

## La Web 2.0 en los sistemas de recuperación de información

Mary Eugenia Luna González

Universidad Central de Venezuela – Venezuela

ANÁLISIS / ANALYSIS

### Resumen

Se realiza un análisis documental para obtener conocimiento sobre el avance de las nuevas tecnologías de la información y comunicación, y la explosión informativa y sobresaturación en la red consecuencia de tal avance, que solo es posible contrarrestar por medio de la utilización de sistemas de recuperación de información combinados con las ventajas ofrecidas en la Web 2.0.

### Palabras clave

*Web 2.0; Redes sociales; Sistemas de información; Recuperación de información*

### The Web 2.0 in the systems of recovery of information

### Abstract

A documentary analysis is realized to obtain knowledge on the advance of the new technologies of the information and communication, and the informative explosion and overcrowding in the network consequence of such a advance, which alone is possible to offset by means of the system utilization of recovery of information combined with the advantages offered in the Web 2.0.

### Keywords

*Web; Web 2.0; Social network ; Information systems; Information retrieval*

## Introducción

Con la aparición de las nuevas tecnologías de la información y comunicación y la facilidad para publicar y transmitir elementos informativos por medio de la web, se obtiene una saturación de contenidos y de conocimiento, que se encuentran en tal medio sin un proceso normalizado de organización.

La posibilidades que brinda la red electrónica en la actualidad, específicamente la Web 2.0, se ven reflejadas cuando los usuarios pueden consultar, descargar y subir información libremente, y en la oportunidad de interactuar con otros usuarios; trayendo como consecuencia que Internet se vea cargada de tanta información y haciendo difícil discernir que es relevante y que no según las necesidades informativas del usuario final del medio electrónico.

Ante tal situación surge la necesidad de contar con sistemas de recuperación de información electrónica que permitan realizar búsquedas mediante un sistema lógico, arrojando contenidos ajustados a la realidad y que satisfagan las necesidades de los usuarios.

## Objetivo

Conocer como se manifiesta el sistema de recuperación de información electrónica dentro de los portales electrónicos y cuales son las ventajas que para tal fin ofrece la Web 2.0.

## Metodología utilizada

Para abordar la temática se inicia con un análisis documental, examinando la literatura especializada en el área, proveniente de fuentes documentales tanto impresas como digitales; con la finalidad construir un marco conceptual relacionado con el tema de la Web 2.0 y las ventajas ofrecidas por la misma.

## La Web como sistema de gestión de información

La web es una invención mecanicista, creada con la finalidad de funcionar como sistema abierto de información, en donde el ingreso y actualización del conocimiento en los sistemas de representación permite la interacción y publicación documental de diversos individuos, tanto especialistas como aficionados a temas específicos, lo que implica una saturación informativa y sobrecarga del sistema digital.

La representación de los contenidos en la web no se hace de una manera homogénea, lo cual dificulta la búsqueda y posterior recuperación de la información contenida en la documentación.

Es por ello que surge la necesidad de procesar la información, clasificándola para ponerla al servicio de los usuarios; para tan fin se llevan a cabo procesos de indización por parte del personal documentalista que labora en las compañías de recuperación de información electrónica a través de Internet.

## Los buscadores de información electrónica

En los sistemas electrónicos de recuperación de información, son utilizados diversos procedimientos para llevar a cabo las búsquedas, uno de ello es la recuperación por medio de la búsqueda lógica, utilizándose para esto operadores booleanos (AND, OR y NOT), relacionando y combinando los términos de una manera lógica.

La relación AND (conjunción o intersección lógica) es utilizada para relacionar dos o más términos que se requiere que aparezcan a la vez en los documentos. Con la relación OR (suma lógica) se buscan documentos que contengan uno de los términos o todos ellos. La relación NOT (negación lógica) excluye los documentos e los que aparezca un término determinado (Lancaster y Pinto, 2001).

A pesar de las ventajas que ofrece la utilización de operadores booleanos en las búsquedas de conocimiento en los sistemas digitales de recuperación de información, se presenta una situación desfavorable cuando no todas las personas que recurren a dicho servicio se encuentran familiarizados con la correcta combinación de los mismos, lo cual hace que los sistemas recuperen información errónea o innecesaria.

Otra forma de recuperación de información se da cuando los usuarios no conocen las búsquedas lógicas por medio de la utilización de los operadores booleanos, tales usuarios recurren a la búsqueda por medio de la utilización del texto libre, usando palabras y frases en lenguaje natural que el usuario considera pertinentes de acuerdo a su necesidad; lo cual arroja la coincidencia de términos en el título del documento o en cualquier parte de este.

Los sistemas de recuperación de información no responden con tanta satisfacción las necesidades cada vez más específicas de los usuarios documentales, ya que los mismos no han superado las brechas presentadas en el uso de las palabras clave y de los descriptores para la recuperación de información, al no ofrecer respuesta a la totalidad de las necesidades se presentan una serie de problemas y dificultades a la hora de acceder al contenido de las páginas.

## Problemas que presentan los buscadores en la web

La principal limitación que presentan los buscadores, es que la mayoría trabajan por patrones de coincidencia textual, dejando de lado las relaciones terminológicas y el contenido documental.

Los buscadores electrónicos carecen de mecanismos de recuperación lógica; por ende las búsquedas informativas de los usuarios ocasionan al sistema enlaces de términos descontextualizados que sobresaturan y obstruyen la recuperación informativa, haciendo que los resultados encontrados carezcan de relación informativa con las temáticas buscadas, y en consecuencia resulta una mezcla de lo fundamental con lo irrelevante. Al respecto (García Gutiérrez, 2001: 5-6) indica *“por el momento, el hipertexto no va más allá de los viejos índices sistemáticos o permutados, no es mas que un índice transversal y caótico con efectos especiales”*

Las computadoras con la información obtenida de las páginas en la web, construyen índices con las palabras del título y algunas del contenido, las cuales han sido incluidas en los metadatos, y cuando el usuario al querer cubrir una necesidad informativa formula una pregunta al sistema, tal sistema se encarga de realizar una comparación de las palabras contenidas en el índice con las de la solicitud, arrojando como respuesta solo los documentos en donde exista una coincidencia terminológica; tales limitaciones hacen que las indagaciones en las bases de datos de las páginas web se vean frustradas en cuanto al contenido informativo obtenido.

Los metadatos o el etiquetado del sistema de indización, son invisibles para el usuario, los mismos son asignados por el creador o el profesional que procesa la información a cada unidad de contenido (título, subtítulo, autor, ciudad, entre otros), a las imágenes y a cualquier otro tipo de archivo. Tales metadatos constituyen un papel esencial en la representación del contenido de las páginas, para ser localizados mediante los sistemas de búsqueda de información electrónica.

Ante lo expuesto con anterioridad se hace necesario mencionar que el sistema de asignación de metadatos no se encuentran normalizados de una manera semántica o sintáctica, que permita que se utilicen las mismas terminologías, y de esta forma el creador del documento o el procesador de las páginas en Internet puedan describir el documento de manera uniforme, facilitando así el proceso de recuperación de información en la web.

También sucede que las personas que elaboran la documentación y los procesadores de las páginas se equivoquen y por ende olviden colocar los metadatos, lo hagan de manera incorrecta, o los coloquen en unas páginas y en otras no; trayendo como consecuencia que los motores de búsqueda no tomen en cuenta tales metadatos para generar los resultados de las indagaciones. (Codina, 2003).

De igual manera muchas veces desde la generación del recurso no existen los metadatos, sino que son agregados después; también se presenta la situación en la que cuando no todos los autores colocan los metadatos, la web los asigna automáticamente; y al no participar en el proceso de indización, las búsquedas no presentan los resultados deseados por el usuario.

Al conocer la problemática presentada por los buscadores electrónicos al momento de responder a los requerimientos específicos de los usuarios, es necesaria la utilización de un recurso que permita la intervención del usuario en los procesos de gestión de información, en este sentido se encuentra la Web 2.0, que ofrece la posibilidad de que los usuarios generen soluciones a las inquietudes del día a día.

## La Web 2.0

La Web 2.0 es un término asociado a la web social, el cual se basa en la interacción de usuarios lograda a partir de diferentes sitios en la red, lo cual facilita el compartimiento de información, la colaboración entre usuarios, y la intervención múltiple en los diseños de información.

Tal término fue utilizado por primera vez por Tim O'Reilly para referirse a una segunda generación de la web, basada en comunidades de usuarios en donde se ofrecían una serie de servicios como, los blogs, los wikis y las redes sociales, fomentando la colaboración y el intercambio de información.

En cuanto a dicho término se refiere, O'Reilly afirma que las aplicaciones Web 2.0:

“son aquellas que sacan partido a las ventajas intrínsecas de la web, ofreciendo un servicio continuamente actualizado que mejora cuanto más gente lo use, utilizando y remezclando los datos de múltiples recursos, incluyendo los usuarios individuales, a la vez que ofrecen sus propios datos y servicios de tal forma que pueden ser utilizados por otros, creando una arquitectura de participación en red, yendo más allá de la página de la web 1.0 para ofrecer experiencias de usuario cada vez más ricas”. (O'Reilly, 2005 y Serrano, 2006, c.p. Margaix Arnal, 2007: 96).

La Web 2.0 permite la autogestión del usuario, la interacción entre personas por medio del aprovechamiento de la inteligencia colectiva contenida en las redes sociales, las wikis, los foros y los blogs; los cuales logran que el usuario comunique sus inquietudes y opiniones, aportando así soluciones en la construcción de una web y sistemas de información más humanos y acordes a la realidad.

Aprovechar las ventajas de la Web 2.0 mejora de manera continua el servicio, ya que mientras los usuarios puedan interactuar con el servicio mostrando cuáles son sus necesidades y tomando en cuenta sus requerimientos, el servicio de consulta informativa y documental puede adaptarse a las mismas; tal proceso es conocido como inteligencia colectiva, en donde el usuario es considerado como fuente de información, ya que el mismo por medio de las interacciones hace posible que se comparta e intercambie el conocimiento contenido en la información.

El software social propone elementos que pueden ser aprovechados al permitir a los usuarios comunicarse unos con otros por medio de la utilización de herramientas de la Web 2.0 como, los blogs, que contienen información con la capacidad de ser ordenada cronológicamente, permitiendo al usuario hacer búsquedas por periodos de tiempo y hacer comentarios de las mismas; de igual manera se encuentran los foros, que permiten a los usuarios crear un tema de discusión en el cual puedan participar el resto de los usuarios, como una conversación; en la misma temática se encuentran las wikis, que permite a los usuarios leer, escribir y modificar mensajes que hayan sido escritos por otros usuarios.

## **Ventajas de utilizar la Web 2.0 para el diseño de sistemas de recuperación de información electrónica**

La interacción usuario sistema en las búsquedas informativas es más productiva cuando el usuario es capaz de plantear sus necesidades informativas, y las mismas al ser tomadas en cuenta por el personal que procesa la información son adaptadas al lenguaje de representación utilizado en el procedimiento de indexación.

Los servicios a través de la web participativa se encuentran orientados a la colaboración de usuarios, y no al uso informativo; en donde los esfuerzos deben estar orientados al diseño de los sistemas a partir de la colaboración y de las opiniones, para crear servicios que sean útiles y respondan a las necesidades de los usuarios, generando a su vez un nuevo conocimiento.

Al respecto O'Reilly (2005) y Serrano (2006), indican que *“la confianza radical y el aprovechamiento de la inteligencia colectiva no son cambios tecnológicos, son cambios de actitud y, junto con otros elementos han puesto la base de lo que se llama la actitud 2.0”* (O'Reilly, 2005 y Serrano, 2006, c.p. Margaix Arnal, 2007)

El aprovechamiento de la inteligencia colectiva por medio de la utilización de las herramientas ofrecidas por la web social, hace que los usuarios en un principio pasivos se conviertan en actores participativos en los procesos de elaboración y gestión de los contenidos informativos de los sistemas de representación de la información.

## Conclusiones

1. La web brinda la oportunidad de la incorporación de metadatos, pero los mismos deben ser generados luego de que es creado el documento, lo cual puede traer como consecuencia que dichos metadatos no sean incluidos en la descripción, o sean incluidos pero de forma manipulada o errónea.
2. La coincidencia de conceptos, mas no de contenidos en el momento de la búsqueda en los sistemas de recuperación electrónica, arroja como resultado información irrelevante, obviando contenidos pertinentes en la investigación o la búsqueda simple; entendiendo que la relevancia y pertinencia vienen dadas por el usuario de acuerdo a su necesidad informativa.
3. Es necesario que los metadatos sean creados a la par de la generación del documento, imposibilitando de esa manera las manipulaciones para obtener los primeros lugares en la búsqueda, y haciendo posible una recuperación exhaustiva de información.
4. Con la intervención del usuario a través de la Web 2.0, se obtienen opiniones acertadas de las necesidades informativas, las cuales pueden ser potenciadas creando lenguajes documentales que permitan humanizar los sistemas y ajustar las búsquedas informativas a las necesidades de la sociedad.
5. El personal que procesa los contenidos documentales en la web no hace distinción entre los usuarios a los que prestan sus servicios por no conocerlos potencialmente; y el usuario al tener libre acceso al contenido de dichas bases de datos, consigue información que no ha sido discriminada según sus necesidades.
6. es necesario que los buscadores electrónicos utilicen un lenguaje codificado, en donde puedan diseñarse etiquetas estructuradas y explícitas, en cuanto a la totalidad de los elementos abarcados, ofreciendo metadatos para la posterior búsqueda en áreas como el título, subtítulo, autor, ciudad, organización y contenido de la publicación.
7. A partir del aprovechamiento de la inteligencia colectiva, los usuarios expresan sus aportes a la organización del conocimiento y manifiestan sus necesidades por medio de las redes sociales, las wikis, los blogs, los foros, entre otros; haciendo posible su adaptación a través de los aportes brindados en las sugerencias expuestas en los medios ofrecidos por la Web 2.0.

## Bibliografía

- CODINA, Luís. (2003). La web semántica: una visión crítica. En: El profesional de la información, vol. 12, nº 2, pp. 149-152.
- GARCIA GUTIERREZ, A. (2001). Redes digitales y exomemoria. Conferencia dictada en el Congreso Ibérico «La sociedad de la Comunicación», Málaga.
- LANCASTER, F.; PINTO M. (2001). Procesamiento de la información científica. Madrid: ArcoLibros.
- MARGAIX ARNAL, Dídac. (2007). Conceptos de web 2.0 y biblioteca 2.0, definiciones y retos para las bibliotecas actuales. En: El profesional de la información, vol. 16, nº 2, pp. 95-106.
- Petróleos de Venezuela. PDVSA. (en línea). Disponible en: <http://www.pdvsa.com/>
- RAMÍREZ CESPEDES, Zulia. (2006). Propuesta de pautas para la construcción de una ontología sobre ciencia de la información. Tesis de maestría. Facultad de Comunicación. Universidad de la Habana.

## Datos del autor

### Mary Eugenia Luna González

Licenciada en Archivología, egresada de la Universidad Central de Venezuela. Archivista en la Oficina de Auditoría Interna del Ministerio del Poder Popular de Minería y Petróleo de Venezuela.

[malunag2000@yahoo.com](mailto:malunag2000@yahoo.com)

**Recibido - Received:** 2011-08-09

**Aceptado - Accepted:** 2011-09-30



*This work is licensed under a [Creative Commons Attribution-NonCommercial-No Derivative Works 3.0 United States License](https://creativecommons.org/licenses/by-nc-nd/3.0/).*



*This journal is published by the [University Library System](https://www.library.pitt.edu/) of the [University of Pittsburgh](https://www.pitt.edu/) as part of its [D-Scribe Digital Publishing Program](https://www.library.pitt.edu/dscribe/) and is cosponsored by the [University of Pittsburgh Press](https://www.press.pitt.edu/).*