

# Diseño de Arquitecturas de Información: Descripción y Clasificación

Yusef Hassan Montero [nosolousabilidad.com]

Ana Núñez Peña [the-cocktail.com]

14/01/2005

[http://www.nosolousabilidad.com/articulos/descripcion\\_y\\_clasificacion.htm](http://www.nosolousabilidad.com/articulos/descripcion_y_clasificacion.htm)

## Organización de contenidos

La función de un arquitecto de información es organizar grandes cantidades de contenidos con el objetivo de que el usuario pueda manejarlos, navegar por ellos fácilmente y satisfacer sus necesidades de información. En otras palabras, organizar el desorden, hacer recuperable, localizable o accesible la información ('findability').

En este artículo revisaremos las distintas etapas, conceptos y herramientas para la organización de contenidos: análisis, descripción, indización, y clasificación. En la figura 1 podemos observar un esquema resumen del proceso de organización de contenidos.

Diseño de Arquitecturas de Información: Organización de Contenidos

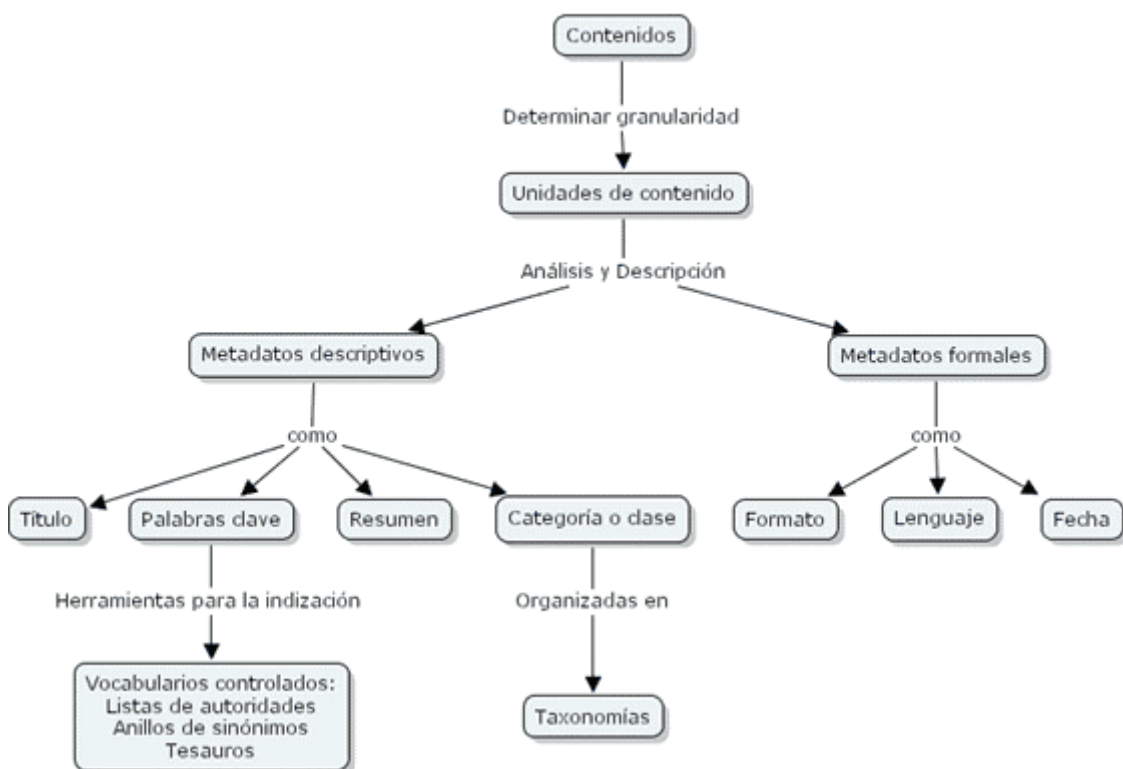


Figura 1: Esquema del proceso de organización de contenidos

## Contenidos y granularidad

Por contenidos se entiende todos aquellos 'trozos' de información u objetos informativos a organizar, estructurar y clasificar por el arquitecto de información: textos, imágenes, videos, etc.

Una primera tarea del arquitecto es analizar la naturaleza de los contenidos a organizar, y determinar su granularidad, que se refiere al nivel de descomposición o grado en que pueden ser divididos estos contenidos.

El nivel mínimo de granularidad de un contenido es aquel grado de descomposición en el que una determinada información sigue manteniendo su significación comunicativa, o el grado en el que físicamente no puede seguir descomponiéndose.

Por ejemplo, habrá casos en los que una misma imagen pueda ser utilizada en contextos diferentes - en asociación con diferentes objetos informativos-, o que por el contrario, se encuentre vinculada de forma persistente a un contexto determinado. Este caso podría ser el de una imagen embebida en un documento con formato hermético, como PDF. Este documento sería entonces el nivel mínimo de granularidad: **una unidad de contenido**.

Una vez determinados los grados de granularidad de los contenidos, la siguiente tarea es catalogar o describir dichas unidades de contenido, aplicarles metadatos.

## Metadatos

Los metadatos son 'información acerca de información'. **Aplicar metadatos significa caracterizar cada unidad de contenido a través de una serie de pares 'atributo-valor/es'**.

Los atributos pueden ser formales o administrativos - lenguaje, formato, fecha de creación, tamaño...-, o *pueden ser descriptivos* – título, resumen, palabras clave... -. Mientras que los administrativos tienen una función relacionada principalmente con la gestión de los contenidos, los descriptivos tienen por objeto facilitar y mejorar la recuperación de información por parte del usuario.

Ejemplo de metadatos:

Atributo	Valor/es
Fecha	30/08/1998
Título	Evaluación de la calidad de uso
Palabras clave	test de usuarios; evaluación heurística; card sorting
Categoría	Desarrollo web

El número y tipos de metadatos a aplicar dependerá de la naturaleza de los contenidos. Si por ejemplo lo que se está catalogando es un conjunto fotográfico, los metadatos a aplicar podrían ser: autor, formato, tamaño, peso, resolución, color, palabras clave, ...

**De cara al usuario y la recuperación de información, los tipos de metadatos más útiles son los descriptivos**, ya que indican explícitamente sobre qué trata la unidad de contenido.

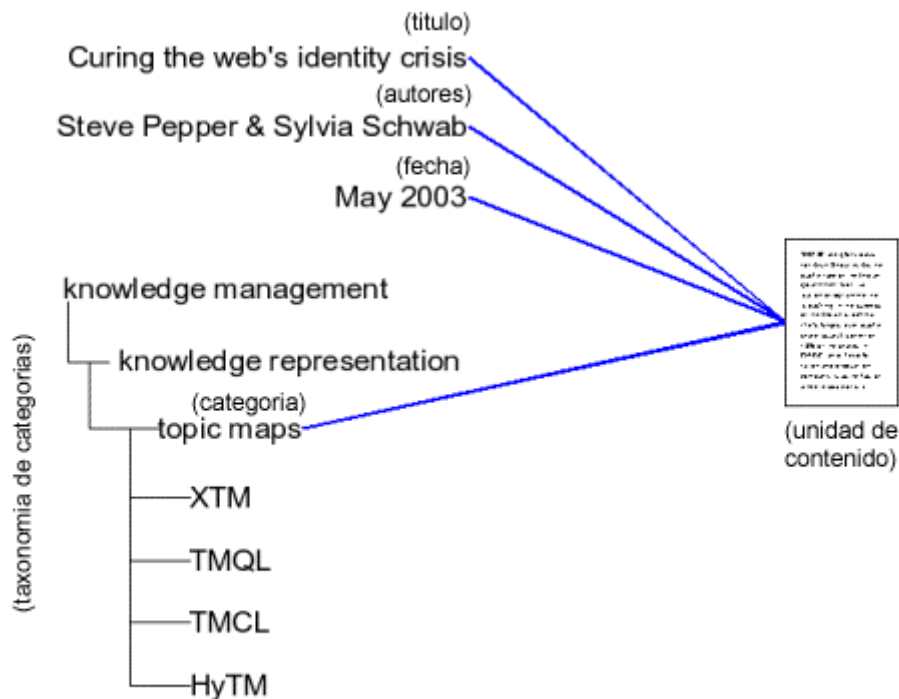


Figura 2. (Garshol, 2004)

En la figura de ejemplo podemos observar el esquema de una unidad de contenido descrita mediante metadatos, donde las líneas azules representan los atributos (título, autores, fecha y categoría) que vinculan cada unidad de contenido con sus correspondientes valores para estos atributos.

## Palabras Clave y Vocabularios Controlados

**Las palabras clave son un conjunto de términos que representan los conceptos más significativos de la unidad de contenido.** Resumen su contenido a través de los conceptos más relevantes tratados.

**El primer inconveniente que presentan** está relacionado directamente con la ambigüedad propia del lenguaje natural: *polisemia y sinonimia*. Un término puede representar más de un concepto diferente y un mismo concepto podría ser representado por varios términos diferentes.

Para evitar estos problemas que podrían repercutir en su utilidad descriptiva, es necesario contar con herramientas o mecanismos que aseguren la consistencia de los valores definidos para estos metadatos, que controlen la terminología empleada.

Un vocabulario controlado es una lista o índice de términos que establece relaciones unívocas y precisas entre ellos, así como con los conceptos representados. Dependiendo del número y tipo de relaciones definidas entre los términos se suele hablar de anillos de sinónimos o listas de autoridades como ejemplos de vocabularios controlados simples, o de tesauros como un tipo de vocabulario controlado más complejo. Mientras que en los primeros las relaciones entre términos son únicamente de sinonimia y de preferencia, en los *tesauros* el número de relaciones entre términos es mayor:

### *Jerárquicas:*

- TG o Término Genérico: Relación del término con un término "padre", del que desciende.
- TE o Término Específico: Relación del término con un término "hijo", más concreto.

### *Equivalencia:*

- USE: Relación entre el término no preferente y el término preferente, o término a usar (más aceptado).
- UP o Usado Por: Relación entre el término preferente y el término no preferente.

### *Asociativas:*

- TR o Término Relacionado: Relación entre el término y otro término afín.

### *Semánticas:*

- NA o Nota de Alcance: Explicación o descripción del significado del término

Indizar comprendería la tarea de catalogar o describir los recursos a través de palabras clave que forman parte de un índice terminológico o vocabulario controlado.

La primera ventaja de usar un vocabulario controlado en la indización de las unidades de contenido tiene que ver con la consistencia de las palabras clave asignadas: cada concepto tendrá una correspondencia unívoca con un término, y viceversa, eliminando así los problemas de polisemia y sinonimia.

Otra ventaja de su uso estará relacionada con la recuperación o búsqueda de estas unidades de contenido por el usuario mediante los sistemas de búsqueda y navegación, tema del que hablaremos posteriormente.

## **Categorías y Taxonomías**

La distinción entre palabras clave y categorías – y por tanto entre indización y clasificación – es una distinción formal. Mientras que las palabras clave asociadas a una unidad de contenido representan sus conceptos más significativos, la categoría asociada describe el tema principal bajo el que se engloba la unidad de contenido.

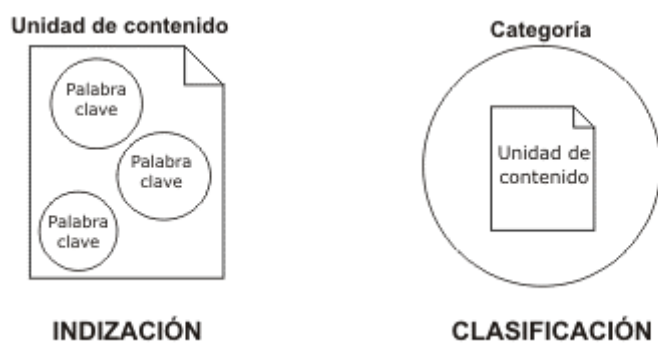


Figura 3: Diferencia entre palabras clave y categorías

Una unidad de contenido podría pertenecer a más de una categoría, aunque esta forma de clasificar haría más difusa, todavía, la diferencia entre categorías y palabras clave.

Las categorías y sus relaciones se organizan en lo que en el entorno de la Arquitectura de la Información se suelen denominar taxonomías de categorías.

El término taxonomía proviene del área de la Biología que estudia la clasificación de los seres vivos en estructuras jerárquicas. Muy utilizado en Arquitectura de la Información, resulta un término bastante ambiguo y vago, y según en qué contexto se utilice se le podría asignar un significado u otro.

De forma general podemos decir que una taxonomía es una estructura jerárquica de elementos o grupos de elementos, por lo que es normal encontrar utilizado este término en relación con los tesauros, ontologías, sistemas de clasificación, etc.

En este artículo el término taxonomía es entendido como la organización jerárquica del conjunto de categorías bajo las que clasificar las unidades de contenido. Un ejemplo de taxonomía de categorías es el que se observa en la figura 2, donde las líneas negras representan las relaciones jerárquicas entre categorías, y la línea azul el vínculo de pertenencia entre unidad de contenido y categoría.

## **Clasificaciones facetadas**

Los principios básicos de las clasificaciones facetadas ya fueron tratados en '[Clasificaciones facetadas y metadatos \(I\): Conceptos básicos](#)'. Por otro lado, en '[Clasificaciones facetadas y metadatos \(II\): XFML](#)' se trató su definición e implementación mediante lenguajes de etiquetado, y en '[Sistemas facetados de exploración de archivos](#)' se analizó un ejemplo de sistema de navegación facetado.

**La clasificación facetada parte de la premisa de que una unidad de contenido puede ser descrita a través de varias dimensiones o facetas**, cada una de las cuales contiene su propia relación de posibles valores o categorías.

Siguiendo el enfoque de este artículo, todos aquellos tipos de metadatos donde más de una unidad de contenido pudiera compartir valor, pueden ser considerados facetas. No tendría sentido, por ejemplo, considerar facetas a tipos de metadatos como 'título' o 'resumen', ya que sus posibles valores serán únicos para cada unidad de contenido. En cambio, 'categoría temática', 'fecha', 'formato' o 'idioma' sí podrían ser consideradas facetas de una misma clasificación.

La clasificación facetada resulta realmente útil cuando las unidades de contenido son lo suficientemente homogéneas como para que todas puedan ser descritas a través de un conjunto de metadatos comunes.

Un ejemplo podría ser la organización de anotaciones en una bitácora, como es el caso de [efimera.org](http://efimera.org):

Clasificado en:

- ▶ Autor: [Juanjo Seixas](#)
- ▶ Espacio: [Galiza](#)
- ▶ Forma: [Sitio web](#)
- ▶ Tópico: [Diseño](#)
- ▶ Tiempo: [Diciembre de 2004](#)

*Figura 4:* Pie de anotación de efímera.org

En este blog, cada anotación o post se encuentra descrito por un conjunto de metadatos facetados comunes: espacio, forma, tópico y tiempo.

#### **Anotaciones ordenadas por tópicos**

[Arte](#) (44) | [Cartografía](#) (0) | [Cine](#) (7) | [Cómico](#) (1) | [Cuerpo](#) (2) | [Dios](#) (9) | [Diseño](#) (142) | [Educación](#) (18) | [Fotografía](#) (34) | [Ilustración](#) (23) | [La ciudad y yo](#) (58) | [Literatura](#) (5) | [Moda](#) (1) | [Música](#) (12) | [Poesía](#) (7) | [Política](#) (64) | [Sociedad](#) (107) | [Tecnología](#) (37) |

#### **Anotaciones ordenadas por su forma**

[Anotación](#) (29) | [Aplicación](#) (5) | [Artículo](#) (14) | [Blog](#) (15) | [Código](#) (1) | [Cómico](#) (5) | [Correo-e](#) (2) | [Disco](#) (3) | [Exposición](#) (8) | [Flash](#) (2) | [Gente](#) (3) | [Ilustración](#) (7) | [Juego](#) (2) | [Libro](#) (21) | [Manifiesto](#) (3) | [Mapa](#) (8) | [Objeto](#) (1) | [Película](#) (16) | [Producto](#) (3) | [Proyecto](#) (12) | [Revista](#) (3) | [Sitio web](#) (48) | [Tipografía](#) (7) | [Tutorial](#) (1) | [TV](#) (3) |

#### **Anotaciones ordenadas espacialmente**

[África](#) (1) | [Alemania](#) (3) | [Argentina](#) (3) | [Australia](#) (2) | [Austria](#) (1) | [Brasil](#) (1) | [Canadá](#) (1) | [Checoslovaquia](#) (1) | [Chile](#) (2) |

*Figura 5:* [Sección de archivos](#) de efímera.org

## **De la Arquitectura a la Interacción**

Describir y categorizar los contenidos de un sitio web no tendría sentido si no facilitara la interacción entre usuario y espacio de información, la vía y forma en que el usuario puede navegar y explorar el sitio web en busca de la información necesitada.

A continuación describimos cómo se verá reflejada la organización de contenidos en los dos ejes principales de interacción entre usuario y sitio web: sistemas de búsqueda y de navegación.

## Sistemas de búsqueda

El sistema de búsqueda o buscador interno es una de las herramientas más útiles para el usuario de cara a la recuperación de información.

Su funcionamiento se basa en la indización automática de la parte textual de las unidades de contenido del sitio. De forma simplificada y muy resumida podemos decir que esta indización se realiza contabilizando frecuencias de ocurrencia o aparición de términos en las unidades de contenido. A partir de estas frecuencias de aparición (entre otras variables) se pondera o estima la capacidad de cada término para representar el contenido de la unidad.

Cuando el usuario introduce una consulta, el sistema equipara la consulta con los índices ponderados de términos o palabras clave, y devuelve al usuario como resultados pertinentes aquellas unidades en las que, con mayor peso, aparecen los términos introducidos.

La indización intelectual o humana – descrita en este artículo- aumentaría la 'recuperabilidad' de las unidades de contenido en términos de precisión y exhaustividad. Por ejemplo, el sistema de búsqueda podría responder a consultas donde el usuario utilizara términos que no aparecieran en el cuerpo textual de la unidad de contenido, pero sí entre las palabras clave asignadas a esta unidad por el arquitecto de información.

Este caso se vuelve más gráfico cuando lo que se describe mediante palabras clave son imágenes o fotografías, donde un buscador por sí solo no podría ofrecer resultados, ya que no tienen base textual que pueda ser indizada automáticamente.

A los beneficios de la indización humana habría que sumar aquellos derivados de la implementación y utilización de vocabularios controlados por parte del sistema de búsqueda.

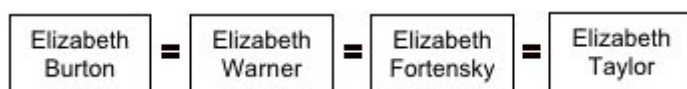


Figura 6: (Fast, Leise, Steckel; 2003b)

Al definir relaciones de sinonimia o preferencia entre términos (figura 6) el sistema de búsqueda podría devolver como resultados unidades de contenido en cuyo texto o palabras clave asignadas apareciera el mismo concepto pero en diferente forma, o también podría sugerir al usuario reformular la consulta.

## Sistemas de navegación

El sistema de clasificación de las unidades de contenido condicionará la navegación del usuario. Si han sido clasificados mediante una taxonomía jerárquica de categorías, la estructuración de los nodos del sitio web seguirá una forma jerárquica, podrá emplear elementos de navegación propios de este tipo de estructuras como son las [migas de pan](#), y permitir al usuario una navegación vertical (figura 7).

El **diseñador de interacción** deberá decidir cómo se verán reflejadas las categorías en la estructura de navegación, si habrá nodos que contengan más de un nivel de la categorización o la forma visual de mostrar estas categorías.

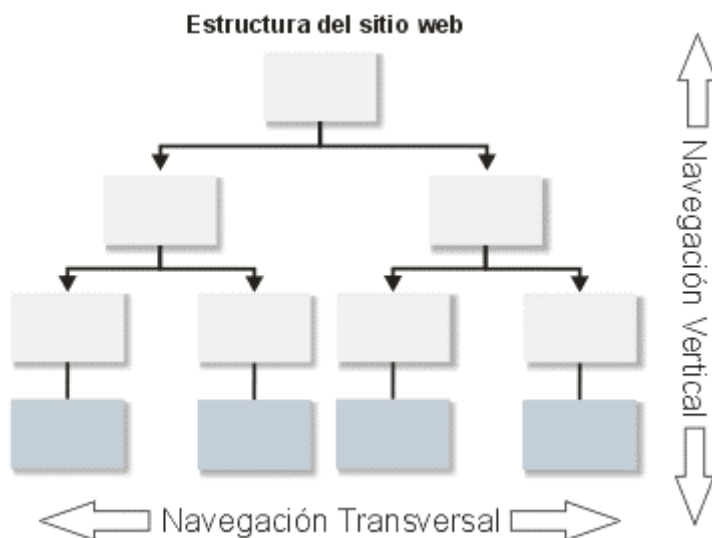


Figura 7: Tipos de navegación

Si a las unidades de contenido le han sido aplicados metadatos en forma de palabras clave, se podría posibilitar al usuario, cuando se encuentre visualizando una unidad de contenido, realizar una navegación 'transversal', mostrando en forma de enlaces las palabras clave asignadas a la unidad. Una vez que el usuario hace 'clic' sobre una de las palabras clave, se le mostraría una lista ordenada – bajo criterios de relevancia, similitud o cronológicos – de aquellas unidades de contenido que comparten en su descripción esa misma palabra clave.

Otra opción de navegación transversal sería hacer invisibles las palabras clave para el usuario, mostrando directamente enlaces a otras unidades de contenido que compartan alguna palabra clave con la unidad que se encuentre visualizando en ese momento.

En el caso de que estas palabras clave hubieran sido asignadas mediante un tesoro, existiría la opción de ofrecer al usuario navegar y explorar el propio tesoro, pero desde el punto de vista de la usabilidad normalmente no resulta recomendable. Los tesauros son índices con muchas palabras clave y muchas relaciones definidas entre ellas, por lo que normalmente resultarán complejos de entender y navegar por el usuario. La verdadera utilidad de los tesauros es como herramienta de control terminológico para el indizador, y no tanto como herramienta de navegación.

Otros metadatos como 'fecha' o 'autor' podrían ser utilizados para generar herramientas de navegación en forma de índices cronológicos o alfabéticos. Igualmente, en base a metadatos como 'lugar' se podrían generar mapas topográficos de navegación.

El caso de las clasificaciones facetadas puede dar lugar a distintos sistemas de navegación. Si bien una navegación facetada 'pura' como la descrita en ([Varela-Tabuyo; 2004](#)) es una navegación por filtrado - donde el usuario va seleccionando valores para cada atributo o faceta, reduciendo así el conjunto de posibles resultados para su búsqueda -, también se

puede optar por definir cada faceta como una vía de navegación independiente del resto – véase figura 5 -.

## Conclusiones

En este artículo se ha intentado describir de forma resumida todo el proceso de descripción, indización y clasificación de contenidos; proceso que dependiendo de la cantidad de contenidos y su naturaleza puede llegar a convertirse en una ardua tarea.

Como podemos comprobar, las posibilidades y herramientas de navegación basadas en metadatos son diversas, y finalmente quedará en manos del diseñador de interacción la decisión de ofrecer unas u otras, en base siempre a las necesidades y habilidades de la audiencia objetiva.

En lo que respecta a la clasificación y descripción de contenidos podemos concluir que a mayor grado de descripción de las unidades de contenido – a mayor número de metadatos diferentes aplicados -, mayores serán las opciones de navegación que se le puedan ofrecer al usuario final.

Para ampliar información sobre el tema recomendamos consultar la bibliografía anexa, donde son tratados con más profundidad muchos de los conceptos mencionados en este artículo.

## Bibliografía

*Denton, W. (2003a).* [Putting Facets on the Web: An Annotated Bibliography](#). Miskatonic University Press.

*Denton, W. (2003b).* [How to Make a Faceted Classification and Put It On the Web](#). Miskatonic University Press.

*Fast, K.; Leise, F.; Steckel, M. (2002a).* [All About Facets & Controlled Vocabularies](#). Boxes & Arrows.

*Fast, K.; Leise, F.; Steckel, M. (2002b).* [What Is A Controlled Vocabulary?](#). Boxes & Arrows.

*Fast, K.; Leise, F.; Steckel, M. (2003a).* [Creating a Controlled Vocabulary](#). Boxes & Arrows.

*Fast, K.; Leise, F.; Steckel, M. (2003b).* [Synonym Rings and Authority Files](#). Boxes & Arrows.

*Fast, K.; Leise, F.; Steckel, M. (2003c).* [Controlled Vocabularies: A Glosso-Thesaurus](#). Boxes & Arrows.

*Garshol, L.M. (2004).* [Metadata? Thesauri? Taxonomies? Topic Maps!](#). Journal of Information Science, Vol. 30, n. 4, pp. 378-391, 2004.

*Hassan, Y.; Martín, F.J.; Martín, O. (2003a).* [Clasificaciones facetadas y metadatos \(I\): Conceptos básicos](#). Nosolousabilidad.com.

*Hassan, Y.; Martín, F.J.; Martín, O. (2003b).* [Clasificaciones facetadas y metadatos \(II\): XFML](#). Nosolousabilidad.com.

*Lider, B.; Mosoiu, A. (2003).* [Building a Metadata-Based Website](#). Boxes & Arrows.

*Morville, P. (2002).* [Bottoms Up: Designing complex, adaptive systems](#). New Architect Magazine.

*Ricci, C. (2004).* [Developing and Creatively Leveraging Hierarchical Metadata and Taxonomy](#). Boxes & Arrows.

*Rosenfeld, L.; Morville, P. (2002).* Information Architecture for the World Wide Web (2nd edition). O'Reilly. ISBN 0-596-00035-9.

*Rovira, C. (2004).* [Mapa conceptual sobre metadatos, taxonomias, ontologias y mapas conceptuales](#). MapasConceptuales.com.

*Varela-Tabuyo, R. (2004).* [Sistemas facetados de exploración de archivos](#). Nosolousabilidad.com.

*Warner, A.J. (2002).* [A Taxonomy Primer](#). Lonomy.