

---

# Metadatos

---

PID\_00271449

Blas Torregrosa García

---

Tiempo mínimo de dedicación recomendado: 1 hora

---



**Blas Torregrosa García**

Ingeniero en Informática y máster universitario en Seguridad de las Tecnologías de la Información y de las Comunicaciones (MISTIC) por la Universitat Oberta de Catalunya (UOC). Especializado en ciberseguridad. Profesor colaborador en el máster de Ciencia de Datos de la UOC y profesor asociado en la Universidad de Valladolid (UVA).

El encargo y la creación de este recurso de aprendizaje UOC han sido coordinados por el profesor: Ferran Prados Carrasco (2020)

Primera edición: febrero 2020  
© Blas Torregrosa García  
Todos los derechos reservados  
© de esta edición, FUOC, 2020  
Av. Tibidabo, 39-43, 08035 Barcelona  
Realización editorial: FUOC

*Ninguna parte de esta publicación, incluido el diseño general y la cubierta, puede ser copiada, reproducida, almacenada o transmitida de ninguna forma, ni por ningún medio, sea este eléctrico, químico, mecánico, óptico, grabación, fotocopia, o cualquier otro, sin la previa autorización escrita de los titulares de los derechos.*

# Índice

<b>Introducción</b> .....	5
<b>1. ¿Qué son los metadatos?</b> .....	7
1.1. Diferencia entre datos y metadatos .....	8
1.2. Tipos y funciones .....	8
<b>2. Estándares de metadatos</b> .....	10
2.1. Tipo y estructura de metadatos .....	10
2.2. Dublin Core Metadata Initiative .....	11
2.3. IEEE/LTSC LOM e IMS Global Learning Meta-Data .....	12
2.4. Más sobre estándares .....	14
<b>Bibliografía</b> .....	15



## Introducción

Los metadatos pasaron a ser conocidos en el mundo cuando en 2013 Edward Snowden, un antiguo empleado de la Agencia Central de Inteligencia (CIA, por sus siglas en inglés), reveló que la Agencia Nacional de Seguridad (NSA, en inglés) estaba llevando a cabo una vigilancia masiva en las actividades de correo electrónico y dispositivos móviles de sus ciudadanos. La información que estaba registrando no era el contenido de las conversaciones o mensajes sino los **metadatos**, es decir, información acerca de esas conversaciones o mensajes.

En este módulo veremos qué son los metadatos, la forma en la que se tratan y algunos estándares sobre metadatos.



## 1. ¿Qué son los metadatos?

Una definición de metadatos frecuentemente usada es «**datos sobre datos**». Esta es una definición literal, aunque no excesivamente aclaratoria.

En 2004 la Organización Nacional de Estándares de Información (NISO, por sus siglas en inglés) definió los metadatos como «información estructurada que describe, explica, localiza o hace más fácil recuperar, usar o administrar un recurso de información».

La característica principal de los metadatos es su naturaleza **referencial**, es decir, se refieren a «otra cosa» que no son ellos mismos.

Hay una gran cantidad de metadatos asociados a cualquier hecho cotidiano, por ejemplo, una llamada de teléfono, particularmente las llamadas desde dispositivos móviles. Los metadatos más obvios sobre una llamada son los números de teléfono de la persona que llama y del destinatario. A continuación está la hora y la duración de la llamada. Y para las llamadas realizadas desde teléfonos inteligentes, están las ubicaciones geográficas de las personas involucradas (la mayoría de estos dispositivos tienen funcionalidad GPS o por medición del alcance de las torres de telefonía móvil en las que se encuentran).

Así, se puede considerar que los metadatos se comportan como una información que **describe un recurso**. De la misma forma que la información nutricional de un alimento envasado: los metadatos son información sobre el contenido.

Los metadatos pueden describir una gran variedad de información, como:

- El asunto del recurso.
- Los creadores del recurso.
- La información técnica necesaria para almacenar y acceder al recurso.
- Los derechos legales del recurso.

Los metadatos se pueden usar para buscar, reunir y mantener recursos durante largos períodos de tiempo. De esta manera se considera toda la información que un usuario necesitará para saber si el recurso satisface sus necesidades o si puede usarlo y cómo. Los metadatos pueden indicar de qué trata la información, cómo usarla, si necesita permisos y dónde obtener esa información. De este modo los metadatos permiten que las búsquedas sean más significativas.

### Metadato

Del griego *meta* 'más allá' y del latín *datum* 'lo dado'.

VALOR NUTRITIVO MEDIO (aprox. por 100 g. de producto)	
Valor Energético	338 kcal 1.437 kJ
Proteínas	7,9 g
Hidratos de Carbono	75,3 g
De los cuales azúcares	0,2 g
Grasas	0,5 g
De las cuales saturadas	0,1 g
Fibra Alimentaria	2,1 g
Sodio	0,0 g

Figura 1. Ejemplo de metadatos: etiqueta nutricional de un alimento envasado

Fuente: [es.wikipedia.org/wiki/Archivo:Etiqueta\\_nutricional\\_de\\_un\\_alimento\\_envasado.JPG](https://es.wikipedia.org/wiki/Archivo:Etiqueta_nutricional_de_un_alimento_envasado.JPG)

## 1.1. Diferencia entre datos y metadatos

Aunque puedan parecer similares incluso en el nombre, los datos y los metadatos no son lo mismo. La explicación más sencilla es que los datos en esencia son el producto de información real, mientras que los metadatos son datos sobre esos datos.

En el contexto de un archivo de vídeo los datos son los fotogramas del vídeo que se visualiza. El nombre del archivo y el número de fotogramas, así como la información sobre derechos de autor, no forman parte de esa presentación, sino que describen los datos. Estos son los metadatos.

Los datos en sí se denominan **datos esenciales** porque son la verdadera esencia de la información que se almacena. Los datos esenciales son los que tienen derechos de propiedad y valor comercial, mientras que los metadatos ayudan a encontrar, administrar y distribuir esos datos esenciales.

## 1.2. Tipos y funciones

Los tipos de metadatos dependen tanto del contexto como del dominio de la aplicación. En 1998 el Getty Research Institute, en una publicación sobre metadatos, identificó cinco tipos de metadatos y sus funciones:

1) **Metadatos administrativos.** Metadatos utilizados para la gestión y administración de recursos de información; por ejemplo, adquisición información, derechos y seguimiento de reproducción, requisitos legales de acceso e información de ubicación.

2) **Metadatos descriptivos.** Metadatos utilizados para identificar y describir recursos de información relacionados; por ejemplo, catalogación, encontrar ayudas o índices especializados.

3) **Metadatos de preservación.** Metadatos relacionados con la gestión de la conservación de recursos de información o acciones para proteger recursos físicos o digitales.

4) **Metadatos técnicos.** Metadatos relacionados con el funcionamiento de sistemas o comportamiento de metadatos; por ejemplo, información sobre hardware y requisitos de software, formatos, tasa de compresión o datos de autenticación y seguridad.

5) **Metadatos de uso.** Metadatos relacionados con el nivel y tipo de uso de recursos de información; por ejemplo, búsqueda y seguimiento de usuarios.



En 2004 la Organización Nacional de Estándares de Información (NISO) de EE.UU. agrupa los tipos anteriores de metadatos en dos categorías y añade los metadatos estructurales:

- 1) **Metadatos descriptivos**, que describen un recurso para fines de identificación y búsqueda.
- 2) **Metadatos estructurales**, que indican cómo se organizan los recursos.
- 3) **Metadatos administrativos**, que proporcionan información de administración de recursos, incluyendo metadatos técnicos, de gestión de derechos y de preservación.

## 2. Estándares de metadatos

Los metadatos están formados por elementos de datos asociados al recurso (elementos de metadatos). Los **esquemas de metadatos** o **modelos de metadatos** son conjuntos de elementos de metadatos diseñados para un propósito específico, como describir un tipo particular de recurso. Cuando un esquema de metadatos está definido y es ampliamente aceptado, es adoptado por la mayoría de la industria. Y cuando una especificación es ampliamente reconocida y adoptada, se convierte en un **estándar de metadatos**.

No hay un único estándar en metadatos; al contrario, existen numerosos estándares de metadatos o especificaciones de metadatos que se pueden adaptar para satisfacer necesidades concretas del contexto.

### 2.1. Tipo y estructura de metadatos

Los registros de metadatos tienen tres componentes: sintaxis, estructura y semántica.

La **sintaxis** son los estándares de codificación que funcionan como un contenedor para la estructura y la semántica, un conjunto de reglas que permiten la interpretación de los contenidos.

Como ejemplos de estándares de codificación tenemos:

- RDF (marco de descripción de recursos)
- XML (lenguaje de marcado extensible)
- JSON-LD (JSON para datos enlazados)

La **estructura** se ajusta dentro de la sintaxis y está compuesta por un esquema de metadatos.

Algunos ejemplos de estructuras son:

- Dublin Core
- Categorías para la Descripción de Obras de Arte (CDWA Lite)
- Descripción de Archivo Codificado (EAD)

Mientras que el estándar de codificación indica a un sistema informático cómo interpretar un registro de metadatos, la estructura indica a un usuario humano cómo interpretar la información contenida en etiquetas individuales (por ejemplo, creador, título, tema, etc.). Los estándares de estructura pueden considerarse como conjuntos de etiquetas (*tags*) preparados.

La pieza final de un registro de metadatos es la **semántica** o los estándares de contenido. Los estándares de contenido definen cómo debe ir la información dentro de las etiquetas o *tags*.

El uso de estándares de contenido garantiza la uniformidad de la información contenida en las etiquetas, es decir, es una pieza fundamental para producir metadatos interoperables.

## 2.2. Dublin Core Metadata Initiative

**Dublin Core Metadata Initiative (DCMI)** es una organización internacional con sede en Dublín (Ohio, EE. UU.) que se financia a través de proyectos y subvenciones y en la que el trabajo se realiza por voluntarios de todo el mundo.

Los objetivos de esta iniciativa son:

- Desarrollar estándares de metadatos para la recuperación de información en Internet entre diversos dominios de información.
- Definir marcos de trabajo para la interoperabilidad entre conjuntos de metadatos.
- Facilitar conjuntos de metadatos especializados para una comunidad o disciplina.

De forma general, los metadatos Dublin Core constan de:

- 15 elementos básicos (ver tabla 1).
- Calificadores para detalles adicionales.
- Tipos de datos (formatos de fecha/tiempo, tipos MIME).

Tabla 1. Conjunto de elementos de Dublin Core

Etiqueta	Descripción
DC. Title	<b>Título:</b> el nombre dado a un recurso.
DC. Creator	<b>Autor:</b> la entidad primariamente responsable de la creación del contenido intelectual del recurso.
DC. Subject	<b>Materias y palabras clave:</b> el tema del contenido del recurso.
DC. Description	<b>Descripción:</b> la descripción del contenido del recurso.

<b>Etiqueta</b>	<b>Descripción</b>
DC. Publisher	<b>Editor:</b> la entidad responsable de hacer que el recurso se encuentre disponible.
DC. Contributor	<b>Colaborador:</b> la entidad responsable de hacer contribuciones al contenido del recurso.
DC. Date	<b>Fecha:</b> una fecha asociada con un evento en el ciclo de vida del recurso.
DC. Type	<b>Tipo:</b> la naturaleza o categoría del contenido del recurso.
DC. Format	<b>Formato:</b> la manifestación física o digital del recurso.
DC. Identifier	<b>Identificación:</b> una referencia no ambigua para el recurso dentro de un contexto dado.
DC. Source	<b>Fuente:</b> una referencia a un recurso del cual se deriva el recurso actual.
DC. Language	<b>Idioma:</b> el idioma del contenido intelectual del recurso.
DC. Relation	<b>Relación:</b> una referencia a un recurso relacionado.
DC. Coverage	<b>Cobertura:</b> la extensión o ámbito del contenido del recurso.
DC. Rights	<b>Derechos:</b> la información sobre los derechos de propiedad y sobre el recurso.

Dublin Core puede, por tanto, considerarse como un pequeño lenguaje para realizar una clase particular de declaraciones sobre recursos. En este lenguaje hay dos clases de términos: **elementos** (nombres) y **calificadores** (adjetivos).

### 2.3. IEEE/LTSC LOM e IMS Global Learning Meta-Data

Se trata del estándar por excelencia para Metadatos de Objetos Educativos y está patrocinado por el Comité de Estandarización de Tecnologías Educativas del IEEE (Institute of Electrical and Electronics Engineers).

LOM (*Learning Object Metadata*) fue aprobado como el estándar IEEE 1484.12.1:2002 sobre metadatos para objetos de aprendizaje como un estándar abierto para la descripción de los «objetos de aprendizaje».

LOM define y especifica un esquema de metadatos con atributos, sus definiciones y una estructura jerárquica que los relaciona. No incluye información acerca de cómo representar estos metadatos o con qué mecanismos se puede transmitir y procesar esta información. Una iniciativa de IMS Global Learning Consortium ha especificado modelos para el uso de los metadatos y una representación en XML que proporcionan los mecanismos para ser usados en la práctica.

El principal objetivo de este estándar es facilitar la búsqueda, la evaluación, la adquisición y el uso de recursos docentes.

Así, los estándares LOM se centran en el conjunto mínimo de propiedades que permiten que los objetos educativos sean gestionados, ubicados y evaluados.

Los metadatos LOM se organizan en una estructura en árbol con un nodo raíz que corresponde al nodo que se está describiendo (LOM). Dependiendo de la raíz tenemos subelementos que pueden contener otros subelementos. Para cada elemento de la jerarquía se indica la definición, el tipo de datos, los valores permitidos y la multiplicidad.

El Esquema de base de LOM se compone de 9 categorías y 47 elementos (ver tabla 2).

Tabla 2. Esquema de base de LOM

Categorías	Elementos
1. General	Identificador Título Entrada de catálogo Idioma Descripción Descriptor Cobertura Estructura Nivel de agregación
2. Ciclo de vida	Versión Estatus Otros colaboradores
3. Meta-metainformación	Identificador Entrada de catálogo Otros colaboradores Esquema de metadatos Idioma
4. Técnica	Formato Tamaño Ubicación Requisitos Comentarios sobre la instalación Otros requisitos para plataformas Duración
5. Uso educativo	Tipo de interactividad Tipo de recurso de aprendizaje Nivel de interactividad Densidad semántica Usuario principal Contexto: nivel educativo Edad Dificultad Tiempo previsto de aprendizaje Descripción Idioma
6. Derechos	Coste Copyright y otras restricciones Descripción

Categorías	Elementos
7. Relación (con otros recursos)	<b>Tipo:</b> naturaleza de la relación con el recurso principal <b>Recurso:</b> recurso principal al que se refiere esta relación
8. Observaciones	Persona Fecha Descripción
9. Clasificación	Finalidad Nivel taxonómico Descripción Descriptor

## 2.4. Más sobre estándares

A continuación se presentan algunas organizaciones que proporcionan estándares de metadatos sobre temas especializados:

- **Creative Commons**, que contiene licencias en las que se pueden usar metadatos sobre derechos de propiedad y otorgar derechos.
- **Biblioteca del Congreso de los Estados Unidos de América** sigue siendo la fuente de referencia para la catalogación en bibliotecas y mantiene la Clasificación de la Biblioteca del Congreso (LCC)<sup>1</sup> y los Epígrafes Temáticos de la Biblioteca del Congreso (LCSH).<sup>2</sup>
- **FGCD** (Federal Geographic Data Committee) fueron de los pioneros en la implantación de metadatos. Los metadatos son especialmente importantes para los datos geoespaciales con los que generar mapas fiables.
- **Ecological Metadata Language** (EML) centrado en datos sobre ecología.
- **Water Markup Language** (WaterML) con especial interés en series temporales sobre observaciones hidrológicas.

<sup>(1)</sup>Del inglés, *Library of Congress Classification*.

<sup>(2)</sup>Del inglés, *Library of Congress Subject Headings*.

## Bibliografía

**Alemu, G.; Stevens, B.** (2015). *An Emergent Theory of Digital Library Metadata*. Cambridge: Chandos Publishing.

**Gidley, S.; Castanedo, F.** (2017). *Understanding Metadata*. O'Reilly Media.

**Lei Zeng, M.; Qin, J.** (2016). *Metadata* (2.<sup>a</sup> ed.). Chicago: American Library Association.

**Lubas, R.; Schneider, I.; Jackson, A.** (2013). *The Metadata Manual*. Cambridge: Chandos Publishing.

**Pomerantz, J.** (2015). *Metadata*. Massachusetts: The MIT Press.

**Wootton, C.** (2009). *Developing Quality Metadata*. Massachusetts: Focal Press.

