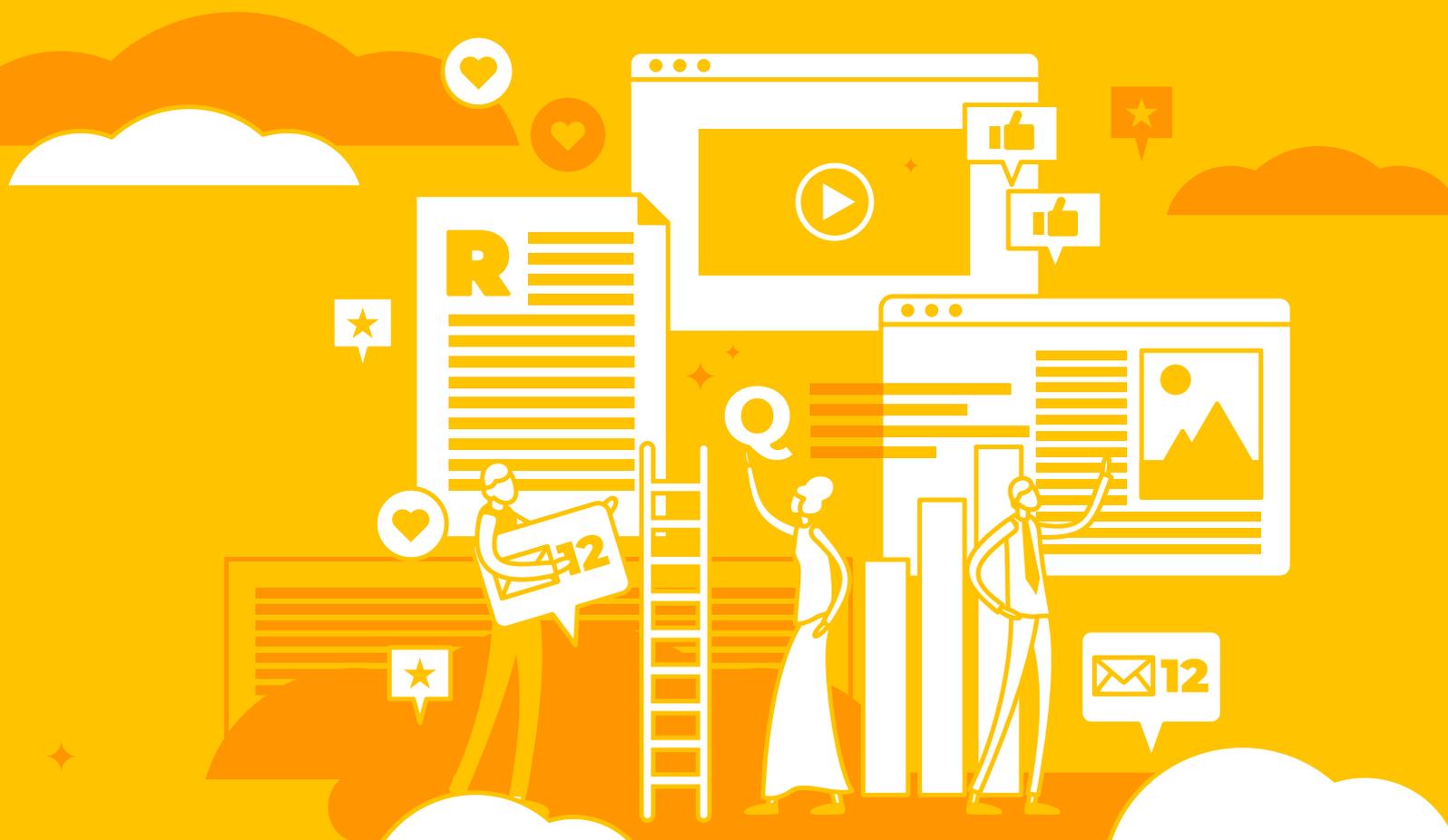




Formación en
Competencias
Digitales

1

Búsqueda y gestión de información y datos



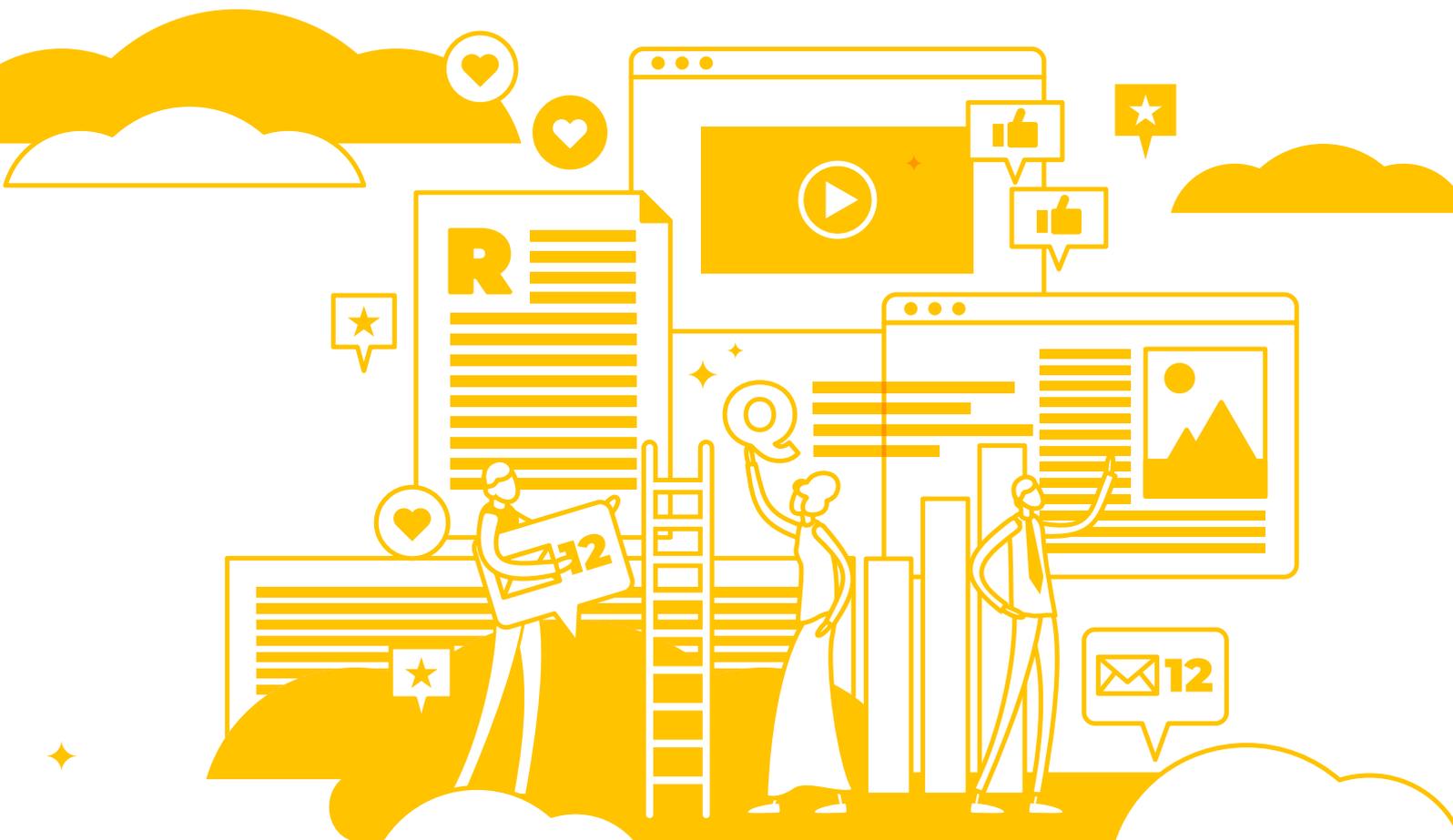


Formación en
Competencias
Digitales



Búsqueda y gestión
de información y datos

Nivel C2





Búsqueda y gestión de información y datos

ÍNDICE

1.1. NAVEGAR, BUSCAR Y FILTRAR DATOS, INFORMACIÓN Y CONTENIDOS DIGITALES

- [*Búsqueda y filtrado de información en buscadores académicos, portales web y bases de datos*](#)
- [*La búsqueda y filtrado en repositorios digitales*](#)

1.2. EVALUAR DATOS, INFORMACIÓN Y CONTENIDOS DIGITALES

- [*Sesgo y estereotipos informativos*](#)
- [*Listas y plataformas informativas*](#)
- [*Perfiles en redes sociales*](#)
- [*Uso ético de la IA para reforzar la opinión*](#)

1.3. GESTIÓN DE DATOS, INFORMACIÓN Y CONTENIDOS DIGITALES

- [*Aplicaciones de procesamiento de datos*](#)
- [*Estrategias avanzadas de recuperación en almacenamiento online*](#)
- [*Inteligencia de datos*](#)



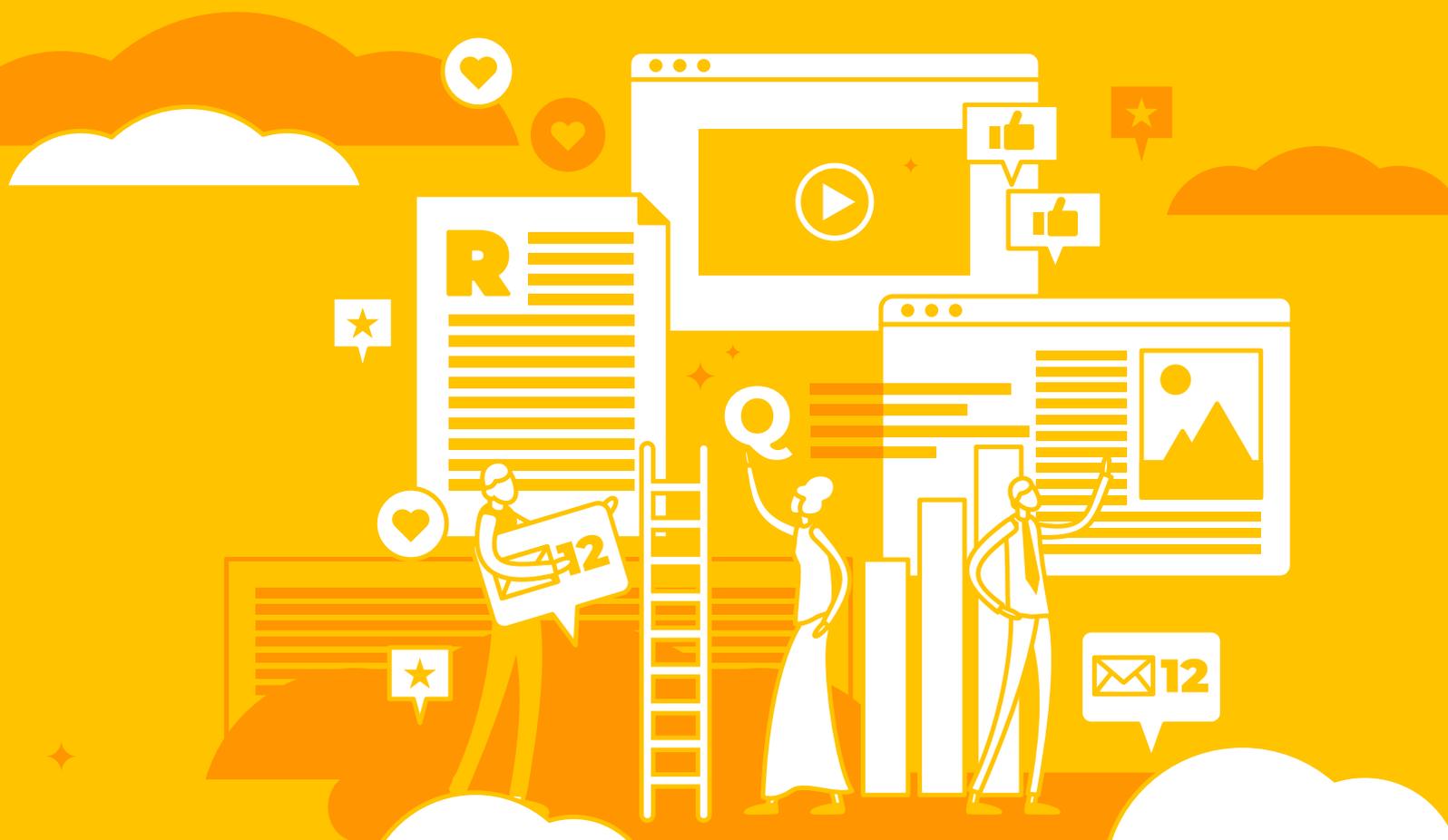


DigitAll

Búsqueda y gestión
de información y datos

1.1

NAVEGAR, BUSCAR Y FILTRAR DATOS, INFORMACIÓN Y CONTENIDOS DIGITALES





Búsqueda y gestión
de información y datos

Nivel C2 1.1 Navegar, buscar y filtrar datos,
información y contenidos digitales

Búsqueda y filtrado de información en buscadores académicos, portales web y bases de datos





Búsqueda y filtrado de información en buscadores académicos, portales web y bases de datos

Introducción

En la actualidad, existen diversas herramientas de búsqueda de información científica en las cuales se puede localizar diversidad de documentos e información como: artículos, revistas, libros, actas de conferencias y patentes. Para la búsqueda y filtrado de información, se aplican varios criterios entre los que se identifican aspectos genéricos (mismos criterios de búsqueda y filtrado) y específicos (según cada herramienta).

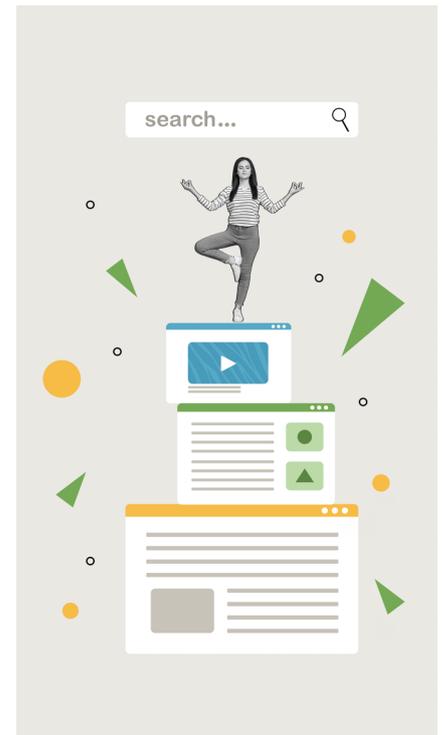
Este documento trata de describir los criterios genéricos y específicos para la búsqueda y filtrado de información, datos y recursos en buscadores académicos, bases de datos bibliográficas, portales bibliográficos, bibliotecas digitales y directorios en línea.

Procesos de búsqueda y filtrado de información

Como se ha indicado, la búsqueda y filtrado puede ser **genérica** o **específica**. La primera se refiere a los aspectos transversales o típicos que pueden utilizarse en cualquiera de estas herramientas y, la segunda, a criterios únicos que se utilizan en cada una de ellas, en tanto que dependen de sus características.

Antes de iniciar, es importante recordar que las principales definiciones analizadas en el vídeo 2 del nivel C2 están relacionadas con las siguientes herramientas:

- Los **buscadores académicos** o **motores de búsqueda académicos** son un tipo de buscador especializado en **localizar** y **mostrar** información científica presente en Internet. Algunos ejemplos serían **Google Académico**, también conocido como **Google Scholar**, **Science Research** y **World Wide Science**.



BUSCADORES ACADÉMICOS

Google Académico/Scholar:
scholar.google.es

Scinece Research:
scienceresearch.com

World Wide Science:
worldwidescience.org



- Las **bases de datos académicas** son sistemas de información que registran documentos que resultan de actividades académico-científicas tales como artículos, actas de congresos, libros y capítulos de libros. Entre las más reconocidas se encuentran, por ejemplo, **Web of Science** o **WoS**, **Scopus** y **ERIC**.
- Los **portales bibliográficos** son plataformas de acceso a bases de datos de diversas temáticas, permitiendo descargar artículos a texto completo. Ejemplo de ello es **DIALNET**, creado por la Universidad Internacional de La Rioja.
- Las **bibliotecas digitales** son colecciones de información científica procedente fundamentalmente de revistas y libros. Por ejemplo, las de las universidades, organismos e instituciones públicas o privadas o las bibliotecas internacionales, como **Scielo España**.
- Los **directorios en línea** permiten localizar documentos publicados por revistas académicas, siendo un ejemplo **DOAJ**.

**BASES DE DATOS ACADÉMICAS****Web of Science/WoS:**webofscience.com/wos**Scopus:**scopus.com**ERIC:**eric.ed.gov**DIALNET**dialnet.unirioja.es**SCIELO ESPAÑA**scielo.isciii.es/scielo.php**DOAJ**doaj.org **NOTA****Recursos libres y privados**

Las herramientas de búsqueda de las bases de datos bibliográficas, buscadores académicos, portales bibliográficos, bibliotecas digitales o los directorios en línea pueden ser privadas o libres. Aunque la mayoría permiten acceder en formato abierto a la información científica, otras bases de datos como Scopus o Web of Science (WoS) requieren pagar una cuota para acceder a su información. Es por ello que las universidades españolas destinan un presupuesto para que sus estudiantes, docentes e investigadores puedan acceder a la WoS y Scopus mediante sus credenciales institucionales.

Procesos de búsqueda y filtrado de información genérica

Para la **búsqueda genérica** de información se pueden utilizar los siguientes criterios:

- **Búsqueda por autor:** en este caso, deberá colocar el nombre y apellidos del autor que desea buscar. Por ejemplo: Isaac Newton. Asimismo, es importante tener en cuenta aspectos como:



1 | Uso del operador comillas (""): permite buscar con exactitud todos los documentos que contengan o mencionen a un texto determinado. Para la búsqueda de un autor se colocaría de la siguiente manera: "Isaac Newton".

2 | Normalización de nombres de autor: ello implica que, en ocasiones, los autores firman sus libros u obras de una manera específica que necesariamente no tiene que coincidir con su nombre y apellidos. Por ejemplo: La Saga de Harry Potter es escrita por Joanne Rowling bajo el seudónimo J. K. Rowling. Además, a veces, las editoriales, revistas y fuentes académicas referencian al autor o autores con la letra inicial del nombre y el primer apellido. Por ejemplo, Gabriel García Márquez puede encontrarse bajo los siguientes criterios de búsqueda: G. García o G.G. Márquez. En diferentes países, especialmente iberoamericanos, emplean nombre, segundo nombre y dos apellidos, para los que se hace uso, como recurso, del guion (-). Por ejemplo: Gabriel García-Márquez. Por tanto, otra forma de buscar el autor sería de esta manera: Gabriel García-Márquez.

- **Búsqueda por palabras clave:** permite buscar documentos que contengan dicha palabra clave, recordando, para ello, el posible uso del operador comillas. En todo caso, debe velar por no cometer errores ortográficos que impidan la localización de la información.
- **Búsqueda por cadena de búsqueda:** puede concatenar varias palabras claves empleando el operador AND o su símbolo +.

De igual manera, el **filtrado de información** presenta aspectos genéricos presentes en todas las herramientas de búsquedas, como:

- **Filtrado cronológico:** te permite acotar o seleccionar los documentos que fueron publicados en un período determinado. Para ello, tiene dos opciones:

1 | Opción 1: solo se selecciona el año mínimo. Por ejemplo: 2018 le mostrará los resultados publicados desde 2018 (enero) hasta el momento que realiza la búsqueda.





2 | Opción 2: selecciona un rango. Por ejemplo: 2018–2022, para lo cual se te devolverán los resultados publicados entre enero de 2018–diciembre de 2022.

- **Filtrado de jerarquía o de orden:** te permite filtrar por orden cronológico (de mayor a menor o menor a mayor) o por relevancia.
- **Filtrado por tipología documental:** te permite seleccionar un tipo de documento específico. Por ejemplo: artículo de revisión, patentes, tesis u otros.

Procesos de búsqueda y filtrado de información específica

Los procesos de búsqueda específica **siempre dependerán** de la herramienta empleada. Para ello, se debe indagar y navegar por todas sus funcionalidades.

A continuación, se ilustrarán algunos ejemplos:

- **Buscadores académicos:** *Google Scholar*, *Science Research* y *World Wide Science*. Podrá acotar la búsqueda de autor utilizando la palabra reservada **autor:** y el nombre del autor (por ejemplo: *author: Penélope Cruz*) o buscar por título utilizando la palabra **title:** y la palabra clave como, por ejemplo, *title: fresas*. En todo caso, en ambos buscadores puede utilizar la opción «avanzado», que te permitirá localizar solo por autor o título.

 **BUSCADORES ACADÉMICOS**

Google Scholar:
scholar.google.es

Science Research:
scienceresearch.com

World Wide Science:
worldwidescience.org

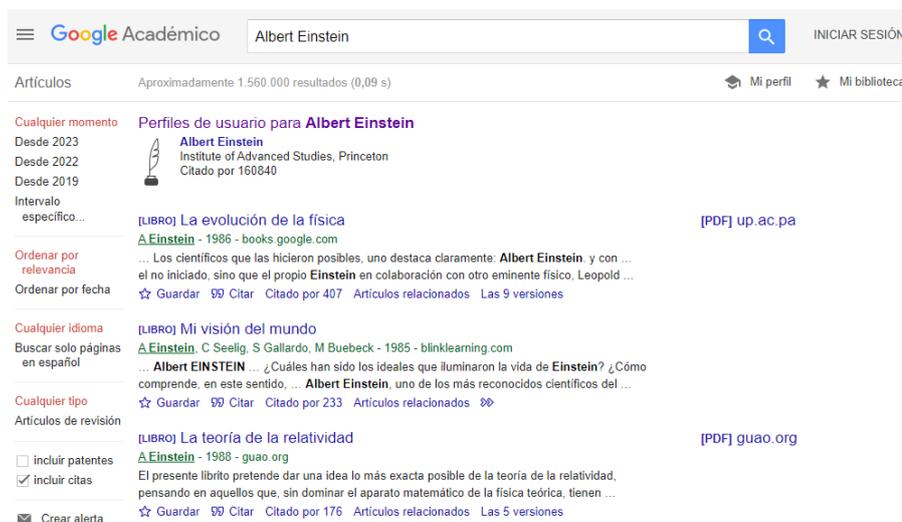


Figura 1. Ejemplo del resultado para la búsqueda de Albert Einstein, en Google Scholar o Google Académico



- **Bases de datos académicas:** (*Web of Science* o *WoS*, *Scopus* y *ERIC*). En el buscador de estas herramientas se facilitan varias opciones para buscar información, entre las que se destacan: búsqueda por título, resumen, autor, afiliación, fuente de publicación (nombre de la revista, libros o actas de conferencias) y DOI.

Para la búsqueda por título, resumen, autor, afiliación y fuente de publicación siempre se utilizan palabras clave que representan la búsqueda a realizar. En el caso del DOI, solo se utilizan los caracteres numéricos. Es decir, de <https://doi.org/10.3916/C70-2022-09> se introduciría exclusivamente **10.3916/C70-2022-09** en el buscador.

NOTA

DOI

DOI (Digital Object Identifier) es un identificador único y permanente para las publicaciones electrónicas. Proporciona información sobre la descripción de los objetos digitales (revistas, artículos...) y su localización en Internet a través de metadatos (autor, título, datos de publicación, etc.). Funciona como una URL única en Internet. Por ejemplo: doi.org/10.3916/C70-2022-09 te dirige a una publicación científica determinada.

Saber más

Búsqueda en WoS, Scopus y ERIC

Para profundizar en la búsqueda de información en dichas bases de datos, puede consultar las siguientes direcciones electrónicas:

- **Scopus:** biblioguias.ucm.es/scopus/buscar
- **Web of Science (WoS):** e.digitall.org.es/wos-buscar
- **ERIC:** eric.ed.gov/?advanced

- **Portales bibliográficos:** *DIALNET*. Además de los criterios genéricos mencionados anteriormente, se puede realizar la búsqueda por resúmenes y títulos. El procedimiento consiste en indicar una o varias palabras clave, y seleccionar la opción correspondiente para iniciar la búsqueda (resúmenes o títulos).
- **Bibliotecas digitales:** *Scielo España*. Incluyen, además, la búsqueda por revistas, título, autor, resumen y área (por ejemplo: medicina clínica).



BASES DE DATOS ACADÉMICAS

Web of Science/WoS:

webofscience.com/wos

Scopus:

scopus.com

ERIC:

eric.ed.gov



DIALNET

dialnet.unirioja.es



SCIELO ESPAÑA

scielo.isciii.es/scielo.php



- **Directorios en línea: DOAJ.** Incluyen dos tipos de búsqueda: la de artículos científicos, cuyas opciones son la búsqueda por título, resumen, área y autor, y por revista, que puede ser por título, ISSN o editorial.



NOTA

ISSN

El ISSN (International Standard Serial Number / Número Internacional Normalizado de Publicaciones Seriadadas) es un código numérico reconocido internacionalmente para la identificación de las publicaciones seriadas o en serie. Por ejemplo, el ISSN: 1988-592X se refiere a la **Revista de Educación** (e.digital.org.es/revista-educacion), publicación científica del Ministerio de Educación y Formación Profesional español.

Como observa, las herramientas de búsqueda de bases de datos bibliográficas, portales bibliográficos, bibliotecas digitales y directorios en línea proveen mayor diversidad de criterios de búsqueda que los buscadores académicos, lo cual permitirá obtener información científica acotada a los objetivos. Cabe destacar que solo algunas de estas herramientas son accesibles en español, como son los casos de **Google Académico**, **DIALNET**, **Scielo España**, las bibliotecas digitales de las universidades iberoamericanas y **Redalyc** (biblioteca digital y base de datos bibliográfica).



Google Académico:
scholar.google.es

DIALNET:
dialnet.unirioja.es

Scielo España:
scielo.isciii.es/scielo.php

Redalyc:
redalyc.org

Procesos de filtrado específicos

Al igual que la búsqueda de información, el filtrado depende de la herramienta empleada. Por tal motivo, a continuación, se muestran algunos de los criterios de filtrado comunes según la tipología de estas herramientas.

- **Buscadores académicos: Google Scholar, Science Research y World Wide Science.** En **Google Scholar**, además de los filtrados genéricos, permite filtrar por idioma y patentes. Mientras, **Science Research** agrupa los resultados de búsqueda por diversos criterios, tales como: autores, fuentes de publicación, tipo de documentos, y tópicos o descriptores en lo que se encuentra la o las palabras clave utilizadas. Por último, el buscador **World Wide Science** funciona de manera similar a **Science Research**, ya que agrupa los resultados de búsqueda según sus mismos criterios, incluyendo país y lenguaje.



Google Scholar:
scholar.google.es

Scinece Research:
scienceresearch.com

World Wide Science:
worldwidescience.org



- **Bases de datos académicas: *Web of Science* o *WoS*, *Scopus* y *ERIC*.** Las bases de datos funcionan generalmente con una misma dinámica. Es decir, utilizan criterios similares. Por ejemplo, ***WoS*, *Scopus* y *ERIC*** incluyen en sus filtrados los criterios de: área del conocimiento o descriptores, autores y título de la fuente de publicación (revista, libros y actas de conferencias). Las dos principales bases de datos académicas (***Web of Science* o *WoS*, y *Scopus***) introducen otros filtros relacionados con: palabras clave (Keywords), afiliaciones (affiliation), países (country), idiomas (language) y niveles de acceso abierto (open access).



BASES DE DATOS ACADÉMICAS

Web of Science/WoS:

webofscience.com/wos

Scopus:

scopus.com

ERIC:

eric.ed.gov

Saber más

Niveles de acceso abierto (Open Access)

A continuación, se ofrece el significado de los diferentes niveles o rutas del acceso abierto cuya información proveen las bases de datos en relación a los resultados de búsqueda:

- **Acceso abierto diamante:** documentos en acceso abierto sin cargos para autores ni lectores.
- **Acceso abierto dorado:** documentos que se encuentran en revistas que solo publican en acceso abierto.
- **Acceso abierto híbrido:** documentos que se encuentran en revistas que brindan a los autores la opción de publicar en acceso abierto.
- **Acceso abierto bronce:** documentos cuya versión publicada del registro o manuscrito aceptado para publicación. El editor ha optado por proporcionar acceso gratuito temporal o permanente.
- **Acceso abierto verde:** documentos cuya versión publicada o manuscrito aceptado para publicación, está disponible en repositorio.

Es relevante señalar que las opciones de filtrado de información en las bases de datos académicas están en inglés (Keywords, affiliation, country, language y open access). En este documento se traducirán para facilitar su comprensión.



Imagen 2. Peter Suber en 2015. Director de proyecto del Harvard's Open Access Project, uno de los precursores del movimiento Open Access. Fotograma del [vídeo de SPARC Europe](#) (14 de abril de 2021). Hablemos de...Infraestructura de Ciencia Abierta con Peter Suber.



- **Portales bibliográficos: DIALNET.** En este portal bibliográfico se destacan los filtrados relacionados con: textos completos, editores, idiomas, descriptores, actas de congresos y países.
- **Bibliotecas digitales: Scielo España y directorios en línea: DOAJ.** En relación a las bibliotecas digitales no se identifican criterios de filtrados específicos ya que se posibilitan los generales (epígrafe 2). Los directorios en línea, al igual que las bibliotecas, presentan pocos o ningún criterio de filtrado, pues solo se destaca por materia (tipo descriptores o áreas del conocimiento) y por revista académica.

Las herramientas de búsqueda tienen una alta coincidencia en relación con los filtrados de información, por lo que se centran, básicamente, en aspectos genéricos. Solo las bases de datos académicas (**Web of Science** o **WoS**, y **Scopus**), por su nivel de complejidad, incluyen criterios de filtrado que ayudan a acotar los resultados de búsqueda.

Se ha de resaltar que, tanto los «descriptores» como «áreas de conocimiento» que proveen las herramientas de búsqueda como criterios para el filtrado de la información, dependen de sus especificaciones. Es decir, puede que la misma búsqueda se realice en varias herramientas y que el resultado del filtrado arroje resultados diferentes.

⚠ ATENCIÓN

Advertencias y errores frecuentes que se deberían evitar

Entre los errores frecuentes en la búsqueda de información en buscadores académicos, bases de datos bibliográficas, portales bibliográficos, bibliotecas digitales y directorios en línea, destacan:

- 1 | Errores ortográficos.
- 2 | Palabras clave solo en español.
- 3 | Uso inadecuado de operadores comilla ("") y AND (+).

i Saber más

Procesos de búsqueda de información científica

Para profundizar en la búsqueda de información científica, recomendamos los siguientes enlaces:

- Recurso audiovisual de la Universidad Miguel Hernández de Elche: e.digitall.org.es/video-elche
- Documento elaborado por la Universidad Complutense de Madrid: e.digitall.org.es/complutense-informacion



DIALNET

dialnet.unirioja.es



SCIELO ESPAÑA

scielo.isciii.es/scielo.php



DOAJ

doaj.org



WEB OF SCIENCE

webofscience.com/wos



SCOPUS

scopus.com



Búsqueda y gestión
de información y datos

Nivel C2 1.1 Navegar, buscar y filtrar datos,
información y contenidos digitales

La búsqueda y filtrado en repositorios digitales





La búsqueda y filtrado en repositorios digitales

Introducción

Como ha aprendido en vídeos y documentos anteriores, se aplican diversos criterios para la búsqueda y filtrado en Internet, en los cuales se identifican aspectos genéricos (mismos criterios de búsqueda y filtrado) y específicos (según cada herramienta). También ha aprendido sobre otras herramientas de búsqueda diferentes como: bases de datos, bibliotecas digitales y buscadores académicos que ofrecen las instituciones gubernamentales y no gubernamentales.

No obstante, existe una herramienta adicional para la búsqueda y localización de información científica: los repositorios digitales con carácter profesional. Por tal motivo, el objetivo de este documento es describir los criterios genéricos o comunes para la búsqueda y filtrado de información en estos repositorios digitales profesionales.



Repositorios digitales: definición e importancia para la búsqueda en Internet

Los repositorios digitales, como se expresa en la *Biblioteca de la Universidad Complutense de Madrid*, son aplicaciones web (software) que almacenan archivos digitales. En ellos, se reúne, preserva, difunde y organiza, en acceso abierto, la producción científica de una institución o de un área temática.

En estos repositorios, los autores y las instituciones comparten, entre otros, sus artículos, informes, datos de investigación, ponencias en congresos científicos, recursos y objetos educativos digitales, etc. Es decir, difunden varios tipos de documentación relacionada con su producción científica y educativa, respetando siempre las políticas de acceso a la información que establecen las fuentes de publicación. Esto es, las editoriales y revistas de publicación si es el caso.

Por tanto, los repositorios digitales utilizan un modelo de acceso abierto para que las personas puedan acceder a la producción. En este caso, científica. Cabe destacar que habitualmente los repositorios digitales son **abiertos**, aunque algunos requieren



BIBLIOTECA
DE LA UNIVERSIDAD
COMPLUTENSE
DE MADRID

e.digital1.org.es/complutense



determinados permisos especiales de acceso. Por ejemplo, aquellos que pertenecen a algunas instituciones privadas pueden solicitar previa autenticación y, por tanto, solo los trabajadores y usuarios de dicha institución pueden acceder con sus credenciales.

La CRUE (Conferencia de Rectores de las Universidades Españolas) resume en la siguiente infografía los principales objetivos y beneficios de los repositorios institucionales:



Figura 1. Fuente: REBIUN. Red de bibliotecas universitarias españolas CRUE (e.digitall.org.es/crue).

Existen diferentes tipos de repositorios digitales según su propósito y contenido. Entre ellos, destacan:

- Repositorios institucionales:** creados por instituciones como las universidades (por ejemplo: **Universidad Complutense de Madrid**, **Universidad de Huelva**, o la **Universidad de Salamanca**), centros de estudios (por ejemplo: **Centro de Estudios Andaluces**) o de investigación (ejemplo: **Repositorio institucional del Consejo Superior de Investigaciones**) para almacenar y difundir su producción científica.

Los repositorios institucionales pueden agruparse según el tipo de recursos digitales que gestionen como, por ejemplo, los centrados en recursos educativos digitales u objetos digitales educativos (ODE). Estos recopilan y comparten materiales educativos como cursos en

REPOSITORIOS INSTITUCIONALES

Universidad Complutense de Madrid:
eprints.ucm.es

Universidad de Huelva:
rabida.uhu.es/dspace

Universidad de Salamanca:
gredos.usal.es

Centro de Estudios Andaluces:
e.digitall.org.es/est-andaluces

Repositorio Institucional del Consejo Superior de Investigaciones:
digital.csic.es



línea, presentaciones, guías, libros electrónicos, videos, objetos de aprendizaje como laboratorios o simuladores, entre otros, destinados a la educación formal e informal. Dos claros ejemplos son: el banco de recursos **EDIA** y el repositorio federado **PROCOMUN** del INTEF (Instituto Nacional de Tecnologías Educativas y de Formación del Profesorado) en el que se incluyen materiales creados y evaluados por equipos de docentes activos que los ofrecen como recursos educativos abiertos (REA).

- **Repositorios temáticos:** recogen la producción científica de un área de conocimiento a nivel nacional o internacional. Por ejemplo: **ArXiv** (física, matemáticas, informática, biología cuantitativa, finanzas cuantitativas, estadística, ingeniería eléctrica y ciencias de sistemas y economía), **BioRxiv** (ciencias de la vida), **CiteSeerX** (Ciencias de la computación y la Información), **Cogprints** (psicología, neurociencias, lingüística o informática), **Pubmed Central** (Biomedicina y Ciencias de la Vida), **RePEC** (Economía) y **SSRN** (ciencias sociales, economía, derecho, gobernanza corporativa y humanidades).
- **Repositorios centralizados:** recopilan la producción científica depositada «en abierto» por investigadores de distintas instituciones y de distintas áreas temáticas. Por ejemplo: **Zenodo**, que pertenece al programa europeo **OpenAIRE** y operado por **CERN**. **Zenodo** permite a los investigadores depositar artículos, datos y software para apoyar al almacenamiento de información científica y compartir dicha información con la comunidad internacional.
- **Repositorios de software:** almacenan y proporcionan acceso a software, programas y herramientas informáticas. Estos repositorios permiten a los desarrolladores compartir, colaborar y distribuir software de código abierto y de código cerrado. Entre los más conocidos están **GitHub** y **GitLab**.

**EDIA**e.digitall.org.es/edia**PROCOMUN**procomun.intef.es**REPOSITORIOS TEMÁTICOS****ArXiv:**arxiv.org**BioRxiv:**biorxiv.org**CiteSeerX:**csxstatic.ist.psu.edu/home**Cogprints:**e.digitall.org.es/cogprints**Pubmed Central:**ncbi.nlm.nih.gov/pmc**RePEC:**repec.org**SSRN:**ssrn.com/index.cfm**REPOSITORIOS CENTRALIZADOS****Zenodo:**zenodo.org**OpenAIRE:**openaire.eu**CERN:**home.cern**REPOSITORIOS DE SOFTWARE****GitHub:**github.com**GitLab:**about.gitlab.com

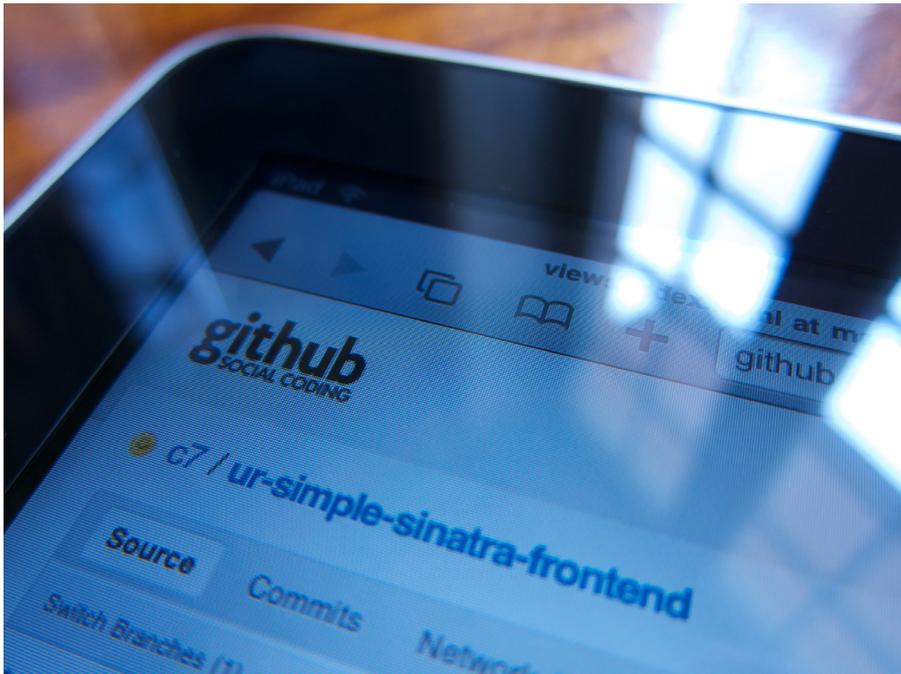


Imagen 2. Desarrollador de software accediendo como usuario al repositorio GitHub desde un iPad. Fuente: Peter Hellberg, 2010, Flickr ([e.digitall.org.es/github-flickr](https://www.flickr.com/photos/peterhellberg/5211111111/)). CC BY-SA 2.0

NOTA

Localización de repositorios

Para localizar repositorios se pueden utilizar directorios, buscadores y agregadores de repositorios. El Directory of Open Access Repositories (OpenDOAR), por la Universidad de Nottingham de Reino Unido y la Universidad de Lund de Suecia, es el más completo de todos.

v2.sherpa.ac.uk/opensoar/

Estrategia de búsqueda y filtrado en repositorios digitales

La búsqueda de información depende de varios factores, entre los cuales se encuentran las palabras clave, las estrategias de búsqueda y los operadores. También dependen de las herramientas de búsqueda, pues estas ofrecen alternativas específicas para búsqueda y filtrado de información. A continuación, se presentan estas características de los repositorios digitales atendiendo a su tipología y criterios de búsqueda transversales.

En los repositorios institucionales como, por ejemplo, los de las universidades y centros de estudios o investigación, se destacan los siguientes criterios de búsqueda.

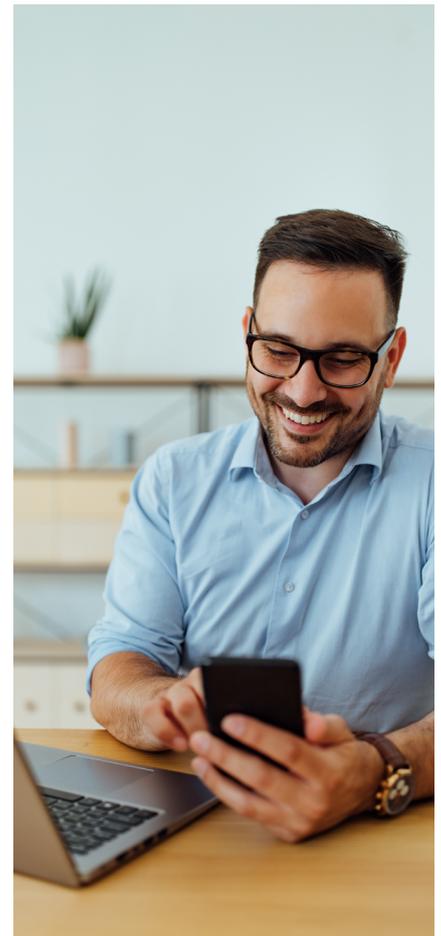


- **Tipología de fuentes:** también se les conoce en algunos repositorios digitales como **Comunidades** o **Colecciones**. En esta clasificación, se dispone de enlaces directos a documentos ya clasificados como: docencia, biblioteca, revistas de la universidad, investigación (tesis, artículos y ponencias producidas por docentes, estudiantes e investigadores de la universidad, pudiendo estar organizado por áreas de conocimiento, departamentos o áreas de la institución).
- **Por fecha de publicación:** para la búsqueda y filtrado de documentos según su fecha de publicación.
- **Por títulos:** en este criterio de búsqueda se pueden emplear dos alternativas: 1) si conoce el título del documento que desea localizar, puede colocarlo en el cuadro de búsqueda o, 2) por palabras clave, las cuales se pueden relacionar mediante el símbolo + del operador AND.
- **Por materias** (también se pueden encontrar como etiquetas, áreas o temas): se trata de escoger la materia o área del conocimiento que se desea buscar, por ejemplo: comunicación, psicología, informática, educación social, etc. Posteriormente, dependiendo de las funcionalidades del repositorio, puede usar otros filtros como año o autor.
- **Por autores:** para ello, deberá buscar el nombre y apellidos del autor. Recuerde que los metadatos de un autor pueden variar (un solo apellido, nombres y segundos nombres, un apellido e inicial del nombre, etc.). Por tanto, se recomienda usar diferentes criterios de búsqueda.

En algunos repositorios se incluyen criterios como **últimos añadidos** para acceder directamente a los documentos de reciente publicación.

En el caso de los **repositorios temáticos**, debido a su alcance internacional, puede que estén en su mayoría en inglés. En ellos puede encontrar y utilizar los siguientes criterios generales de búsqueda:

- **Búsqueda básica:** a través de palabras clave y operadores, fundamentalmente mediante el símbolo + del operador AND.





- **Búsqueda avanzada:** además de buscar por palabras clave y su interrelación con operadores, acotar la búsqueda por fechas, área de conocimiento, autores, título, resúmenes o DOI específico.

- **Búsquedas específicas:**

1 | Buscar y navegar por temas (*Subject search and browse*). Le permite seleccionar un área de conocimiento específico como física, química, ciencias sociales, ciencias de la vida, etc. Allí localizar documentos y ordenarlos en función de la fecha de producción.

2 | Búsqueda por fecha. Mediante un rango determinado.

En el caso de los buscadores centralizados, como **Zenodo**, se pueden realizar las dos primeras búsquedas descritas anteriormente acerca de los buscadores temáticos (básica y avanzada) con la especificidad de que, en la avanzada, se puede filtrar la información por tipología de acceso abierto y por tipo de documento (.mp4, .html, .pdf y .doc, entre otros formatos).

La búsqueda de información en repositorios digitales es muy similar a la de otros tipos de herramientas de búsqueda relacionadas como, por ejemplo, las bases de datos académicas, bibliotecas digitales y buscadores académicos.

En todos ellos debe recordar siempre que el éxito radica fundamentalmente en la estrategia de búsqueda y diversidad de fuentes, por lo que podrá comprobar la información que le ofrece cada resultado de búsqueda.

⚠ ATENCIÓN

Advertencias y errores frecuentes que se deberían evitar

Los errores frecuentes en la búsqueda de información en repositorios digitales se relacionan con:

- 1 | Errores ortográficos.
- 2 | Palabras clave solo en español.
- 3 | Uso inadecuado de operadores comilla ("") y AND (+).

⚠ ATENCIÓN

Advertencias y errores frecuentes que se deberían evitar

Recuerde que un documento electrónico puede tener un DOI (Digital Object Identifier), el cual es un identificador único y permanente para las publicaciones electrónicas y funciona como una URL (ubicación de una página web o de un archivo en Internet).

doi.org





Saber más

Recolectores y directorios de repositorios

A continuación, le sugerimos algunos enlaces con acceso a repositorios digitales en los cuales puedes localizar diversos documentos de libre acceso.

- Agregador nacional de repositorios científicos de acceso abierto (**RECOLECTA o Recolector de Ciencia Abierta**, recolecta.fecyt.es) creado por el Ministerio de Ciencia e Innovación de España. En este se encuentran más de 170 repositorios y más de 2.000.000 de documentos.
- Repositorios institucionales de las universidades españolas. rebiun.org/directorio-repositorios-institucionales
- Portal de acceso abierto a colecciones de patrimonio digital de archivos, bibliotecas y museos españoles (Hispana). hispana.mcu.es
- Biblioteca Virtual Miguel de Cervantes. Recopilación de libros en acceso abierto y otros materiales bibliográficos en idioma español. cervantesvirtual.com



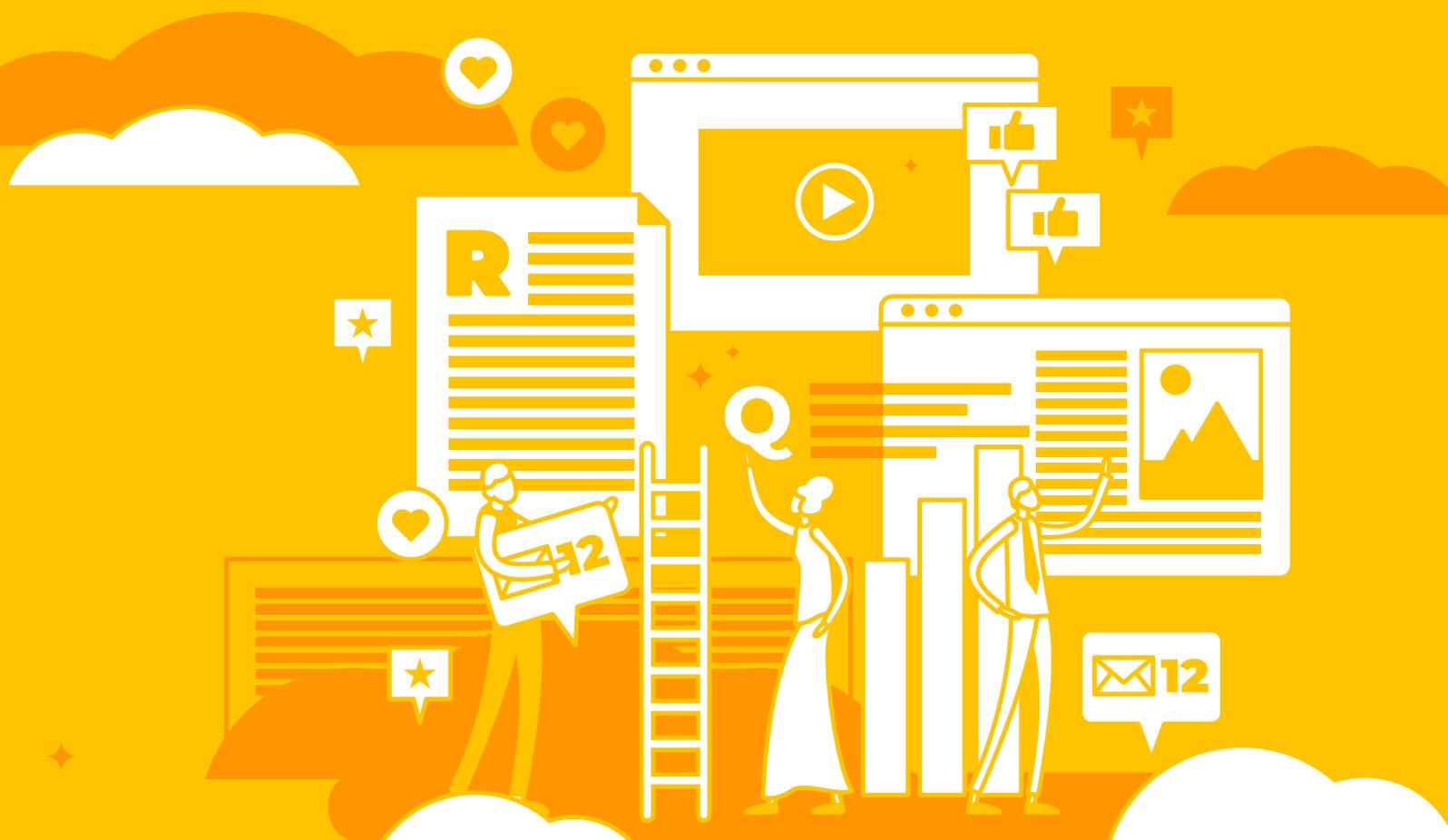


DigitAll

Búsqueda y gestión
de información y datos

1.2

**EVALUAR DATOS,
INFORMACIÓN
Y CONTENIDOS
DIGITALES**





Búsqueda y gestión
de información y datos

Nivel C2 1.2 Evaluar datos, información
y contenidos digitales

Sesgo y calidad de la información





Sesgo y calidad de la información

Visualización del sesgo informativo

Varios vídeos han explicado en qué consiste el sesgo informativo, cuáles son sus objetivos y consecuencias y cómo reconocerlo. En este documento se desea mostrar un ejemplo práctico de aparición de sesgo en contenidos reales y como reconocerlo.

▶ SESGO INFORMATIVO (AVANZADO)

Evaluación y determinación del sesgo informativo. Identificación y validación de fuentes no sesgadas. Sesgo informativo en canales de difusión (TV, radio, periódicos, RRSS). Herramientas para la valoración del sesgo.

e.digitall.org.es/A1C12C2V02

A finales de julio de 2023 una serie de graves incendios se extendió por múltiples países del Mediterráneo como Italia, Grecia o Argelia. Se puede considerar que el cambio climático que ha incrementado la temperatura del planeta es la causa principal de estas tragedias. Sin embargo, este tema tiene asociada una importante carga ideológica, por lo que el tratamiento de la información en los medios es diferente, tal y como se muestra en la figura 1.



Figura 1. Número y tamaño de titulares en referencia a los incendios en el Mediterráneo del julio de 2023.



En la página principal de un diario, el asunto se trata con varios titulares en los que se marca la extensión geográfica del problema. En otro, la información se limita a un solo país y aparece en un tamaño mucho menor. En ambos casos, el lenguaje utilizado es bastante neutral, aunque la expresión “causa estragos” puede considerarse un poco sensacionalista. En el ejemplo de la figura 2, sin embargo, el titular incluye un apoyo sin reservas a la teoría del cambio climático, como causa directa de los incendios.

🏠 > Noticias > Mundo

Incendios en Grecia | El cambio climático aviva el fuego en el Mediterráneo



Figura 2. Afirmación directa de que el cambio climático es la causa de los incendios en el Mediterráneo.

Ejemplo de comprobación de la calidad de la información

En los vídeos también se ha trabajado en describir pautas para evaluar la calidad de la información recogida en un determinado contenido. A continuación, se muestra un ejemplo de aplicación.

Al realizar una búsqueda sobre los incendios en el Mediterráneo en Facebook, se comprueba como en los primeros puestos aparecen vídeos publicados por diferentes medios de comunicación. Se puede considerar que todos ellos proceden de fuentes contrastadas, aunque no estarán libres de un posible sesgo informativo asociado a su línea editorial. También aparece la publicación de una usuaria que se identifica por su nombre y apellido (figura 3).

A la hora de valorar esta aportación, se debería comenzar con la autoridad de su creadora. Los datos de su biografía



son escasos. No se indica si se dedica a la información o si tiene relación con el escenario del suceso. A continuación, se observan el resto de sus publicaciones y se comprueba un claro sesgo ideológico. No aparecen amigos en su cuenta que es pública.

El siguiente paso es contrastar el contenido. Se realiza una búsqueda por texto y por imagen y se comprueba que coinciden con lo publicado dos días antes por la **web de National Geographic** (e.digitall.org.es/incendio). Se puede afirmar que el contenido es veraz, pero lo correcto hubiera sido indicar de donde se ha obtenido la información, o simplemente haber enlazado a la publicación original.



Figura 3. Publicación en Facebook sobre los incendios del Mediterráneo.

Además, el pie de figura en ese caso indica: *“Las altas temperaturas en Grecia provocan incendios cada año. En la imagen se muestra una zona de Evia afectada por el fuego en agosto de 2021”*, por lo que la imagen ni siquiera se corresponde con lo que estaba ocurriendo en ese momento.

Saber más

AllSides. Media Bias. allsides.com/media-bias

How do you tell when the news is biased? It depends on how you see yourself. e.digitall.org.es/biased

Evaluating sources. Aurian Library. guides.auraria.edu/evaluatingsources



Búsqueda y gestión
de información y datos

Nivel C2 1.2 Evaluar datos, información
y contenidos digitales

Listas y plataformas informativas





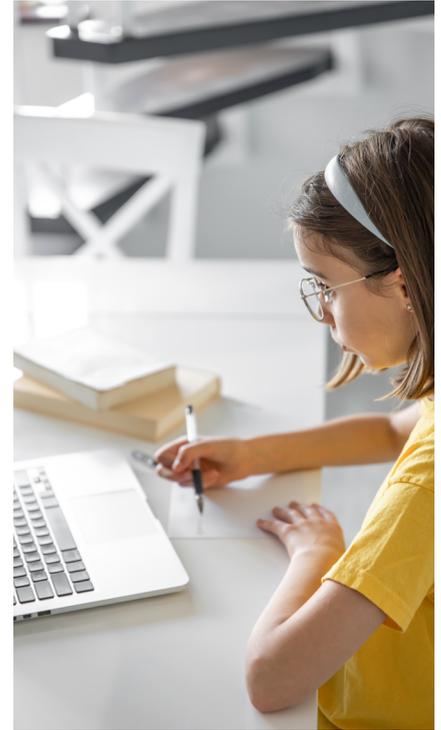
Listas y plataformas informativas

Introducción

Identificar listas y plataformas sesgadas o con información errónea es esencial en la era actual de sobrecarga de información.

El sesgo puede influir en la percepción del público, moldeando opiniones, creencias y comportamientos en función de una perspectiva particular, en lugar de hechos objetivos.

Una plataforma con inclinaciones puede, no solo omitir información vital, sino también presentar datos de manera que respalden una narrativa específica, lo que puede llevar a malentendidos y decisiones mal informadas



⚠ ATENCIÓN

En un mundo donde se confía cada vez más en las plataformas digitales para obtener información sobre eventos mundiales, políticas, salud y más, reconocer y entender estos sesgos es esencial para mantener una sociedad bien informada y capaz de tomar decisiones críticas y fundamentadas.

En el vídeo de Suscripción a listas y plataformas informativas se expuso una metodología para trabajar tanto con las listas como con las plataformas informativas, mediante la cual se minimizaba la creación de listas con información errónea o plataformas sesgadas.



SUSCRIPCIÓN A LISTAS Y PLATAFORMAS INFORMATIVAS

Listas y plataformas informativas. Utilidad. Listado y descripción de las principales. Exactitud y fiabilidad de listas y plataformas. Metodología de trabajo con las principales listas y plataformas.

e.digitall.org.es/A1C12C2V04



Metodología para identificar listas o plataformas sesgadas

Con las herramientas y estrategias adecuadas, es posible navegar en el vasto océano de información y discernir lo que es fiable de lo que no lo es. Para ello, se propone a continuación una metodología para identificar listas o plataformas que pueden conducir a interpretaciones erróneas:

- 1 | Investigación de la Fuente.** En este punto hay que tener en consideración tres aspectos: *¿Cuál es el origen de la fuente?* *¿Reputación de la fuente?* *¿Financiación de la fuente?* Una fuente desconocida o que oculta puede ser una señal de alerta. Por otro lado, si la fuente tiene un historial de difusión de información falsa o sesgada, es probable que continúe haciéndolo. Además, si existen fuentes de financiación, éstas pueden influir en el contenido, especialmente si hay conflictos de intereses.
- 2 | Verificación Cruzada.** Es necesario comparar la información presentada con otras fuentes. Si la mayoría se contradicen o no corroboran la información, es probable que esté sesgada o sea incorrecta.
- 3 | Análisis del Contenido.** Se deben considerar tres aspectos fundamentales, la emocionalidad, la generalidad y la falta de evidencia. Las fuentes sesgadas a menudo utilizan un lenguaje cargado emocionalmente o titulares sensacionalistas para atraer la atención. Además, las declaraciones que generalizan sin evidencia concreta, o que emplean lógica defectuosa pueden indicar sesgo o desinformación.
- 4 | Verificación de Hechos.** Se deben usar plataformas y organizaciones de verificación de hechos, como *Snopes* o *FactCheck.org*, para comprobar la veracidad de las afirmaciones realizadas.
- 5 | Evaluación de Imágenes y Medios.** Las imágenes, videos o gráficos pueden ser manipulados o sacados de contexto. Herramientas como *Google Reverse Image Search* o *TinEye* pueden ayudar a verificar la originalidad de las imágenes.



6 | Revisa los Comentarios y Discusiones. En muchas plataformas, especialmente en línea, los comentarios o discusiones pueden ofrecer perspectivas adicionales, correcciones o verificaciones de la comunidad.

7 | Evaluación de Algoritmos. Las plataformas que utilizan algoritmos para presentar información, como redes sociales, pueden favorecer cierto tipo de contenido. Ser consciente de cómo funcionan estos algoritmos puede ayudar a identificar sesgos inherentes.

8 | Educación Continua. Es necesario mantenerse actualizado sobre las técnicas comunes de desinformación y las tácticas utilizadas por fuentes sesgadas.

9 | Mantenerse Escéptico. Si algo parece demasiado bueno (o malo) para ser verdad, probablemente lo sea. Es recomendable no tomar la información al pie de la letra y cuestionar las fuentes, el contexto y la intención detrás del contenido.

El uso de esta metodología permitirá detectar, no solo lo que es inmediatamente falso, sino reconocer las sutiles inclinaciones que pueden desviar la comprensión clara de un tema. Al adoptar un enfoque crítico y sistemático, se fortalece la capacidad para navegar por el paisaje informativo, asegurando una interpretación más equilibrada y auténtica de los acontecimientos y datos presentados.

Saber más

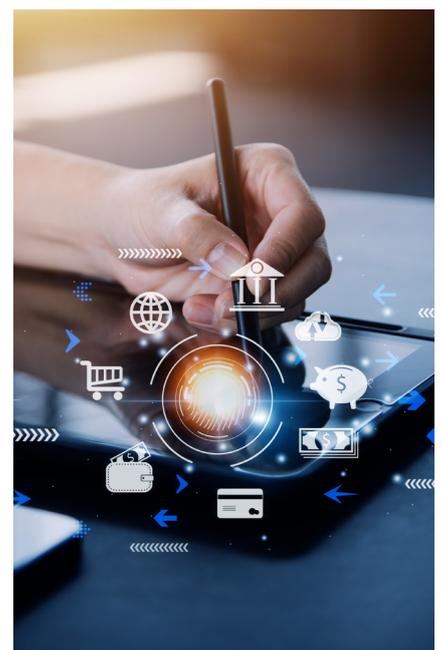
Google Reverse Image Search, o Búsqueda inversa de imágenes de Google, es una función del popular motor de búsqueda Google que permite a los usuarios buscar en la web utilizando una imagen en lugar de texto. Esta función es útil para identificar el origen de una imagen, encontrar versiones de mayor resolución, ver dónde más se ha usado una imagen o descubrir contenido relacionado.

e.digitall.org.es/reverse-search

Saber más

TinEye es una herramienta avanzada de búsqueda inversa de imágenes creada por una compañía canadiense. Destaca por ser uno de los primeros motores de búsqueda en internet que se basa en tecnología de reconocimiento de imágenes para realizar búsquedas, en lugar de depender de palabras clave, metadatos o marcas de agua.

services.tineye.com





Búsqueda y gestión
de información y datos

Nivel C2 1.2 Evaluar datos, información
y contenidos digitales

Perfiles en redes sociales





Perfiles en redes sociales

Evaluación de perfiles en redes sociales

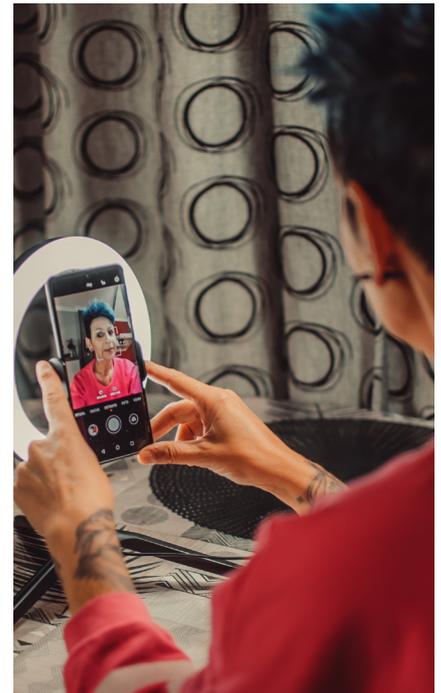
El seguimiento de perfiles en redes sociales es importante para garantizar la calidad de la información que se obtiene a través de ellas y para evitar malas experiencias con perfiles falsos. En el vídeo sobre herramientas de seguimiento de perfiles en redes sociales, se han presentado una serie de pautas para analizar la fiabilidad de las cuentas en estas aplicaciones.



SEGUIMIENTO Y MANEJO DE PERFILES EN REDES SOCIALES

En este vídeo se explica la identificación y seguimiento de perfiles en redes sociales, así como la realización de procesos de rastreo, trazabilidad y verificación de los perfiles.

e.digitall.org.es/A1C12C2V05



En este documento, se va a describir cómo aplicar estas pautas en tres de las principales redes sociales: Facebook, Instagram y X (antigua Twitter) localizando en ellas la información necesaria para ello.

- La **foto de perfil** que puede ser utilizada en una búsqueda en caso de resultar sospechosa.
- El **nombre de usuario**. Se debe comprobar si se corresponde con el generado automáticamente al crear la cuenta. En ese caso suele coincidir con las iniciales del nombre real y una cifra numérica.
- Los **datos biográficos** proporcionados al crear la cuenta. La ausencia puede ser sospechosa, así como una redacción extraña.
- El **número de seguidores y personas seguidas**. Hay que desconfiar si el segundo es muy superior al primero. Si, por el contrario, el número de seguidores es excesivamente elevado, se puede comprobar el número de interacciones asociados a los contenidos. Si se recibe un número escaso de "Me gusta" se puede sospechar que estos seguidores no son reales.
- **Número y tipo de interacciones** con contenidos no propios.
- **Datos temporales** como la fecha de creación de la cuenta o el ritmo de publicación.
- Las características de los **contenidos** y su coherencia a lo largo del tiempo.



Se parte del supuesto de que a una cuenta llega una solicitud de seguimiento de una persona desconocida y se quiere localizar toda la información mencionada antes de aprobarla o rechazarla.

Verificación de perfiles en Facebook

La foto de perfil, como en cualquier red social, es fácilmente localizable. Debajo de ella se localiza un menú (figura 1) donde localizar la mayor parte de la información.



Figura 1. Menú con la información incluida en un perfil de Facebook.

Sobre cada una de las *Publicaciones* se puede localizar el **tipo** y **número de interacciones**, así como la **fecha de publicación**. También se incluyen los **comentarios** recibidos. La pestaña *Información* proporciona los datos de la biografía y la de *Amigos*, los seguidores de esa cuenta, siempre que se haya decidido hacer pública esa información. La pestaña *Registro de visitas*, recoge las ubicaciones que se deseen compartir y la pestaña *Más*, aporta información sobre gustos, páginas que se siguen o interacciones realizadas. Estos últimos datos pueden ayudar a reconocer el tipo de usuario detrás de esa cuenta.

Verificación de perfiles en Instagram

Para obtener información de un perfil de Instagram, basta con pinchar sobre su foto y observar la principal información que aparece junto a ella (figura 2).



Figura 2. Información de cuenta en Instagram.



La información se puede ampliar pulsando los tres puntos que aparecen a la derecha y seleccionando **“información sobre esta cuenta”**. Se abrirá una ventana como la que aparece en la (figura 3).

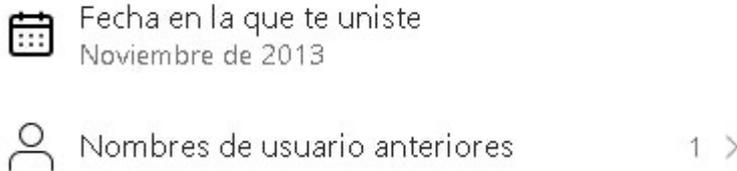


Figura 3. Más información sobre una cuenta de Instagram.

Aparece la fecha de creación y las veces que ha cambiado de nombre dicha cuenta. Demasiados cambios de nombre pueden informar de que una cuenta es falsa o robada.

Para obtener información sobre la interacción con los contenidos de la cuenta, basta pasar el ratón per encima de las imágenes para ver el número de interacciones. Para comprobar las fechas de publicación y los comentarios recibidos, se debe acceder a cada contenido. El análisis de las fotografías de la cuenta indicará si ha habido cambios drásticos en el tipo de publicaciones o si éstas no se inician en la fecha de creación de la cuenta, que puede indicar el borrado de contenidos antiguos para evitar detectar cambios de usuario sospechosos. En algunos casos, las publicaciones son numerosas solo en el momento de la creación para aparentar de que se trata de una cuenta normal y luego dejan de producirse.

Verificación de perfiles en X

En las cuentas de X la información básica de los perfiles también aparece debajo de la fotografía de perfil y el nombre de usuario (figura 4).



Figura 4. Datos básicos de un usuario en X.

Toda la información relacionada con los contenidos o post se localiza fácilmente junto a ellos. La fecha de publicación al inicio y la interacción con ellos en la parte inferior (figura 5).



Figura 5. Interacciones asociadas a un contenido de X.

También es fácil localizar toda la información de la interacción de la cuenta con sus propios posts y con los de otros en el menú de pestañas que aparece bajo la biografía (figura 6).



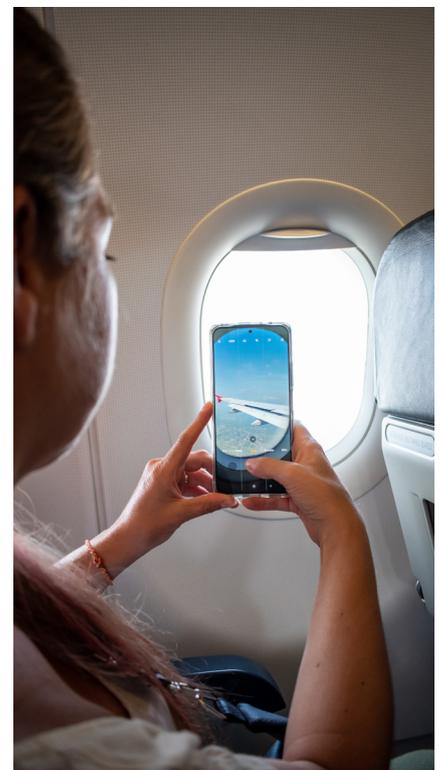
Figura 6. Actividad de un usuario en X

Conclusión

La interacción con usuarios desconocidos a través de las redes sociales debe siempre realizarse con prudencia. Como se menciona en los vídeos, las pautas proporcionadas no son infalibles, pero permiten confirmar posibles sospechas. En este documento se ha descrito cómo localizar información biográfica, temporal y de actividad de las cuentas en tres redes sociales.

Los certificados de verificación  de cuenta que pueden obtenerse en estas redes sociales son una garantía de la persona o marca que hay detrás

Estas pautas de comportamiento pueden ser similares a las de cuentas abandonadas. Este tipo de cuentas no suele realizar solicitudes de seguimiento y menos a desconocidos. En el caso de dudar si se conoce a ese usuario, todas las plataformas habilitan el envío de mensajes privados que permiten comprobarlo.



Saber más

Pequeña guía para detectar perfiles falsos en redes sociales. RETINA tendencias. e.digitall.org.es/guia-perfiles-falsos

Verificar tus cuentas en Facebook y en Instagram. e.digitall.org.es/verificar-cuentas

Four steps to identify fake accounts on social media. e.digitall.org.es/fake-account



Búsqueda y gestión
de información y datos

Nivel C2 1.2 Evaluar datos, información
y contenidos digitales

Uso ético de la IA para reforzar la opinión





Uso ético de la IA para reforzar la opinión

La ética en la inteligencia artificial

La inteligencia artificial ha irrumpido en la sociedad como una herramienta que proporciona beneficios a sus ciudadanos, oportunidades de crecimiento para las empresas y mejoras para el medio ambiente. Sin embargo, un mal uso de esta tecnología puede conducir a obtener resultados opuestos.

En un vídeo anterior hemos visto cómo se puede garantizar que los sistemas de inteligencia artificial se comporten de manera responsable y se alineen con los valores de la sociedad. En este documento se van a presentar diferentes ejemplos de cómo se avanza hacia este objetivo en diferentes ámbitos.



BUEN USO DE LA IA PARA REFORZAR LA OPINIÓN

Aprendizaje supervisado, no supervisado y aprendizaje por refuerzo. Método de prueba y error. Objetivo. Finalidad. El algoritmo Q learning. Open AI Gym. Aplicación ética del aprendizaje por refuerzo: administración pública, periodismo, fuentes de datos, rastreadores.

e.digitall.org.es/A1C12C2V06

⚠ ATENCIÓN

Los problemas de sesgo y falta de transparencia pueden dar lugar a un incremento de las desigualdades y a la pérdida de biodiversidad.

En el desarrollo de la inteligencia artificial ningún territorio o grupo social debe quedarse atrás por falta de acceso a la tecnología o por el desarrollo de soluciones que no sean suficientemente justas.

Uso ético de la inteligencia artificial en las empresas

Muchas empresas asumen una serie de principios básicos para garantizar la **aplicación responsable de la inteligencia artificial** (e.digitall.org.es/etica-ia) en su actividad. Estos principios incluyen:

- Reflejar la diversidad actuando contra el sesgo.
- Trasparencia que garantice asumir responsabilidades sobre los resultados de su funcionamiento.
- Garantizar la igualdad de condiciones.





Esta visión es compartida por las empresas participantes en el **LabS IA Responsable e Inclusiva** (e.digitall.org.es/labs-ia). Por ejemplo, **El Corte Inglés** y **Hoteles Meliá Internacional** aplican inteligencia artificial para personalizar la experiencia del usuario y maximizar las capacidades de sus trabajadores, o la empresa **Tendam**, que no identifica a los clientes a la hora de analizar sus datos, para adaptarse a sus preferencias sin poner en riesgo su privacidad. Todos estos ejemplos se recogen en el **Decálogo LabS IA responsable e inclusiva** (e.digitall.org.es/decalogo-labs).

Uso ético de la inteligencia artificial en instituciones públicas

Una de las principales obligaciones de las administraciones respecto a la inteligencia artificial es la de legislar para garantizar que su uso sea justo y no produzca daños de ningún tipo a la sociedad o el medio ambiente.

La legislación con relación a la inteligencia artificial debe incluir aspectos de anticipación, protección efectiva, seguimiento de impacto, ejecución y reparación.

Además, las administraciones en su actividad pueden aplicar soluciones de inteligencia artificial en beneficio de la comunidad. Por ejemplo, desarrollando **chatbots como el del Gobierno de Aragón** (e.digitall.org.es/chatbot-aragon) que faciliten información en abierto y el uso de herramientas de IA a todos los ciudadanos o iniciativas como **Policy Cloud** (policycloud.eu) para facilitar la digitalización y gestión de grandes cantidades de datos o uso de la nube en el ámbito público y privado.

Uso ético de la inteligencia artificial en los medios de comunicación

Como se mostró en el vídeo sobre “Automatización mediante IA del sesgo y refuerzo de opinión”, el uso de la inteligencia artificial está muy extendido en los medios de comunicación, pudiendo tener graves consecuencias para la imparcialidad en





la información y los puestos de trabajo. Sin embargo, este tipo de herramientas permiten un más rápido acceso a la noticia y aprovechar al máximo las capacidades creativas y de reflexión de los profesionales del periodismo.



AUTOMATIZACIÓN MEDIANTE IA DEL SESGO Y REFUERZO DE OPINIÓN

Automatización avanzada del sesgo informativo mediante el uso de la IA. Identificación del sesgo introducido. Análisis de técnicas de refuerzo de opinión mediante IA.

e.digitall.org.es/AIC12CIV06

Es por ello que se está trabajando en las redacciones para una implantación ética de esta tecnología, lo que se refleja en documentos como **Algoritmos en las redacciones: retos y recomendaciones para dotar a la Inteligencia Artificial de los valores éticos del periodismo** (e.digitall.org.es/redacciones), publicado por el **Consell de la Informació de Catalunya**.

Uso ético de la inteligencia artificial en educación

El ámbito de la educación tiene la responsabilidad de preparar a los futuros desarrolladores de la inteligencia artificial lo que incluye las consideraciones éticas. También de formar sobre ella a todos los ciudadanos mostrando sus ventajas, pero también sus riesgos, y así avanzar en la alfabetización digital de la ciudadanía.

Además, la inteligencia artificial permite avanzar en la personalización del aprendizaje, adaptando los currículos formativos a cada estudiante y así, incrementar sus oportunidades.

Sin embargo, la educación debe afrontar graves riesgos como la gestión responsable de los datos de sus estudiantes, así como evitar el sesgo en la selección de participantes en programas de estudios o en los contenidos impartidos.





Uso ético de la inteligencia artificial en la justicia

La *Carta Europea sobre el Uso Ético de la Inteligencia Artificial en los Sistemas Judiciales y su Entorno* (e.digitall.org.es/carta-etica) aprobada en diciembre de 2018 por la Comisión Europea para la Eficiencia de la Justicia, aborda los peligros del uso de herramientas de inteligencia artificial en los juzgados y tribunales. Sin embargo, estas herramientas son útiles a la hora de analizar gran cantidad de documentación y pueden ayudar a que los tiempos de preparación y desarrollo de un proceso se reduzcan, solucionando uno de los problemas más importantes del sistema judicial. Entre las pautas recogidas, destaca que en todo momento los ciudadanos deben ser conscientes de si alguna de las decisiones ha sido tomada por la inteligencia artificial y conocer el razonamiento seguido.

Uso ético de la inteligencia artificial por parte de las grandes compañías tecnológicas

Para terminar, es importante recordar que son las grandes compañías tecnológicas las que lideran el desarrollo de las aplicaciones de inteligencia artificial. Cuentan con una gran cantidad de datos que se incrementa día a día mediante el rastreo de la información que les proporcionan sus clientes a través de su actividad en internet y redes sociales. Por tanto, es una buena noticia que compañías como Google publiquen una serie de *principios a seguir en sus desarrollos* (e.digitall.org.es/principios-ia). Sin embargo, se debe también vigilar que ciertos usos como la eliminación de contenidos que puedan resultar ofensivos no se traduzcan en censura.





i *Saber más*

Aplicación responsable de la inteligencia artificial en las empresas.

e.digitall.org.es/etica-ia

Decálogo Labs IA responsable e inclusiva. SERES Fundación sociedad y empresa responsable. Everis. e.digitall.org.es/decalogo-labs

Recomendación sobre la ética en la inteligencia artificial. Acuerdo de la UNESCO adoptado el 23 de noviembre de 2021. e.digitall.org.es/ai-ethics

Ética de la Inteligencia Artificial en el periodismo: nos jugamos el futuro de la profesión. e.digitall.org.es/ia-periodismo

Algoritmos en las redacciones: retos y recomendaciones para dotar a la Inteligencia Artificial de los valores éticos del periodismo. Consell de la Informació de Catalunya. e.digitall.org.es/redacciones

Carta Europea sobre el Uso Ético de la Inteligencia Artificial en los Sistemas Judiciales y su Entorno. e.digitall.org.es/carta-etica

IA en Google: nuestros principios. e.digitall.org.es/principios-ia

Mark Zuckerberg: la inteligencia artificial censurará “los mensajes de odio” antes de que se publiquen. e.digitall.org.es/ia-mensajes-odio

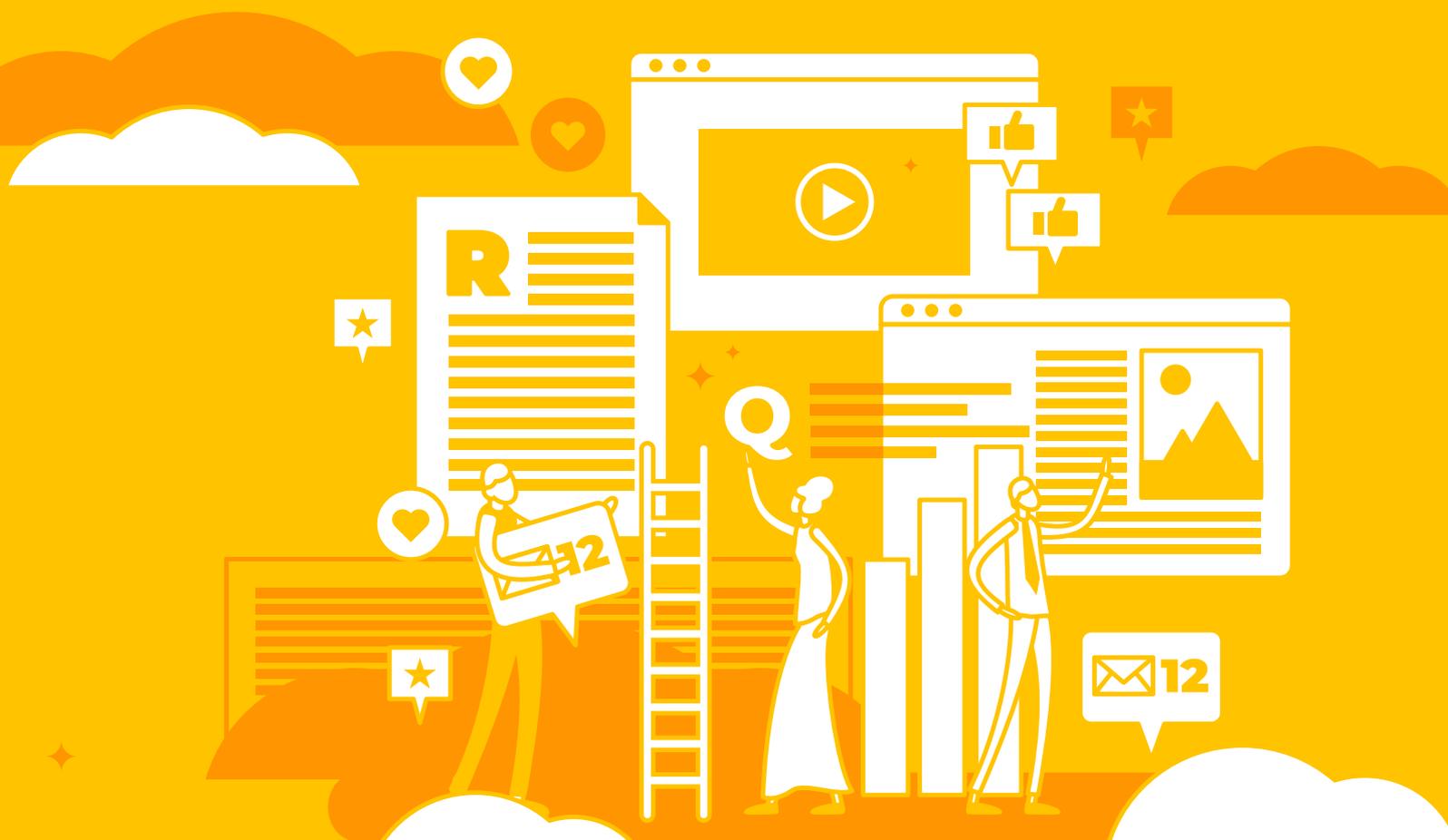


DigitAll

Búsqueda y gestión
de información y datos

1.3

GESTIÓN DE DATOS, INFORMACIÓN Y CONTENIDOS DIGITALES





Búsqueda y gestión
de información y datos

Nivel C2 1.3 Gestión de datos, información
y contenidos digitales

Aplicaciones de procesamiento de datos





Aplicaciones de procesamiento de datos

Herramientas para el procesamiento y organización de datos

En la siguiente tabla se indican las principales herramientas disponibles para el procesamiento y organización de datos de uso generalizado.

Concretamente, se plantean herramientas para la creación de encuestas y formularios, compartición de archivos y sincronización de ficheros entre dispositivos, videoconferencia y mensajería y, finalmente, gestión de calendarios.



GENERACIÓN DE DATOS EN ENTORNOS ONLINE

Explicar la existencia de herramientas para recopilar datos como formularios online (Google Forms, Formularios Microsoft ...), programación de reuniones (Doodle), compartición de archivos (Google Drive, Apple iCloud, JumShare) y plataformas de trabajo colaborativo (Google Chat, Microsoft Teams,...).

e.digitall.org.es/A1C13B2V02

TABLA 1. HERRAMIENTAS PARA EL PROCESAMIENTO Y ORGANIZACIÓN DE DATOS.

Herramienta	Utilidad y contexto	Dominio de aplicación
Google Forms	<ul style="list-style-type: none"> Nueve tipos de formularios para respuestas cortas, elección múltiple, y otros Visualización de datos en tiempo real Herramienta intuitiva y sencilla 	Creación de encuestas o cuestionarios en el navegador web o móvil sin necesidad de instalar ningún software especial
Doodle	<ul style="list-style-type: none"> Herramienta online gratuita Permite el envío automático y recepción de respuestas por correo electrónico de los participantes Permite visualizar a todos los participantes las respuestas de los demás intervinientes 	Herramienta online para establecer un conjunto de fechas y horas posibles para celebrar una reunión
Google Drive Apple iCloud Microsoft OneDrive	<ul style="list-style-type: none"> Sincronización automática de ficheros en la nube Multiplataforma en Windows y MacOS Interfaz web 	Compartición de ficheros entre dispositivos de un mismo usuario y también de carpetas con otros usuarios.
Google Workspace Microsoft 365	<ul style="list-style-type: none"> Edición en la nube, de forma colaborativa en tiempo real, de documentos de texto, hojas de cálculo y presentaciones Multiplataforma en Windows y MacOS Control de cambios 	Herramientas de trabajo colaborativo en la nube
Google Chat Microsoft Teams	<ul style="list-style-type: none"> Herramientas en la nube Integración con Google Workspace y Microsoft 365 Versiones gratuitas y de uso empresarial 	Interacción directa entre usuarios mediante videoconferencia y mensajería por Internet
Google Calendar Outlook Calendar	<ul style="list-style-type: none"> Gestión de múltiples calendarios compartidos Funcionamiento online y offline Priorización de tareas Interfaz web 	Gestión de calendarios de eventos, recordatorios, tareas y citas



Otras herramientas para el procesamiento y organización de datos

En la tabla siguiente se presenta un resumen de otras herramientas que se plantean como alternativas a las anteriores para el procesamiento y organización de datos. Algunas son de tipo genérico y otras presentan características atractivas por disponer de funcionalidades más avanzadas que las anteriores.

Concretamente, se plantean otras herramientas para la creación de encuestas y formularios, compartición de archivos de gran volumen, sincronización de ficheros entre dispositivos, videoconferencia y mensajería y, finalmente, gestión de calendarios.

TABLA 2. OTRAS HERRAMIENTAS PARA LA EDICIÓN DE CONTENIDOS DIGITALES.

Herramienta	Utilidad y contexto	Dominio de aplicación
Microsoft Forms	<ul style="list-style-type: none"> • Test de conocimientos, encuestas y cuestionarios. • Visualización de datos en tiempo real • Compartición de formularios anónima incluso fuera de una organización • Exportación a Excel 	Creación de encuestas o cuestionarios en el navegador web o móvil sin necesidad de instalar ningún software especial
HubSpot Forms	<ul style="list-style-type: none"> • Crear formularios de forma sencilla sin conocimientos técnicos • Integración de formularios en una o más páginas o mediante ventanas pop-up 	Creación de encuestas orientadas al mundo empresarial, para la captura de clientes potenciales
SurveyMonkey	<ul style="list-style-type: none"> • Herramientas avanzadas para el procesamiento de las respuestas • Herramientas avanzadas de monitorización del progreso e integración con otras herramientas 	Creación de encuestas avanzadas
Jumpshare WeTransfer MyAirBridje	<ul style="list-style-type: none"> • Herramientas online con interfaz web • Posibilidad de notificación sobre la descarga de los ficheros por parte de los destinatarios 	Compartición de ficheros de gran volumen usando un almacenamiento temporal en red
Box DropBox	<ul style="list-style-type: none"> • Sincronización automática de ficheros en la nube • Multiplataforma • Interfaz web 	Compartición de ficheros entre dispositivos de un mismo usuario y también de carpetas con otros usuarios.
Teamwork Quip Confluence	<ul style="list-style-type: none"> • Edición en la nube, de forma colaborativa, de documentos de texto, hojas de cálculo y presentaciones • Control de cambios 	Herramientas de trabajo colaborativo en la nube
Slack Skype WhatsApp Webex Meetings	<ul style="list-style-type: none"> • Herramientas en la nube • Versiones gratuitas y de uso empresarial 	Interacción directa entre usuarios mediante videoconferencia o mensajería por Internet
iCloud Calendar	<ul style="list-style-type: none"> • Entorno de gestión de calendarios de Apple • Integración con todo el ecosistema Apple 	Gestión de calendarios asociado al entorno de productividad Apple
Zoho Calendar	<ul style="list-style-type: none"> • Sincronización multiplataforma • Integración con otros gestores de calendarios 	Gestión empresarial avanzada de calendarios



Búsqueda y gestión
de información y datos

Nivel C2 1.3 Gestión de datos, información
y contenidos digitales

Estrategias avanzadas de recuperación en almacenamiento online





Estrategias avanzadas de recuperación en almacenamiento online

Introducción

En el contexto del almacenamiento online, estas estrategias avanzadas permiten a las organizaciones mantener la continuidad del negocio y la confianza del cliente, facilitando el revertir y restaurar con precisión.

La implementación de estas estrategias avanzadas en el almacenamiento online no solo brinda seguridad y confianza, sino que también ofrece resiliencia en el dinámico paisaje digital, asegurando que los datos estén siempre disponibles y protegidos.

A lo largo de las siguientes secciones, se detallarán las principales estrategias.



ALMACENAMIENTO, RECUPERACIÓN Y VERSIONADO DE LA INFORMACIÓN

Configuración del almacenamiento online para facilitar la recuperación de los datos. Tipos de recuperación de datos (enlazar con copias de seguridad). Importancia de la latencia. Versionado. Funciones avanzadas de gestión de versiones y recuperación.

e.digitall.org.es/A1C13C2V04

Puntos de Recuperación Objetivos (RPO) y Tiempos de Recuperación Objetivos (RTO)

Estos son dos indicadores clave en la recuperación de datos. El **RPO** se refiere a la cantidad máxima de datos que se pueden perder, mientras que el **RTO** alude al tiempo que se tarda en restaurarlos. Ambos deben definirse claramente según las necesidades del negocio.

Por ejemplo, si el RPO es de 2 horas, eso significa que puedes permitirte perder hasta 2 horas de datos en caso de un desastre. Si el RTO es de 4 horas, entonces necesitas ser capaz



de recuperar tus sistemas y datos dentro de ese período de tiempo después de una interrupción.

Los principales mecanismos y estrategias para el RPO son los siguientes:

1| Frecuencia de copias de seguridad. Si el RPO es corto, las copias de seguridad deben realizarse con mayor frecuencia. Por ejemplo, un RPO de 15 minutos requeriría backups cada 15 minutos.

2| Replicación continua. Para RPOs extremadamente bajos, las organizaciones podrían emplear replicación continua, donde los cambios en los datos se replican en tiempo real o casi en tiempo real a un sitio secundario.

Para el caso de los RTO, nos encontramos los siguientes:

1| Soluciones de recuperación de alta disponibilidad. Estas soluciones permiten a las organizaciones minimizar el tiempo de inactividad al tener sistemas redundantes en funcionamiento.

2| Recuperación en la nube. En algunos casos, las organizaciones pueden aprovechar la nube para una rápida recuperación, arrancando máquinas virtuales en un entorno en la nube mientras se soluciona el problema en sus instalaciones principales.

Es fundamental comprender que RPO y RTO son métricas que deben equilibrarse en función del coste y la importancia de los datos.

⚠ ATENCIÓN

Un RPO y RTO bajos suelen ser más costosos porque requieren tecnología más avanzada y soluciones de alta disponibilidad.

Así mismo, la estrategia en torno a RPO y RTO es esencial para la planificación de la continuidad del negocio y la recuperación de desastres. Las organizaciones deben determinar cuánto tiempo y datos están dispuestas a perder y planificar en consecuencia, invirtiendo en las tecnologías y procesos adecuados para garantizar que puedan cumplir con esos objetivos. Es un ejercicio que equilibra el riesgo, el impacto operativo y el coste.





Recuperación Granular

En lugar de restaurar sistemas completos, esta técnica permite recuperar elementos individuales (como un correo electrónico específico o un documento). Es especialmente útil cuando solo se necesita restaurar una pequeña parte de los datos.

Sin embargo, hay que tener en cuenta una serie de consideraciones técnicas que se muestran en la siguiente tabla:

CONSIDERACIONES TÉCNICAS SOBRE LOS COMPONENTES DE LA RECUPERACIÓN GRANULAR.

Componentes	Consideraciones técnicas
Backup	No todas las soluciones de <i>backup</i> admiten recuperación granular. Es esencial seleccionar herramientas que ofrezcan esta capacidad si es crítica para la operación.
Indexación	Para permitir una recuperación granular eficiente, los sistemas suelen requerir una indexación adecuada de los datos respaldados, lo que puede requerir recursos y tiempo adicionales durante las operaciones de backup
Almacenamiento	Las soluciones de recuperación granular a menudo requieren más espacio de almacenamiento, ya que pueden almacenar múltiples versiones del mismo archivo o datos para permitir puntos de restauración específicos.
Pruebas	Es esencial realizar pruebas regulares de recuperación granular para garantizar que los datos se puedan recuperar según lo previsto y que se comprenden los procedimientos

La recuperación granular aporta una serie de beneficios que podemos ver a través de aplicaciones típicas:

- **Bases de datos.** En contextos de grandes bases de datos, no siempre es práctico o eficiente restaurar toda la base de datos para recuperar un registro o tabla específica. La recuperación granular permite a los administradores de bases de datos recuperar elementos específicos sin afectar a toda la base.
- **Correo electrónico.** En sistemas de correo como Exchange, la recuperación granular podría permitir la restauración de un único correo electrónico o carpeta, sin afectar a todo el buzón o sistema.
- **Sistemas de archivos.** En sistemas de archivos, la recuperación granular puede permitir restaurar un solo archivo o directorio en lugar de toda una partición o unidad.
- **Eficiencia y tiempo.** Evita la necesidad de largos tiempos de restauración para grandes conjuntos de datos cuando sólo se necesita una pequeña parte.



- **Minimiza el impacto.** Reduce el riesgo de sobrescribir o afectar otros datos durante el proceso de recuperación.

Restauración desde *snapshots*

Algunos sistemas de almacenamiento online crean “instantáneas” o snapshots de los datos en momentos específicos.

Estas capturas permiten una restauración rápida a un punto exacto en el tiempo.

Funciona capturando el estado y la disposición de los datos sin necesidad de realizar una copia completa. Son especialmente útiles en entornos donde los datos cambian con frecuencia y pueden ser restaurados rápidamente en caso de fallo o corrupción.

Con la correcta implementación y gestión, puede ser un componente esencial de una estrategia de continuidad del negocio y recuperación de desastres.

A la hora de implementar esta estrategia, es fundamental tener en cuenta las siguientes consideraciones técnicas:

- 1 | Tecnología subyacente.** Los snapshots pueden basarse en hardware (a nivel de array de almacenamiento) o en software (como en soluciones de virtualización o sistemas de archivos específicos). La elección puede influir en la performance y la eficiencia.
- 2 | Dependencia de la cadena.** Los snapshots incrementales crean una cadena de dependencia. Si uno de los snapshots en la cadena se corrompe o se pierde, puede afectar la capacidad de restaurar desde snapshots posteriores.
- 3 | Gestión del espacio.** Aunque los snapshots son eficientes, el almacenamiento que requieren puede crecer con el tiempo a medida que se capturan más cambios. Es fundamental gestionar y, a veces, consolidar o eliminar snapshots antiguos.
- 4 | Coherencia de datos.** Es vital garantizar que los snapshots sean coherentes desde el punto de vista de la aplicación, especialmente para aplicaciones con transacciones, como bases de datos.

NOTA

La restauración desde snapshots es una herramienta poderosa en el mundo del almacenamiento y la recuperación de datos. Ofrece restauraciones rápidas y eficiencia en el almacenamiento, pero conlleva desafíos técnicos que deben ser manejados con cuidado.



5 | Impacto en el rendimiento y prestaciones del sistema.

Tomar snapshots puede afectar el rendimiento del sistema, especialmente si se toman con mucha frecuencia o en sistemas con alta I/O.

6 | Implementación:

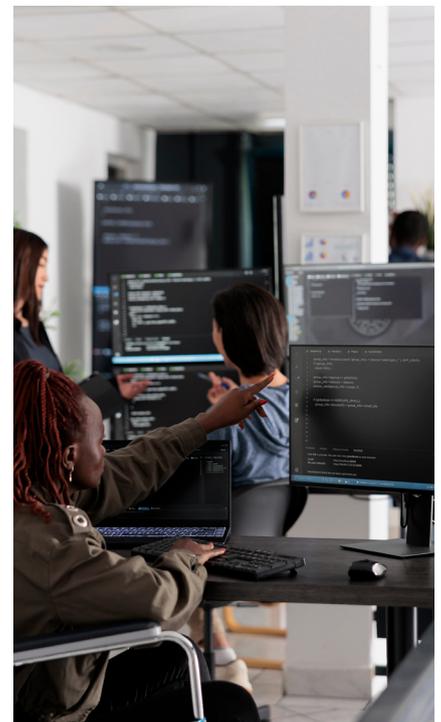
- **Planificación.** Decidir la frecuencia de los snapshots basándose en RPO (Punto de Recuperación Objetivo) y el valor de los datos.
- **Automatización.** Muchas soluciones modernas permiten la automatización de snapshots, garantizando que se tomen en intervalos regulares sin intervención manual.
- **Pruebas.** Al igual que con cualquier estrategia de backup, es esencial probar regularmente la restauración desde snapshots para garantizar que funcionen como se espera.

A cambio, la restauración desde snapshots nos aporta una serie de beneficios notables:

- **Restauración rápida.** Dado que los snapshots contienen información sobre los cambios desde la última instantánea, la restauración suele ser mucho más rápida en comparación con las restauraciones tradicionales.
- **Eficiencia de almacenamiento.** Los snapshots no requieren duplicar todos los datos cada vez que se realiza uno. En su lugar, sólo se almacenan las diferencias (o deltas) entre el estado actual y el anterior.
- **Frecuencia.** Debido a su eficiencia, se pueden tomar snapshots con mucha frecuencia, permitiendo puntos de restauración más granulares.
- **Protección contra fallos.** Pueden ser utilizados para revertir rápidamente los sistemas a un estado anterior en caso de fallos de software, corrupción de datos o ataques maliciosos.

Pruebas de recuperación

No basta con tener un plan de recuperación; es vital probarlo regularmente. Estas pruebas garantizan que el proceso de recuperación funcione como se espera y permite a las





organizaciones ajustar sus estrategias según sea necesario, siendo una parte esencial del proceso de continuidad del negocio y gestión de desastres.

Las Pruebas de Recuperación son ensayos técnicos planificados y ejecutados con el objetivo de evaluar la eficacia de los planes, protocolos y herramientas de recuperación implementados por una organización.

Estas pruebas garantizan que los sistemas, aplicaciones y datos pueden ser restaurados con éxito desde una copia de seguridad o punto de recuperación específico en caso de un fallo o desastre.

Entre sus principales beneficios, hay que destacar que las pruebas permiten validar que las estrategias de recuperación actuales son efectivas y pueden cumplir con los objetivos de nivel de servicio (SLA) definidos. Además, ayudan a identificar fallos, vulnerabilidades o carencias en los procesos de recuperación, y ofrecen una oportunidad práctica para que los equipos técnicos se familiaricen y entrenen en procedimientos de recuperación.

A la hora de implementar esta estrategia, hay que tener en cuenta las siguientes consideraciones técnicas:

- 1 | Entorno de pruebas aislado.** Es crucial contar con un entorno de pruebas que replique el entorno de producción pero que esté completamente aislado de él, para evitar interrupciones o corrupciones accidentales.
- 2 | Documentación detallada.** Cada prueba debe estar acompañada de una documentación clara que detalle los pasos, objetivos y criterios de éxito de la prueba.
- 3 | Revisión de dependencias.** Es esencial revisar y considerar todas las dependencias entre aplicaciones, bases de datos y servicios al diseñar escenarios de prueba.
- 4 | Automatización de pruebas.** En entornos complejos, es beneficioso utilizar herramientas que automatizan ciertos aspectos de la prueba, garantizando la consistencia y eficiencia.



⚠ ATENCIÓN

Los tipos más comunes de pruebas de recuperación son:

- **Prueba de restauración.** Verifica que los datos pueden ser restaurados desde una copia de seguridad.
- **Prueba de failover.** Evalúa la capacidad de cambiar a sistemas o sitios de recuperación secundarios.
- **Prueba de carga.** Evalúa la capacidad del sistema de recuperación para manejar cargas de trabajo similares a la producción.
- **Prueba de acceso remoto.** Verifica que los usuarios pueden acceder a aplicaciones y datos desde ubicaciones remotas en caso de un desastre.

Automatización de procesos de recuperación

Las soluciones más avanzadas permiten automatizar procesos de recuperación basados en criterios definidos, garantizando respuestas rápidas y reduciendo el margen de error humano.

La automatización de procesos de recuperación se refiere a la utilización de herramientas, scripts y sistemas que permiten restaurar datos y aplicaciones de forma rápida y consistente sin intervención humana significativa.

Esta estrategia se centra en minimizar el tiempo de inactividad y garantizar una recuperación eficiente. Es una inversión en resiliencia y eficiencia. Al garantizar respuestas rápidas y consistentes ante fallos, las organizaciones pueden proteger sus operaciones y su reputación en el mercado.

Su importancia técnica es notable, ya que la intervención manual en procesos críticos aumenta el riesgo de errores. La automatización elimina o reduce este factor. Además, permite una recuperación más rápida al activar procesos predefinidos al detectar fallos o problemas, y garantiza que cada recuperación sigue los mismos pasos y procedimientos, asegurando resultados predecibles.

Los beneficios tangibles que ofrece son una disponibilidad mejorada (los sistemas y datos vuelven a estar operativos en menos tiempo, reduciendo pérdidas económicas y de reputación) y una mayor eficiencia (la automatización permite liberar al personal técnico para que se concentre en otros desafíos críticos).

⚠ ATENCIÓN

La clave está en combinar tecnología, pruebas regulares y formación continua para mantenerse al día con las crecientes demandas y desafíos del mundo tecnológico.



Saber más

Los principales mecanismos y herramientas para implementar la automatización de los procesos de recuperación son:

- **Scripts de recuperación.** Escritos para ejecutar secuencias de comandos que restauran bases de datos, aplicaciones y sistemas operativos a un estado previo.
- **Orquestadores de Procesos.** Como Ansible, Puppet o Chef, que permiten diseñar flujos de trabajo automatizados para la recuperación.
- **Plataformas DRaaS (*Disaster Recovery as a Service*).** Ofrecen recuperación automatizada en la nube, activando réplicas de sistemas y datos cuando se detecta una interrupción.

Para conseguir una implementación efectiva, se recomienda:

- **Pruebas regulares.** Una estrategia automatizada no probada es tan mala como no tener estrategia alguna. Es fundamental realizar pruebas periódicas.
- **Monitorización.** Incorporar sistemas de monitoreo que detecten fallos y activen automáticamente los procesos de recuperación.
- **Documentación.** Aunque la recuperación esté automatizada, es esencial tener documentación detallada para el equipo técnico. Esto facilita la solución de problemas y la adaptación a situaciones imprevistas.
- **Escalabilidad.** La automatización debe ser capaz de adaptarse a crecimientos en datos y aplicaciones. Es vital considerar la escalabilidad al diseñar procesos automatizados.
- **Notificaciones.** Aunque el proceso sea automático, debe haber mecanismos para alertar al personal técnico sobre fallos y el estado de la recuperación.





Búsqueda y gestión
de información y datos

Nivel C2 1.3 Gestión de datos, información
y contenidos digitales

Inteligencia de datos



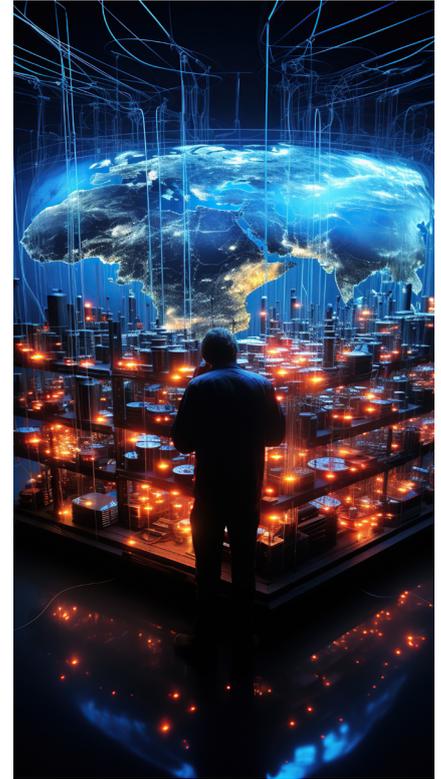


Inteligencia de datos

Introducción

Cualquier Organización produce datos de manera intensiva en su operativa diaria. Aprovechar correctamente los mismos almacenándolos, analizándolos y construyendo modelos predictivos puede aportar una visión en tiempo real del desempeño de la organización, medir su eficacia, optimizar procesos y tomar decisiones de planificación que mejoren sus carencias y aprovechen sus fortalezas de una Organización.

La Inteligencia de negocio permite la definición de estrategias empresariales a corto, medio y largo plazo basadas en la información externa e interna que una empresa puede almacenar, procesar y analizar. Este tipo de software se conoce como Sistemas de Ayuda a la Toma de Decisiones.



EXPLORANDO LA INTELIGENCIA DE DATOS

Definición de Sistema de Información, obtención de datos, principales tipos de Sistema de Información, formatos de exportación y crawling de datos.

e.digitall.org.es/A1C13C2V05

Los Sistemas de Ayuda a la toma de decisiones se apoyan en los Sistemas de Información de las Organizaciones. El siguiente vídeo recoge información sobre los Sistemas de Información.



TOMA DE DECISIONES BASADA EN DATOS

Importancia de la toma de decisiones y su automatización. Toma de decisiones en tiempo real vs puntual. Herramientas de toma de decisiones. Modelado de flujos de procesamiento de datos y aplicaciones.

e.digitall.org.es/A1C13C2V06



Herramientas de Escritorio

Algunas de las herramientas de toma de decisiones más empleadas para escritorio son KNIME y ProM Tools. Describimos, a continuación, algunas de sus características.

KNIME

KNIME (KoNstanz Information MinEr) es una plataforma gratuita de escritorio para minería de datos que, en su versión comercial, ha evolucionado a un entorno desplegable en cloud. La plataforma permite la automatización de flujos de tareas en multitud de ámbitos de aplicación.

KNIME permite construir modelos de aprendizaje utilizando un entorno visual desarrollado con el Interfaz de Desarrollo de Aplicaciones JAVA-Eclipse. La herramienta permite la confección de flujos de trabajo mediante nodos (que encapsulan distintos tipos de algoritmos) y flechas (que representan flujos de datos) que se despliegan y se pueden combinar interactivamente. Una vez leídos los datos con nodos específicos para fuentes muy variadas como XML, CSV o conexiones al API se pueden realizar con otros nodos especializadas manipulaciones de los datos para su preprocesamiento, análisis estadístico y creación de modelos predictivos.

El sencillo interfaz de KNIME ha facilitado su uso a profesionales no relacionados con el ámbito de la informática y se utiliza en investigación farmacéutica e inteligencia de negocio.

ProM Tools

La herramienta ProM es un entorno para minería de procesos orientado a la extracción de conocimiento sobre un proceso de negocio a partir de los datos que se producen (logs) en su ejecución.

La herramienta es extensible, gratuita, y está implementada en lenguaje Java. Esto facilita su extensión y aplicación a dominios concretos mediante el desarrollo de plugins.





Herramientas Cloud

La computación en la nube proporciona un ecosistema para el despliegue de aplicaciones que proporciona muchos beneficios en cuanto a capacidad computacional, escalabilidad, ratio coste beneficio y reducción de la huella de carbono. Los proveedores más importantes son Amazon AWS, Azure y Google. Para más información puedes consultar el siguiente video.



ALMACENAMIENTO ONLINE Y EN LA NUBE

Definición de almacenamiento en la nube. Tipos de nube. Tipos de provisionamiento de servicios y aplicaciones según el nivel de control que alcanza el proveedor (IaaS, SaaS, PaaS). Almacenamiento online. Principales proveedores y diferencias en los modelos que ofrecen.

e.digitall.org.es/A1C13C2V03

En este contexto, existen servicios proporcionados por los proveedores para el despliegue de aplicaciones que permitan escalar de forma adecuada sistemas de ayuda a la toma de decisiones. Comentamos a continuación los más relevantes en cuanto a impacto

Google DataProc

Google Dataproc es un servicio para ejecutar Apache Hadoop, Apache Spark, Apache Flink, Presto y multitud de herramientas software libre. El despliegue de soluciones para la ayuda a la toma de decisiones se puede realizar de forma segura a escala mundial e integrada con Google Cloud con coste reducido.



Google DataProc

cloud.google.com/dataproc



Figura 1. Google DataProc. Fuente: Imagen con licencia CCC obtenida de Wikimedia.
e.digitall.org.es/dataproc



BigQuery

BigQuery es un entorno cloud para el almacenamiento escalable de datos y que tiene como principal dominio de aplicación los Sistemas de Ayuda a la Toma de Decisiones de tipo empresarial. El entorno utiliza funciones integradas de Inteligencia artificial orientadas a la Inteligencia de Negocio y permite despliegues de estos sistemas a gran escala.



BigQuery

cloud.google.com/bigquery

Saber más

La industria de la computación en la nube ha crecido con una enorme rapidez. La elección de la solución más adecuada puede resultar compleja, dado que existen multitud de opciones y, además, cada proveedor utiliza una terminología distinta, con lo que la toma de decisiones acerca de qué plataforma cloud utilizar y qué tipo de despliegue realizar puede resultar un proceso muy complejo.

Puedes tener más información en el siguiente enlace:

kinsta.com/blog/aws-vs-azure/





DigitAll

Formación en
Competencias
Digitales



Coordinación General

Universidad de Castilla-La Mancha
Carlos González Morcillo
Francisco Parreño Torres

Coordinadores de área

Área 1. Búsqueda y gestión de información y datos

Universidad de Zaragoza
Francisco Javier Fabra Caro

Área 2. Comunicación y colaboración

Universidad de Sevilla
Francisco Javier Fabra Caro
Francisco de Asís Gómez Rodríguez
José Mariano González Romano
Juan Ramón Lacalle Remigio
Julio Cabero Almenara
María Ángeles Borrueco Rosa

Área 3. Creación de contenidos digitales

Universidad de Castilla-La Mancha
David Vallejo Fernández
Javier Alonso Albusac Jiménez
José Jesús Castro Sánchez

Área 4. Seguridad

Universidade da Coruña
Ana M. Peña Cabanas
José Antonio García Naya
Manuel García Torre

Área 5. Resolución de problemas

UNED
Jesús González Boticario

Coordinadores de nivel

Nivel A1

Universidad de Zaragoza
Ana Lucía Esteban Sánchez
Francisco Javier Fabra Caro

Nivel A2

Universidad de Córdoba
Juan Antonio Romero del Castillo
Sebastián Rubio García

Nivel B1

Universidad de Sevilla
Francisco de Asís Gómez Rodríguez
José Mariano González Romano
Juan Ramón Lacalle Remigio
Montserrat Argandoña Bertran

Nivel B2

Universidad de Castilla-La Mancha
María del Carmen Carrión Espinosa
Rafael Casado González
Víctor Manuel Ruiz Penichet

Nivel C1

UNED
Antonio Galisteo del Valle

Nivel C2

UNED
Antonio Galisteo del Valle

Maquetación

Universidad de Salamanca
Fernando De la Prieta Pintado
Pilar Vega Pérez
Sara Alejandra Labrador Martín

Creadores de contenido

Área 1. Búsqueda y gestión de información y datos

1.1 Navegar, buscar y filtrar datos, información y contenidos digitales

Universidad de Huelva

Ana Duarte Hueros (coord.)
Arantxa Vizcaíno Verdú
Carmen González Castillo
Dieter R. Fuentes Cancell
Elisabetta Brandi
José Antonio Alfonso Sánchez
José Ignacio Aguaded
Mónica Bonilla del Río
Odriel Estrada Molina
Tomás de J. Mateo Sanguino (coord.)

1.2 Evaluar datos, información y contenidos digitales

Universidad de Zaragoza

Ana Belén Martínez Martínez
Ana María López Torres
Francisco Javier Fabra Caro
José Antonio Simón Lázaro
Laura Bordonaba Plou
María Sol Arqued Ribes
Raquel Trillo Lado

1.3 Gestión de datos, información y contenidos digitales

Universidad de Zaragoza

Ana Belén Martínez Martínez
Francisco Javier Fabra Caro
Gregorio de Miguel Casado
Sergio Ilarri Artigas

Área 2. Comunicación y colaboración

2.1 Interactuar a través de tecnología digitales

Iseazy

2.2 Compartir a través de tecnologías digitales

Universidad de Sevilla

Alién García Hernández
Daniel Agüera García
Jonatan Castaño Muñoz
José Candón Mena
José Luis Guisado Lizar

2.3 Participación ciudadana a través de las tecnologías digitales

Universidad de Sevilla

Ana Mancera Rueda
Félix Biscarri Triviño
Francisco de Asís Gómez Rodríguez
Jorge Ruiz Morales
José Manuel Sánchez García
Juan Pablo Mora Gutiérrez
Manuel Ortigueira Sánchez
Raúl Gómez Bizcocho

2.4 Colaboración a través de las tecnologías digitales

Universidad de Sevilla

Belén Vega Márquez
David Vila Viñas
Francisco de Asís Gómez Rodríguez
Julio Barroso Osuna
María Puig Gutiérrez
Miguel Ángel Olivero González
Óscar Manuel Gallego Pérez
Paula Marcelo Martínez

2.5 Comportamiento en la red

Universidad de Sevilla

Ana Mancera Rueda
Eva Mateos Núñez
Juan Pablo Mora Gutiérrez
Óscar Manuel Gallego Pérez

2.6 Gestión de la identidad digital

Iseazy

Área 3. Creación de contenidos digitales

3.1 Desarrollo de contenidos

Universidad de Castilla-La Mancha

Carlos Alberto Castillo Sarmiento
Diego Cordero Contreras
Inmaculada Ballesteros Yáñez
José Ramón Rodríguez Rodríguez
Rubén Grande Muñoz

3.2 Integración y reelaboración de contenido digital

Universidad de Castilla-La Mancha

José Ángel Martín Baos
Julio Alberto López Gómez
Ricardo García Ródenas

3.3 Derechos de autor (copyright) y licencias de propiedad intelectual

Universidad de Castilla-La Mancha

Gabriela Raquel Gallicchio Platino
Gerardo Alain Marquet García

3.4 Programación

Universidad de Castilla-La Mancha

Carmen Lacave Rodero
David Vallejo Fernández
Javier Alonso Albusac Jiménez
Jesús Serrano Guerrero
Santiago Sánchez Sobrino
Vanesa Herrera Tirado

Área 4. Seguridad

4.1 Protección de dispositivos

Universidade da Coruña

Antonio Daniel López Rivas
José Manuel Vázquez Naya
Martíño Rivera Dourado
Rubén Pérez Jove

4.2 Protección de datos personales y privacidad

Universidad de Córdoba

Aida Gema de Haro García
Ezequiel Herruzo Gómez
Francisco José Madrid Cuevas
José Manuel Palomares Muñoz
Juan Antonio Romero del Castillo
Manuel Izquierdo Carrasco

4.3 Protección de la salud y del bienestar

Universidade da Coruña

Javier Pereira Loureiro
Laura Nieto Riveiro
Laura Rodríguez Gesto
Manuel Lagos Rodríguez
María Betania Groba González
María del Carmen Miranda Duro
Nereida María Canosa Domínguez
Patricia Concheiro Moscoso
Thais Pousada García

4.4 Protección medioambiental

Universidad de Córdoba

Alberto Membrillo del Pozo
Alicia Jurado López
Luis Sánchez Vázquez
María Victoria Gil Cerezo

Área 5. Resolución de problemas

5.1 Resolución de problemas técnicos

Iseazy

5.2 Identificación de necesidades y respuestas tecnológicas

Iseazy

5.3 Uso creativo de la tecnología digital

Iseazy

5.4 Identificar lagunas en las competencias digitales

Iseazy



El material del proyecto DigitAll se distribuye bajo licencia CC BY-NC-SA 4.0. Puede obtener los detalles de la licencia completa en: <https://creativecommons.org/licenses/by-nc-sa/4.0/deed.es>