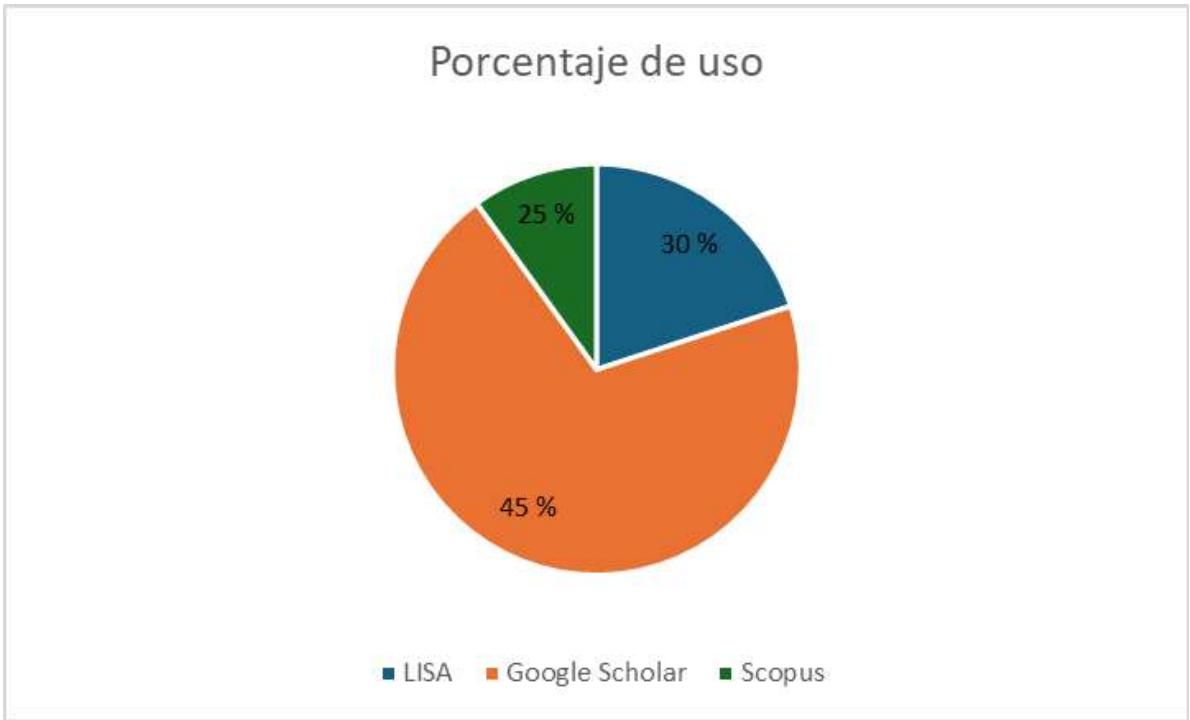
Incorrecto

Retroalimentación: Exceso de datos y colores. Visualización ilegible.



Incorrecto

Retroalimentación: Porcentajes incorrectos, suman más de 100%.



Correcto

Retroalimentación: Las secciones suman el 100% sobre un todo.

DISEÑO DE UNA INFOGRAFÍA CON PIKTOCHART

[Actividad propuesta 2]

Tiempo aproximado de realización de la actividad:
1 hora / 1 hora y 30 minutos.

1. Objetivo

El objetivo de esta actividad es llevar a cabo una representación gráfica basada en datos en la herramienta de creación de infografías en la web Piktochart con una licencia gratuita.

2. Enunciado explicativo de la actividad

Tras registrarnos en la aplicación Piktochart (<https://piktochart.com/es/>) mediante correo electrónico o cuenta de Google, vamos a crear una infografía partiendo de una plantilla gratuita, bien en blanco o con un diseño previo que podamos editar.

3. Desarrollo de la actividad

En la infografía que realicemos vamos a procurar insertar una serie de elementos diversos que nos ofrece la aplicación:

1. Al menos un componente gráfico de los disponibles: iconos, ilustraciones, 3D o fotografías.
2. Al menos dos estilos de texto diferentes.
3. Un componente de diseño de la categoría de “Diagramas y gráficos” o “Procesos y diagramas de flujo”.
4. Un gráfico basado en datos que tengamos en una tabla en formato .XLSX o .CSV y, de esta manera, importar los datos directamente a Piktochart.
5. Un mapa en el que incluyamos una variable cuantitativa.
6. Un vídeo de Youtube.
7. Firma o créditos de la infografía.

4. Visualización de la representación gráfica final

Durante el desarrollo de la actividad podemos ir comprobando el resultado final mediante la previsualización, además de utilizar la función de grabación de la presentación.

También podremos compartir la infografía final, además de exportarla, aunque únicamente en formato .PNG.

LA ALFABETIZACIÓN EN DATOS EN EL CONTEXTO
UNIVERSITARIO DE LAS CIENCIAS SOCIALES
(ALFADATAUniv_CCSS)

LA GESTIÓN DE DATOS DE INVESTIGACIÓN

Yolanda Martín González
Ana Iglesias Rodríguez



LA GESTIÓN DE DATOS DE INVESTIGACIÓN

ÍNDICE

1. La gestión de datos de investigación
2. El proceso de gestión de datos de investigación (fases, principios FAIR, ...)
3. El Plan de Gestión de Datos
 - 3.1. Qué es
 - 3.2. Por qué es necesario elaborar un Plan de Gestión de Datos
 - 3.3. Requerimientos de las agencias de financiación relativos a la gestión de datos

Referencias

1. LA GESTIÓN DE DATOS DE INVESTIGACIÓN

La Ciencia abierta constituye una nueva manera de hacer ciencia (Abadal, 2021), ya que se trata de un modelo de investigación científica que impulsa, entre otras cosas, la apertura de todas las fases del proceso científico lo que conlleva facilitar la utilización, reutilización y distribución de los datos de investigación y promover prácticas para su compartición (*data sharing*). Esto supone que un investigador no sólo permite el acceso libre a sus publicaciones sino también a los datos subyacentes a estas, favoreciendo, de este modo, la replicabilidad de sus investigaciones, así como una mayor difusión e impacto de las mismas. Sin embargo, para lograr este propósito, es necesario que los datos se ajusten a los criterios FAIR, se faciliten en formatos y licencias abiertas, y se difundan en repositorios que favorezcan su acceso y preservación.

Se denomina “dato de investigación” a todo aquel material digital, generado o recolectado durante la realización de una investigación, y que permite a un investigador validar o certificar los resultados obtenidos. Por su parte, un *dataset* es un conjunto organizado de datos relacionados.

En el entorno científico los datos presentan una naturaleza variada en función del ámbito disciplinar donde se contextualiza la investigación. Así, se consideran “datos de investigación” a conjuntos de informaciones textuales, numéricas y multimedia, lo que incluye estadísticas, resultados de experimentos, mediciones, observaciones, grabaciones e imágenes de entrevistas¹, pero también fotografías, cuadernos de laboratorio, cuestionarios, diarios, protocolos y algoritmos, entre otros.

Los datos deben someterse a un proceso que posibilite gestionarlos desde el inicio del ciclo de vida de los mismos. Asimismo, han de estar descritos conforme a una serie de metadatos y debidamente documentados de modo que pueda conocerse su origen, el por qué y el para qué fueron creados, y en qué contexto. La correcta gestión de los datos posibilita encontrarlos fácilmente, mantenerlos seguros y tener suficiente documentación para poder utilizarlos durante y posteriormente a la investigación.

Este proceso permite dar respuesta a las siguientes cuestiones: (1) Qué datos se van a recopilar; (2) Cómo se van a organizar; (3) Dónde se almacenarán; (4) Cómo se van a documentar; (5) Cómo van a preservarse; (6) Cómo y dónde van a difundirse; (7) Qué requisitos éticos y legales se han de considerar; y (8) Quién va a ser el responsable de llevar a cabo la gestión de datos.

1 Directiva 2019/1024 del Parlamento Europeo y del Consejo de 20 de junio de 2019, relativa a los datos abiertos y a la reutilización de la información. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=uriserv:OJ.L_.2019.172.01.0056.01.SPA

2. EL PROCESO DE GESTIÓN DE DATOS DE INVESTIGACIÓN

El proceso de gestión de los datos científicos comprende una serie de acciones interrelacionadas y consecutivas que pasamos a enumerar y a describir a continuación.

1. Asignar roles y responsabilidades.
2. Identificar y organizar los datos.
3. Describir y documentar los datos.
4. Diseñar una estrategia de almacenamiento y preservación de los datos.
5. Considerar los aspectos éticos y legales relacionados con los datos.
6. Seleccionar los datos que se van a difundir y dónde se compartirán.
7. Identificar los recursos (técnicos, financieros, etc.) que demandará la gestión de los datos.
8. En el caso de que la investigación reciba financiación comprobar cuáles son los requerimientos de la agencia financiadora en materia de gestión de datos.

La primera tarea comprende la asignación de la persona o grupo de personas que se van a encargar de llevar a cabo, de manera adecuada, este proceso. Lo más recomendable es que sea un profesional de la gestión de la información, sin embargo, este perfil no siempre está presente en los grupos de investigación por lo que se aconseja que recaiga esta tarea en uno o varios investigadores que pueden contar con el asesoramiento del personal bibliotecario.

Una vez asignada esta responsabilidad y los distintos roles que cada uno ha de desempeñar, el siguiente paso consiste en identificar los datos producidos y/o recopilados durante la investigación, así como su naturaleza, esto es, si se trata de datos textuales, numéricos, multimedia, etc. De igual modo, se han de tener en cuenta los formatos en los que se disponen y el volumen de la información que se ha de gestionar. También conviene realizar un registro donde se documente la procedencia de los datos y se establezca cómo se van a disponer en los distintos ficheros y carpetas. En este sentido, han de primar los formatos estándar y abiertos (no propietarios) que favorecen el uso, reutilización y distribución de los datos.

Los datos deben almacenarse en distintos ficheros y carpetas utilizando denominaciones y códigos que permitan su fácil identificación y localización. Los nombres asignados han de ser descriptivos y breves, se han de separar los términos mediante guiones bajos, y, en el caso de incluir fechas, se aconseja seguir el modelo americano AAAAMMDD. Para facilitar la identificación de los ficheros se seguirá un sistema secuencial y se podrá añadir información que refleje el estado del mismo (borrador, final, provisional, definitiva, etc.). Se deberán renombrar siempre las distintas versiones de los ficheros que se vayan generando y eliminar las versiones obsoletas.

Los ficheros se organizan en carpetas, por lo general, en función del tipo de datos que contengan o de la actividad de la investigación a la que se refieran. No deben crearse más de tres o cuatro niveles de carpetas ni incluirse más de diez ficheros en cada una de ellas.

Describir y documentar los datos constituye una de las acciones más relevantes en el proceso de gestión, ya que permitirá conocer el origen de los datos, contextualizarlos para posteriormente preservarlos, visualizarlos y recuperarlos adecuadamente. En esta fase hemos de asegurarnos de que los datos cumplen con los denominados principios FAIR que garantizarán que los datos sean recuperables (*findables*), accesibles (*accessibles*), interoperables (*interoperables*) y reutilizables (*reuses*).

Para que un dataset pueda ser recuperable, tanto por las personas como por las máquinas, se le ha de asignar un identificador único y persistente, se ha de describir mediante la asignación de metadatos y se ha de almacenar en algún repositorio o base de datos que pueda ser rastreada por un motor de búsqueda. Una vez localizado el dataset, el usuario debe poder acceder a él y, para ello, se han de utilizar *software* o programas informáticos abiertos (no propietarios) al tiempo que se establecerán distintos permisos de acceso. Asimismo, los datos se integran con otros datos por lo que han de ser interoperables con otras aplicaciones.

Esto se consigue mediante el uso de lenguajes y esquemas informáticos estandarizados capaces de ser leídos por máquinas. Por último, los datos han de ser reutilizables y, para ello, han de acompañarse de licencias donde se indiquen los derechos y limitaciones de uso que tienen, pero, también, han de organizarse según estándares de descripción, almacenamiento e intercambio.

Decidir dónde se almacenan los datos constituye otra de las acciones que han de tenerse en cuenta en el proceso de gestión de datos científicos. Se ha de escoger aquella plataforma o depósito que ofrezca las mayores garantías en términos de seguridad y preservación. De este modo, se puede optar por conservar los datos en la Nube, en un disco duro, etc.

Si se pretende difundir los conjuntos de datos generados y/o recabados durante la investigación es importante tener en cuenta que estos no se hallen sujetos a ninguna restricción por razones de confidencialidad, privacidad y copyright. Además, cuando el *dataset* contenga datos personales, estos deberán anonimizarse o contar con el consentimiento explícito de las personas que han participado en la investigación. En esta fase del proceso de gestión es imperativo considerar lo estipulado en la normativa vigente en materia de protección de datos y de propiedad intelectual, y, al mismo tiempo, asegurar que se realiza un uso legal y ético de los mismos.

La siguiente acción consiste en decidir qué datos se preservarán y difundirán y, para ello, se aconseja seguir la recomendación de la Comisión Europea que promueve que los datos de investigación se proporcionen de manera “tan abierta como sea posible, tan cerrada como sea necesario”. En este sentido, se han de identificar: los propósitos que pueden cumplir los dataset, aquellos que han de preservarse por mandato y los que, en nuestra opinión, deberían conservarse. En este punto también debería tenerse en cuenta el coste que puede conllevar su preservación.

Una vez identificados los conjuntos de datos que se quiere difundir, se procederá a seleccionar el repositorio más adecuado. En la actualidad contamos con varias plataformas que permiten la compartición de datos de investigación como son los propios repositorios de las universidades y/o centros de investigación u otras plataformas de ámbito internacional como son Zenodo, Figshare, Dryad o Mendeley Data, entre otras.

Hoy día, las agencias de financiación de la investigación demandan que este proceso quede reflejado en un documento formal y vivo conocido como Plan de Gestión de Datos, que acompaña al proceso de investigación y describe de qué manera se gestionan los datos generados y/o recopilados en el transcurso del mismo, de modo que estén debidamente documentados, preservados y sean recuperables.

El Plan de Gestión de Datos debe elaborarse específicamente para la investigación que se va a realizar.

3. EL PLAN DE GESTIÓN DE DATOS

3.1. Qué es

El Plan de Gestión de Datos (PGD) es un documento en el que se detalla y describe cómo se van a gestionar los datos que se generan y/o se recopilan en el transcurso de una investigación. De este modo, el plan permite mantener los datos seguros, garantizando su integridad y la adecuada custodia de los mismos desde su generación y, especialmente, al concluir un proyecto de investigación, además de su preservación y reutilización posterior. En definitiva, este documento permite gestionar los datos de manera adecuada y eficiente.

Los aspectos que debe considerarse a la hora de elaborar un PGD son los siguientes:

- Qué datos de investigación se pretende crear y/o recoger.
- Cómo se van a organizar los datos y dónde se almacenarán de manera segura.
- Cómo se describirán y documentarán los datos durante el proceso de recopilación y análisis de los mismos.
- Cómo se aplicará la normativa de protección de datos y cómo se gestionarán de manera ética y legal.

- Qué política y prácticas se aplicarán que aseguren la preservación a lo largo plazo de los datos.
- Sobre quién o quiénes recaerá la responsabilidad de la gestión de los datos de investigación.
- Qué recursos (financieros, técnicos, etc.) requiere la gestión adecuada de los datos de investigación.

3.2. Por qué es necesario elaborar un Plan de Gestión de Datos

En términos generales, el Plan de Gestión de Datos constituye una herramienta esencial para garantizar la disponibilidad, la preservación y la reutilización de los datos en el ámbito de la investigación y en otros contextos donde la gestión de la información es clave. Y, más concretamente, resulta esencial por las siguientes razones:

1. Un PGD garantiza la integridad y la fiabilidad de los datos generados durante la investigación, ya que reduce la posibilidad de cometer errores y mejora la calidad de los resultados porque en él se establecen prácticas y estándares claros de actuación.
2. Cada vez más, las instituciones, organismos y agencias financiadoras establecen como requisito obligatorio que los proyectos de investigación cumplan con normativas específicas de gestión de datos. Y, el PGD es la herramienta que permite cumplir no sólo con este requisito, sino también con las regulaciones y políticas institucionales establecidas al respecto, evitando posibles sanciones y asegurando la integridad científica y la ética de la investigación.
3. El PGD garantiza la transparencia y la reproducibilidad de los datos, puesto que proporciona detalles acerca de la metodología que se ha empleado en la recopilación, procesamiento y análisis de los datos. Este hecho, además de incrementar la transparencia de la investigación y facilitar la reproducción de los resultados por parte de otros investigadores, fortalece la validez y la confiabilidad de los hallazgos obtenidos en la investigación.
4. El PGD permite preservar los datos a largo plazo, asegurando que la información sea accesible y utilizable no sólo durante el transcurso de la investigación, sino también después de que ésta concluya. Esto es esencial para la replicabilidad de los resultados y para desarrollar nuevas investigaciones basadas en los mismos datos.
5. El PGD ayuda a planificar y a gestionar de forma eficiente los recursos relacionados con la gestión de datos, como el almacenamiento, la seguridad y las herramientas tecnológicas necesarias, facilitando la optimización del uso de estos recursos y evitando problemas potenciales.
6. El PGD incluye medidas para proteger la información sensible y reducir los riesgos de pérdida o acceso no autorizado a los datos, siendo esta una de las grandes preocupaciones en el momento actual. Estas medidas, además de asegurar que se cumplan las normativas y directrices relevantes, ayudan a mantener la confidencialidad y la integridad de los datos.
7. El PGD, al describir de forma detallada los métodos de recopilación y procesamiento de los datos, facilita la reutilización de los datos por parte de otros investigadores, favoreciendo y contribuyendo a la transferencia del conocimiento en la comunidad científica.
8. El PGD también facilita la publicación de resultados, teniendo en cuenta que, cada vez en mayor medida, las revistas científicas y las plataformas de publicación solicitan la disponibilidad de los datos para respaldar los resultados publicados. Un buen Plan de Gestión de Datos ayuda a preparar y proporcionar estos datos de manera efectiva.

3.3. Requerimientos de las agencias de financiación relativos a la gestión de datos

Como venimos indicando a lo largo de este módulo, los Planes de Gestión de Datos se han convertido en un componente esencial para garantizar una gestión adecuada de los datos de investigación. Su elaboración e integración en los procesos de investigación constituye un requerimiento, cada vez más frecuente, de las diferentes agencias financiadoras que han convertido este documento en un requisito obligatorio para poder optar a proyectos de investigación financiados. Así, por ejemplo, la primera vez que el Ministerio de Ciencia e Innovación solicitó incorporar un Plan de Gestión de Datos en los Proyectos I+D+i fue en la convocatoria del año 2020, siguiendo lo establecido en el Plan Estatal de Investigación Científica y Técnica y de Inno-

vación 2017-2018. Y, actualmente, entre las medidas de actuación que establece el Ministerio de Ciencia e Innovación (2023) en su “Estrategia Nacional de Ciencia Abierta (ENCA) 2023-2027”, cabe destacar, dado el tema que estamos abordando, la medida B2, que es una de las establecidas en el Eje B. Gestión de datos de investigación siguiendo los principios FAIR.

Esta medida, entre otras cuestiones indica que,

(...) Se hará obligatoria la realización de un **plan de gestión de datos** (PGD) como parte integral de los proyectos de investigación financiados con fondos públicos, incluyendo la evaluación de este plan dentro de las actividades de seguimiento de los proyectos, siempre que la disciplina objeto de la investigación conlleve la existencia de datos. Este PGD tendrá un estándar que facilite su evaluación y seguimiento, de acuerdo con un procedimiento y protocolo acorde. (p.17)

En lo que respecta a las políticas de Ciencia Abierta, es importante mencionar que la Ley 17/2022, de 5 de septiembre, por la que se modifica la Ley 14/2011, de 1 de junio, de la Ciencia, la Tecnología y la Innovación, reconoce la importancia de la ciencia como un recurso público y defiende el acceso abierto a los resultados de la investigación financiados con fondos públicos. El artículo 37 “Ciencia abierta” promueve el desarrollo de infraestructuras y plataformas abiertas, obliga al depósito en abierto de publicaciones, **datos**, códigos y métodos en repositorios y fomenta la participación abierta de la sociedad civil en los procesos científicos.

También, la Ley Orgánica 2/2023 del 22 de marzo, del Sistema Universitario LOSU dedica el artículo 12 al “Fomento de la Ciencia abierta y Ciencia ciudadana” asegurando el mandato de depósito en abierto de los resultados de investigación en el ámbito universitario, destacando la importancia de los repositorios institucionales para que se cumpla este propósito.

Por su parte, la UNESCO (2021), en su Recomendación sobre la Ciencia Abierta², plantea la necesidad de contar con una ciencia abierta, accesible, inclusiva y basada en la colaboración. Para ello, es necesario que científicos e ingenieros utilicen licencias abiertas a través de las cuales compartan sus publicaciones, sus **datos**, sus materiales pedagógicos, sus *software* y su equipamiento.

En este documento, la UNESCO, ofrece una serie de directrices y principios clave para favorecer y fomentar la apertura en la investigación científica. Entre estos mandatos y recomendaciones cabe destacar dos relativas a la temática a la que nos referimos: (1) la adopción de políticas que fomenten la apertura de los datos científicos, lo que conlleva la libre disponibilidad y accesibilidad de conjuntos de datos generados en el contexto de la investigación científica; (2) la transparencia y la reproducibilidad en la investigación científica. A este respecto, pide a los investigadores que compartan no sólo los resultados de sus investigaciones, sino también los datos obtenidos de la misma y los métodos utilizados. Esta recomendación se enmarca en la Recomendación sobre la Ciencia y los Investigadores Científicos de 2017³, en la Estrategia sobre la contribución de la UNESCO a la promoción del libre acceso a la información y la investigación científicas de 2011⁴, en la Recomendación de los Recursos Educativos Abiertos (REA) (2019)⁵ y en la nueva Recomendación sobre la Ética de la Inteligencia Artificial⁶.

A nivel europeo, es importante señalar que en el Programa Horizonte 2020 se estableció como obligatorio para todas las áreas disciplinarias, la implementación de un proyecto piloto para el acceso abierto a las publicaciones, así como la ejecución de un proyecto piloto relacionado con la gestión de datos de investigación. Este último requería que ciertas disciplinas presentaran un **Plan de Gestión de Datos** y pusieran a disposición del público los datos resultantes de sus investigaciones. Este piloto se extendió a todas las áreas del programa y a todos los proyectos financiados a partir del 1 de enero de 2017 (European Commission, 2016), convirtiéndose en un requisito obligatorio para todos los proyectos que se soliciten. Este Plan debe incluir, al menos, información sobre los datos que se generarán, dónde se compartirán y almacenarán, y, en el caso de que sucediera, cuáles son las razones que justificarían la decisión de no compartir los datos de manera abierta. Este mismo requisito se mantiene como obligatorio en Horizon Europe (2021-2027) que, al igual que en el Plan anterior, obliga a redactar y a mantener un PGD, que deberá actualizarse a lo largo de la ejecución del proyecto.

2 Recomendación de la Unesco sobre la Ciencia Abierta. <https://doi.org/10.54677/YDOG4702>

3 https://unesdoc.unesco.org/ark:/48223/pf0000263618_spa

4 https://unesdoc.unesco.org/ark:/48223/pf0000211047_spa

5 <https://unesdoc.unesco.org/ark:/48223/pf0000373755/PDF/373755eng.pdf.multi.page=20>

6 https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa

Horizon Europe (2021-2027) promueve, como uno de sus principales retos, la adopción de la ciencia abierta como un modo de trabajo para todos los investigadores. Para ello, insta a los investigadores a poner en acceso abierto los datos y los resultados de sus investigaciones; a que faciliten la puesta en común abierta y en una fase temprana de sus investigaciones, mediante un registro previo, informes, prepublicaciones, búsqueda de soluciones a problemas específicos, entre otros; a que utilicen infraestructuras abiertas de investigación en las que compartan conocimientos y datos; a poner en práctica medidas que garanticen la reproducibilidad de los resultados; a facilitar la colaboración abierta en el ámbito de la ciencia y con otros agentes del conocimiento (ciudadanos, sociedad civil y usuarios finales), favoreciendo con ello la ciencia ciudadana. Añadido a esto, Horizon Europe exige y obliga a los investigadores, a que todas las publicaciones científicas resultantes de los proyectos financiados por este programa, estén en acceso abierto de forma inmediata; a que hagan una gestión responsable de los datos de investigación, cumpliendo con los principios FAIR, y a que se dé acceso a terceros a los resultados de las investigaciones para que puedan verificar o validar las publicaciones.

Con respecto a las propuestas de investigación, estos programas marco incluyen la obligación de elaborar y presentar, como ya hemos indicado anteriormente, un **Plan de Gestión de Datos** que garantice la preservación y el acceso a los datos de investigación. En concreto, en la Guía del Participante Horizonte Europa (Ministerio de Ciencia e Innovación (MCIN), CDTI, E.P.E. y Fundación Española para la Ciencia y la Tecnología (FECYT), 2022), se dice que: “el Plan de Gestión de Datos (Data Management Plan - DMP) es obligatorio para todos los proyectos que generen o reutilicen datos y debe estar alineado con el Plan de difusión, explotación y comunicación” (p.112).

¿Cuáles son los pasos que establece la Comisión Europea en la elaboración de un Plan de Gestión de Datos en estos casos?

- Paso 1: Redactar el PGD (<https://dmponline.dcc.ac.uk/>). Se entregará a los 6 meses, se efectuará una revisión intermedia y una revisión final.
- Paso 2: Encontrar un repositorio. Ha de ser coherente con las necesidades de los datos. Repositorios de datos: (i) disciplina/institucional; (ii) www.re3data.org; (iii) Zenodo.
- Paso 3: Depósito de datos. Datos abiertos, Metadatos, otras herramientas: Formatos de archivo estándar, esquema estándar de metadatos, licencias abiertas.
- Soporte: Infraestructura e información de apoyo: Plan Piloto de Datos de Investigación en Abierto (Open Research Data Pilot) (Plan Piloto Horizonte 2020): Directrices de la Comisión Europea, OpenAire.eu, www.dcc.ac.uk

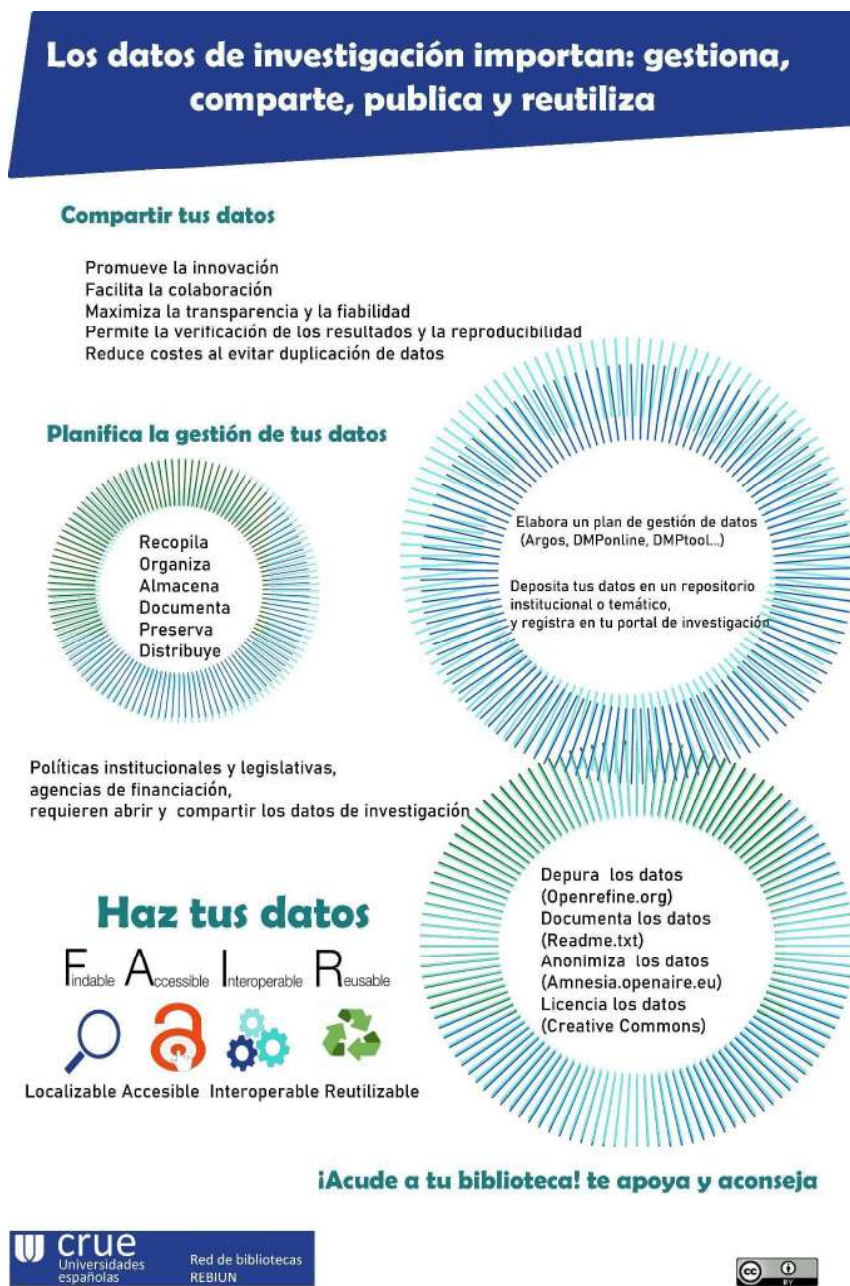
En lo que respecta al mandato de acceso abierto a los datos de investigación (artículo 29.3), el texto indica que los beneficiarios deben depositar los datos en un repositorio y permitir el acceso, el análisis (minería de datos), la explotación, la reproducción y la difusión de los mismos, de forma gratuita para cualquier usuario. Se deben proporcionar herramientas necesarias para validar los resultados. Pero, existen excepciones para proteger ciertos datos si ponerlos a disposición podría comprometer el objetivo principal de la acción.

Cabe destacar que la Comisión Europea suele hacer de forma regular un seguimiento del grado de cumplimiento de estos mandatos, conllevando, el incumplimiento de los mismos, consecuencias económicas, la reducción de la subvención (art. 43, p.91) u otras medidas (European Commission, 2017, p.90).

A modo de resumen, presentamos la infografía elaborada por la CRUE (2023), en la que se pone de manifiesto que “los datos de investigación importan: gestiona, comparte, publica y reutiliza” (Figura 1).

Figura 1.

Los datos de investigación importan: gestiona, comparte, publica y reutiliza.



Fuente: CRUE (2023). <https://hdl.handle.net/20.500.11967/1318>

Referencias

Abadal, E. (2021). Ciencia abierta: un modelo con piezas por encajar. *Arbor*, 197(799), a588. <https://doi.org/10.3989/arbor.2021.799.003>

Comisión Europea, Dirección General de Investigación e Innovación, (2021). *Horizon Europe, ciencia abierta: puesta en común de conocimientos y datos en una fase temprana y colaboración abierta*. Oficina de Publicaciones de la Unión Europea. <https://data.europa.eu/doi/10.2777/926051>

European Commission. (2016). *H2020 Programme. Guidelines on FAIR Data Management in Horizon 2020*. https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

European Commission. (2017). *H2020 Programme. Multi-Beneficiary General Model Grant Agreement*. https://ec.europa.eu/research/participants/data/ref/h2020/mga/gga/h2020-mga-gga-multi_en.pdf

Ministerio de Ciencia e Innovación, CDTI, E.P.E. y Fundación Española para la Ciencia y la Tecnología (FECYT). (2022). *Guía del participante HORIZONTE EUROPA*. <https://www.horizonteeuropa.es/>

Ministerio de Ciencia e Innovación (2023). *Estrategia Nacional de Ciencia Abierta [ENCA] 2023-2027*. Secretaría General Técnica del Ministerio de Ciencia e Innovación. <https://www.ciencia.gob.es/Estrategias-y-Planes/Estrategias/ENCA.html>

LA ALFABETIZACIÓN EN DATOS EN EL CONTEXTO
UNIVERSITARIO DE LAS CIENCIAS SOCIALES
(ALFADATAUniv_CCSS)

LA GESTIÓN DE DATOS DE INVESTIGACIÓN

Actividad

Yolanda Martín González
Ana Iglesias Rodríguez



EL PLAN DE GESTIÓN DE DATOS

[Actividad propuesta]

Tiempo aproximado de realización de la actividad:
1 hora.

1. Objetivo

El objetivo de esta actividad es diseñar un Plan de Gestión de Datos de Investigación (PGD) aplicable a un ámbito de interés particular.

2. Enunciado explicativo de la actividad

La gestión eficaz de los datos de investigación es fundamental para garantizar la integridad, la accesibilidad, la preservación y la reutilización de los mismos. En esta actividad, elaborarás un plan detallado que aborde los diversos aspectos involucrados en la gestión de datos, desde la asignación de roles hasta la consideración de los aspectos éticos y legales.

3. Desarrollo e instrucciones para realizar la actividad

Una vez leído el documento explicativo en torno a la temática *La Gestión de datos de investigación* y después de haberte familiarizado con las pautas proporcionadas en la plantilla DMPOnline para elaborar un Plan de Gestión de Datos, sigue los pasos descritos a continuación:

- Selecciona un área de interés para tu Plan de Gestión de Datos. Puede ser cualquier campo de investigación que te resulte relevante o significativo.
- Utiliza la plantilla proporcionada por DMPOnline para elaborar tu Plan de Gestión de Datos. Asegúrate de abordar los siguientes aspectos:
 1. Asigna roles y responsabilidades dentro del equipo de investigación en relación con la gestión de datos.
 2. Identifica y organiza los datos que serán generados y/o recopilados durante la investigación.
 3. Describe y documenta de forma adecuada los datos, incluyendo su naturaleza, formato y estructura.
 4. Diseña una estrategia sólida para el almacenamiento y la preservación a largo plazo de los datos, asegurando su accesibilidad e integridad.
 5. Considera los aspectos éticos y legales relacionados con la recolección, uso y divulgación de los datos, garantizando el cumplimiento de las normativas y regulaciones pertinentes.
 6. Selecciona qué datos serán compartidos y dónde serán difundidos, teniendo en cuenta las plataformas y repositorios más adecuados para su publicación.

7. Identifica los recursos (técnicos, financieros y humanos) necesarios para llevar a cabo la gestión de datos de manera efectiva.
 8. En el caso de recibir financiación para la investigación, verifica y cumple con los requisitos específicos de la agencia financiadora en materia de gestión de datos.
- Puedes consultar ejemplos de Planes de Gestión de Datos previamente elaborados para orientarte en el proceso.

4. Recursos requeridos para realizar la actividad

- Documento explicativo sobre *La Gestión de datos de investigación*.
- Registrarse en la plataforma DMPOnline (<https://dmponline.dcc.ac.uk/>).
- Acceder a la plantilla que proporciona DMPOnline, así como a los diferentes ejemplos de Planes de Gestión de Datos que proporciona.

5. Autoevaluación de la actividad

- ¿He asignado roles y responsabilidades dentro del equipo de investigación para la gestión de datos?
 Sí No
- ¿He identificado y organizado adecuadamente los datos que serán recopilados durante la investigación?
 Sí No
- ¿He utilizado nombres breves y descriptivos para denominar las carpetas creadas?
 Sí No
- ¿He establecido un sistema de versionado para los conjuntos de datos, facilitando su seguimiento y actualización?
 Sí No
- ¿He descrito y documentado de manera detallada los datos, incluyendo su naturaleza, formato y estructura?
 Sí No
- ¿He diseñado una estrategia sólida para el almacenamiento y la preservación a largo plazo de los datos?
 Sí No
- ¿He realizado copias de seguridad periódicas de los datos, garantizando su seguridad y disponibilidad en caso de pérdida o daño?
 Sí No

- ¿He considerado los aspectos éticos y legales relacionados con la recolección, uso y divulgación de los datos?
 Sí No

- ¿He tenido en cuenta posibles formas de reutilización de los datos a la hora de seleccionar los que se van a preservar?
 Sí No

- ¿He identificado el criterio/s a seguir para decidir el repositorio más adecuado para compartir los datos?
 Sí No

- ¿He identificado los recursos técnicos, financieros y humanos necesarios para llevar a cabo la gestión de los datos de manera efectiva?
 Sí No

- ¿He verificado y cumplido con los requisitos específicos de la agencia financiadora en materia de gestión de datos, en caso de recibir financiación para la investigación?
 Sí No

- ¿He revisado y completado todos los metadatos y campos requeridos en la plantilla DMPOnline?
 Sí No

- ¿He realizado una revisión final del Plan de Gestión de Datos para garantizar su coherencia y adecuación a los objetivos de la investigación?
 Sí No

- ¿He incluido metadatos detallados para cada conjunto de datos, como descripción, autor, fecha de creación, etc.?
 Sí No

LA ALFABETIZACIÓN EN DATOS EN EL CONTEXTO
UNIVERSITARIO DE LAS CIENCIAS SOCIALES
(ALFADATAUniv_CCSS)

REUTILIZACIÓN Y CITACIÓN DE DATOS

Ana Iglesias Rodríguez
Yolanda Martín González



REUTILIZACIÓN Y CITACIÓN DE DATOS

ÍNDICE

1. Qué es la reutilización de información
2. Normativa europea y estatal sobre reutilización
3. Datos abiertos y reutilización de la información
 - 3.1. Datos abiertos y reutilización en el sector público
 - 3.2. Datos abiertos de investigación y su reutilización
 - 3.3. Citación de los datos

Referencias

1. QUÉ ES LA REUTILIZACIÓN DE INFORMACIÓN

Tomando como referencia el estudio efectuado por van de Sandt et al. (2019), cabe destacar que, como dicen estos autores, en el ámbito de la investigación, la reutilización de la información y de los datos se está convirtiendo en un concepto esencial que está cobrando, cada vez, en mayor medida, más atención e importancia. Este término hace alusión a la práctica de utilizar los datos y la información generados en una investigación previa para llevar a cabo nuevas investigaciones o para otros fines científicos. Es decir, implica aprovechar los recursos ya existentes para obtener nuevos conocimientos o para respaldar y validar estudios futuros.

En los últimos años, se está abogando por la Ciencia Abierta, donde diversos actores y disciplinas se están sumando a la idea de compartir los datos de investigación de manera más accesible y transparente. Este movimiento no sólo promueve la difusión del conocimiento, sino que también facilita la reutilización de los datos, lo que puede conllevar importantes beneficios tanto en términos de ahorro de costes económicos y esfuerzos como en la validación y reproducción de hallazgos científicos.

En este contexto, es fundamental entender que la reutilización de la información y de los datos de investigación además de implicar un cambio en la forma en que se manejan y comparten los datos científicos, también influyen en la manera en que se planifica y gestiona la investigación en su conjunto. Este enfoque, cada vez más adoptado por agencias de financiamiento y organizaciones expertas, resalta la importancia de desarrollar planes de gestión de datos que garanticen la disponibilidad y accesibilidad de los datos de investigación para su posterior reutilización.

Si bien es cierto que “la reutilización es un concepto utilizado para describir fenómenos del mundo real” (van de Sandt et al., 2019, p.2), parece que aún no existe un consenso entre los autores a la hora de ofrecer una definición estándar del término que sea comprensible para el grueso de la población.

Sin embargo, el concepto de “reutilización de la información” no surge inicialmente en el ámbito investigador sino en el contexto del sector público, a finales de los años 90 del siglo XX. Este movimiento viene a destacar el valor estratégico que poseen los datos generados, custodiados y/o conservados por las administraciones públicas, y cómo a partir de ellos, se pueden generar nuevos productos, servicios y/o aplicaciones que conlleven beneficios sociales y económicos en la ciudadanía. Es por ello que nosotros tomaremos como referencia la definición establecida en el Artículo 64.3, del Real Decreto-Ley 24/2021, de 2 de noviembre (p. 133295), que indica que **se entiende por reutilización** el uso por personas físicas o jurídicas de documentos elaborados o custodiados por la Administración General del Estado, las Administraciones de las Comunidades Autónomas y las Entidades que integran la Administración Local; los organismos y entidades del sector público institucional creados para satisfacer necesidades de interés general, que no tengan carácter industrial o mercantil; y las sociedades mercantiles pertenecientes al sector público institucional, con fines

comerciales o no comerciales, siempre que dicho uso no constituya una actividad de servicio público o de prestar servicios de interés general distintos para el que se produjeron.

De igual forma, el Artículo 1.5, del Capítulo I “Disposiciones generales”, del RD 1495/2011, de 24 de octubre, por el que se desarrolla la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público, para el ámbito del sector público estatal, decreta que “se entiende por «agente reutilizador» toda persona, física o jurídica que reutilice información del sector público, ya sea para fines comerciales o no comerciales, siempre que dicho uso no constituya una actividad administrativa pública” (p.5).

2. NORMATIVA EUROPEA Y ESTATAL SOBRE REUTILIZACIÓN

En el momento actual, la información que es de dominio público se ha convertido en un recurso valioso y necesario para impulsar lo que se conoce como Economía de los datos. Al reutilizar esta información producida por entidades gubernamentales y al hacerla accesible en gran medida se logran importantes ventajas, tanto económicas como sociales, contribuyendo este proceso a la transparencia de las acciones llevadas a cabo por las administraciones públicas. Toda esta información generada, recopilada y distribuida por las entidades del sector público puede, a su vez, ser reutilizada por empresas y por ciudadanos, favoreciendo con ello, por un lado, la transparencia, la eficiencia, la participación ciudadana y el desarrollo económico; y, otorgando, por otro lado, un mayor valor a los datos reutilizables provenientes del sector público (conocidos como datos abiertos u Open data), dando lugar este proceso a la creación de nuevos negocios, aplicaciones, servicios y productos.

En el año 1998, la Comisión Europea publicó el Libro Verde sobre la Información del Sector Público en la Sociedad de la Información, y en él se puso en valor el papel estratégico de los datos que están en poder de las instituciones públicas (gobiernos centrales, gobiernos de las administraciones regionales, de los ayuntamientos, etc.).

Las propuestas recogidas en este informe han dado lugar a la promulgación de normativa específica, tanto de carácter europeo como estatal, en torno a la reutilización de la información, destacando especialmente la que se detalla a continuación (para un mayor desarrollo de la misma, véase Anexo I).

La **Directiva 2003/98/CE del Parlamento Europeo y del Consejo, del 17 de noviembre de 2003**, sobre reutilización de la información del sector público, surge con el propósito de establecer un marco común para la reutilización de la información generada y mantenida por organismos del sector público en los Estados miembros de la UE. Su objetivo es promover la transparencia, la eficiencia y la innovación al permitir que los ciudadanos, las empresas y otras organizaciones accedan y utilicen la información del sector público para diversos fines comerciales y no comerciales. Además, busca eliminar barreras y promover la competencia en el mercado de datos, estimulando, de este modo, el crecimiento económico y facilitando la participación democrática en la sociedad. Esta Directiva se transpuso al ordenamiento jurídico español con la **Ley 37/2007, de 16 de noviembre**, sobre reutilización de la información del sector público; y su desarrollo reglamentario, el **Real Decreto 1495/2011, de 24 de octubre**, por el que se desarrolla la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público, para el ámbito del sector público estatal, siendo su objetivo principal, por un lado, establecer un marco legal que permitiera el acceso y la reutilización de la información generada por las entidades del sector público en España; y, por otro lado, promover la transparencia, la eficiencia y la innovación al facilitar el acceso a la información del sector público para diversos fines comerciales y no comerciales. Asimismo, con su desarrollo reglamentario, se buscaba proporcionar las directrices y los procedimientos específicos para la aplicación de la ley en el ámbito del sector público estatal. Este desarrollo reglamentario establece normas detalladas sobre la reutilización de la información del sector público, incluyendo aspectos como los requisitos para la solicitud de acceso a la información, las condiciones de reutilización, las tarifas aplicables, y los procedimientos para resolver conflictos y reclamaciones relacionadas con la reutilización de la información.

Posteriormente, surge la **Ley 18/2015, de 9 de julio**, por la que se modifica la Ley 37/2007, de 16 de noviembre, que transpone la **Directiva 2013/37/UE**, que modifica la Directiva 2003/98/CE, actualizando el régimen de reutilización de la información, cuyo propósito principal fue actualizar y mejorar el régimen de reutilización de la información del sector público en España. Esta modificación tuvo como fin adaptar

la legislación española a los cambios y avances en el ámbito de la gestión de datos y la reutilización de la información, así como garantizar su coherencia con las normativas europeas pertinentes. Al actualizar el régimen de reutilización de la información, la Ley 18/2015 buscaba promover la transparencia, la eficiencia y la innovación al facilitar el acceso y la utilización de la información del sector público para diversos fines comerciales y no comerciales. Además, la modificación tenía como objetivo reforzar la protección de los derechos de privacidad y datos personales de los ciudadanos, así como establecer mecanismos más efectivos para resolver disputas y conflictos relacionados con la reutilización de la información.

Más recientemente, la **Directiva (UE) 2019/1024** del Parlamento Europeo y del Consejo, de 20 de junio de 2019, relativa a los datos abiertos y la reutilización de la información del sector público, establece un marco normativo para promover la apertura y la reutilización de datos del sector público en toda la Unión Europea. Esta directiva trata de fomentar la transparencia, la innovación y el crecimiento económico al facilitar el acceso y la reutilización de la información generada por las entidades del sector público para diversos fines comerciales y no comerciales. Su transposición al ordenamiento jurídico español mediante el **Real Decreto-Ley 24/2021, de 2 de noviembre**, posibilita adecuar la legislación española a los requisitos y estándares establecidos en la Directiva europea. Esto incluye la promoción de la apertura y la reutilización de datos del sector público en España, así como el establecimiento de mecanismos para garantizar la interoperabilidad, la accesibilidad y la disponibilidad de dichos datos. Además, la transposición, en coherencia con los objetivos de la Directiva europea, tenía como objetivo mejorar la coordinación y la colaboración entre las distintas administraciones públicas en el ámbito de la gestión y el intercambio de datos abiertos.

Finalmente, cabe mencionar la **Ley de Datos** (Data Act o DA) que entró en vigor el 11 de enero de 2024. Esta Ley proporciona un marco normativo sólido y coherente para la gestión de datos en la Unión Europea, con el objetivo de promover la innovación, la competencia y la protección de los derechos digitales en el mercado digital europeo. Además, es un pilar fundamental de la Estrategia Europea de Datos, que complementa otras normativas como son el **Reglamento (UE) 2023/2854** sobre normas armonizadas para un acceso justo a los datos y su utilización (DA); la **Directiva (UE) 2019/1014** relativa a los datos abiertos y la reutilización de la información del sector público; y el **Reglamento (UE) 2022/868** relativo a la gobernanza europea de datos (DGA).

A modo de resumen, en la Figura 1, se muestran los contenidos clave de la Ley 37/2007 sobre reutilización de la información del sector público, tras la incorporación de las previsiones de la Directiva (UE) 2019/1024 relativa a los datos abiertos y la reutilización de la información del sector público. Y, a nivel europeo, en la Figura 2, se recogen los aspectos más relevantes de la Ley de Gobernanza de Datos (Comisión Europea, 2020).

Figura 1. Regulación relativa a los datos abiertos y la reutilización de la información del sector público en España.

datos.gob.es

REGULACIÓN RELATIVA A LOS DATOS ABIERTOS Y LA REUTILIZACIÓN DE LA INFORMACIÓN DEL SECTOR PÚBLICO EN ESPAÑA

Contenidos clave de la Ley 37/2007 sobre reutilización de la información del sector público, tras la incorporación de las previsiones de la Directiva (UE) 2019/700 relativa a los datos abiertos y la reutilización de la información del sector público.

OBJETIVO: Autorización generalizada de reutilización de los datos del sector público aplicando el principio de datos abiertos desde el diseño y por defecto.

ÁMBITO DE APLICACIÓN:

- Datos producidos por la Administración General del Estado, las Comunidades Autónomas y entidades que integran la Administración Local.
- Datos generados por organismos y entidades del sector público institucional, incluidas bibliotecas, museos y archivos.
- Datos y servicios referidos en la Directiva INSPIRE.
- Datos de investigación financiada con fondos públicos.
- Datos de determinadas sociedades mercantiles.

PRINCIPALES OBLIGACIONES DE LA LEY 37/2007 DE 16 DE NOVIEMBRE (TEXTO CONSOLIDADO)

01 OFRECER LOS DATOS:

- Formatos abiertos y legibles por máquinas
- Con el máximo nivel de precisión y desagregación posible
- Accesibles
- Fáciles de localizar y reutilizables
- Accesibles mediante descarga masiva
- A través de interfaces de programación de aplicaciones (API)
- En tiempo real para los datos dinámicos
- Publicados conjuntamente con sus metadatos

02 IMPULSAR LA APERTURA DE DATOS DE ALTO VALOR:

Focalizar esfuerzos en abrir datos de alto valor establecidos a nivel europeo y a nivel nacional.

¿Qué son los datos de alto valor (HVD)? Son aquellos... cuyo uso está asociado a considerables beneficios para la sociedad, el medio ambiente y la economía, en particular debido a su idoneidad para la creación de servicios de valor añadido, aplicaciones y puestos de trabajo potencialmente útiles para un alto número de beneficiarios, especialmente del ámbito de las PYMES.

Requisitos legales y técnicos obligatorios de los HVD:

- Uso de licencia de datos abiertos (CC BY) o su equivalente
- Disponibilidad de documentación pública sobre los datos
- Formatos legibles por máquinas
- Disponibilidad vía APIs y descargables de forma masiva
- Gratuidad
- Con una descripción exhaustiva mediante metadatos

03 DISPONER DE UN CATALOGO DE DATOS: Todas las AADP disponen de catálogos propios que interoperarán con el Catálogo Nacional (datos.gob.es) siguiendo la NIT-IBSD.

04 CONTRIBUIR AL CATALOGO NACIONAL: Las AADP deben contribuir a su enriquecimiento mediante procesos de federación con datos.gob.es.

05 OFRECER UN ACCESO GRATUITO: Excepcionalmente pueden aplicarse costes marginales de reproducción, puesta a disposición, difusión o anonimización.

06 FOMENTAR LA REUTILIZACIÓN: Esta no estará sujeta a condiciones a menos que estas sean objetivas, proporcionales, no discriminatorias y estén justificadas por un objetivo de interés público. Se pueden fijar utilizando licencias tipo.

07 PONER EN MARCHA LA UNIDAD RESPONSABLE DE INFORMACIÓN: Todas las AADP deben disponer de una unidad encargada de coordinar las actividades, resolver solicitudes, promover que los datos se ofrezcan con la calidad precisa y llevar a cabo las actuaciones que fomenten su uso.

red.es | Oficina del Dato | aporta

Fuente: Ministerio para la Transformación Digital y de la función pública (s.f). <https://datos.gob.es/es/blog/las-claves-de-la-ley-sobre-reutilizacion-de-la-informacion-del-sector-publico-en-espana>

Figura 2. Ley de Datos (Data Act o DA).

LEY DE DATOS (DATA ACT O DA)

¿QUÉ ES? Es una regulación de la Unión Europea para facilitar la accesibilidad de los datos de forma horizontal. Se define estableciendo principios y directrices para todos los sectores, haciendo hincapié en el acceso justo a los datos. No derechos de los usuarios y la protección de los datos personales.

Reglamento (UE) 2023/1541 sobre normas armonizadas para un acceso justo a los datos y su utilización (DA)

Se centra en:

- Impulsar la puesta a disposición de los datos generados por productos y servicios conectados.
- Promover esfuerzos de interoperabilidad para la utilización de los datos, no siendo aplicable a los datos de tráfico de datos.
- Dejar la puesta a disposición de datos a favor de entidades públicas en situaciones excepcionales.

PRINCIPALES OBJETIVOS DE LA LEY DE DATOS

- Aumentar la transparencia y equidad jurídica respecto al acceso y uso de los datos, en beneficio de la economía.
- Promover el intercambio equitativo de datos y permitir a las empresas competir con igualdad en el mercado.
- Establecer condiciones para que las entidades privadas faciliten datos a las empresas públicas en situaciones excepcionales.
- Promover un marco para la interoperabilidad eficiente de los datos que impulse el intercambio entre empresas y sectores.
- Establecer garantías mínimas que permitan que los usuarios de servicios de tratamiento de datos los accedan aquellos que realizan cualquier operación con los datos como su recogida, estandarización u otro uso poder cambiar de proveedor.

LAS CLAVES DE LA DA

Disponibilidad de datos just

- Se establece un principio de acceso a los datos de forma equitativa, debiendo darse a conocer a los usuarios los datos que pueden acceder, utilizar y compartir con terceros los datos generados.
- Se detallan que las empresas informen sobre qué datos se van a generar, cómo y quién los va utilizar.
- Se eliminan las obligaciones de un cambio efectivo de proveedor.
- Se concretan los aspectos en que una cláusula es abusiva.
- Se obliga a que la Comisión desarrolle cláusulas contractuales modelo para incluir y eliminar y registrar contratos equivalentes de intercambio de datos.

Interoperabilidad

- Se establece que los estándares técnicos y jurídicos para el intercambio de datos.
- Se establecen condiciones específicas para los estándares técnicos, contemplándose un sistema de declaración de conformidad europeo y criterios de normalización.

Datos a favor de AADP

- Se define la posibilidad de que los organismos públicos puedan facilitar datos relevantes para responder con rapidez y seguridad ante emergencias o situaciones de interés público que una carga mínima para las empresas.

¿A QUIÉN AFECTA? Al igual que el Reglamento, su aplicación directa es obligatoria en toda la Unión Europea.

¿CUÁNDO ENTRÓ EN VIGOR? El 11 de enero de 2024. En el ámbito europeo entró en vigor el 11 de febrero de 2024.

¿QUIERES SABER MÁS SOBRE LA LEY DE DATOS? Haz clic sobre los siguientes enlaces:

- La Ley de Datos.
- La regulación europea en materia de datos frente al resto de una aplicación armonizada que incluye la protección de datos.
- La Ley de Datos, una norma pionera en el marco de la Activa Digital de Datos.

red.es | aporta | datos.gob.es

Fuente: Ministerio para la Transformación Digital y de la función pública (s.f). https://datos.gob.es/sites/default/files/blog/file/ley-datos-es_1.pdf

3. DATOS ABIERTOS Y REUTILIZACIÓN DE LA INFORMACIÓN

Como ya se ha explicado anteriormente, un dato es el elemento primario de la información y puede hallarse conformado por números, letras, símbolos, signos, etc., que reunidos adquieren un significado. Por tanto, un dato sólo o aislado no tiene significado por lo que es necesario procesarlo y contextualizarlo. Asimismo, los datos son objetivos, mediante ellos no se expresan juicios de valor, sino que son un recorte de la realidad observable. Así, los datos personales (un nombre, una foto, su DNI, una dirección de correo, etc.) sirven para la identificación de una persona. Por último, y dado que muchos de ellos surgen de la observación y/o la experiencia, los datos son, además, un punto de partida para generar conocimiento y, en este sentido, poseen un gran valor potencial. Mayer-Schönberger y Cukier (2013) afirman que los datos son a la sociedad de la información lo que el combustible a la economía industrial: el recurso esencial que alimenta las innovaciones que usa la gente.

A la hora de hablar de reutilización de la información o de los datos se ha de considerar previamente lo que hemos denominado el ecosistema “Open”. Este ecosistema (Figura 3) se encuentra conformado, a su vez, por varios movimientos o paradigmas interrelacionados y que son: *Open Science* o Ciencia Abierta; *Open Access* o acceso abierto; y *Open Data* o datos abiertos.

Figura 3.
Ecosistema Open.



Fuente: Elaboración propia

La ciencia abierta constituye “un cambio de paradigma en la manera de hacer ciencia que supone realizar con una visión “abierta” (open) todas las etapas o fases de la investigación científica (diseño, recolección de datos, revisión, publicación, etc.)” (Abadal y Anglada, 2020, p.1), de ahí que comprenda, además de otras acciones, por un lado, el acceso abierto, y, por otro lado, los datos abiertos.

En la Declaración de Budapest (2001) sobre acceso abierto (BOAI, 2002), se definió éste como,

(...) la disponibilidad gratuita en Internet, permitiendo a cualquier usuario leer, descargar, copiar, distribuir, imprimir, buscar o usarla con cualquier propósito legal, sin ninguna barrera financiera, legal o técnica, fuera de las que son inseparables del hecho mismo de acceder a Internet. La única limitación en cuanto a reproducción y distribución y el único rol del copyright en este dominio deberá ser dar a los autores el control sobre la integridad de sus trabajos y el derecho de ser adecuadamente reconocidos y citados.

Por su parte, la filosofía de los datos abiertos es una práctica que persigue que determinados datos estén disponibles de forma libre para todo el mundo, sin restricciones de copyright, registro previo, patentes u otros mecanismos de control. Estos datos deben publicarse en bruto (sin procesar), bien estructurados y en formatos conocidos de modo que cualquier persona pueda utilizarlos, redistribuirlos y reutilizarlos creando nuevos productos, servicios, aplicaciones y conocimiento derivados de ellos.

En este contexto cabe distinguir entre los datos abiertos del sector público (open government data) y los datos científicos abiertos (open research data) (Figura 4). En el primer caso, los conjuntos de datos son creados, recopilados y/o conservados por las administraciones públicas en el ejercicio de sus funciones; en el segundo caso, los datos son recogidos y/o creados en el contexto de una investigación. Cada grupo presenta unos rasgos concretos y son recuperables mediante recursos diferentes, aunque, a la hora de ser liberados, todos ellos han de cumplir los mismos requisitos en

términos de formatos en los que se disponen y de licencias bajo las cuales se comparten.

Figura 4.
Datos abiertos del sector público y de investigación.



Fuente: Elaboración propia

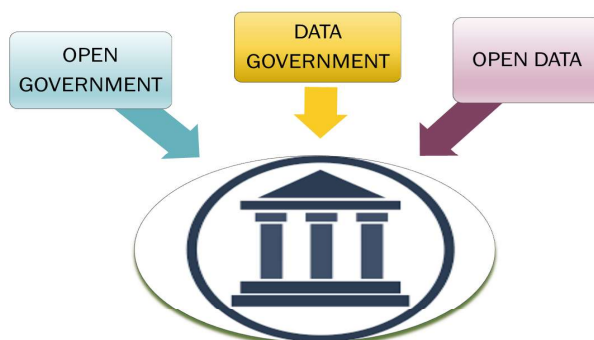
En todos los casos señalados, el fin principal perseguido con la apertura de los datos es su reutilización, ya que esta conlleva una serie de beneficios a nivel económico (aumenta el volumen de negocio de las empresas, la empleabilidad del personal cualificado, etc.), social (procura un mayor acceso a la información y al conocimiento) y para las administraciones (favorece la transparencia, la colaboración y la participación ciudadana en el diseño y evaluación de las acciones y políticas públicas).

La reutilización permite añadir un valor a los datos liberados con el fin de generar nuevas aplicaciones, productos y servicios en base a ellos. Por ejemplo, en el ámbito de la información cartográfica, muchos de los sistemas de navegación existentes se basan en el uso de datos abiertos; y en el ámbito delincriminal, los denominados mapas del delito (crime map) son producto de la reutilización de datos liberados por las fuerzas y cuerpos de seguridad que son quienes los recopilan y registran.

3.1. Datos abiertos en el sector público

La filosofía *Open Data* o de datos abiertos defiende la apertura de datos en poder del sector público con el fin de que estos sean utilizados, reutilizados y redistribuidos. El objetivo principal no es, por tanto, la liberación de los datos *per se* sino facilitar su reutilización posterior mediante la aportación de un valor (Figura 5).

Figura 5.
Movimientos que confluyen en el Open Government Data.



Fuente: Elaboración propia

Un dato es una representación de hechos, conceptos o instrucciones de un modo formalizado, y adecuado para la comunicación, interpretación o procesamiento por medios automáticos o humanos (datos.gob.es, s.f). En el ámbito del sector público se entiende como tal, por ejemplo, el número de estudiantes matriculados en estudios oficiales de grado en las universidades públicas de Castilla y León.

Por su parte, un dataset o conjunto de datos son datos relacionados, convenientemente estructurados y organizados, de forma que puedan ser tratados (procesados) apropiadamente para obtener información (datos.gob.es, s.f.). Siguiendo con el ejemplo anterior, el historial de matrículas en las universidades públicas de Castilla y León.

La apertura de los datos debe realizarse bajo las siguientes premisas:

- Todos los datos públicos deben estar disponibles, siempre que no estén sujetos a restricciones de privacidad o seguridad.
- Los datos deben provenir directamente de la fuente, sin modificación o agregación alguna.
- La publicación debe realizarse a tiempo para preservar el valor de la información, así como mantenerse actualizados.
- Los datos deben estar disponibles de tal manera que cualquiera sin discriminación pueda usarlos, modificarlos y compartirlos para cualquier fin.
- Los datos deben presentar una estructura mínima que permita su tratamiento automatizado.
- La publicación debe estar libre de discriminación y disponible para cualquiera sin necesidad de registro previo.
- Los formatos de publicación deben ser libres, abiertos, accesibles y reutilizables. Ninguna entidad tiene control exclusivo sobre ellos.
- La información debe ser de uso libre, es decir, no estar sujeta a derechos de autor, patentes, marcas ni a regulaciones de secreto industrial o comercial. Tampoco han de estar sujetas a las restricciones de privacidad, seguridad o privilegio reguladas por las normas.

Los principales actores en el ámbito de la reutilización de datos del sector público (Figura 6) son: (1) las administraciones públicas que son quienes generan, recopilan y/o conservan los datos; (2) los reutilizadores que conforman el denominado sector infomediario que surge a la luz de este movimiento y que son quienes añaden un valor a los datos liberados y generan productos, aplicaciones y servicios nuevos que aportan innovación, crecimiento económico y riqueza; (3) los ciudadanos, profesionales y/o usuarios que se benefician del producto generado a partir de la reutilización de información.

Figura 6.
Principales actores en el proceso de reutilización de información del sector público.



La localización y recuperación de los conjuntos de datos liberados por las administraciones públicas se realiza mediante el catálogo de datos abiertos comprendido, a su vez, en el portal de datos abiertos.

El portal provee servicios de descubrimiento avanzados que complementan la experiencia de búsqueda en el catálogo como, por ejemplo, información y acceso a los productos y/o aplicaciones resultado de la reutilización de los datos. Por su parte, el catálogo ofrece una colección de metadatos relevantes -obligatorios y complementarios- sobre los conjuntos de datos, así como una serie de opciones que permiten la delimitación de los resultados obtenidos.

En España contamos con un catálogo de datos abiertos nacional, <https://datos.gob.es/>, así como con catálogos similares a nivel local y autonómico a través de los cuales se permite la consulta de los conjuntos de datos liberados por las distintas administraciones.

Además de estos recursos, contamos con buscadores generales de datos como son Google Data Search o Google Public Data, además de otros de carácter especializado que proporcionan algunas instituciones que cuentan con su propio catálogo de datos abiertos. Este es el caso de la Biblioteca Nacional de España (<https://datos.bne.es>), la Organización de Naciones Unidas (Undata, <http://data.un.org>) o HealthData (<https://www.healthdata29.org/platform/datasets>), por poner algunos ejemplos.

3.2. Datos abiertos de investigación

El concepto *open research data* comprende los datos generados y/o recopilados durante una investigación, disponibles en abierto y que pueden ser utilizados, reutilizados y redistribuidos libremente por cualquier persona, siempre que se reconozca la autoría.

La apertura y compartición de datos científicos que surgen, en su mayoría, de la observación y/o la experiencia, se ha de realizar siguiendo los denominados Principios FAIR, que aseguran que estos sean:

- Localizables: los datos de investigación han de poder encontrarse fácilmente, tanto por las personas como por los medios tecnológicos.
- Accesibles: los datos y metadatos deben poder ser recuperados mediante un identificador, utilizando para ello un protocolo de comunicación abierto y estandarizado.
- Interoperables: los datos deben poder usarse y combinarse con otros datos o herramientas.
- Reutilizables: los datos deben poder reutilizarse y, para ello, es necesario publicarlos bajo licencias de reutilización claras y accesibles.

El acceso y recuperación de estos datos se realiza mediante lo que se suele denominar el repositorio de datos de investigación, aunque nosotros preferimos denominarlo “repositorio de ciencia abierta” puesto que, además de datos, permite la consulta de resultados de investigaciones que se hayan desarrollado.

Estos recursos están aún bastante lejos de ofrecer las prestaciones que ofrecen, por ejemplo, las bases de datos científicas, por lo que resulta un tanto rudimentario el servicio de búsqueda y recuperación que proporcionan.

Por el momento, la reutilización de datos de investigación no es una práctica habitual entre los investigadores, pero sí una tendencia al alza en determinadas áreas de conocimiento como el tecnológico o el de ciencias de la salud (Tena-Parera, 2019), tal como se puso de manifiesto durante la pandemia provocada por el COVID-19. En este sentido, se están fomentando prácticas conocidas como *data sharing* o de compartición de datos que favorecen el intercambio de datos surgidos en el marco de una investigación de campo, o de la recolección de fuentes primarias (entrevistas, investigación participante...) o secundarias (análisis de contenidos...) con otros investigadores para que, a partir de esos datos, sea posible replicar la investigación o continuar investigando otros objetos de estudio. Así pues, nuestros datos “en bruto” estarán disponibles para ser reutilizados por otros investigadores.

Además de por la normativa vigente y los planes de acción europeos y españoles en materia de reutilización de la información, en el ámbito investigador se promueve la compartición de datos científicos, en abierto, por medio de los mandatos y requerimientos de las distintas agencias financiadoras, las políticas editoriales de las revistas científicas, las estrategias e iniciativas en materia de ciencia abierta promovidas por las

universidades y centros de investigación, así como por distintas sociedades científicas y profesionales que apoyan la apertura de este tipo de datos. Todas ellas abogan por el depósito de los datos generados y/o recopilados durante la investigación con el fin de que sean reutilizados por otros investigadores o interesados.

La ya mencionada Directiva (UE) 2019/1024 del Parlamento Europeo y del Consejo de 20 de junio de 2019, relativa a los datos abiertos y la reutilización de la información del sector público, promueve la compartición de datos con la finalidad de que sean reaprovechados por otros. En concreto, estipula:

El volumen de datos de investigación generados está creciendo exponencialmente y pueden ser reutilizados por otros usuarios ajenos a la comunidad científica. Para poder afrontar con eficacia y de manera global los recientes desafíos sociales, la posibilidad de acceder a datos de diferentes fuentes y de todo tipo de sectores y disciplinas, de combinarlos y reutilizarlos ha pasado a ser algo crucial y urgente. (Consideración 27)

Por su parte, el Real Decreto-ley 24/2021, de 2 de noviembre de 2021 que modifica la Ley 37/2007 sobre reutilización de la información del sector público, que traspone la norma europea al ordenamiento jurídico español, en su artículo 3 bis, referido a los datos de investigación, contempla que,

las entidades incluidas en el ámbito de aplicación del artículo 2 de la presente Ley y que realicen actividades de investigación o financien la investigación adoptarán medidas para apoyar que los datos de investigaciones financiadas públicamente sean plenamente reutilizables, interoperables y de acceso abierto.

Al abogar por la apertura de los datos, las políticas públicas transmiten la idea de que los datos son una entidad con un alto potencial de reutilización. Para justificar la compartición de datos, proponen dos categorías de argumentos.

La primera categoría se basa en una concepción acumulativa del conocimiento: la puesta en común de datos contribuye a una mayor transparencia en la investigación, a la reutilización de los resultados por las comunidades científica y docente, pero también por las empresas y los ciudadanos, y al desarrollo de nuevas formas de investigación basadas en el uso de datos masivos. La segunda categoría de argumentos es económica: la puesta en común de datos fomenta la innovación y el crecimiento económico, y contribuye a rentabilizar las sumas comprometidas por los organismos públicos para financiar la investigación. (Rebouillat, 2021, p. 7)

3.3. Citación de los datos

En los apartados anteriores hemos incidido en el potencial que tienen los datos abiertos y cómo estos pueden ser objeto de reutilización, tanto en el sector público como en el ámbito científico. Pues bien, cuando se reutilizan conjuntos de datos, al igual que ocurre cuando referenciamos otros recursos de información, estos deben ser citados con el fin de que se reconozca su autoría y se favorezca la transparencia de la investigación científica.

Para citar datos se ha de seguir una norma que permita identificarlos y recuperarlos de manera adecuada, por lo que irán acompañados de un identificador único y persistente (DOI, Handle, etc.), así como de un identificador de autor (ORCID). En la sección de referencias bibliográficas se citará, de manera independiente, cada conjunto de datos. Además, siempre que sea posible, se enlazarán los datos con los resultados de investigación correspondientes.

A la hora de elaborar la correspondiente referencia bibliográfica deberán combinarse distintos tipos de elementos, unos se incluirán de manera obligatoria y otros es recomendable su incorporación (CRUE-Rebiun, 2018). Entre los primeros se encuentran el autor, la fecha, el título, el identificador persistente, el tipo de recurso y la versión. Entre los recomendados se incluyen, el identificador de autor, el nombre del repositorio de datos, la publicación, el productor, el ámbito geográfico y el ámbito cronológico.

La referencia bibliográfica se redactará siguiendo lo establecido en relación con los datos en cualquier norma (APA, Chicago, etc.) o según lo que propongan los principales repositorios de datos como son Dataverse o Dryad, entre otros.

A continuación, se muestra un ejemplo de cita redactada atendiendo a la norma APA:

Martín González, Y., Iglesias Rodríguez, A. (2021). *Alfabetización en datos en bibliotecas-CRAI españolas curso académico 2020-2021* [Dataset]. Versión de 12 de junio de 2021. Zenodo. <https://doi.org/10.5281/zenodo.4434086>

Referencias

Abadal Falgueras, E. y Anglada Ferrer, LL. M. (2020). Ciencia Abierta: cómo han evolucionado la denominación y el concepto. *Anales de Documentación* 23(1). <http://dx.doi.org/10.6018/analesdoc.378171>

Arias Pou, M. (2008). Reutilización de la información y protección de datos personales. *Revista de Estudios Locales. Cunal*, 107, 28-43. <https://www.navarra.es/NR/rdonlyres/DCF7A548-3551-481C-8CE4-6CB-D251352BC/162691/5ReutilizacionInformacion1.pdf>

datos.gob.es. (10 de marzo de 2023). Las claves de la Ley sobre reutilización de la información del sector público en España. *datos.gob.es reutiliza la información pública*. Recuperado el 18 de abril de 2024 de <https://datos.gob.es/es/blog/las-claves-de-la-ley-sobre-reutilizacion-de-la-informacion-del-sector-publico-en-espana>

Declaración de Budapest sobre el Acceso Abierto (BOAI, 2001). Recuperado el 18 de abril de 2024 de <https://www.budapestopenaccessinitiative.org/read/spanish-translation/>

Diario Oficial de la Unión Europea. (2019). *Directiva (UE) 2019/1024 del Parlamento Europeo y del Consejo de 20 de junio de 2019 relativa a los datos abiertos y la reutilización de la información del sector público (versión refundida)*. L 172/56. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=uriserv:OJ.L_2019.172.01.0056.01.SPA

Directiva 2013/37/UE del Parlamento Europeo y del Consejo, de 26 de junio de 2013, por la que se modifica la Directiva 2003/98/CE relativa a la reutilización de la información del sector público. *DOUE*, núm. 175, de 27 de junio de 2013, páginas 1 a 8 (8 págs.). <https://www.boe.es/buscar/doc.php?id=DOUE-L-2013-81251>

Directiva 2014/24/UE del Parlamento Europeo y del Consejo, de 26 de febrero de 2014, sobre contratación pública y por la que se deroga la Directiva 2004/18/CE. *DOUE*, núm. 94, de 28 de marzo de 2014, páginas 65 a 242 (178 págs.). <https://www.boe.es/buscar/doc.php?id=DOUE-L-2014-80598>

Ley 18/2015, de 9 de julio, por la que se modifica la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público. *BOE núm. 164*, de 10 de julio de 2015, páginas 57436 a 57450 (15 págs.). <https://www.boe.es/eli/es/l/2015/07/09/18>

Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público. *BOE núm. 276*, de 17 de noviembre de 2007. <https://www.boe.es/eli/es/l/2007/11/16/37/con>

Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público. *BOE*, núm. 236, de 02 de octubre de 2015. <https://www.boe.es/eli/es/l/2015/10/01/40/con>

Ley 9/2017, de 8 de noviembre, de Contratos del Sector Público, por la que se transponen al ordenamiento jurídico español las Directivas del Parlamento Europeo y del Consejo 2014/23/UE y 2014/24/UE, de 26 de febrero de 2014. *BOE núm. 272*, de 09 de noviembre de 2017. <https://www.boe.es/eli/es/l/2017/11/08/9/con>

Mayer-Schönberger, V.; Cukier, K. (2013). *Big data. La revolución de los datos masivos*. Turner Noema.

Real Decreto 1495/2011, de 24 de octubre, por el que se desarrolla la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público, para el ámbito del sector público estatal. *BOE núm. 269*, de 08 de noviembre de 2011. <https://www.boe.es/eli/es/rd/2011/10/24/1495/con>

Real Decreto-ley 24/2021, de 2 de noviembre, de transposición de directivas de la Unión Europea en las materias de bonos garantizados, distribución transfronteriza de organismos de inversión colectiva, datos abiertos y reutilización de la información del sector público, ejercicio de derechos de autor y derechos afi-

nes aplicables a determinadas transmisiones en línea y a las retransmisiones de programas de radio y televisión, exenciones temporales a determinadas importaciones y suministros, de personas consumidoras y para la promoción de vehículos de transporte por carretera limpios y energéticamente eficientes. *BOE* núm. 263, de 3 de noviembre de 2021, páginas 133204 a 133364 (161 págs.). <https://www.boe.es/eli/es/rdl/2021/11/02/24>

Reglamento (UE) 2022/868 del Parlamento Europeo y del Consejo de 30 de mayo de 2022 relativo a la gobernanza europea de datos y por el que se modifica el Reglamento (UE) 2018/1724 (Reglamento de Gobernanza de Datos). *DOUE*, núm. 152, de 3 de junio de 2022, páginas 1 a 44 (44 págs.). <https://www.boe.es/buscar/doc.php?id=DOUE-L-2022-80835>

Reglamento (UE) 2023/2854 del Parlamento Europeo y del Consejo, de 13 de diciembre de 2023, sobre normas armonizadas para un acceso justo a los datos y su utilización, y por el que se modifican el Reglamento (UE) 2017/2394 y la Directiva (UE) 2020/1828 (Reglamento de Datos). *DOUE*, núm. 2854, de 22 de diciembre de 2023, páginas 1 a 71 (71 págs.). <https://www.boe.es/buscar/doc.php?id=DOUE-L-2023-81895>

Tena-Parera, D. (2019). El valor de los datos. *Questiones publicitarias*, 24, 71-76. <https://doi.org/10.5565/rev/qp.338>

Valero, J. (13 de diciembre de 2021). Convalidación del Real Decreto-ley sobre reutilización de la información del sector público y datos abiertos. *datos.gob.es reutiliza la información pública*. Recuperado el 18 de abril de 2024 de <https://datos.gob.es/es/blog/convalidacion-del-real-decreto-ley-sobre-reutilizacion-de-la-informacion-del-sector-publico-y>

Van de Sandt, S., Dallmeier-Tiessen, S., Lavasa, A., & Petras, V. (2019). The Definition of Reuse. *Data Science Journal*, 18(1), 1-19. <http://doi.org/10.5334/dsj-2019-022>

Anexo I. Normativa Europea y Estatal sobre reutilización de la información. Aspectos más relevantes

Tabla 1.

Normativa Europea y Estatal sobre reutilización de la información.

Normativa Europea Estatal	Aspectos relevantes	Transposición de normativa	Aspectos relevantes
La Directiva 2003/98/CE del Parlamento Europeo y del Consejo, del 17 de noviembre de 2003 , sobre reutilización de la información del sector público	<ul style="list-style-type: none"> - Estableció un conjunto de normas mínimas y criterios homogéneos para el tratamiento de la información susceptible de ser reutilizada por personas físicas o jurídicas 	Esta directiva, se trasladó al ordenamiento jurídico español con la Ley 37/2007, de 16 de noviembre , sobre reutilización de la información del sector público; y su desarrollo reglamentario: Real Decreto 1495/2011, de 24 de octubre , por el que se desarrolla la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público, para el ámbito del sector público estatal	<ul style="list-style-type: none"> - Establece el régimen jurídico general para la reutilización de la información pública. - Garantiza que esta reutilización de la información se lleve a cabo en el marco de unas condiciones claras, transparentes y no discriminatorias. - Destaca la autorización potestativa de la reutilización como eje central. - Destaca el potencial de la información pública para el desarrollo de la sociedad de la información y la participación democrática.
La Ley 18/2015, de 9 de julio , por la que se modifica la Ley 37/2007, de 16 de noviembre	<ul style="list-style-type: none"> - Se ajusta el ámbito subjetivo, objetivo y las tarifas aplicables a la ley 	Esta Ley, transpone la Directiva 2013/37/EU , que modifica la Directiva 2003/98/CE, actualizando el régimen de reutilización de la información	<ul style="list-style-type: none"> - Considera los cambios en la generación de información, avances tecnológicos y creciente conciencia sobre el valor de la información pública. - Muestra un especial interés por la reutilización como herramienta para fomentar el crecimiento económico, el compromiso social y la transparencia.
La Ley 9/2017, de 8 de noviembre , de Contratos del Sector Público	<ul style="list-style-type: none"> - Tercera modificación de la Ley 37/2007, de 16 de noviembre. - Ajusta el ámbito subjetivo a las disposiciones de la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público, su ámbito objetivo y las tarifas aplicables. 	Se transponen al ordenamiento jurídico español las Directivas del Parlamento Europeo y del Consejo 2014/24/UE, de 26 de febrero de 2014	<ul style="list-style-type: none"> - Su objetivo es mejorar la eficiencia, la transparencia y la equidad en los procesos de contratación pública en la UE, al mismo tiempo que promueve la sostenibilidad y la igualdad de acceso al mercado para las empresas europeas

Normativa Europea Estatal	Aspectos relevantes	Transposición de normativa	Aspectos relevantes
<p>La Directiva (UE) 2019/1024 del Parlamento Europeo y del Consejo, de 20 de junio de 2019, relativa a los datos abiertos y la reutilización de la información del sector público</p>	<ul style="list-style-type: none"> - <i>Ámbito de aplicación:</i> La Directiva se aplica a los datos generados, recolectados o financiados por las autoridades del sector público de los Estados miembros de la UE. - <i>Principios de apertura y reutilización:</i> La Directiva establece que los datos del sector público deben estar disponibles para su reutilización y uso, y deben ser fácilmente accesibles, procesables y utilizables. - <i>Licencias y condiciones de reutilización:</i> Se insta a los Estados miembros a establecer licencias estándar que permitan la reutilización de los datos del sector público. Estas licencias deben ser transparentes, no discriminatorias y no restrictivas. - <i>Excepciones y restricciones:</i> La directiva establece algunas excepciones y restricciones para la reutilización de datos del sector público, como la protección de datos personales, la seguridad nacional y los derechos de propiedad intelectual. - <i>Portal Europeo de Datos Abiertos:</i> La directiva prevé el establecimiento de un portal europeo de datos abiertos, que sirva como punto de acceso centralizado a los conjuntos de datos del sector público de toda la UE. - <i>Cooperación e intercambio de buenas prácticas:</i> La Directiva promueve la cooperación entre los Estados miembros y la Comisión Europea para fomentar la apertura de los datos y compartir buenas prácticas en la reutilización de la información del sector público. - Busca fomentar la <i>transparencia, la innovación y el crecimiento económico.</i> - <i>Busca mejorar el acceso en tiempo real a los datos dinámicos y aumentar el suministro de datos públicos valiosos para la reutilización</i>, incluidos los de las empresas públicas, organizaciones que financian la investigación y organizaciones que realizan actividades de investigación. - Establece que los <i>datos del sector público deben ser liberados en formatos que sean fáciles de utilizar y reutilizar</i>, preferiblemente en formatos electrónicos y de manera gratuita o a un coste mínimo. 	<p>Esta directiva, se ha trasladado al ordenamiento jurídico español con el Real Decreto-Ley 24/2021, de 2 de noviembre, de transposición de directivas de la Unión Europea en las materias de bonos garantizados, distribución transfronteriza de organismos de inversión colectiva, datos abiertos y reutilización de la información del sector público, ejercicio de derechos de autor y derechos afines aplicables a determinadas transmisiones en línea y a las retransmisiones de programas de radio y televisión, exenciones temporales a determinadas importaciones y suministros, de personas consumidoras y para la promoción de vehículos de transporte por carretera limpios y energéticamente eficientes.</p>	<ul style="list-style-type: none"> - Dedica el Libro tercero a la Transposición de la Directiva (UE) 2019/1024, del Parlamento Europeo y del Consejo de 20 de junio de 2019, relativa a los datos abiertos y la reutilización de la información del sector público. <p>Las novedades importantes que introduce a este respecto son:</p> <ul style="list-style-type: none"> - <i>Apertura de datos:</i> Establece medidas para impulsar la apertura de datos del sector público, lo que significa facilitar el acceso y la disponibilidad de la información generada por las entidades públicas para su uso por parte de ciudadanos, empresas y organizaciones. - <i>Actualización del régimen de reutilización de la información:</i> Se actualiza y moderniza el marco normativo existente sobre la reutilización de la información del sector público para adaptarlo a los avances tecnológicos y a las necesidades actuales. - <i>Fomento de la transparencia y la innovación:</i> Se promueve la transparencia en las administraciones públicas al poner a disposición más información de manera accesible. Además, se busca fomentar la innovación y el desarrollo de nuevos servicios y aplicaciones basados en datos abiertos. - Y, recoge en el Artículo 64, todos los aspectos relativos a la modificación de la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público. - Su <i>propósito</i> es promover la transparencia, la innovación y el desarrollo económico a través de la apertura de datos.

Normativa Europea Estatal	Aspectos relevantes	Transposición de normativa	Aspectos relevantes
	<ul style="list-style-type: none"> - <i>Las personas y organizaciones pueden acceder a los datos y utilizarlos para diferentes propósitos</i>, como el desarrollo de aplicaciones, la realización de investigaciones o la creación de servicios innovadores. - <i>Promueve la interoperabilidad de los datos</i>, es decir, los datos de diferentes instituciones y organismos del sector público deben poder combinarse y utilizarse juntos de manera efectiva, ya que esto facilita la colaboración y el intercambio de información entre diferentes entidades. 		<ul style="list-style-type: none"> - Ampliación del ámbito de aplicación: Se amplía el alcance de la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público, para adaptarlo a las disposiciones de la Directiva de la UE. Esto implica que más tipos de información y entidades públicas estarán sujetas a las normativas de datos abiertos y reutilización. - Nuevas obligaciones para las administraciones públicas, en cuanto a la apertura y reutilización de datos. Esto puede incluir la obligación de publicar ciertos conjuntos de datos de manera accesible y en formatos abiertos. - Establecimiento de condiciones de reutilización: Se definen condiciones claras para la reutilización de la información del sector público, como tarifas aplicables, procedimientos de autorización y requisitos de calidad de los datos. - Fomento de la interoperabilidad y la calidad de los datos: Se promueve la interoperabilidad de los datos y se establecen estándares de calidad para garantizar que la información sea útil y fácilmente accesible para los usuarios. - Exclusión de documentos sensibles: Se excluye la reutilización de documentos cuyo acceso esté limitado por razones de protección de información sobre infraestructuras críticas. - Especificación de datos de alto valor: Se establece la posibilidad de especificar conjuntos de datos adicionales, además de los establecidos por la Comisión Europea, por parte del Ministerio de Asuntos Económicos y Transformación Digital.

Normativa Europea Estatal	Aspectos relevantes	Transposición de normativa	Aspectos relevantes
			<ul style="list-style-type: none"> - Exención de ciertas entidades del sometimiento a normas de procedimiento: sociedades mercantiles, centros de enseñanza y entidades de investigación quedan exentos de ciertas normativas de procedimiento. - Designación de unidades responsables: se establece la obligación de que cada entidad designe una unidad responsable de garantizar la disponibilidad de su información, con funciones relacionadas con la coordinación, actualización y promoción de la reutilización de datos.



Normativa Europea Estatal	Aspectos relevantes	Transposición de normativa	Aspectos relevantes
<p>Ley de Datos (Data Act o DA). Entró en vigor el 11 de enero de 2024</p>	<ul style="list-style-type: none"> - Entró en vigor el 11 de enero de 2024 y será aplicable con carácter general a partir de septiembre de 2025. - <i>Su objetivo:</i> Establecer un marco normativo coherente a nivel europeo para la gestión de datos, con el fin de promover la innovación, la competencia y la protección de los derechos digitales en el mercado digital. - Facilita la accesibilidad de los datos de forma horizontal. - Establece principios y directrices para todos los sectores. - Pone énfasis en el acceso justo a los datos, los derechos de los usuarios y la protección de los datos personales. - Aclara quién puede crear valor a partir de los datos y bajo qué condiciones. - Disponibilidad de los datos en el marco de Internet de las cosas (IoT). - <i>Promoción de la interoperabilidad y la portabilidad de datos:</i> Busca fomentar la interoperabilidad entre sistemas y servicios de datos, así como facilitar la portabilidad de datos entre diferentes proveedores y plataformas, lo que permite a los usuarios tener un mayor control sobre sus datos personales. - <i>Protección de datos y privacidad:</i> Incorpora disposiciones para proteger la privacidad y los datos personales de los individuos, garantizando que el tratamiento de datos se realice de manera transparente, segura y respetando los derechos de los usuarios. - <i>Acceso y reutilización de datos:</i> Establece normas para facilitar el acceso y la reutilización de datos, tanto para fines comerciales como no comerciales, promoviendo la apertura y la transparencia en el uso de datos en la UE. - <i>Fomento de la competencia y la innovación</i> en el mercado digital para facilitar el acceso a los datos y promover la creación de nuevos productos y servicios basados en datos. 	<p>Es un pilar clave de la Estrategia Europea de Datos, que complementa a otras normativas:</p> <ul style="list-style-type: none"> - Reglamento (UE) 2023/2854 sobre normas armonizadas para un acceso justo a los datos y su utilización (DA). - Directiva (UE) 2019/1014 relativa a los datos abiertos y la reutilización de la información del sector público. - Reglamento (UE) 2022/868 relativo a la gobernanza europea de datos (DGA) 	<p>El Reglamento (UE) 2022/868 relativo a la gobernanza europea de datos (DGA):</p> <ul style="list-style-type: none"> - <i>Su objetivo</i> es establecer un marco normativo que promueva la gestión segura y eficiente de los datos en la UE. - <i>Su propósito</i> es fomentar la confianza, la innovación y el desarrollo económico en el mercado digital europeo. - Regula el intercambio seguro y voluntario de conjuntos de datos que están en poder de organismos públicos, los servicios de intermediación de datos y su cesión altruista. - <i>Creación de un marco de gobernanza:</i> el reglamento establece un marco de gobernanza para la gestión de datos en la UE, con el objetivo de garantizar la interoperabilidad, la transparencia y la seguridad en el intercambio de datos entre los Estados miembros. - <i>Facilitación del acceso a los datos:</i> Busca facilitar el acceso a los datos en toda la UE, promoviendo la reutilización de datos de forma segura y compatible con la normativa de protección de datos. - <i>Promoción de la confianza:</i> Fomenta la confianza en el uso de datos al establecer estándares y buenas prácticas para la gestión de la privacidad, la seguridad y la calidad de los datos. - <i>Establecimiento de un mercado único de datos:</i> Su objetivo es crear un mercado único de datos en la UE, donde los datos puedan circular libremente entre los Estados miembros y las empresas puedan aprovechar el potencial de los datos para la innovación y el desarrollo económico. - <i>Protección de datos y derechos digitales:</i> Incorpora disposiciones para proteger los derechos de privacidad y los derechos digitales de los ciudadanos europeos en el contexto de la gestión y el intercambio de datos.

Normativa Europea Estatal	Aspectos relevantes	Transposición de normativa	Aspectos relevantes
	<ul style="list-style-type: none"> - Gobernanza de datos: Establece mecanismos de gobernanza para garantizar la transparencia y la responsabilidad en la gestión de datos, así como para proteger los derechos de los usuarios y promover la colaboración entre los diferentes actores del ecosistema de datos. - Datos a favor de la Administración Pública 		<ul style="list-style-type: none"> - Impulso a la economía de datos: Busca impulsar la economía de datos en la UE al facilitar la creación de valor a partir de los datos y promover la innovación y la competitividad en el mercado digital. - Directiva (UE) 2019/1014 relativa a los datos abiertos y la reutilización de la información del sector público. - Establece el marco jurídico para la reutilización de la información del sector público puesta a disposición del público como datos abiertos. - El Reglamento (UE) 2023/2854 sobre normas armonizadas para un acceso justo a los datos y su utilización (DA) - Su objetivo es crear un entorno propicio para el acceso y la utilización de datos en la UE. - Su propósito: Promover la competencia, la innovación y el desarrollo económico en el mercado digital. - Acceso justo a los datos: Establece normas para garantizar un acceso justo y no discriminatorio a los datos, especialmente aquellos que son considerados como recursos clave para la innovación y el desarrollo económico. - Promoción de la competencia: Busca fomentar la competencia en el mercado de datos al facilitar el acceso a los mismos para una amplia gama de usuarios, incluidas pequeñas y medianas empresas (PYME) y las startups. - Interoperabilidad y estándares: Promueve la interoperabilidad entre sistemas y servicios de datos, así como la adopción de estándares comunes para facilitar el intercambio y la utilización de datos en toda la UE. - Gobernanza de los datos: Establece mecanismos de gobernanza para garantizar la transparencia y la responsabilidad en la gestión de los datos, así como para proteger los derechos de los usuarios y los proveedores de datos. - Facilitación de la innovación: Busca fomentar la innovación al eliminar barreras al acceso y uso de los datos, lo que puede impulsar el desarrollo de nuevos productos y servicios digitales. - Protección de datos y seguridad: Incorpora disposiciones para garantizar la protección de datos personales y la seguridad de la información en el contexto del acceso y la utilización de datos.

LA ALFABETIZACIÓN EN DATOS EN EL CONTEXTO
UNIVERSITARIO DE LAS CIENCIAS SOCIALES
(ALFADATAUniv_CCSS)

REUTILIZACIÓN Y CITACIÓN DE DATOS

Actividad

Ana Iglesias Rodríguez
Yolanda Martín González



REUTILIZACIÓN DE DATOS DE INVESTIGACIÓN

[Actividad propuesta 1]

Tiempo aproximado de realización de la actividad:

1 hora.

1. Objetivo

El objetivo de esta actividad es determinar qué conjuntos de datos creados y/o recogidos durante el proceso de investigación pueden ser objeto de reutilización, así como proponer posibles formas de utilización de los datos abiertos.

2. Enunciado explicativo de la actividad

Durante el proceso de investigación científica, se generan y recopilan conjunto de datos que poseen un potencial considerable más allá del estudio inicial. Esta actividad busca identificar estos conjuntos de datos y explorar oportunidades para su reutilización, fomentando, con ello, la generación de nuevo conocimiento.

3. Desarrollo e instrucciones para realizar la actividad

Una vez leído el documento explicativo en torno a la temática *Reutilización y citación de datos*, lleva a cabo las siguientes acciones:

- Identifica los conjuntos de datos generados y/o recopilados durante la investigación que estés desarrollando o planeando llevar a cabo.
- Selecciona e indica aquellos dataset que puedan ser objeto de reutilización.
- Dispón o convierte los conjuntos de datos seleccionados en formatos abiertos y asegúrate de que estén disponibles bajo licencias Creative Commons u otras licencias abiertas.
- Propón alguna otra investigación o estudio, o en su defecto, algún producto, aplicación o servicio que pueda crearse a partir de la reutilización de estos conjuntos de datos.

4. Conocimientos y recursos requeridos para realizar la actividad

Acceso a la colección “datos de investigación” del repositorio institucional de tu universidad, así como de otros similares.

Acceso a repositorios de datos de investigación generalistas como Zenodo, Fighshare, Mendeley Data u otros similares.

5. Autoevaluación de la actividad

- ¿He revisado y comprendido el documento explicativo sobre *Reutilización y citación de datos*?
 Sí No
- ¿He identificado los conjuntos de datos generados y/o recogidos durante mi investigación?
 Sí No
- ¿He seleccionado y señalado aquellos conjuntos de datos que podrían ser objeto de reutilización por otros investigadores?
 Sí No
- ¿He dispuesto los conjuntos de datos seleccionados en formatos abiertos?
 Sí No
- ¿He comprobado que los conjuntos de datos están disponibles bajo licencias Creative Commons u otras licencias abiertas?
 Sí No
- ¿He pensado de qué modo se pueden reutilizar, en futuras investigaciones, los conjuntos de datos generados y/o recopilados o, si se puede crear algún producto, aplicación o servicio a partir de la reutilización de los conjuntos de datos?
 Sí No
- ¿He consultado ejemplos de conjuntos de datos de investigación compartidos en repositorios como GREDOS, Zenodo, u otros?
 Sí No
- ¿He adquirido conocimiento sobre licencia de datos y licencias Creative Commons? (En el caso de no tenerlos al inicio de la actividad)
 Sí No
- ¿He considerado la importancia de la reutilización de datos de investigación y su potencial para generar nuevo conocimiento?
 Sí No
- ¿He explorado posibles colaboraciones con otros investigadores o instituciones para facilitar la reutilización de los conjuntos de datos?
 Sí No

REUTILIZACIÓN DE DATOS DEL SECTOR PÚBLICO

[Actividad propuesta 2]

Tiempo aproximado de realización de la actividad:
1 hora.

1. Objetivo

El objetivo de esta actividad es profundizar en el conocimiento de la apertura de datos en el ámbito de las Administraciones públicas e identificar posibles formas de reutilización de dichos datos.

2. Enunciado explicativo de la actividad

Esta actividad se centra en la búsqueda y localización de productos, servicios y/o aplicaciones que hayan surgido de la reutilización de conjuntos de datos del sector público, en el ámbito de tu interés.

3. Desarrollo e instrucciones para realizar la actividad

Una vez leído el documento explicativo en torno a la temática *Reutilización y citación de datos*, sigue los siguientes pasos:

- Accede al portal datos.gob.es
- Localiza ejemplos de reutilización de datos de las administraciones públicas referidos a una temática concreta de tu interés.
- Rellena una ficha descriptiva con la siguiente información:
 - Agente reutilizador: El individuo, organización o empresa que ha reutilizado los datos del sector público para crear un servicio, producto o aplicación.
 - Organismo publicador: La entidad gubernamental o administrativa que ha publicado los datos originales.
 - Catálogo de origen de los datos.
 - Breve descripción del producto o servicio.
 - Fecha de publicación del producto o servicio surgido de la reutilización de los datos.
 - Fecha de creación: fecha de creación del dataset o conjunto de datos.
 - Fecha de actualización: Última fecha en la que se actualizó el conjunto de datos, en caso de disponerse.
 - Distribuciones en los que se dispone: Formatos de archivos en los que están disponibles los datos, en caso de disponerse.
 - Licencias de reutilización: Términos y condiciones de uso bajo los cuales se disponen los datasets, en caso de disponerse.

A continuación, se proporciona un **ejemplo** extraído del portal datos.gob.es para que te sirva de referencia.

- Agente reutilizador: Rediam (Red de Información Medioambiental de Andalucía).
- Organismo publicador: Demarcaciones Hidrográficas del Guadalquivir, Cuencas Mediterráneas Andaluzas, Guadalete-Barbate y Tinto-Odiel-Piedras.
- Catálogo de origen de los datos: <https://portalrediam.cica.es/geonetwork/srv/spa/catalog.search#/home>.
- Breve descripción del producto o servicio: Esta aplicación permite visualizar datos de Rediam (Red de Información Medioambiental de Andalucía) mediante un sistema de consulta rápido y sencillo para la obtención de la situación diaria de la reserva de agua en embalses del territorio andaluz, tanto a nivel de embalse como de demarcación hidrográfica, provincia o sistema de explotación.
- Fecha de publicación del producto o servicio surgido de la reutilización de los datos: 13-05-2024.
- Fecha de creación del dataset: 23-02-2024.
- Fecha de actualización: Diaria.
- Distribuciones en los que se dispone: formatos pdf y xml.
- Licencias de reutilización: CC by 4.0.

4. Recursos requeridos para realizar la actividad

El portal datos.gob.es

5. Autoevaluación de la actividad

¿He accedido al portal datos.gob.es para buscar ejemplos de reutilización de datos del sector público?

Sí No

¿He localizado ejemplos de reutilización de datos de las administraciones públicas referidos a una temática específica de mi interés?

Sí No

¿He completado una ficha descriptiva para cada ejemplo de reutilización, incluyendo todos los campos requeridos?

Sí No

¿He identificado el origen de los conjuntos de datos reutilizados?

Sí No

¿He identificado al agente reutilizador de los datos?

Sí No

¿He indicado el organismo publicador de los datos?

Sí No

¿He verificado las fechas de creación y actualización de los conjuntos de datos?

Sí No

¿He identificado las licencias de reutilización bajo las cuales se disponen los datos?

Sí No