

# Temario de Ayudante de Biblioteca de la Administración General del Estado

Este temario ha sido elaborado por un opositor, para presentarse al proceso selectivo de Ayudante de Bibliotecas de la Administración General del Estado en la [convocatoria de 2021](#).

Incluye todos los temas, de legislación y específicos de bibliotecas, del programa correspondiente a la convocatoria de la Administración General del Estado para cubrir plazas de Ayudante de Bibliotecas en el Ministerios de Cultura y Deporte, Ministerio de Defensa, Ministerio de Asuntos Exteriores, Unión Europea y Cooperación y Ministerio de la Presidencia, Relaciones con las Cortes y Memoria Democrática. «BOE» núm. 149, de 23 de junio de 2021.

---

**Temario completo disponible en:**

<https://www.bibliopos.es/>

---



Temario de Ayudante de Biblioteca de la Administración General del Estado, cedido por su autor a [Bibliopos.es](https://www.bibliopos.es) para su publicación bajo licencia [Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional License](#).

Bajo esta licencia puedes utilizar libremente el temario para uso personal y compartirlo siempre que [cites la fuente](#) y proporciones un enlace a la [licencia](#). No puedes hacer uso comercial del documento.

# B10 Metabuscadores y gestores de enlaces

## Recolectores OAI-PMH.

La gestión de los recursos electrónicos es uno de los grandes retos que bibliotecas y centros de documentación han tenido que afrontar en los últimos años. La incorporación de los recursos-e trajo consigo la necesidad de replantearse la forma en que las bibliotecas los seleccionan, adquieren, describen y ponen a disposición de sus usuarios.

### Metabuscadores

Los *web crawlers* (arañas web, indexadores web o rastreadores web) son bots (programas informáticos para realizar tareas concretas y repetitivas, que actúan de forma autónoma y automática, sin participación o supervisión humana) que atraviesan la estructura de la Red buscando datos, analizan el contenido basándose en las palabras que contienen las páginas web y guardan la información en índices y bases de datos para mejorar el rendimiento de los buscadores. Los **buscadores**, también conocidos como motores de búsqueda, son un interfaz de interrogación textual que busca archivos almacenados en servidores web gracias a su araña web. Cuando un usuario introduce una palabra en un campo de búsqueda, el buscador consulta su índice y le devuelve un listado de direcciones web en los que se mencionan temas relacionados con las palabras buscadas. Cada buscador está configurado bajo una serie de parámetros (filtros de configuración y rastreo) que cada compañía considera los más adecuados para ordenar las búsquedas y mostrar los resultados. Los buscadores más conocidos son *Google, Bing, Yahoo!* y *Ask*.

La metabúsqueda, búsqueda federada o búsqueda cruzada supuso el primer intento de integrar diferentes sistemas de información. Un **metabuscador** (meta-search) es una potente herramienta que da acceso, permite ejecutar una búsqueda y realizar rastreos por los motores de búsqueda más usados, proporcionando una combinación de los mejores resultados que ha devuelto cada buscador. La diferencia con los **multibuscadores** es que, mientras que los multibuscadores muestran los resultados de cada buscador por separado, lo que permite evaluar la relevancia de cada web mostrada, los metabuscadores ofrecen una única lista de resultados.

Para los usuarios, las ventajas de la metabúsqueda se basan en la simplicidad de la interfaz y la facilidad para navegar. La posibilidad de buscar simultáneamente en muchas bases de datos desde una única interfaz es la principal aportación de los sistemas de metabúsqueda (utilizan, entre sus componentes, el Z39.50). Además amplían de forma notoria el ámbito de las búsquedas, proporcionando mayor cantidad de resultados. Si cada buscador sólo es capaz de rastrear una búsqueda entre el 15% de las páginas web que existen en la Red, utilizar un metabuscador supone realizar una búsqueda más exhaustiva. La forma de combinar los resultados depende del metabuscador empleado. A la hora de utilizar los metabuscadores hay que tener en cuenta que cada buscador utiliza su propia estrategia a la hora de recoger la información y ordenar los resultados de las búsquedas, por lo que las páginas de mayor relevancia en un buscador no tienen por qué coincidir con las del resto, aportando puntos de vista distintos.

Desde la perspectiva de las bibliotecas (y también para muchos usuarios), los problemas se derivan de la incompatibilidad de muchas bases de datos con los metabuscadores, las dificultades en la navegación, las opciones de búsqueda limitadas (en comparación con las bases de datos nativas, en muchas ocasiones no ofrecen búsquedas avanzadas), la lentitud en la respuesta y la forma de presentación de los resultados. Una de las desventajas importantes es que, mientras que cada buscador dispone de su propia sintaxis de búsqueda, los metabuscadores no distinguen entre las diferentes sintaxis. En cuanto a la gestión de los resultados, muchas veces el orden de no tiene en cuenta la relevancia y los muestra según otros criterios como la coincidencia de palabras claves,

velocidad de carga de la página, temática... El modelo de funcionamiento implica que los resultados vayan llegando desde las distintas fuentes por diferentes vías y en distintos tiempos. Este sistema de recuperación ofrece problemas para presentar los resultados de una búsqueda según un ranking de relevancia, que basa la importancia de un sitio web en los enlaces que recibe. Puesto que los resultados van llegando secuencialmente, habría que esperar a que estén todos para poder ordenarlos en base a un criterio, y eso alargaría el tiempo del proceso. De esta forma, se asegura que el usuario espere menos tiempo antes de obtener alguna respuesta, aunque estos resultados no son necesariamente los más relevantes.

Es necesario que se desarrolle un modelo de búsqueda centralizado. Este modelo debe basarse en recoger los recursos de información de un universo determinado e indizarlos para que puedan proporcionar los resultados precisos. Es el modelo de la búsqueda integrada, que desarrollan las **herramientas de descubrimiento**. Los metabuscadores buscan en las bases de datos en vivo, es decir, lanzan las consultas a la base de datos en el momento en que el usuario las formula. No crean su propia base de datos, sino que utilizan las bases de datos de otras fuentes. Esto conlleva problemas en cuanto a las capacidades de búsqueda, los tiempos de respuesta y la pertinencia de los documentos recuperados. Las herramientas de descubrimiento, por el contrario, crean su propio índice central con los metadatos y/o el texto completo de las diferentes bases de datos y otros sistemas de información de la biblioteca.

A continuación se exponen algunos de los metabuscadores más utilizados.

- **WebCrawler, Metacrawler y Dogpile.** De aspecto similar (pertenecen a la compañía System1), son metabuscadores que combinan los resultados de *Google, Yahoo!, Bing, Ask* y otros motores de búsqueda populares. Su misión es priorizar esos resultados mostrando los más relevantes primero. Su funcionamiento es muy sencillo, indicando sugerencias de búsquedas relacionadas y con apartados para búsquedas de imágenes, vídeos y noticias. El buscador está en inglés y prioriza este idioma, si bien muestra resultados en distintos idiomas si la búsqueda lo requiere.
- **DuckDuckGo** es un metabuscador que se basa en la privacidad: no almacena la dirección IP de sus usuarios ni guarda ninguna información relacionada con ellos (sólo se almacena información relacionada con las palabras clave que se están utilizando). Esto quiere decir que no tiene capacidad para personalizar los resultados. Los resultados que muestra han sido obtenidos de dos maneras: por una parte, tiene su propio rastreador web (DuckDuckBot), que examina Internet recopilando páginas web; pero por otra parte, también obtiene resultados de cerca de 400 fuentes diferentes, incluyendo buscadores como *Bing* o *Yahoo!* (se excluye de sus fuentes a *Google*), y servicios como *Wikipedia*.
- **Starpape** utiliza *Google* (por lo que los resultados son prácticamente los mismos), pero proporciona una navegación anónima, no guarda ni comparte datos de búsqueda, obviando cookies, perfiles de usuario y otras herramientas de seguimiento de la navegación.

## Gestores de enlaces

En un entorno tan variable como el digital, las páginas web desaparecen, los alojamientos son dinámicos, las marcas modifican su nombre... Por tanto, las URL no son enlaces persistentes. Esto hace que cada día se pierdan miles de enlaces y esto hace que documentos referenciados en motores de búsqueda y bases de datos no se encuentren. Por eso la identificación persistente de un recurso es un requisito indispensable para su identificación unívoca y acceso garantizado. Para resolver estos problemas se hace necesario generar métodos de identificación permanente de los recursos electrónicos (como *DOI, Digital Object Identifier, y OpenURL*) y crear sistemas que permitan la interconexión y la integración de recursos (como *CrossRef* y *SFX*).

El **DOI (Digital Object Identifier)** surge de una iniciativa de editores internacionales para desarrollar un sistema de identificación de documentos. Nace con la intención de hacer permanentes los enlaces de las citas de los artículos de las revistas electrónicas y facilitar la navegación por citas

entre sus artículos. Tiene un doble objetivo: identificar de forma única un objeto de contenido electrónico y servir como vínculo persistente y estable a la localización. DOI, cuya primera versión de la norma fue aprobada por ANSI/NISO en 2000, consiste en una cadena alfanumérica única formada por un prefijo (corresponde a la identificación de la institución asignado por una agencia de registro) y un sufijo (cada institución o revista puede decidir cómo estructurarlo). Una vez registrado, el editor asigna a un artículo electrónico un DOI y vincula la URL del artículo al *DOI resolver* (gestor de enlaces DOI), una aplicación que al añadir el DOI concreto, permite acceder a la información del artículo y también automáticamente a su citación. La funcionalidad principal del sistema DOI es resolver el DOI como una URL registrada y actualizada.

*CrossRef* es una iniciativa de los grandes editores de publicaciones científicas que consiste en la creación de una base de datos con los metadatos básicos de los artículos (título de la revista, ISSN, año, volumen, número, autor principal y números de página), con la finalidad de obtener los DOIs para las bibliografías de los artículos que publican. En el sistema *CrossRef*, cada DOI se asocia con un conjunto de metadatos básicos y la URL del texto completo y así se identifica de forma única el contenido y se proporciona un enlace persistente de su localización en Internet. Cuando un autor entrega al editor un artículo con su bibliografía, el editor busca en la base de datos los DOIs de las referencias del artículo y enlaza cada cita con la versión electrónica de ese artículo. Cuando el usuario lee una referencia que le resulta interesante y pincha sobre el botón correspondiente; su petición llega al servidor de DOI y lo redirige a la página web del artículo indicada por el editor. Entonces surge el problema de cuál es la copia adecuada para ese usuario.

Así fue como se confeccionó la norma *ANSI/NISO Z39.88-2004 (R2010) The OpenURL framework*, estándar que ayuda a construir direcciones web con información descriptiva sobre un registro, a fin de que pueda ser localizado el texto completo del mismo, o información relacionada. *OpenURL* es una herramienta para construir enlaces dinámicamente entre la información de referencia recuperada y su contenido. El funcionamiento del *OpenURL* es sencillo: primero se codifica la referencia bibliográfica dentro de la URL con los metadatos del registro, esta sintaxis es decodificable por un **gestor de enlaces** (*OpenURL Resolver*), una aplicación, que, a partir de datos que ofrecen las citas bibliográficas en formato *OpenURL*, proporciona información relevante sobre el registro deseado, ya sea un artículo a texto completo, una reseña, etc. DOI y OpenURL son tecnologías complementarias y trabajan juntas de varias maneras (por ejemplo, al hacer clic en un enlace DOI, se utiliza como base para extraer los metadatos necesarios para crear OpenURL).

Aunque no hay una traducción aceptada para el término “link resolver” al castellano, las empresas que los desarrollan y la comunidad profesional se refiere a los mismos como “resolvedores de enlaces”. También podría traducirse como servidor o gestor de enlaces, es decir, una aplicación que permite integrar los recursos electrónicos de una biblioteca mediante una tecnología de enlaces basada en *OpenURL*. Un gestor de enlaces proporciona de forma ágil y rápida, todas las opciones de acceso que una biblioteca tiene para un recurso informativo determinado. Su funcionamiento se basa en una base de datos de cobertura, que contiene información actualizada sobre los números de publicaciones seriadas a las que se tiene acceso. Cuando un usuario está consultado un recurso digital, es frecuente que las referencias contenidas en el mismo ayuden a identificar nuevas referencias que pueden ser de gran utilidad. El sistema de resolución de enlaces puede consultar en tiempo real esta base de datos de cobertura, e informarle si hay otros recursos digitales disponibles que contienen el texto completo para dicha referencia. Gracias a esta tecnología, la colección digital no es percibida por el usuario como un conjunto de recursos aislados entre sí, sino que están interconectados entre ellos. La visibilidad que obtienen los recursos de la biblioteca incrementa y se mejora la rentabilidad de las inversiones en recursos digitales. El resultado es que el usuario no debe preocuparse de quién ostenta la propiedad del contenido o dónde está alojado.

Estas herramientas, dentro de las posibilidades que ofrecen, permiten: el acceso a artículos a texto completo (ya sea por suscripción o por acceso abierto), el acceso al Servicio de Obtención de Documentos (SOD), la búsqueda en un catálogo bibliográfico y la exportación de la información del registro a un gestor bibliográfico (*RefWorks*). Esto genera una serie de ventajas:

- Facilitar el acceso al artículo a texto completo desde donde se ha encontrado la referencia
- Aumentar el uso del acceso electrónico a las publicaciones periódicas
- Es cómodo para los usuarios (ya que en pocos pasos acceden al artículo)
- Reduce las peticiones al Servicio de Obtención de Documentos
- La configuración puede ser gestionada por los propios bibliotecarios

**SFX**, creado en la Universidad de Gante y desarrollado por Exlibris, es el servidor de enlaces más utilizado. Proporciona acceso a los servicios que la biblioteca tiene contratados y configurados, como el texto completo, el sumario o el resumen, así como otros servicios de valor añadido, como por ejemplo el servicio de obtención de documentos, la posibilidad de descargar la referencia a *RefWorks* o saber si la revista tiene factor de impacto en *Journal Citation Reports*.

Otro gestor de enlaces es **360 Link**, implementado junto con el buscador de recursos electrónicos *Summon* (ambos productos pertenecientes a Ex Libris). Según como esté configurado este sistema en cada biblioteca, puede ofrecer otros servicios de interés como la posibilidad de rehacer las búsquedas en el catálogo de la biblioteca, en *Summon*, o solicitar el documento por préstamo interbibliotecario, en el caso de no estar incluido en una colección impresa o electrónica.

AbsysNET incorpora la herramienta **c+i/link**, un generador automático de enlaces: sobre una búsqueda determinada, el sistema responde con los ítems coincidentes del catálogo, pero también puede ofrecer un enlace a una búsqueda en herramientas externas.

## Recolectores OAI-PMH

Un **recolector** es un servidor externo a los repositorios de acceso abierto que permite que el usuario busque, localice y acceda a través de un único interfaz y de forma simultánea a la información y a los materiales depositados en varios de ellos. El recolector no contiene documentos originales, sólo sus metadatos, y dirige al usuario al repositorio correspondiente para la consulta del documento original completo. Su importancia está en que, al permitir las búsquedas en todo el conjunto de documentos de varios repositorios, facilitan el acceso y el análisis de la información.

De forma regular el recolector se conecta con sus repositorios y recoge los metadatos de sus documentos. Se suele hacer una recopilación inicial completa y posteriormente sólo se cargan los nuevos registros y las modificaciones, aunque algunos hacen recolecciones completas cada cierto tiempo. Posteriormente se revisa, almacena, organiza e indiza la información recopilada para ofrecerla a los usuarios, junto con otros servicios. Así estos pueden hacer búsquedas en los metadatos y acceder al contenido y a los objetos digitales de varios repositorios de forma simultánea, utilizando para ello un único punto de acceso.

Cada recolector trabaja únicamente con los repositorios con los que ha establecido vínculos, generalmente previa petición de registro y recolección por parte de la institución gestora de estos últimos. Además, el repositorio ha de cumplir con los requisitos técnicos establecidos por el recolector respecto a normalizaciones, metadatos, etc. Planteando la necesidad de poder consultar y trabajar con estos depósitos de información de forma conjunta y buscando mejorar el acceso a la información, facilitar su uso e incrementar su visibilidad en Internet, se llegó al desarrollo de estándares que permitiesen la interoperabilidad y el acceso a contenidos procedentes de varios proveedores y con diferentes formatos. Este es el origen del protocolo para recolección de metadatos de la Iniciativa de Archivos Abiertos (*Open Archives Initiative Protocol for Metadata Harvesting*, **OAI-PMH**), que dio sus primeros pasos en 1999. Proporciona un marco de interoperatividad independiente de la aplicación basado en recolección de metadatos (*metadata harvesting*). Dublin Core se ha convertido en el estándar de facto para todo tipo de repositorios, lo que no impide que cada servidor pueda ofrecer los registros en otros formatos adicionales.

El recolector está soportado por un proveedor de servicio con el objetivo de obtener metadatos de los repositorios. Además, es gestionado por un proveedor de datos para exponer los metadatos para su recolección. En cuanto a los recolectores **OAI-PMH**, es decir, los metabuscadores

de proveedores de datos OAI, se pueden destacar las siguientes características: el usuario consulta un único recolector obteniendo la información de múltiples repositorios; y el acceso a los repositorios y los recursos es inmediata mediante hipervínculos incluidos en la respuesta.

Desde los orígenes de *OAI-PMH*, se han ido creando una serie de herramientas software que soportan distintos aspectos, roles y funciones de la arquitectura del protocolo, todas ellas con el objetivo final de facilitar la distribución de información científica y académica, así como de cualquier tipo de contenido electrónico, a través de la Web. Los programas se caracterizan por ser software libre y haber sido desarrollados por implementadores de la comunidad OAI, principalmente en el seno de proyectos de investigación en universidades pero también en centros de investigación, bibliotecas y consorcios bibliotecarios, proyectos de bibliotecas digitales e incluso por programadores individuales.

De entre los principales recolectores, destacan los siguientes:

- **OAIster**. Es un catálogo colectivo cuya meta es crear una colección de recursos digitales académicos a texto completo, gratuitos, útiles, de fácil acceso (en la actualidad contiene más de 50 millones de registros). Este catálogo se creó a través de la recopilación de material de colecciones de libre acceso de todo el mundo mediante el OAI-PMH. Existe un sitio web de libre acceso disponible para buscar sólo registros de OAIster, aunque se puede acceder a todos los registros de OAIster a través de WorldCat.
- **OpenDOAR** (*Directory of Open Access Repositories*). Es el directorio autorizado de repositorios de acceso abierto académicos y de investigación, tanto institucionales como temáticos, además de los creados por otras organizaciones que financian proyectos, desarrollado por la Universidad de Nottingham (Reino Unido) y la Universidad de Lund (Suecia). Permite delimitar por regiones o países e incluso por el tipo de software.
- **ROAR** (*Registry of Open Access Repositories*). Es un registro de los repositorios de acceso abierto de todo el mundo que, además de facilitar su localización y acceso, proporciona información sobre el crecimiento y la actividad de cada uno de ellos. Creado en 2003 por la Universidad de Southampton, del Reino Unido.
- **Recolecta** o **Recolector de Ciencia Abierta**. Es un repositorio desarrollado en España por la FECYT (Fundación Española para la Ciencia y Tecnología, fundación del sector público que depende del Ministerio de Ciencia e Innovación), en colaboración con REBIUN, que comenzó su andadura en el año 2007. En esta plataforma se agrupan a todas las infraestructuras digitales españolas en las que se publican y/o depositan resultados de investigación en acceso abierto, garantizando que sean interoperables según los estándares de la comunidad mundial, buscando mejorar la visibilidad tanto nacional como internacional, el acceso y la aplicación de las investigaciones que se realizan en España. Promueve la adopción del acceso abierto por todos los investigadores de las universidades y centros de I+D españoles. Pertenece al proyecto europeo **OpenAIRE** (*Open Access Infrastructure for Research in Europe*, que se fusionó con **DRIVER**, *Digital Repository Infrastructure Vision for European Research*) que apoya la aplicación de la política de acceso abierto en Europa.
- **Hispana**. En funcionamiento desde marzo de 2006, es el recolector y directorio de recursos digitales del patrimonio bibliográfico español de archivos, bibliotecas y museos españoles, que gestiona el Ministerio de Cultura y Deporte, en colaboración tanto de Administraciones públicas como de universidades, fundaciones y entidades privadas. Con casi 9 millones de objetos digitales de 233 repositorios, es el agregador nacional de contenidos a **Europeana** (en 2008 inició su andadura con la iniciativa de la Comisión Europea, con la pretensión de ser la gran biblioteca digital de Europa, multilingüe y multicultural, con objetos digitales procedentes de repositorios de diferentes instituciones culturales europeas).