

# 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.

## B09 Los metadatos

### Concepto, objetivos y tipos. Principales modelos de metadatos aplicados a las bibliotecas. La Dublin Core Metadata Initiative y los estándares de la Biblioteca del Congreso y Europeana. Schema.org. ONIX.

#### Los metadatos: concepto, objetivos y tipos

Encontrar información relevante en el World Wide Web se ha convertido en un problema que se incrementa día a día debido al crecimiento explosivo de los recursos electrónicos. Internet, a través de los motores de búsqueda, recupera gran cantidad de documentos pero con muy poca precisión. Como resultado, ofrece mucho ruido en la recuperación y muchas veces la información se convierte en inaccesible. Debido a esta gran diversidad y volumen de recursos en Internet, un amplio colectivo, que incluía profesionales de la información, investigadores de bibliotecas digitales, informáticos de redes, especialistas en lenguajes de marcas, diseñadores de software, etc. se planteó la necesidad de crear un mecanismo para etiquetar, describir e identificar los recursos electrónicos presentes en Internet que permitiera facilitar una posterior búsqueda y una recuperación más eficaz de la información. El objetivo primordial era llegar a un consenso interdisciplinar e internacional para diseñar un código de descripción de recursos. Este mecanismo los constituyen los llamados metadatos o metadata, término acuñado por Jack Myers en la década de los años 60.

Un **metadato** no es más que un dato estructurado sobre la información, o de forma más simple, datos acerca de datos. El concepto de metadatos se puede entender en un sentido amplio o en un sentido más estricto. En un sentido amplio, si es un término que se utiliza para describir datos, se puede considerar que el catálogo de una biblioteca o un repertorio bibliográfico son tipos de metadatos. Sin embargo, si se acota la definición de metadatos dándole un sentido más estricto, los metadatos sólo serían posibles en un contexto digital y en red ya que sólo dentro de este contexto se pueden utilizar los metadatos con la función que les caracteriza, que es la de la localización, identificación, clasificación y descripción de recursos web, legibles e interpretables por máquina y que, por tanto, también sirven para su recuperación. En definitiva no es otra cosa que catalogar, pero en un nuevo contexto, donde los objetos no son físicos sino electrónicos. Los metadatos se forman y almacenan para que puedan ser leídos por los motores de búsqueda. Además, sirven igualmente para recursos que no tienen únicamente la morfología de texto, sino para cualquier tipo de morfologías tales como vídeo, audio o imágenes. Eva Méndez y José A. Senso definen los metadatos, incidiendo en su funcionalidad, como “toda aquella información descriptiva sobre el contexto, calidad, condición o características de un recurso, dato u objeto que tienen la finalidad de facilitar su recuperación, autenticación, evaluación, preservación y/o interoperatividad”.

Según la revisión literatura realizada, los metadatos se marcan como **objetivos** básicos para responder a las necesidades tanto de usuarios como de aplicaciones finales: la localización de recursos, la evaluación de su idoneidad para los propósitos específicos, el acceso u obtención de los datos y, finalmente, el uso de los mismos.

Las **clasificaciones** por parte de los diferentes autores e instituciones son muy variadas, y se establecen atendiendo a los distintos aspectos a los que se dé prioridad a la hora de establecer dichas clasificaciones: distintos aspectos como su forma, funcionalidad, nivel de estructuración de los datos, persona o entidad que los origina, etc. De acuerdo con la naturaleza de los datos que describen, existen tres tipos de metadatos:

- Metadatos descriptivos: Su función es la descripción e identificación de recursos de información, para permitir la búsqueda y la recuperación.

- **Metadatos estructurales:** Permiten identificar las distintas partes o componentes que configuran un recurso de información, la estructura lógica de la información y los archivos que conforman la publicación. Facilitan la navegación y presentación de recursos electrónicos, proporcionan información sobre la estructura interna de los documentos (incluyendo página, capítulo, numeración, índices, etc.) y describen la relación entre ellos.
- **Metadatos administrativos:** Su función es gestionar y administrar los recursos digitales (localización, institución o autor que genera, alberga y mantiene los recursos, fecha de creación y actualización, etc.), incluyen datos técnicos sobre la creación y el control de calidad (informan sobre el funcionamiento del sistema, de los requisitos de hardware y software...), sobre la gestión de derechos y requisitos de control de acceso y utilización e información sobre acción de preservación (actualización de los datos, migración de los contenidos...). Por tanto, la categoría de metadatos administrativos puede subdividirse en metadatos de gestión de derechos, metadatos de preservación y metadatos de tipo técnico.

En cuanto a la forma de asignar metadatos, existen varios modos de asociar metadatos con recursos digitales: incrustando los metadatos dentro del propio documento (lo que implica que los metadatos deben ser creados al mismo tiempo que se crea el recurso, generalmente embebidos y codificados dentro de la etiqueta <meta> en la cabecera del documento HTML); asociando los metadatos por medio de archivos acoplados a los recursos a los que describen (se utiliza, sobre todo, para material multimedia, imágenes, etc. a través del elemento <link>); o metadatos independientes (se mantienen en un depósito separado, generalmente una base de datos mantenida por una organización que puede o no tener acceso al contenido del recurso).

Los metadatos facilitan la interoperatividad entre sistemas con el fin de facilitar la eficiente difusión de los contenidos. Entre los diferentes estándares destaca el protocolo técnico llamado **OAI-PMH** (Open Archive Initiative-Protocol for Metadata Harvesting) que permite automatizar la captura y agregación de metadatos de recursos distribuidos en la web. Constituye uno de los pilares técnicos de los llamados archivos abiertos y de los repositorios institucionales.

## Principales modelos de metadatos aplicados a las bibliotecas

Existen distintos modelos de metadatos, cada uno de ellos con distintos esquemas de descripción. En los distintos modelos, cada objeto se describe por medio de una serie de atributos y el valor de estos atributos es el que puede servir para recuperar la información. Las iniciativas de metadatos siguientes son las más conocidas aplicadas a las bibliotecas.

Además de todos estos esquemas de metadatos se ha desarrollado **RDF** (Resource Description Framework, Marco de Descripción de Recursos), una infraestructura, basada en la sintaxis XML, capaz de representar metadatos que ayudan a la identificación, clasificación o localización de un documento o recurso web. Ofrece una arquitectura de metadatos para la Web que supone un gran avance para construir la infraestructura de la Web Semántica.

## La Dublin Core Metadata Initiative

Existen múltiples iniciativas para describir recursos electrónicos mediante metadatos normalizando y estandarizando las estructuras y arquitecturas de los recursos de Internet, el ejemplo más significativo es el Dublin Core, creado con el patrocinio de OCLC (Online Computer Library Center). Actualmente Dublin Core es el esquema de metadatos más utilizado a nivel mundial. La Iniciativa de Metadatos Dublin Core (DCMI, Dublin Core Metadata Initiative, organización dedicada a fomentar la adopción extensa de los estándares interoperables de los metadatos y a promover el desarrollo de los vocabularios especializados de metadatos para describir recursos que permitan la identificación de recursos mediante sistemas más inteligentes) comenzó en 1995 con la

convocatoria en Dublin (Ohio) de bibliotecarios, investigadores sobre la biblioteca digital, distribuidores de contenidos y expertos en marcado textual para mejorar la elaboración de estándares y normas relacionadas con la recuperación de información aplicables a los recursos. El Dublin Core original surgió en 1998 como un pequeño conjunto de descriptores que rápidamente suscitó el interés general de una amplia variedad de proveedores de información. Los metadatos Dublin Core fueron aprobados como norma nacional en Estados Unidos por el American National Standards Institute como la norma *ANSI/NISO Z39.85-2001*, que se convirtió en estándar internacional mediante la norma ISO 15836-2003 (la última revisión, *ISO 15836-1:2017 Elementos principales* y *ISO 15836-2:2019 Propiedades y clases*, que se transformó en estándar nacional a través de la norma *UNE-ISO 15836:2011 Conjunto de elementos de metadatos Dublin Core*).

El Dublin Core es el esquema de metainformación más utilizado a nivel mundial. En los últimos años su uso se ha generalizado, concretamente en los repositorios institucionales y como medio para la agregación de metadatos en los archivos abiertos. Algunas de las fortalezas de este esquema de metadatos son: su simplicidad, la independencia sintáctica, el alto nivel de normalización formal y el crecimiento y evolución del estándar a través de una institución formal consorciada (la DCMI). Además, el conjunto de elementos DC se ha convertido en una infraestructura operacional del desarrollo de la Web Semántica. La sencillez del Dublin Core puede ser tanto una fortaleza como una debilidad. La simplicidad reduce el coste de la creación de metadatos y fomenta la interoperabilidad. Por otro lado, la sencillez de Dublin Core no se ajusta a la riqueza funcional y semántica que proporcionan esquemas de metadatos complejos, pero se pueden crear equivalencias entre esos esquemas más sofisticados y el Dublin Core para facilitar la exportación y las búsquedas entre diferentes sistemas.

Se estableció un modelo de descripción básico formado por 15 descriptores o campos (inicialmente fueron 13), denominado conjunto de elementos de metadatos Dublin Core (o DCMES, Dublin Core Metadata Element Set), independientes de cualquier método de codificación, que resultaron del esfuerzo por alcanzar un consenso interdisciplinar e internacional. Aunque no obliga a codificar de una forma particular, para facilitar el uso de las distintas alternativas, la DCMI ha publicado recomendaciones sobre cómo codificar los metadatos Dublin Core en HTML y en XML. A pesar de que el Dublin Core se desarrolló originalmente con la mirada puesta en la descripción de documentos entendidos como objetos de información, los metadatos DC pueden aplicarse también a otros recursos. Esos 15 elementos básicos para describir cualquier objeto de información, se presentan habitualmente divididos en tres grupos que indican la clase o alcance de la información: elementos relacionados con el contenido del recurso (título, tipo del recurso, cobertura, etc.), elementos relacionados con el recurso cuando es visto como una propiedad intelectual (creador/autor, editor, derechos, etc.) y elementos relacionados con la instanciación (fecha, formato, el identificador del recurso y la lengua).

## Los estándares de la Biblioteca del Congreso: MODS, MADS y METS

Pero Dublin Core no es el único sistema de metadatos surgido en torno a la Web. Otra iniciativa similar es **MODS** (Metadata Object Description Schema, Esquema para la Descripción de Objetos de Metadatos), elaborado en 2002 por la Network Development and MARC Standards Office (NDMSO, Oficina de Desarrollo de Redes y Normas MARC) de la Library of Congress. Frente a las críticas que recibía Dublin Core por la extrema simplicidad de su sistema, el sistema MODS ofrece un conjunto de metadatos más amplio, compatible en mayor medida con el MARC21 Format for Bibliographic Data. Es un esquema de metadatos descriptivo de 20 elementos que permite crear la descripción de recursos originales o seleccionar los registros existentes en MARC21. Establece una representación de los metadatos en forma de documentos XML, es decir, además de la información que debe indicarse en las descripciones, especifica la forma en la que éstas deben codificarse en XML.

El sistema de metadatos MODS fue complementado con un segundo esquema llamado

**MADS** (Metadata Authority Description Schema), cuya primera versión es de 2004. Su propósito es ofrecer un esquema XML para codificar registros de autoridad que puedan vincularse a las descripciones bibliográficas basadas en MODS. Esto supone una ventaja que le permite una fácil interoperabilidad con cualquier formato de descripción bibliográfica e incluso archivística. Las descripciones elaboradas conforme al modelo de metadatos MADS son compatibles con los principales modelos de metadatos y formatos de descripción bibliográfica, como MARCXML (esquema XML que facilita la representación de registros MARC en formato XML) y Dublin Core.

Son frecuentes los documentos digitales que reúnen distintos archivos de imagen, texto e incluso audio, y que deben presentarse al lector de forma coordinada. Estos documentos se caracterizan por una estructura que debe ser conocida y gestionada por las aplicaciones informáticas que los procesan. Para codificar y representar la estructura de documentos digitales complejos que reúnen distintos archivos, se emplean los metadatos estructurales, que pueden ser codificados a través de la especificación **METS** (Metadata Encoding and Transmission Standard). METS es una norma para codificar los metadatos descriptivos, administrativos y estructurales de objetos de una biblioteca digital utilizando el lenguaje XML schema y para expresar las relaciones entre estos tipos de metadatos. Ofrece los mecanismos necesarios para codificar y representar la estructura de documentos digitales complejos. El estándar fue creado en 2001 y es mantenido por la NDMSO de la Library of Congress. Los metadatos METS constan de su propio espacio de nombres o namespace. El elemento <mets> es el elemento raíz.

## **Los estándares de Europeana: Europeana Data Model (EDM)**

Europeana es un portal que da acceso a millones de objetos de todo tipo del patrimonio cultural europeo. Sin embargo, cada uno de los sectores del patrimonio representados en Europeana utiliza diferentes estándares de datos: LIDO (Lightweight Information Describing Objects) para museos, EAD (Encoded Archival Description) para archivos y METS para bibliotecas digitales. Para salvar este problema se creó un modelo de datos básico con el que se lanzó Europeana en 2008: **Europeana Semantic Elements (ESE)**, un esquema de metadatos que se basa en Dublin Core. Esto conllevaba una serie de problemas: obligaba a la interoperatividad (perdiendo lo metadatos originales), impedía pasar a un modelo semánticamente rico y funcional (no permitía enlazar a recursos externos) e impedía que pudiera incluirse en un proyecto con Linked Open Data.

El **Modelo de datos de Europeana (EDM, Europeana Data Model)** es una propuesta para la estructuración y representación de los datos que Europeana ingiere, gestiona y publica, que supone una mejora con respecto a ESE). EDM no se fundamenta en ningún estándar de una comunidad en particular, sino que más bien adopta un marco de referencia abierto, interdisciplinar y basado en los principios fundamentales de la Web Semántica y Linked Open Data, que pueda dar cabida a la variedad y riqueza de las normas de las comunidades particulares. EDM no sólo es compatible con toda la riqueza de los metadatos de los proveedores de contenidos sino que también permite el enriquecimiento de los datos de una variedad de fuentes de terceros. Este modelo, construido sobre estándares del World Wide Web Consortium (W3C) como RDF, OAI-ORE (Open Archives Initiative-Object Reuse and Exchange, para la reutilización e intercambio de objetos digitales), la ontología SKOS (Simple Knowledge Organization System, apoyado por) y Dublin Core, es más expresivo y flexible que ESE.

## **PREMIS**

Los documentos electrónicos presentan un reto adicional: su preservación. Transcurrido un período largo de tiempo, incluso los formatos de fichero más comunes pueden convertirse en obsoletos, lo que significa que las aplicaciones actuales no pueden reproducirlos. Los metadatos de preservación tienen como finalidad especificar los requisitos técnicos para su acceso y lectura

mediante el registro del tipo y el año del soporte de almacenamiento y las fechas en las que se refrescaron los ficheros por última vez. En los últimos años se han lanzado distintas iniciativas, entre las que destaca **PREMIS** (Preservation Metadata: Implementation Strategies, Metadatos de preservación: estrategias de ejecución). En junio del 2003, OCLC (Online Computer Library Center) y RLG (Research Libraries Group), patrocinaron la formación del grupo de trabajo PREMIS, compuesto por expertos internacionales en la utilización de metadatos aplicados a actividades de preservación digital: mantenimiento de la viabilidad, la disponibilidad, la claridad, la autenticidad y la identidad en el contexto de la preservación. En 2005 la tarea del grupo concluyó con la publicación del *Diccionario de datos PREMIS de metadatos de preservación*, un material práctico y exhaustivo para las puesta en práctica de metadatos de preservación en sistemas de gestión de objetos digitales. La tercera versión se publicó en 2015. El *Diccionario de datos PREMIS* organiza los metadatos en torno a cinco entidades: entidad intelectual (se refiere a un conjunto de contenidos que se considera una única unidad intelectual a efectos de gestión y descripción, por ejemplo, un libro o un mapa), objeto (unidad discreta de información en formato digital), eventos (actividades que se realizan para la preservación de un objeto), agentes (personas, entidades o aplicaciones informáticas relacionadas con un evento) y derechos (permisos para realizar ciertas acciones sobre un objeto para su preservación).

## Schema.org

**Schema.org** es una iniciativa conjunta lanzada en 2011 y mantenida por Bing, Google, Yahoo! y Yandex (los motores de búsqueda más grandes del mundo) para unificar la forma de etiquetar y enriquecer la información de una página web para que las arañas de los motores de búsqueda (*bots*) puedan indexarla, clasificarla y mostrarla, con el fin de mejorar la visibilidad y posicionamiento de la página web. Este lenguaje de marcado es un conjunto de etiquetas HTML que se utilizan para describir el contenido de una página web y dotarlo de lenguaje semántico enriquecido que los buscadores pueden reconocer, comprender, ordenar y mostrar en sus resultados. Desde 2012, WorldCat está añadiendo a las páginas web de los registros bibliográficos los datos descriptivos codificados con Schema.org. Proporciona una colección de vocabularios (o esquemas de datos estructurados) que definen las propiedades (el tipo de dato que puede contener y su alcance) con las que se puede codificar el contenido de las páginas web para generar resultados de búsqueda más ricos y relevantes.

## ONIX

**ONIX** (ONIX for Books) es un estándar internacional abierto, basado en XML, diseñado para la codificación y el intercambio de información bibliográfica y comercial en formato digital orientada a la industria del libro. El estándar (su primera versión es de 2000, la versión 3.0 actual es de 2009) creado y mantenido por EDItEUR, grupo internacional formado con organizaciones asociadas de 25 países, entre ellos la Federación de Gremios de Editores de España (FGEE), que coordina el desarrollo de estándares para el comercio electrónico en el sector editorial. Permite a los editores de libros y libros electrónicos crear y administrar un corpus de metadatos enriquecidos sobre sus productos e intercambiarlos con sus clientes (distribuidores y minoristas). El estándar ONIX for Books proporciona una forma de comunicar información sobre el autor, el editor, el precio, la fecha de publicación, las dimensiones físicas, la sinopsis y muchos otros detalles de un libro (resúmenes, críticas, premios, biografías, cubierta, vídeos...).