

Organización de los contenidos en los sitios Web: las taxonomías

Lic. Anisleiby Fernández Hernández¹

RESUMEN

Se abordan las taxonomías en la organización y representación de los contenidos en el Web. Se abordan las taxonomías en la arquitectura de información, la función del arquitecto de información; así como los principales elementos a considerar para construir una taxonomía.

Palabras clave: Taxonomías, arquitectura de información.

ABSTRACT

The taxonomies in the organization and representation of the contents in the Web are considered, as well as the taxonomies in information architecture, the function of the information architect, and the main elements to take into consideration to make a taxonomy.

Key words: Taxonomies, information architecture.

Copyright: © ECIMED. Contribución de acceso abierto, distribuida bajo los términos de la Licencia Creative Commons Reconocimiento-No Comercial-Compartir Igual 2.0, que permite consultar, reproducir, distribuir, comunicar públicamente y utilizar los resultados del trabajo en la práctica, así como todos sus derivados, sin propósitos comerciales y con licencia idéntica, siempre que se cite adecuadamente el autor o los autores y su fuente original.

Cita (Vancouver): Fernández Hernández A. Organización de los contenidos en los sitios Web: las taxonomías. Acimed 2007;15(5). Disponible en:

http://bvs.sld.cu/revistas/aci/vol15_5_07/aci12507.htm [Consultado: día/mes/año].

La sociedad humana siempre ha generado un conocimiento sobre su entorno y la realidad, física y social que lo rodea. Ello surge de la propia necesidad del hombre de tratar de explicar o comprender el medio que habita. Pero, en las formas de organización social actuales, ha adquirido una dimensión y un lugar en la sociedad, que va más allá de su función histórica. Hoy, la función del conocimiento en las organizaciones ha cambiado drásticamente como resultado del surgimiento de nuevos paradigmas, asociados a la denominada sociedad de la información y a una nueva economía, basada en el conocimiento.

El desarrollo de las llamadas nuevas tecnologías de la información y la comunicación (NTICs) cambia la sociedad y si el medio se transforma, entonces también cambia el acceso a la información, que es hoy mucho más rápido, amplio y preciso. Existe una mayor demanda de información y una participación más directa en la búsqueda de información por parte de los usuarios.

Bajo esta dinámica, las NTICs pueden facilitar parte de las condiciones relacionadas con la implementación de una gestión del conocimiento, al soportar y potenciar las capacidades de cada uno de los miembros de una organización.

En los últimos años, diversos sistemas de organización del conocimiento (SOC) se han utilizado con éxito para ordenar la información en la World Wide Web: desde los tradicionales sistemas de clasificación y tesauros, hasta las más novedosas taxonomías, ontologías y las redes semánticas.

Un sitio que contiene información bien estructurada y clasificada puede ser fácilmente entendido por otros sitios Web y por los buscadores, así como cumplir mejor sus objetivos de diseminar la información lo más ampliamente posible, además de permitir al usuario encontrar lo que busca.

Describir y categorizar los contenidos de un sitio Web no tendría sentido si no facilitara la interacción entre el usuario y el espacio de información, la vía y forma en que el usuario puede navegar y explorar el sitio Web en busca de la información necesitada. En este sentido, las taxonomías tienen un lugar fundamental, porque permiten crear una jerarquía y organización de los contenidos que se presentan en un sitio.

A principios de los años 90 del siglo XX, el concepto de taxonomía se incorpora a diversos ámbitos del conocimiento, como la psicología, las ciencias sociales y la informática, para designar casi todos los sistemas de acceso a la información que intentan establecer coincidencias entre la terminología del usuario y del sistema. Los primeros especialistas que desarrollaron sistemas de organización de contenidos para la Web formaban parte del área de consultoría en gestión del conocimiento, y procedían de ambientes próximos a la informática y la ingeniería -gestión de contenidos y arquitectura de la información; ellos no conocían la tradición de los lenguajes documentales de la Bibliotecología y la Documentación, y entonces, asignaron el término taxonomía para los sistemas que desarrollaban. Este término se mantiene en uso actualmente para designar los sistemas de organización de contenidos en el contexto de Internet, aunque la teoría y la práctica de los lenguajes documentales se aplicando de forma intensiva en este contexto.

Por ello, en esta contribución, se pretende valorar el uso de las taxonomías para la organización de contenidos en sitios Web en forma general

Las taxonomías y la organización de los contenidos en sitios Web

Debido a la popularidad que la Web ha adquirido en los últimos años, y al constante progreso en este terreno, quienes organizan los conocimientos han desarrollado una serie de habilidades, experiencias y métodos para realizar este tipo de tareas. Estos métodos han apoyado sustancialmente el perfeccionamiento de la organización de los contenidos expuestos en una Web, sobre la base de algunos criterios de organización, entre los cuales se encuentran las taxonomías; ellos han posibilitado que el tiempo de desarrollo de los sitios Web cada vez sea menor, y por consiguiente, que se construya un Web con mayor celeridad.

Una buena organización de los contenidos expuestos en un Web facilitará el uso del producto al usuario final, mejorar la eficacia de la recuperación de la información y por

ende lograr una eficiente gestión de la información y del conocimiento; mientras mayor sea la estructuración y organización de la información, más fácil será al usuario hallar lo que realmente busca, además de que existen pruebas de que los sitios con un alto nivel de organización de sus contenidos son mucho más visitados en la Word Wide Web.

Muchas veces los usuarios saben exactamente lo que buscan, conocen las etiquetas y términos que deben utilizar para encontrar lo que buscan, sólo desean encontrarlo e irse. Otros, por el contrario, no saben lo que están indagando. Vienen al sitio con una idea vaga de la información que necesitan; no pueden siquiera conocer las etiquetas para describir lo que desean o incluso si existen o no. Esto tiene sus ventajas porque mientras exploran ocasionalmente el sitio para encontrar lo que necesitan, pueden aprender sobre productos y salir del sitio con conocimientos (o productos) que antes no conocían, ni sabían que los necesitaban, y esto se puede lograr porque el sistema está bien diseñado, existe una estructura y una organización de los contenidos.

El desarrollo de las NTICs apoya, en cierto sentido, esta actividad, pero son sólo los especialistas quienes son capaces y deben saber seleccionar y organizar los contenidos. El contenido es una cuestión exclusiva de los individuos, las tecnologías sólo facilitan su gestión.

Crear una página Web fácil de usar no es tarea fácil. Una experiencia interactiva positiva se deriva en muchos casos de la correcta organización de los contenidos de un Web, y de taxonomías que así lo permitan y faciliten.

Si deseamos diseñar una página que el usuario final sepa utilizar, genere experiencia interactiva positiva y, a su vez, resulte un proyecto rentable, no debe perderse de vista la función fundamental que desempeñan las taxonomías en este sentido. Como estructuras predeterminadas que se emplean para dividir un área temática -en este caso, el área temática se refiere al contenido del sitio-, y estas áreas en otras más pequeñas y así sucesivamente, ellas posibilitan organizar y estructurar los contenidos en clases, a partir de propiedades de conceptos, por su similitud. Se requiere entonces de un análisis conceptual que diferencie estos conceptos (clases) y a su vez, los agrupe según sus características.

Muchos son los autores que han abordado el tema de las taxonomías, por ejemplo, *Abed Gregio* afirma que:

“Una taxonomía es un proceso científico (o un sistema particular) de categorizar entidades, es decir, de organizarlas en grupos. Un sistema taxonómico de ser claro y consistente, flexible, exhaustivo y práctico”.¹

El mismo autor plantea además que la base para el desarrollo de una buena taxonomía son las características taxonómicas, que son las propiedades o características de los objetos que se categorizarán; estas características taxonómicas se denominan comúnmente atributos y deben satisfacer o cumplir los siguientes requisitos:¹

- **Objetividad:** cuando la característica se identifica sobre la base de un conocimiento objetivo exclusivamente. El objeto (atributo) debe observarse claramente.

- Determinación: cuando existe un proceso claro que pueda seguirse para identificar (extraer) la característica.
- Reproducibilidad: cuando varias personas describen (extraen) de forma independiente las características de un mismo objeto y coincidan en el valor observado.
- Mutuamente excluyente: cuando la inclusión de un grupo en una categoría lo excluye de la categorización en cualquier otra.
- Exhaustivos: cuando los grupos incluyen todas las posibilidades.
- Aceptable: lógica e intuitivamente: de forma que la comunidad acepte las categorías.
- Útil: cuando puede utilizarse para obtener conocimiento en el campo de investigación.

De ello, se deduce que una taxonomía aplicada a un sitio Web es una organización jerarquizada, que se emplea para estructurar contenidos, ella divide estos en clases y los agrupa según sus características. Por tanto, favorecen la navegabilidad y facilitan además acceder al capital intelectual de una determinada organización (figura 1).



Fig. 1. Taxonomías.

Por supuesto que esta definición no es la única. De hecho, existen muchísimas ofrecidas por diferentes autores, sobre todo, basadas más en los procesos que se han de seguir para la construcción de cualquier taxonomía.

En realidad, el objetivo último, a la hora de desarrollar una taxonomía, es mejorar la organización de los contenidos mediante la representación de la información.

Lo que puede parecer una tarea “rápida y fácil”, definir la estructura de la navegación y la organización de los contenidos de un sitio Web, es también uno de los factores más críticos del éxito. La estructura del sitio, también designada como la arquitectura o la taxonomía de la información, es el fundamento del sitio Web. Crear una taxonomía eficaz es tanto un arte como una ciencia. Si se utilizan términos incorrectos, los usuarios no podrán navegar por intuición. Crear un sitio demasiado profundo generará como consecuencia que al usuario le cueste más trabajo encontrar el contenido que busca.

Las taxonomías, la arquitectura de información y el profesional de la información

En la disciplina de *Arquitectura de la Información* (AI) se emplea frecuentemente el término “taxonomías”, algo muy lógico si se considera que la AI se ocupa de estructurar y organizar los contenidos (información) en productos Web.

Si nos circunscribimos exclusivamente a la AI en el campo del Web, una de las definiciones ofrecidas por *Rosenfeld* y *Morville* puede resultarnos de fácil comprensión:

“... Es el arte y la ciencia de estructurar y clasificar los contenidos de los sitios Web e intranets con el fin de ayudar a los usuarios a encontrar y manejar la información”.²

En la práctica, el “arquitecto de información” crea una gran parte de los cimientos para la organización del contenido de un sitio, sin importar dónde resida el contenido de este, sea en una multimedia, en campos de una base de datos o en sitios complejos y portales; él crea la estructura o mapa de información que posibilita a otros encontrar sus propios caminos hacia el conocimiento. No se ocupa de los aspectos formales del diseño de las interfaces, desde el punto de vista de la AI, la interfaz tiende a ser transparente, lo que interesa es el contenido y la forma en que el usuario accede a este, de manera que pueda manejarlo, navegar por ellos fácilmente y satisfacer sus necesidades de información. El “arquitecto de información” ayuda a evitar que la falta de organización convierta la experiencia del usuario en una pesadilla.

En el esquema de la organización de contenidos que presenta *Montero* evidencia claramente cómo en el campo de la AI se emplean las taxonomías para la organización de los contenidos a desarrollar en un producto Web (figura 2).



Fig. 2. Esquema de la organización de los contenidos.

Su autor ilustra cómo el arquitecto de información organiza y estructura los contenidos, analiza su naturaleza, determina su nivel de granularidad, hasta que físicamente no pueden descomponerse más, cómo introduce los metadatos descriptivos y así sucesivamente hasta llegar a las categorías o clases organizadas en taxonomías.

En los pasos metodológicos que se aplican para crear las arquitecturas de información se encuentra muy presente el profesional de la información en su función de arquitecto de información para la organización de los contenidos de un producto Web.

En el paso tres y cuatro específicamente de la guía propuesta por Montes de Oca puede apreciarse:³

En el paso tres, “Determinación de los contenidos y los requerimientos funcionales”, el autor afirma que al llegar a este punto, existe suficiente información para poder determinar los contenidos. “Y que a su vez, con el análisis de la competencia, (realizado en pasos anteriores), puede identificarse un grupo de contenidos que tal vez se adecuen al sitio que se creará, sea por tratarse de productos similares, de organizaciones con misiones similares o por contar con un universo de usuarios potenciales con similares necesidades informativas y características...”. Deben analizarse, y lo expresa muy bien el autor, las necesidades informativas identificadas en el estudio de los usuarios para determinar cuáles contenidos son los más apropiados para satisfacer las necesidades existentes y las fuentes de información de dónde se extraerán dichos contenidos, porque

para cada grupo temático, existen elementos particulares de contenidos y requerimientos funcionales específicos.

En el paso cuatro: “Definir la estructura del sitio”, el autor propone que se debe realizar el etiquetado de los contenidos seleccionados en el paso anterior, estas deben ser comprensibles para los usuarios y, a la vez, representativas del contenido que encierran.

En un nivel más profundo, al entrar a analizar las etiquetas que representarán el contenido, se encuentran las taxonomías o taxones que permiten lograr una mejor estructuración y un nivel mayor de organización de los contenidos que se desean representar.

Una vez determinadas las etiquetas que designarán al contenido, se pasa a la creación de los esquemas y estructuras de organización de la información a lo que *Josef Hassan* denomina en su propuesta de guía metodológica, diseño conceptual. En este sentido, afirma: “ la mejor forma de asegurar la adaptabilidad de un sitio Web es mediante la separación del contenido, presentación, estructura lógica e interacción”.⁴

Como puede apreciarse, 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 y en este sentido, las taxonomías tienen un lugar fundamental.

PRINCIPALES ELEMENTOS PARA CONSTRUIR UNA TAXONOMÍA

Construcción de la taxonomía

Según *Centelles*, la construcción de las taxonomías supone la realización de cuatro procesos:⁵

1. Delimitación de la realidad -entidad, área de conocimiento, sector industrial, etc.- que representará la taxonomía.
2. Extracción del conjunto de términos o categorías que representan dicha realidad.

Para realizar este proceso es necesario establecer, cuáles son las fuentes prioritarias y los mecanismos de extracción idóneos para cada una de ellas.

Existen tres tipos: las fuentes personales, integradas por usuarios del Web y especialistas en el dominio del Web; las fuentes documentales, integradas por documentos representativos de los tipos de contenidos identificados en la fase de planificación estratégica y las taxonomías e instrumentos disponibles para la representación del conocimiento – que van desde las nomenclaturas de las unidades y recursos existentes en una entidad a los cuadros de clasificación de la documentación administrativa.

El resultado de este proceso es un registro de términos o categorías representativas.

3. Control terminológico de los términos o categorías.

Se identifican los diferentes términos que designan un mismo concepto; en caso de que sean dos o más es necesario determinar cuál se considera preferente y cuáles se consideran no preferentes. En segundo lugar, es necesario dar una forma correcta y consistente a todos los términos de la taxonomía, con independencia de si son o no preferentes.

El resultado de este proceso es el establecimiento de una relación de equivalencia entre los términos de la taxonomía.

4. Establecimiento del esquema y la estructura de organización de los términos o categorías.

El esquema de organización incluye el criterio o criterios utilizados para dividir y agrupar las categorías.... los criterios son ilimitados y su idoneidad depende del objeto que se desea deba representar mediante la taxonomía.

Algunos ejemplos de los criterios más utilizados son: los temas, las materias y disciplinas, las personas, las entidades, los destinatarios, los procesos, tareas y funciones, los tipos de documentos, etcétera.

El modelo estructural define el tipo de relación que se establece entre las agrupaciones de categorías derivadas del esquema de organización.

Tradicionalmente, se han distinguido dos técnicas para el desarrollo de la estructura de la taxonomía: la técnica de arriba hacia abajo ("up to down") y la técnica de abajo hacia arriba ("down to up"):

- La aplicación de la técnica de arriba hacia abajo supone la identificación inicial de un número limitado de categorías superiores, y la agrupación del resto de categorías en niveles sucesivos de subordinación hasta alcanzar los niveles de categorías más específicas.
- La aplicación de la técnica de abajo hacia arriba se basa en la identificación inicial de las categorías más específicas, que se agrupan en niveles sucesivos de superordenación hasta alcanzar el nivel de categorías superiores.

A todo esto, pudiera agregarse que es fundamental, a la hora de construir una taxonomía, conocer toda la terminología que se utilizará, tanto en el caso de una taxonomía para un sitio Web como para una organización específica. Nadie mejor que una persona bien calificada, y con un sólido conocimiento del dominio para el cual trabaja para saber como construir una taxonomía.

CONSIDERACIONES FINALES

Saber organizar los contenidos electrónicos, analizarlos, procesarlos y representarlos de forma sintética es cada día más necesario en el entorno Web. En la medida que ellos se encuentren más organizados y mejor estructurados, los usuarios recuperarán y utilizarán con más éxito los contenidos disponibles, y esto consecuentemente redundará en una eficaz gestión del conocimiento.

Las taxonomías, cuando se pretende desarrollar un sistema de gestión de contenidos, tienen un lugar privilegiado, y por ello, es esencial conocer para qué sirven, cómo se utilizan, cómo funcionan, cómo construirlas, quiénes se ocupan de su creación y mantenimiento y, sobre todo, cómo validarlas.

La arquitectura de la información es una de las disciplinas que actualmente se encarga de construir taxonomías para la organización y representación de los contenidos en el Web y ella son una clase de herramienta fundamental para ello. El arquitecto de información responsable de cumplir con esta tarea debe siempre considerar que a mayor grado de descripción de las unidades de contenido, mayores serán las opciones para la navegación para el usuario.

REFERENCIAS BIBLIOGRÁFICAS

1. Abed Gregio AR, Barbato LGC, Duarte LO, Montes A, Hoepers C, Stedding-Jessen K. Taxonomias de Vulnerabilidades: Situação Atual. Disponible en: <http://www.ppgia.pucpr.br/~maziero/pesquisa/ceseg/sbseg05/artigos/12463.pdf> [Consultado: 3 de enero del 2007].
2. Rosenfeld L, Morville P. Information Architecture for the World Wide Web: Designing Large-Scale Web Sites. Sebastopol (Canada): O'Reilly; 2002.
3. Montes de Oca Sánchez de Bustamante A. Arquitectura de información y usabilidad: nociones básicas para los profesionales de la información. *Acimed* 2004;12(6). Disponible en: http://bvs.sld.cu/revistas/aci/vol12_6_04/aci04604.htm [Consultado: 7 de febrero del 2007].
4. Hassan Montero Y, Martín Fernández F. Diseño inclusivo: Marco metodológico para el desarrollo de sitios Web accesibles. *NSU* 2003;(2). Disponible en: <http://www.nosolousabilidad.com/articulos/disenoinclusivo.htm> [Consultado: 11 de febrero del 2007].
5. Centelles M. Taxonomías para la categorización y la organización de la información en sitios Web. *Hipertext.net* 2005;(3). Disponible en: <http://www.hipertext.net> [Consultado: 30 de enero del 2007].

Recibido: 30 de marzo del 2007. Aprobado: 15 de abril del 2007.

Lic. *Anisleiby Fernández Hernández*. Dirección de Información. Universidad de las Ciencias Informáticas. Carretera a San Antonio de los Baños Km2 ½. Reparto Lourdes. Boyeros. La Habana. CP19 370. Correo electrónico: ani@uci.cu

¹Licenciada en Bibliotecología y Ciencias de la Información. Dirección de Calidad de Software. Universidad de las Ciencias Informáticas. Cuba.

Ficha de procesamiento

Clasificación: Artículo original.

Términos sugeridos para la indización

Según DeCS¹

CIENCIAS DE LA INFORMACION; INTERNET; INTERFASE USUARIO-COMPUTADOR; SISTEMAS DE COMPUTACION.

INFORMATION SCIENCE; INTERNET; USER-COMPUTER INTERFACE;
COMPUTER SYSTEMS.

Según DeCI²

INFORMACION/organización y administración; NAVEGACION; SITIO
WEB/organización y administración; INTRANET; INTERFASE USUARIO-
COMPUTADORA/desarrollo.

INFORMATION/organization and administration; NAVIGATION; WEB
SITE/organization and administration; INTRANET; USER-COMPUTER
INTERFACE/development.

¹BIREME. Descriptores en Ciencias de la Salud (DeCS). Sao Paulo: BIREME, 2004.
Disponible en: <http://decs.bvs.br/E/homepagee.htm>

²Díaz del Campo S. Propuesta de términos para la indización en Ciencias de la Información.
Descriptores en Ciencias de la Información (DeCI). Disponible en: <http://cis.sld.cu/E/tesauro.pdf>