Los **metadatos** (del [griego](https://es.wikipedia.org/wiki/Lengua_griega) *μετα*, *meta*, 'después de, más allá de'[1](https://es.wikipedia.org/wiki/Metadato#cite_note-1)​ y [latín](https://es.wikipedia.org/wiki/Lengua_latina) *datum*, 'lo que se da', «[dato](https://es.wikipedia.org/wiki/Dato)»[2](https://es.wikipedia.org/wiki/Metadato#cite_note-2)​), literalmente «sobre datos», son datos que describen otros datos. En general, un grupo de metadatos se refiere a un grupo de datos que describen el contenido informativo de un objeto al que se denomina *recurso*.[3](https://es.wikipedia.org/wiki/Metadato#cite_note-3)​ El concepto de metadatos es análogo al uso de [índices](https://es.wikipedia.org/wiki/%C3%8Dndice) para localizar objetos en vez de datos. Por ejemplo, en una biblioteca se usan fichas que especifican autores, títulos, casas editoriales y lugares para buscar libros. Así, los metadatos ayudan a ubicar datos.[4](https://es.wikipedia.org/wiki/Metadato#cite_note-4)​

Para varios campos de la [informática](https://es.wikipedia.org/wiki/Inform%C3%A1tica), como la [recuperación de información](https://es.wikipedia.org/wiki/Recuperaci%C3%B3n_de_informaci%C3%B3n) o la [web semántica](https://es.wikipedia.org/wiki/Web_sem%C3%A1ntica), los metadatos en [etiquetas](https://es.wikipedia.org/wiki/Etiqueta_%28metadato%29%22%20%5Co%20%22Etiqueta%20%28metadato%29)son un enfoque importante para construir un puente sobre el [intervalo semántico](https://es.wikipedia.org/wiki/Intervalo_sem%C3%A1ntico), ya que cualquier recurso tiene, cuando está almacenado conjuntamente con otros, la necesidad de ser descrito para facilitar las búsquedas que pudieran tratar de encontrarlo a partir de sus características distintivas. Esto es cierto para cualquier tipo de recurso, tanto para un vídeo como para un libro en una biblioteca o para un hueso en el armario de un paleontólogo.

El concepto de metadatos es anterior a Internet y a la web, si bien es cierto que las nuevas necesidades de búsqueda de información han suscitado un interés por las normas y prácticas de metadatos hasta entonces desconocido.

**Índice**

  [ocultar]

* [1Definiciones](https://es.wikipedia.org/wiki/Metadato#Definiciones)
	+ [1.1Distinción entre datos y metadatos](https://es.wikipedia.org/wiki/Metadato#Distinci%C3%B3n_entre_datos_y_metadatos)
	+ [1.2Metadatos sobre metadatos](https://es.wikipedia.org/wiki/Metadato#Metadatos_sobre_metadatos)
* [2Objetivos](https://es.wikipedia.org/wiki/Metadato#Objetivos)
* [3Clasificación](https://es.wikipedia.org/wiki/Metadato#Clasificaci%C3%B3n)
* [4Ciclo de vida](https://es.wikipedia.org/wiki/Metadato#Ciclo_de_vida)
	+ [4.1Creación](https://es.wikipedia.org/wiki/Metadato#Creaci%C3%B3n)
	+ [4.2Manipulación](https://es.wikipedia.org/wiki/Metadato#Manipulaci%C3%B3n)
	+ [4.3Destrucción](https://es.wikipedia.org/wiki/Metadato#Destrucci%C3%B3n)
* [5Metadatos en la informática](https://es.wikipedia.org/wiki/Metadato#Metadatos_en_la_inform%C3%A1tica)
* [6Almacenamiento](https://es.wikipedia.org/wiki/Metadato#Almacenamiento)
	+ [6.1Codificación](https://es.wikipedia.org/wiki/Metadato#Codificaci%C3%B3n)
* [7Vocabularios controlados y ontologías](https://es.wikipedia.org/wiki/Metadato#Vocabularios_controlados_y_ontolog%C3%ADas)
* [8Metadatos y objetos de aprendizaje](https://es.wikipedia.org/wiki/Metadato#Metadatos_y_objetos_de_aprendizaje)
* [9Paradatos](https://es.wikipedia.org/wiki/Metadato#Paradatos)
	+ [9.1Almacenamiento de paradatos](https://es.wikipedia.org/wiki/Metadato#Almacenamiento_de_paradatos)
* [10Crítica y problemas asociados al uso de los metadatos](https://es.wikipedia.org/wiki/Metadato#Cr%C3%ADtica_y_problemas_asociados_al_uso_de_los_metadatos)
* [11Formatos y estándares](https://es.wikipedia.org/wiki/Metadato#Formatos_y_est%C3%A1ndares)
* [12Véase también](https://es.wikipedia.org/wiki/Metadato#V%C3%A9ase_tambi%C3%A9n)
* [13Referencias](https://es.wikipedia.org/wiki/Metadato#Referencias)

Definiciones[[editar](https://es.wikipedia.org/w/index.php?title=Metadato&action=edit&section=1)]

El término «**metadatos**» no tiene una definición única. Según la definición más difundida de metadatos es que son «datos sobre datos». También hay muchas declaraciones como «informaciones sobre datos»,[5](https://es.wikipedia.org/wiki/Metadato#cite_note-5)​ «datos sobre informaciones»[6](https://es.wikipedia.org/wiki/Metadato#cite_note-6)​ e «informaciones sobre informaciones».[7](https://es.wikipedia.org/wiki/Metadato#cite_note-7)​

Otra clase de definiciones trata de precisar el término como «descripciones estructuradas y opcionales que están disponibles de forma pública para ayudar a localizar objetos»[8](https://es.wikipedia.org/wiki/Metadato#cite_note-Bultermann-8)​ o «datos estructurados y codificados que describen características de instancias conteniendo informaciones para ayudar a identificar, descubrir, valorar y administrar las instancias descritas».[9](https://es.wikipedia.org/wiki/Metadato#cite_note-9)​ Esta clase surgió de la crítica de que las declaraciones más simples son tan difusas y generales que dificultarán la tarea de acordar estándares, pero estas definiciones no son muy comunes.

Podemos también considerar los metadatos, en las áreas de telecomunicaciones e informática, como información no relevante para el usuario final pero sí de suma importancia para el sistema que maneja los datos. Los metadatos son enviados junto a la información cuando se realiza alguna petición o actualización de la misma.

En el campo biológico los metadatos se han convertido en una herramienta fundamental para el descubrimiento de datos e información. En este contexto se pueden definir los metadatos como «una descripción estandarizada de las características de un conjunto de datos» con esto se incluye la descripción del contexto en el cual los datos fueron coleccionados y además se refiere al uso de estándares para describirlos [[1]](http://www.siac.net.co/sib/metadatos/).

**Distinción entre datos y metadatos**[[editar](https://es.wikipedia.org/w/index.php?title=Metadato&action=edit&section=2)]

La mayoría de las veces no es posible diferenciar entre datos y metadatos. Por ejemplo, un poema es un grupo de datos, pero también puede ser un grupo de metadatos si está adjuntado a una canción que lo usa como texto.

Muchas veces, los datos son tanto "datos" como "metadatos". Por ejemplo, el título de un texto es parte del texto como a la vez es un dato referente al texto (dato como metadato).

**Metadatos sobre metadatos**[[editar](https://es.wikipedia.org/w/index.php?title=Metadato&action=edit&section=3)]

Debido a que los metadatos son datos en sí mismos, es posible crear metadatos sobre metadatos. Aunque, a primera vista, parece absurdo, los metadatos sobre metadatos pueden ser muy útiles. Por ejemplo, fusionando dos imágenes y sus metadatos distintos puede ser muy importante deducir cuál es el origen de cada grupo de metadatos, registrando ello en metadatos sobre los metadatos.

Objetivos[[editar](https://es.wikipedia.org/w/index.php?title=Metadato&action=edit&section=4)]

El uso de los metadatos mencionado más frecuentemente es la refinación de consultas a [buscadores](https://es.wikipedia.org/wiki/Buscador). Usando información adicional los resultados son más precisos, y el usuario se ahorra filtraciones manuales complementarias.

El intervalo semántico plantea el problema de que el usuario y el ordenador no se entiendan porque este último no comprenda el significado de los datos. Es posible que los metadatos posibiliten la comunicación declarando cómo están relacionados los datos. Por eso la [representación del conocimiento](https://es.wikipedia.org/wiki/Representaci%C3%B3n_del_conocimiento) usa metadatos para categorizar informaciones. La misma idea facilita la [inteligencia artificial](https://es.wikipedia.org/wiki/Inteligencia_artificial) al deducir conclusiones automáticamente.

Los metadatos facilitan el [flujo de trabajo](https://es.wikipedia.org/wiki/Flujo_de_trabajo) convirtiendo datos automáticamente de un formato a otro. Para eso es necesario que los metadatos describan contenido y estructura de los datos.[10](https://es.wikipedia.org/wiki/Metadato#cite_note-Wroe-10)​

Algunos metadatos hacen posible una [compresión de datos](https://es.wikipedia.org/wiki/Compresi%C3%B3n_de_datos) más eficaz. Por ejemplo, si en un vídeo el software sabe distinguir el primer plano del fondo puede usar algoritmos de compresión diferentes y así mejorar la cuota de compresión.[11](https://es.wikipedia.org/wiki/Metadato#cite_note-Hosch-11)​

Otra idea de aplicación es la presentación variable de datos. Si hay metadatos señalando los detalles más importantes, un programa puede seleccionar la forma de presentación más adecuada. Por ejemplo, si un teléfono móvil sabe dónde está localizada una persona en una imagen, tiene la posibilidad de reducirlo a las dimensiones de su pantalla. Del mismo modo un navegador puede decidir presentar un diagrama a su usuario ciego en forma táctil o leída.[12](https://es.wikipedia.org/wiki/Metadato#cite_note-12)​

Clasificación[[editar](https://es.wikipedia.org/w/index.php?title=Metadato&action=edit&section=5)]

Los metadatos se clasifican usando tres criterios:

**Contenido**

Subdividir metadatos por su contenido es lo más común. Se puede separar los metadatos que describen el *recurso mismo* de los que describen el *contenido del recurso*. Es posible subdividir estos dos grupos más veces, por ejemplo para separar los metadatos que describen el *sentido del contenido* de los que describen la *estructura del contenido* o los que describen el *recurso mismo* de los que describen el *ciclo vital del recurso*.

**Variabilidad**

Según la variabilidad se puede distinguir metadatos *mutables* e *inmutables*. Los inmutables no cambian, no importa qué parte del recurso se vea, por ejemplo el nombre de un fichero. Los mutables difieren de parte a parte, por ejemplo el contenido de un vídeo.[13](https://es.wikipedia.org/wiki/Metadato#cite_note-Smith-13)​

**Función**

Los datos pueden ser parte de una de las tres capas de funciones: subsimbólicos, simbólicos o lógicos. Los datos subsimbólicos no contienen información sobre su significado. Los simbólicos describen datos subsimbólicos, es decir añaden sentido. Los datos lógicos describen cómo los datos simbólicos pueden ser usados para deducir conclusiones lógicas, es decir añaden comprensión.[14](https://es.wikipedia.org/wiki/Metadato#cite_note-Stamou-14)​

Ciclo de vida[[editar](https://es.wikipedia.org/w/index.php?title=Metadato&action=edit&section=6)]

El [ciclo de vida](https://es.wikipedia.org/wiki/Ciclo_de_vida) de los metadatos comprende las fases *creación*, [*manipulación*](https://es.wikipedia.org/wiki/Manipulaci%C3%B3n) y [*destrucción*](https://es.wikipedia.org/wiki/Eliminar_%28inform%C3%A1tica%29). El análisis minucioso de cada una de las etapas saca a la luz asuntos significativos.

**Creación**[[editar](https://es.wikipedia.org/w/index.php?title=Metadato&action=edit&section=7)]

Se pueden crear metadatos *manualmente*, *semiautomáticamente* o *automáticamente*. El proceso manual puede ser muy laborioso, dependiente del formato usado y del volumen deseado, hasta un grado en el que los seres humanos no puedan superarlo. Por eso, el desarrollo de utillaje semiautomático o automático es más que deseable.

En la producción automática el software adquiere las informaciones que necesita sin ayuda externa. Aunque el desarrollo de algoritmos tan avanzados está siendo objeto de investigación actualmente, no es probable que la computadora vaya a ser capaz de extraer todos los metadatos automáticamente. En vez de ello, se considera la producción semiautomática más realista; aquí un servidor humano sostiene algoritmos autónomos con la aclaración de inseguridades o la proposición de informaciones que el software no puede extraer sin ayuda.

Hay muchos expertos que se encargan del diseño de herramientas para la creación de metadatos pero que ignoran cuestionar este proceso. Según los que no evitan el asunto, la generación no debe comenzar después de la terminación de un recurso sino que debe hacerse durante la fabricación: hay que archivar los metadatos tan pronto como se originan, con los conocimientos especiales del productor, para evitar una laboriosa reconstrucción posterior. Por eso, se tiene que integrar la producción de metadatos en el procedimiento de fabricación del recurso.[13](https://es.wikipedia.org/wiki/Metadato#cite_note-Smith-13)​

**Manipulación**[[editar](https://es.wikipedia.org/w/index.php?title=Metadato&action=edit&section=8)]

Si los datos cambian, los metadatos tienen que cambiar también. Aquí se hace la pregunta: ¿quién va a adaptar los metadatos? Existen modificaciones que pueden ser manejadas de forma sencilla y automática, pero hay otras donde la intervención de un servidor humano es indispensable.

La *metaproducción*, el reciclaje de partes de recursos para crear otros recursos, demanda atención particular. La fusión de los metadatos afiliados no es trivial, especialmente si se trata de información con relevancia jurídica, como por ejemplo la [gestión digital de derechos](https://es.wikipedia.org/wiki/Gesti%C3%B3n_digital_de_derechos).

**Destrucción**[[editar](https://es.wikipedia.org/w/index.php?title=Metadato&action=edit&section=9)]

Además hay que investigar la [destrucción de metadatos](https://es.wikipedia.org/wiki/Eliminar_%28inform%C3%A1tica%29). En algunos casos es conveniente eliminar los metadatos junto con sus recursos, en otros es razonable conservar los metadatos, por ejemplo para supervisar cambios en un documento de texto.

Metadatos en la informática[[editar](https://es.wikipedia.org/w/index.php?title=Metadato&action=edit&section=10)]

Los metadatos han cobrado gran relevancia en el mundo de [Internet](https://es.wikipedia.org/wiki/Internet), por la necesidad de utilizar los metadatos para la clasificación de la enorme cantidad de datos. Además de la clasificación los metadatos pueden ayudar en las búsquedas. Por ejemplo, si buscamos un artículo sobre vehículos, este dato tendrá sus correspondiente metadatos clave adjuntos, como 4 ruedas, motor, etc.

Otros ejemplos de usos de metadatos en la informática:

* [Metatags](https://es.wikipedia.org/wiki/Metatags) en [HTML](https://es.wikipedia.org/wiki/HTML): etiquetas con [información](https://es.wikipedia.org/wiki/Informaci%C3%B3n) sobre el propio documento web: autor, editor, [codificación](https://es.wikipedia.org/wiki/Codificaci%C3%B3n_de_caracteres), etc.
* Información en el propio [sistema de ficheros](https://es.wikipedia.org/wiki/Sistema_de_ficheros): [HFS](https://es.wikipedia.org/wiki/HFS) o [ReiserFS](https://es.wikipedia.org/wiki/ReiserFS%22%20%5Co%20%22ReiserFS), por nombrar dos. Se ven completados por los buscadores inteligentes ([Beagle](http://web.archive.org/web/http%3A/beaglewiki.org/) o [Spotlight](https://web.archive.org/web/20090512202004/http%3A/www.apple.com/es/macosx/features/spotlight/)) que saben reconocer estos metadatos.
* Clasificaciones de fotos: [F-Spot](http://www.gnome.org/projects/f-spot/), [Picasa](http://www.picasa.com/) o [iPhoto](http://www.apple.com/ilife/iphoto/), por ejemplo.
* Clasificadores de canciones: contienen metadatos sobre las canciones, bien en [MP3](https://es.wikipedia.org/wiki/MP3) o en [CD](https://es.wikipedia.org/wiki/CD) de audio, en un formato llamado [ID3](https://es.wikipedia.org/wiki/ID3). Por ejemplo: [iTunes](http://www.apple.com/es/itunes/) y [Rhythmbox](http://www.rhythmbox.org/).

Almacenamiento[[editar](https://es.wikipedia.org/w/index.php?title=Metadato&action=edit&section=11)]

Hay dos posibilidades para almacenar metadatos: depositarlos *internamente*, en el mismo documento que los datos, o depositarlos *externamente*, en su mismo recurso. Inicialmente, los metadatos se almacenaban internamente para facilitar la administración.

Hoy, por lo general, se considera mejor opción la localización externa porque hace posible la concentración de metadatos para optimizar operaciones de búsqueda. Por el contrario, existe el problema de cómo se liga un recurso con sus metadatos. La mayoría de los estándares usa [URIs](https://es.wikipedia.org/wiki/URI%22%20%5Co%20%22URI), la técnica de localizar documentos en la [World Wide Web](https://es.wikipedia.org/wiki/World_Wide_Web), pero este método propone otras preguntas, por ejemplo qué hacer con documentos que no tienen URI.

**Codificación**[[editar](https://es.wikipedia.org/w/index.php?title=Metadato&action=edit&section=12)]

Los primeros y más simples formatos de los metadatos usaron texto no cifrado o la codificación binaria para almacenar metadatos en ficheros.

Hoy, es común codificar metadatos usando [XML](https://es.wikipedia.org/wiki/XML). Así, son legibles tanto por seres humanos como por computadoras. Además este lenguaje tiene muchas características a su favor, por ejemplo es muy simple integrarlo en la World Wide Web. Pero también hay inconvenientes: los datos necesitan más espacio de memoria que en formato binario y no está claro cómo convertir la estructura de [árbol](https://es.wikipedia.org/wiki/%C3%81rbol_%28teor%C3%ADa_de_grafos%29) en una corriente de datos.

Por eso, muchos estándares incluyen utilidades para convertir XML en codificación binaria y viceversa, de forma que se unen las ventajas de los dos.

Vocabularios controlados y ontologías[[editar](https://es.wikipedia.org/w/index.php?title=Metadato&action=edit&section=13)]

Para garantizar la uniformidad y la compatibilidad de los metadatos, muchos sugieren el uso de un *vocabulario controlado*fijando los términos de un campo. Por ejemplo, en caso de sinónimos o interlenguaje hay que acordarse qué palabras se usan para evitar que el buscador localice «español» pero no «española».

Una [ontología](https://es.wikipedia.org/wiki/Ontolog%C3%ADa) además define las relaciones de los términos del vocabulario para que la computadora puede evaluarlas automáticamente. Así es posible presentar una página web sobre «[Vincent Van Gogh](https://es.wikipedia.org/wiki/Vincent_Van_Gogh)» aunque el usuario tecleó «pintores neerlandeses»; usando una ontología adecuada el buscador comprende que Van Gogh fue un pintor [neerlandés](https://es.wikipedia.org/wiki/Neerland%C3%A9s).

Un concepto muy similar a las ontologías son las [folksonomías](https://es.wikipedia.org/wiki/Folksonom%C3%ADa%22%20%5Co%20%22Folksonom%C3%ADa). Las ontologías son definidas por expertos del campo que ordenan los términos, pero las folksonomías son definidas por los mismos usuarios.

Metadatos y objetos de aprendizaje[[editar](https://es.wikipedia.org/w/index.php?title=Metadato&action=edit&section=14)]

En el *e-learning* los metadatos se emplean para describir los objetos y recursos de aprendizaje con el objetivo de facilitar las búsquedas en los repositorios. A menudo la información de metadatos de los repositorios de objetos de aprendizaje cumple el **estándar**[**IEEE LOM**](https://es.wikipedia.org/wiki/Learning_Object_Metadata), el cual define un conjunto de nueve categorías de información, que permiten describir los recursos tanto desde el punto de vista didáctico como técnico, lo que permite hacer búsquedas mucho más ajustadas que permitirán a los usuarios (profesores que componen un nuevo curso con materiales existentes, o alumnos interesados en aprender sobre un cierto tema) obtener resultados más ajustados a los criterios de búsqueda introducidos.

Los metadatos son una parte esencial del paradigma de los objetos de aprendizaje, puesto que:

* La reutilización de objetos de aprendizaje se basa en la creación y uso de metadatos, descripciones a menudo externas a los propios recursos.
* Los metadatos, si se proporcionan en los lenguajes adecuados, permiten desarrollar novedosas herramientas tecnológicas que facilitan la búsqueda y manipulación de los objetos de aprendizaje.
* Facilitan la recuperación de la información, al describir el contenido y sus relaciones con otros recursos.
* Facilitan la interoperabilidad, ya que hacen más fácil compartir e intercambiar información.
* Simplifican la gestión y el almacenamiento, pues permiten guardar información sobre el ciclo de vida de los recursos.
* Ayudan a gestionar adecuadamente y proteger los derechos de propiedad intelectual.

Los metadatos son por tanto un elemento de valor fundamental. Un recurso digital con un diseño pedagógico excelente no es *per se* un buen objeto de aprendizaje, sino que lo será en la medida en que los metadatos que lo describan sean también de calidad.

Paradatos[[editar](https://es.wikipedia.org/w/index.php?title=Metadato&action=edit&section=15)]

Se conoce como **paradatos** al conjunto de datos que se genera durante la interacción entre un usuario y un recurso o servicio de un escenario educativo (un entorno virtual de aprendizaje, repositorio, red social, etcétera). En función del recurso o servicio accedido y las operaciones realizadas con el mismo, los paradatos generados contendrán más o menos información. Esta información se puede almacenar para ser posteriormente analizada, con el objetivo de comprender mejor cómo los usuarios interactúan en dicho escenario educativo, detectando posibles problemas, así como oportunidades de mejora tanto del propio escenario educativo y de las herramientas utilizadas como del entendimiento del proceso que siguen los usuarios en el mismo. A partir del análisis realizado, es posible entonces construir, entre otros, sistemas de recomendación, esquemas de reputación, visualizaciones de la interacción, etc.

Por ejemplo, si un usuario se descarga un documento de un repositorio de objetos de aprendizaje, será posible saber que el usuario *U* ha descargado el documento *D* en el momento *T*. Esta información puede usarse para detectar aquellos recursos más descargados o, por el contrario, aquellos que nunca son accedidos por los usuarios, pero también conocer en qué épocas del semestre académico se produce una mayor utilización del repositorio. Otra posibilidad es si un usuario *U* evalúa un recurso *R* con una valoración *X* en el momento *T*, por ejemplo, un comentario en un grupo de Facebook. Esto permite conocer los comentarios mejor o peor valorados por los usuarios, así como los usuarios más activos.

En general, el objetivo es guardar la interacción de los usuarios con el sistema en forma de tuplas *(U, T, S, R, X)*: un usuario *U* en el momento *T* utiliza un servicio *S* sobre un recurso *R* con un resultado *X*. Esta es la mínima información que debe ser almacenada para su posterior análisis.

**Almacenamiento de paradatos**[[editar](https://es.wikipedia.org/w/index.php?title=Metadato&action=edit&section=16)]

Dado que la interacción en un entorno virtual de aprendizaje suele realizarse mediante el uso de un navegador web, un usuario que visita los espacios ofrecidos por el mismo deja un rastro en forma de accesos a páginas web que quedan recogidos en los ficheros de registro (en inglés, [log files](https://en.wikipedia.org/wiki/Server_log)) de los servidores web que soportan el sistema. Por lo tanto, parece factible analizar los ficheros de registro para extraer la información relativa a la interacción. El problema, no obstante, es que estos ficheros contienen muchas más entradas relativas a la disposición de los elementos que forman una página web, que realmente relacionadas con el propio resultado de la interacción de los usuarios, la cual debe ser calculada a partir de la secuencia de saltos capturada en los ficheros de registro. El coste computacional de analizar ficheros de registro es muy elevado (contienen millones de líneas de registro) y no es sencillo cuando se trata de sistemas complejos con múltiples servidores donde el rastro dejado por un usuario puede estar fragmentado en diferentes ficheros.

Por lo tanto, si es necesario recoger paradatos para su análisis posterior, es mejor que el sistema sea diseñado específicamente con un servicio de recogida enfocado a almacenar solamente la información requerida para el análisis, reduciendo el tamaño necesario así como el tiempo de proceso. En función del objetivo del análisis, los paradatos pueden almacenarse dentro de los metadatos de un recurso, dentro del perfil del usuario o bien, lo más habitual, en una tabla (o base de datos, en función de su complejidad) separada.

Crítica y problemas asociados al uso de los metadatos[[editar](https://es.wikipedia.org/w/index.php?title=Metadato&action=edit&section=17)]

Algunos expertos critican fuertemente el uso de metadatos. Sus argumentos más sustanciosos son:

* Los metadatos son costosos y necesitan demasiado tiempo. Las empresas no van a producir metadatos porque no hay demanda y los usuarios privados no van a invertir tanto tiempo.
* Los metadatos son demasiado complicados. La gente no acepta los estándares porque no los comprende y no quiere aprenderlos.
* Los metadatos dependen del punto de vista y del contexto. No hay dos personas que añadan los mismos metadatos. Además, los mismos datos pueden ser interpretados de manera totalmente diferente, dependiendo del contexto.
* Los metadatos son ilimitados. Es posible adherir más y más metadatos útiles y no hay fin.
* Los metadatos son superfluos. Ya hay buscadores potentes para textos, y en el futuro la técnica *query by example*(«búsqueda basada en un ejemplo») va a mejorarse, tanto para localizar imágenes como para música y vídeo.

Algunos estándares de metadatos están disponibles pero no se aplican: los críticos lo consideran una prueba de las carencias del concepto de metadatos. Hay que notar que este efecto también puede ser causado por insuficiente compatibilidad de los formatos o por la enorme diversidad que amedrenta a las empresas. Fuera de eso hay formatos de metadatos muy populares.[8](https://es.wikipedia.org/wiki/Metadato#cite_note-Bultermann-8)​

Si bien la inclusión de metadatos es necesaria, para facilitar y potenciar un buen número de tareas importantes, también existen problemas asociados al uso de los mismos. Algunos de los problemas más citados en la literatura científica se resumen en los siguientes puntos:

* Falta de compleción: introducir metadatos es a menudo una tarea ingrata y que requiere un esfuerzo considerable. Esto lleva a una cierta tendencia a no completar (o a completar deficientemente) los registros de metadatos, ya que las organizaciones bien no perciben adecuadamente la necesidad de ofrecer registros de metadatos completos o no pueden hacer frente a los altos costes, en términos de esfuerzo, de completar sus colecciones. Este problema es más patente cuando el número de objetos de aprendizaje a anotar es importante (colecciones de miles de objetos de aprendizaje) o cuando el número de elementos de metadatos a rellenar es elevado (por encima de la veintena).
* Dificultades de interoperabilidad: algunas informaciones de metadatos descansan en el uso de vocabularios, colecciones de términos cerradas cuyos elementos se han de emplear obligatoriamente para consignar la información relativa al elemento de metadatos en cuestión. Los estándares permiten utilizar distintos vocabularios para un elemento de metadatos, por lo que estos pueden variar de una institución a otra causando dificultades para intercambiar recursos, para que sistemas externos operen sobre metadatos con distinto origen, etc. Así por ejemplo, la información para el elemento *5.8.Dificultad* del estándar IEEE LOM ha de escogerse (según el estándar IEEE LOM) de la siguiente lista de términos: *muy fácil* / *fácil* / *medio*/ *difícil* / *muy difícil*. Sin embargo, el propio estándar IEEE LOM permite utilizar otro vocabulario si se estima oportuno, por lo que una cierta institución podría simplificar el número de categorías a 3, ampliarlo a 10 para ofrecer una escala más detallada, o elegir otra escala más adecuada a su contexto tal como por ejemplo: *básico* / *básico-necesita-apoyo* / *imprescindible-apoyo-externo*/ *complejo*.
* Inconsistencia semántica y otros problemas derivados de los estándares: no es infrecuente observar cómo dos instituciones distintas ofrecen informaciones diferentes para el mismo elemento de metadatos, a menudo por falta de claridad del estándar de metadatos utilizado. Tomando como ejemplo el estándar IEEE LOM, algunos elementos como el elemento *5.8.Dificultad* dependen en gran medida de la opinión subjetiva de la persona que crea el registro de metadatos y, en consecuencia, están destinados a ser inconsistentes con registros creados por otras personas pues uno podría considerar *difícil* lo que para otro es *muy difícil*. Otro problema similar es la incompletitud semántica, es decir, la introducción de información incompleta para un cierto elemento de metadatos, no proporcionando toda la información que sería posible -y deseable- ofrecer.
* Están "orientados a humanos": los metadatos son descripciones textuales que las personas interpretamos fácilmente. Sin embargo, a los sistemas y aplicaciones informáticas no les resulta sencillo procesar esta información para ofrecer servicios añadidos, ya que los metadatos no han sido escritos para que los entiendan "máquinas". Así, resulta difícil programar un buscador que priorice los resultados de una búsqueda de objetos de aprendizaje en un repositorio, en función de, por ejemplo, la información sobre derechos de acceso (elemento LOM *6.2.Derechos*) o de su cobertura geográfica o temporal (elemento *1.8.Cobertura*), sencillamente porque dicha información está en un texto escrito en un lenguaje humano del que es difícil extraer la información buscada: es preciso procesar el texto mediante complejas técnicas de análisis lingüístico, dividirlo en sus partes esenciales, analizar cada parte y extraer la información, etc., teniendo siempre en cuenta que en textos como la descripción de la cobertura puede haber localismos, omisiones obvias para las personas pero incomprensibles para una máquina, o cualesquiera otras complejidades propias de los lenguajes humanos.

Formatos y estándares[[editar](https://es.wikipedia.org/w/index.php?title=Metadato&action=edit&section=18)]

Hay dos grupos que impulsan el desarrollo de formatos de metadatos: la *técnica multimedia* y la *web semántica*. El destino de la técnica multimedia es describir un singular recurso de multimedia, el de la web semántica la descripción de recursos de cada tipo y además el encadenamiento de los conocimientos. Los formatos más populares y grandes son:

* [**ID3**](https://es.wikipedia.org/wiki/ID3) hace posible la notación de metadatos muy sencillos, tales como título e intérprete, en ficheros de audio MP3. El formato es muy popular y demuestra que los metadatos pueden ser útiles.
* [MPEG-7](https://es.wikipedia.org/wiki/MPEG-7)
* [MPEG-21](https://es.wikipedia.org/wiki/MPEG-21)
* [TV-Anytime](https://es.wikipedia.org/wiki/TV-Anytime)
* [EXIF](https://es.wikipedia.org/wiki/EXIF)
* [Dublin Core](https://es.wikipedia.org/wiki/Dublin_Core)
* [LOM](https://es.wikipedia.org/wiki/Learning_Object_Metadata) utilizado en los objetos de aprendizaje
* [Marco de descripción de recursos (RDF)](https://es.wikipedia.org/wiki/Marco_de_descripci%C3%B3n_de_recursos)
* [RDF Schema](https://es.wikipedia.org/wiki/RDF_Schema)
* [OWL](https://es.wikipedia.org/wiki/OWL)
* [NewsML](https://es.wikipedia.org/w/index.php?title=NewsML&action=edit&redlink=1)
* [SportsML](https://es.wikipedia.org/w/index.php?title=SportsML&action=edit&redlink=1)

Véase también[[editar](https://es.wikipedia.org/w/index.php?title=Metadato&action=edit&section=19)]

* [Dato](https://es.wikipedia.org/wiki/Dato)
* [Infraestructura de Datos Espaciales](https://es.wikipedia.org/wiki/Infraestructura_de_Datos_Espaciales)
* [Objeto de aprendizaje](https://es.wikipedia.org/wiki/Objeto_de_aprendizaje)
* [Estandarización](https://es.wikipedia.org/wiki/Estandarizaci%C3%B3n)