

## 4. LA BÚSQUEDA DE INFORMACIÓN

### Cinco formas de buscar información

La distinción entre información y documentos es importante. Siempre necesitamos información, pero a veces buscamos más directamente la información misma donde o como aparezca y otras veces buscamos documentos concretos que suponemos la contienen o que nos interesa en particular conocer.

Podríamos hablar de cinco grandes tipos o formas de buscar información / documentos. Evidentemente, en la práctica, suelen darse reunidos, entremezclados, o aparecer unos como etapas para avanzar hacia los otros en una pesquisa más o menos compleja, combinada.

- 1) **Aproximarnos a un tema.** Mediante fuentes secundarias como tratados, manuales, libros de texto, enciclopedias, diccionarios, *handbooks*, *reviews*, buscadores y portales de Internet, etc
- 2) **Obtener datos concretos, específicos.** Consiste en conseguir "información factual". Mediante: enciclopedias, *handbooks*, listas de términos, diccionarios, vocabularios, catálogos comerciales, directorios, anuarios, fuentes numéricas, repertorios estadísticos, buscadores, directorios y portales de Internet, etc. En el mundo impreso, estas fuentes suelen llamarse "de referencia rápida"
- 3) **Identificar correctamente un documento, localizar dónde está publicado y/o disponible y conseguirlo.** Consiste en confirmar los datos de una obra, buscar dónde está en una biblioteca, completar datos imprecisos o insuficientes de la cita de un artículo, averiguar en qué revista se ha publicado y en qué biblioteca está, identificar normativa legal o técnica y conseguirla, averiguar dónde se ha presentado una tesis y conseguirla, precisar los datos o situación administrativa de una patente que conocemos por indicios, etc.... No se trata de una búsqueda por temas sino de ir por un documento concreto: identificarlo correctamente, localizar dónde está publicado y/o depositado y conseguirlo. Mediante : bases de datos, catálogos de bibliotecas, webs y similares.
- 4) **Averiguar qué documentos hay sobre un tema publicados o disponibles hasta la fecha.** Consiste en hacer una búsqueda temática para conocer con cierta profundidad la información que hay al respecto y poder luego conseguirla y estudiarla si es necesario. Es la forma más profunda y

exhaustiva de buscar documentación. Se lleva a cabo, típicamente, antes de iniciar una investigación, como primer paso, para estar al corriente del estado de conocimientos disponibles y no perder el tiempo repitiendo un trabajo o practicando el seguidismo, para controlar el contexto de nuestra indagación. A esta técnica, como servicio de documentación, se la conoce como Búsqueda retrospectiva de información (BRI). Los documentos primarios que buscamos con más frecuencia son artículos de revista, pero también informes, monografías, congresos, patentes, tesis, etc. Mediante bases de datos documentales especializadas, de índices y resúmenes, generalmente, y de documentos especiales, como tesis, patentes, informes, etc

- 5) **Permanecer informados regularmente de documentos o datos nuevos.** Consiste en conseguir mediante un servicio de documentación o mayorista de información, a través de procedimientos más o menos automatizados, datos de las novedades que aparezcan sobre un asunto o con unas características, generalmente en forma de listas de documentos, por ejemplo: innovaciones en un tipo de procesos industriales, artículos de un tema, contenidos nuevos de unas revistas, etc.. También se conoce esta técnica o servicio como **alertas** o Difusión selectiva de información (**DSI**). El **RSS** es un nuevo formato de documentos utilizado para la sindicación de contenidos que se ha popularizado muy deprisa como alternativa a las alertas en la red.

### **La estrategia de búsqueda, nueve pasos en búsquedas temáticas complejas**

Un paso previo: aclarar bien qué es lo que necesitamos realmente, en todos sus aspectos, centrar o definir bien la necesidad de información, la pregunta que hemos de formular a un sistema de información o BD:

- el tema que nos interesa en general
- bajo qué aspectos nos interesa y bajo cuáles no
- lengua de los documentos
- el tipo de documentos primarios que buscamos
- la época de los documentos que buscamos

Un vez que tenemos claro lo que buscamos podríamos seguir estos nueve pasos:

1. **Identificar conceptos** importantes que describan la información que buscamos

2. Sintetizarlos en **palabras clave**. Traducirlas al inglés.
3. Determinar si hay **sinónimos, términos relacionados u otras variaciones** que sea conveniente introducir en la búsqueda
4. Determinar cómo vamos a **combinar los términos**: Operadores booleanos y truncamiento
5. Elegir la **base de datos o el motor de búsqueda**.
6. Familiarizarse con las **características de la base de datos o motor de búsqueda** elegida. : ayudas, búsqueda avanzada, FAQ
7. **Expresar la búsqueda** en función de las posibilidades de la Base de datos
8. **Evaluar** los resultados
9. **Replantear** la estrategia en los pasos 2-4

Encontrar una referencia que se ajusta a nuestra demanda es de hecho el mejor modo de precisar y mejorar nuestra búsqueda, usándola como ejemplo podemos encontrar las que tienen los mismos términos descriptores, clasificaciones, etc.

Es importante llevar un registro de las búsquedas, anotando la fecha, la base de datos o el motor de búsqueda utilizado, la formulación de la búsqueda, el número de resultados obtenidos.

### **Buscar en Internet - GOOGLE**

**Internet y la World Wide Web han tenido un profundo efecto en la forma en que se busca información y se investiga.**

Hay acceso a más información y más rápido. Pero surgen nuevos problemas: encontramos demasiada información, una gran parte de esta información no es la que buscamos, no sabemos encontrar información disponible en la red o tardamos demasiado en localizarla. Algunas veces los resultados que ofrece la búsqueda son frustrantes. Antes de buscar hay que plantearse que tipo de información buscamos, quizás la Biblioteca tenga una suscripción a bases de datos especializadas en el tema que nos ocupa, o puede que haya seleccionado en su página web direcciones interesantes que nos faciliten la búsqueda.

- **Motores de búsqueda y directorios:**

Existen dos tipos diferentes de buscadores en Internet: los buscadores de índices temáticos y los servidores o motores de búsqueda. Los primeros contienen la información ordenada y clasificada por temas, buscamos recursos navegando en un índice

jerarquizado. Se elaboran manualmente. Los motores de búsqueda son sistemas de búsqueda por palabras clave, incorporan automáticamente páginas web mediante "robots" de búsqueda en la red. Google es motor de búsqueda pero dispone de un directorio de índices temáticos.

K. Hartman y E. Ackermann en *Searching and researching on the Internet & the www* recomiendan unas **pautas para elegir la mejor herramienta de búsqueda en la red**:

<b>Directorios y bibliotecas virtuales</b>	<ul style="list-style-type: none"> <li>▪ Contienen listados de temas de recursos seleccionados, normalmente clasificados jerárquicamente. Suelen haber sido evaluados cuidadosamente.</li> <li>▪ Suelen contener <b>enlaces a bases de datos especializadas</b></li> <li>▪ <b>Más útiles para encontrar:</b> una visión de conjunto sobre un tema recursos evaluados datos. Bases de datos especializadas</li> </ul>
<b>Motores de búsqueda</b>	<ul style="list-style-type: none"> <li>▪ Intentan indizar toda la web, la mayoría son bases de datos de texto completo, enormes y actualizadas regularmente</li> <li>▪ <b>Más útiles para encontrar:</b> Información multidisciplinar, con múltiples aspectos Información poco clara o poco conocida Grandes cantidades de información sobre un tema particular desde diferentes perspectivas</li> </ul>

**¿Cómo funcionan los motores de búsqueda?** Recolectan información. Utilizan programas – robots o arañas – que localizan en Internet recursos y los cargan en una base de datos, en la que se puede buscar utilizando el motor de búsqueda correspondiente. Google recorre todas las páginas web utilizando los enlaces para ir de un sitio a otro, recopila la información y la incluye en sus índices. Un programa robot rastrea la red y actualiza periódicamente todo su índice con sitios nuevos y eliminando los obsoletos.

Ordenan los resultados. La ordenación de resultados de la búsqueda es fundamental para que un motor de búsqueda sea eficaz. Google utiliza un sistema de clasificación de páginas web denominado PageRank, que emplea una extensa estructura de vínculos como un indicador del valor de una página individual para priorizar su importancia.

**Internet invisible.** Los motores no indexan toda la web. Hay información inaccesible para ellos (invisible web o *hidden Internet*), por ejemplo el contenido de las grandes bases de datos bibliográficas especializadas.

- **GOOGLE.**

Con el tiempo Internet se ha convertido en algo más complejo, millones y millones de páginas se han acumulado en la red. Los primeros buscadores se volvieron lentos y apareció una nueva generación de motores de búsqueda capaz de gestionar esta ingente masa de información. Google rápidamente tuvo la aceptación de los internautas y se convirtió en el número uno, hoy evoluciona a velocidad de vértigo ofreciendo constantemente nuevos servicios y herramientas. Hay otros motores, pero nos vamos a ceñir a Google, sin duda hoy es el más eficaz y universal, mañana quién sabe, tal vez sea historia.

### **Explotando las posibilidades de Google para buscar información científica:**

Google ofrece la sencillez de uso como una de sus mejores cualidades, ¿quién no es capaz de buscar en Google?, pero un poco de lectura de lectura puede ser muy provechoso para explotar todas sus posibilidades. La búsqueda será más eficaz si explotamos todas las opciones de búsqueda de las herramientas que manejamos.

<b>Búsqueda de libros</b> <b>Buscador académico</b> <b>Búsqueda avanzada</b> <b>Ayuda general de Google</b> <b>Búsqueda de blogs</b> <b>Google maps</b> <b>Alertas de Google</b> <b>Noticias</b> <b>Directorio</b>	<b>Son 9 herramientas de Google que es conveniente conocer</b>
--	--

Google, y otros buscadores, extiende sus redes hacia el territorio de la información bibliográfica. En algunos casos colabora con las bases de datos más importantes, por ejemplo:

- **OPen worldCat.** Proyecto de acceso abierto al catálogo de la red OCLC de Bibliotecas, el mayor catálogo bibliográfico del mundo. Con la colaboración de Google y Yahoo.
- **CrossRef search.** Proyecto en colaboración con editores de revistas académicas electrónicas.

Google ha desarrollado dos nuevas herramientas cada vez más competitivas y especialmente interesantes en el ámbito académico que nos ocupa:

**Búsqueda de libros** <http://books.google.es>. Los libros que se encuentran mediante la Búsqueda de libros de Google provienen de dos fuentes: editores y bibliotecas. Google proyecta incluir las colecciones de algunas bibliotecas en la Búsqueda de libros de Google y ofrecer información sobre ellos, como la ficha de catálogo y, en determinados casos, algunos fragmentos, es decir, frases donde aparece su término de búsqueda en contexto. Dos grandes Bibliotecas universitarias españolas se han sumado a este proyecto.

**Búscador académico** <http://scholar.google.es>. Es un servicio de búsquedas tipo académico, es decir, permite realizar búsquedas específicas en el campo académico para encontrar información relevante y material relacionado con la educación, universidades, investigación, tesis, bibliografías, etc. Este servicio incluye el análisis de citas y la exportación de registros a refworks y otros gestores de referencias bibliográficos.

### **¿Buscar en Google o en la Biblioteca?**

No son opciones excluyentes, combinemos una nueva y potentísima herramienta de búsqueda de información con los servicios que ofrecen las Bibliotecas. De momento son entidades que conviven, cada vez más relacionadas: Google utiliza los recursos de las Bibliotecas en su estrategia de conseguir el acceso a todo y los bibliotecarios utilizan Google para acceder a recursos propios y externos. Desechemos la idea de la Biblioteca tradicional, hoy en las universidades españolas las Bibliotecas combinan los recursos electrónicos con los fondos impresos, desde hace años están presentes en Internet y sus servicios y colecciones son cada vez más virtuales. Sumemos, no restemos, conozcamos las ventajas y carencias de ambos, utilizando en cada caso el mejor recurso para encontrar información de calidad en el tiempo adecuado.

El factor tiempo es importante. Es posible que en muchos casos a través de Google podamos encontrar todo lo que necesitemos. Pero, cuando se conoce su funcionamiento, una base de datos especializada, INSPEC por ejemplo, permite resolver un problema de información con mucha más rapidez. Con otros factores añadidos: calidad y cantidad de resultados, facilidades para desarrollar la búsqueda, enlaces a texto completo suscritos por la Biblioteca...

***Is Google enough? Comparison of an internet search engine with academic library resources*** Jan Brophy and David Bawden. Department of Information Science, City University, London, UK

The purpose of the study was to compare an internet search engine, Google, with appropriate library databases and systems, in order to assess the relative value, strengths and weaknesses of the two sorts of system.

**Findings – Google is superior for coverage and accessibility. Library systems are superior for quality of results. Precision is similar for both systems. Good coverage requires use of both, as both have many unique items. Improving the skills of the searcher is likely to give better results from the library systems, but not from Google**

Cop. Emerald Group Publishing Limited

[www.emeraldinsight.com/0001-253X.htm](http://www.emeraldinsight.com/0001-253X.htm)

## **Buscar en Bases de datos especializadas**

### **• Conocer las bases de datos**

Hay cientos de bases de datos accesibles en la red. Pero entre todas podemos decir que normalmente sólo una, dos a lo sumo son las adecuadas para resolver con eficacia y en el menor tiempo posible una búsqueda de información en un área temática o para un tipo de documentos. Debemos informarnos sobre cual utilizar, acudiendo a un experto, o a una guía sobre fuentes de información.

Hay cierta información sobre las bases de datos que conviene conocer:

- Entidad responsable de su producción.
- Prestigio en su campo. Buscamos la mejor disponible : INSPEC o Mathscinet son buenos ejemplos en su campo.
- Qué información contiene : ¿registros bibliográficos, sumarios, texto completo?
- Materia y tipo de documentos que cubre
- Cobertura temporal y actualización
- Acceso. ¿En línea, gratis o a través de redes como las bibliotecarias o universitarias...?

Con esta información sabremos si la base de datos es la adecuada. Una vez elegida es importante conocer cómo funcionan las herramientas de búsqueda del producto. **La recuperación**

**electrónica de información** consiste en seleccionar, aislar, un conjunto de referencias de documentos que cumplen unas determinadas condiciones de entre el universo total de referencias incluidas en la base de datos. Esas condiciones son la ecuación o estrategia de búsqueda y consisten en la presencia de unas secuencias de caracteres, términos de búsqueda, relacionados de determinadas maneras en unos determinados campos de las referencias. **Es un problema de Álgebra de Boole, lógica de conjuntos**

- **Tres gigantes en el mundo de las bases de datos**

**INSPEC.** Imprescindible. Inspec es el servicio de información bibliográfica líder en el campo de la literatura científica y técnica de Física, ingeniería eléctrica, comunicaciones e informática. Proporciona referencias de artículos de publicaciones especializadas de todo el mundo, fundamentalmente revistas (4.000), además de congresos, libros, informes y tesis. Monumento histórico de la documentación científica producido por la británica Institution of Electrical Engineers (IEE), ahora Institution of Engineering & Technology (IET). INSPEC (que significa Information Service in Physics, Electrotechnology and Control) es hoy una base de datos con ocho o diez millones de referencias. Pero lo más interesante es que su primera versión impresa, como Science Abstracts, empezó a publicarse en 1898. Por tanto, recoge gran parte de la ciencia e ingeniería contemporáneas: todo lo centrado en la Física y en la tecnología basada en la Física o relacionada con ella (electricidad, electrónica, comunicaciones, informática, etc.)

Impresiona pensar que cuando INSPEC nació reinaba la Física clásica de Newton, a la que se había agregado sólo veinticinco años antes el electromagnetismo de Maxwell. Se estaba investigando la radiactividad (Röntgen había descubierto los rayos X, etc.) y se acababa de descubrir el electrón (Thomson, 1897). Todavía tenía que llegar 1900 para que Planck inaugurara la mecánica cuántica, y sobre todo esperaba el 1905 de Einstein... **Es decir, INSPEC nació con la nueva Física.**

**MATHSCINET.** Imprescindible. Si Inspec es líder en el campo de la física Mathscinet lo es el de las matemáticas. Versión en línea de la prestigiosa base de datos de la American Mathematical Society, MathSci, la de sus publicaciones Current Mathematical Publications (índice temático de datos bibliográficos de publicaciones recientes y de próxima aparición) y Mathematical Reviews (reseñas de esas publicaciones). Contiene resúmenes firmados de artículos, libros,



comunicaciones de congresos, etc de Matemáticas, Ciencias de la computación, Estadística, Econometría.

Mathscinet cubre las publicaciones más importantes, con más de dos millones de registros fundamentalmente de libros y artículos. Cuenta con una característica excepcional, continuando con la tradición de la publicación impresa *Mathematical review* , publicada por primera vez en 1940, las referencias bibliográficas incluyen una reseña crítica, elaborada por un revisor, seleccionado por la revista entre matemáticos expertos en el tema.

Para buscar información sobre literatura anterior a 1940 una alternativa es la base de datos editada por la European Mathematical Society **Zentralblatt Math**. Su cobertura se extiende hasta 1868.

**Bases de datos del Web of Knowledge.** El ISI Web of Knowledge es una pasarela que integra el acceso a algunas de las más importantes bases de datos de información bibliográfica en cualquier ámbito de la *ciencias, ciencias sociales y humanidades* : Science Citation Index, Current Contents, Journal Citation Reports, Derwent entre otras. No llegan a cubrir la literatura en física y matemática de forma exhaustiva como lo hacen las bases de datos especializadas – Inspec y Mathscinet - pero ofrecen herramientas avanzadas muy útiles para rastrear y organizar la información. Las bases de datos del Institute for Scientific Information siguen siendo líderes en el campo del análisis de citas, y única referencia de momento en valoración del impacto de las revistas científicas.

Acceso a la Web of Knowledge  
para todas las instituciones  
públicas de investigación

Un servicio público a la comunidad  
científica e investigadora española



### **Buscar en repositorios institucionales. Acceso abierto a las publicaciones científicas**

En el más amplio sentido de la expresión y con respecto a las publicaciones científicas, el OA (Open Access) está vinculado a las

iniciativas o proyectos que favorezcan y promuevan el acceso abierto, libre y sin restricciones a los trabajos publicados por la comunidad científica. Los términos "libre" (free) y "abierto" (open) no son equivalentes; el primero es sinónimo de gratuito, mientras que "abierto" incluye el acceso sin barreras económicas y reivindica los derechos del autor sobre sus artículos

Los Repositorios institucionales son archivos digitales accesibles a través de Internet que reúnen la producción intelectual de una Institución como puede ser un centro de investigación o una Universidad. Incluyen artículos de revistas, ponencias, capítulos de libros o cualquier otra forma de publicación científica.

- **Acceso abierto**

**El problema.** El acceso a todo el creciente volumen de publicaciones científicas necesarias para la investigación y el aprendizaje está en una situación de riesgo. Las tendencias de la edición científica - especialmente los modelos de negocio de los editores comerciales- están limitando la capacidad de mantener, no ya de aumentar, la amplitud y profundidad de las colecciones de las bibliotecas porque ello es insostenible para sus presupuestos.

**La Asociación de Editores Americanos (AAP) contrata una campaña contra el Open Access**

La revista **Nature** acaba de publicar (en. 2007) en su último número una noticia que no presagia nada bueno: la Asociación de Editores Americanos, a la que pertenecen, entre otras editoriales Elsevier, Wiley y the "American Chemical Society", ha contratado a una empresa de relaciones públicas para que diseñe una campaña en contra del acceso abierto a la información científica financiada con fondos públicos

**El reto.** El reto es explorar y experimentar con modelos alternativos de publicación y comunicación científica basadas en el acceso abierto aprovechando las oportunidades creadas por las tecnologías de Internet.

**Respuestas estratégicas.** En la actualidad están en marcha acciones para desarrollar modelos alternativos de publicación y comunicación científica. Las principales son:

- Competir con los editores comerciales de revistas científicas : alianza entre universidades, bibliotecas de investigación y otras organizaciones
- **Revistas de "acceso abierto"** que cubren los costes de publicación con modelos económicos alternativos a las suscripción o "pasarelas de acceso comerciales".
- **Archivos o repositorios** de versiones electrónicas de los resultados de la investigación, accesibles en línea a través de Internet. Cada vez más comunidades científicas e Instituciones están creando estos archivos digitales para difundir los resultados de su investigación.

- **La búsqueda de "documentos de acceso abierto"**

Para que los documentos que se publican en repositorios sean fácilmente recuperables se han desarrollado herramientas como el Protocolo para la transmisión de contenidos en Internet denominado **OAI-PMH** (Open Archives Initiative–Protocol for Metadata Harvesting).

Los buscadores cómo Google recuperan esta información, además hay buscadores especializados en información científica como Scirus, o el propio Google Scholar, y algunas bases de datos tienen herramientas para localizarlos. El **Web of Science** de ISI tiene una herramienta de búsqueda que indiza los documentos publicados en archivos y repositorios temáticos e institucionales. **OAister** es un catálogo colectivo con más de 10 millones de registros de más de 700 repositorios que incorporan el protocolo OAI-PMH

Además hay la posibilidad de acceder directamente a revistas o repositorios.

### **Lista de la UNED de archivos abierto de Matemáticas y Física:**

- CERN Document Server: ofrece acceso al texto completo de documentos (preprints, artículos, libros, revistas, fotografías) de interés para los investigadores en física de las partículas y áreas relacionadas.
- PhysicsWeb: enlaces a un gran número de archivos de preprints de física.
- ArXiv.org
- Mathematical Physics Preprint Library
- Mathematics Preprint Server
- MathematicsWeb
- Physics Documents Worldwide
- Physnet

**ArXiv.** Imprescindible. Arxiv es un servicio de búsqueda y acceso al texto de documentos electrónicos en los campos de la física, matemática, *nonlinear-science*, informática y biología. Puesto en marcha en 1991 , gestionado por la Cornell University Library, es probablemente el repositorio de archivos abiertos más conocido y utilizado en la literatura que nos ocupa. Se estima que en física de partículas un 90% de la literatura que se publica se puede encontrar en el ArXiv. Muchos de estos documentos son versiones *preprint* de artículos que luego se publican formalmente en revistas científicas.