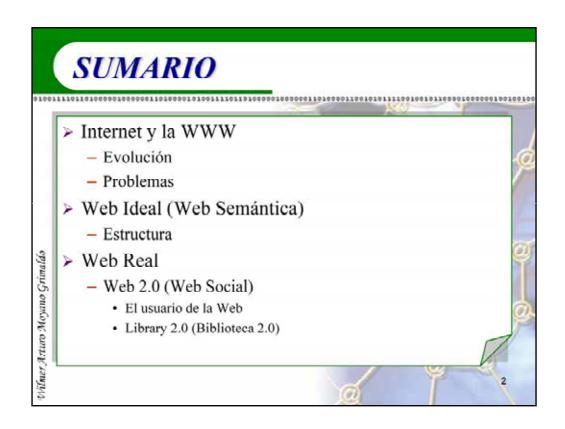
# Web Semántica y Web Social como tendencias profesionales

### Wilmer Arturo Moyano Grimaldo

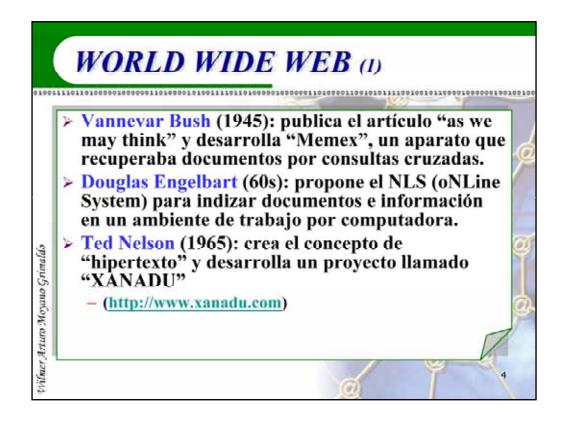
Profesional en Ciencia de la Información – Bibliotecólogo Salamanca, 29 de marzo de 2007



```
INTERNET: hechos históricos

1545-1555 – Bibliotheca universalis (Conrad Gesner)
1969 – Arpanet
1970 – 1979: Proyecto Gutenberg, correo electrónico, proyecto Internetting, protocolo TCP/IP... ¿Microsoft? [USMARC]
1980 – 1989: Internet, protocolo FTP, SGML, WWW...
1990 – 1999: Consorcio de la Web (W3C), XML, RDF, , Dublin Core, Internet 2, NAPSTER...
2000... OAI, Archivos abiertos, Web Semántica, GOOGLE, [Web 2.0 (weblogs, rss, flickr, wikis, You Tube...)]
```

- •Bibliotheca Universalis: Proyecto de Conrad Gesner, padre de la bibliografía, para organizar todo el conocimiento existente en su tiempo y que hoy en día está tratando de retomar el G7 con ayuda de la Internet
- •Proyecto Gutenberg, ideado por Michael Hart para permitir el acceso a literatura gratuita, hoy en día cuenta con mas de 8 mil títulos.
- •Proyecto Internetting: conexión de diferentes redes entre si
- •TCP/IP:
  - •IP: Regula en tráfico entre los nodos de origen y el destino
  - •TCP: asegura que toda la información que sale del origen llegue a su destino



#### Vannevar Bush:

- •Articulo: "As We May Think" expresa la preocupación por el incremento en la cantidad de información que existía y los pocos e ineficientes sistemas para encontrarla, además de ser un artículo un poco visionario sobre lo que acontecería en varios aspectos a futuro con la información, lo cual justifica en cierta forma nuestra labor y el por qué nuestra profesión es tan importante hoy en día.
- •desarrolla un aparato llamado MEMEX. Este aparato permitiría a cada individuo almacenar su información en microfilmes, consultarlos rápidamente y, lo que es más importante, crear vínculos entre unos documentos y otros, de modo que durante la lectura de un documento se recordara al lector qué documentos contenían información relacionada

DOUGLAS ENGELBART: padre del Mouse, crea un sistema que permitía acceder a documentos en un ambiente de computadora, de una manera mas eficiente para los usuarios.

#### TED NELSON:

•Hipertexto: es la representación de la información como nodos de redes unidos entre si de tal forma que permiten una navegación no lineal entre los nódulos.

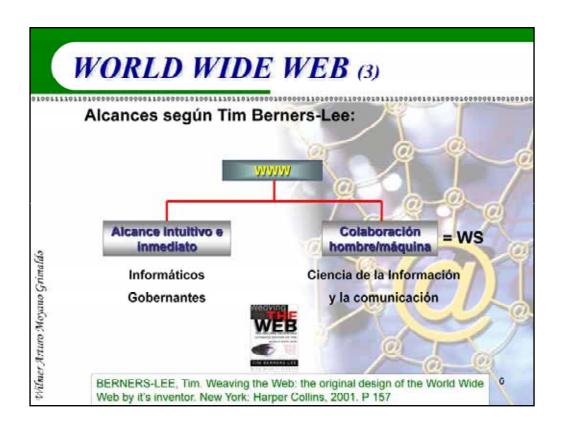


#### TIM BERNERS-LEE:

Padre de la WWW, basado en las investigaciones anteriores, pero especialmente basado en el hipertexto.

La WWW, en realidad era el primer navegador o Browser de la Historia.

En 1993 ya se había masificado el uso de la Web, con 50 servidores, a 1997 ya existían 650 mil.





Hoy en día como es sabido, las consultas de información en internet se ven truncadas por mucho ruido lo que genera que las respuestas no sean lo que se espera.

# WWW: RETOS EN LA WEB

- Datos distribuidos sobre diferentes plataformas
- > Datos volátiles
- > Gran volumen de datos
- > Datos redundantes y no estructurados
- > Calidad de los datos
- Datos heterogéneos

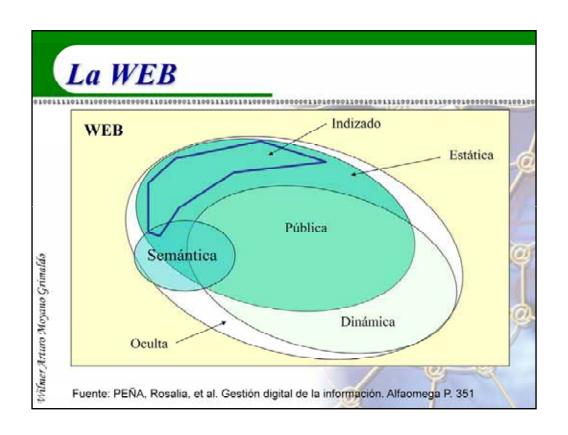
Wilner Arturo Moyano Grimaldo

### Factores críticos de éxito en la WWW

#### Según Berners-Lee

- Links ilimitados desde cualquier lugar de la Web a cualquier otro
- Normas técnicas abiertas como base para el crecimiento continuo de aplicaciones innovadoras
- Separación de capas de red, facilitando innovaciones independientes para el transporte, vías y aplicaciones de información en la red

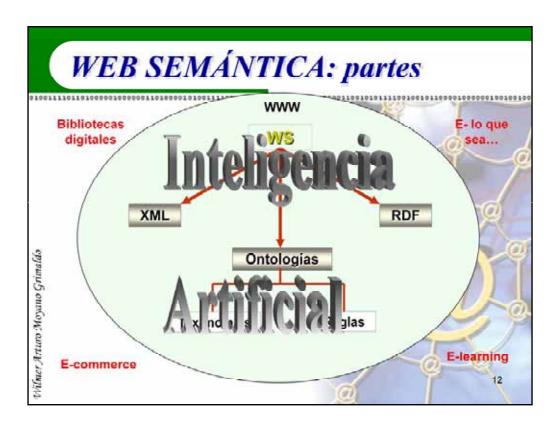
Wilner Arturo Moyano Grimaldo



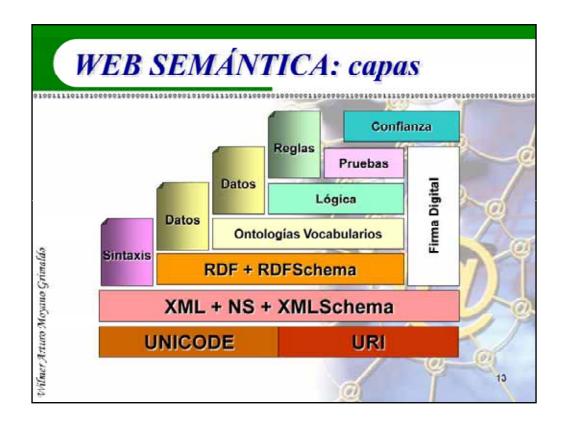
# WEB SEMÁNTICA: ¿qué es?

- Es una extensión de la WWW donde la información estará bien definida para facilitar la cooperación entre computadores y hombres.
  - Semántica = Significativo
- "Es una telaraña de conexiones entre diferentes tipos de datos que le permiten a las máquinas hacer algo que no podrían hacerlo directamente"

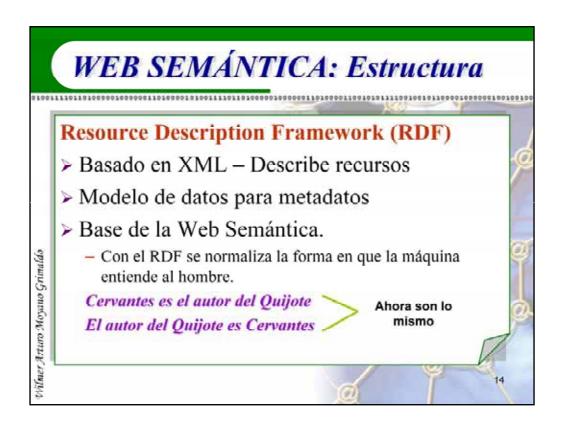
Wilmer Arturo Moyano Grimaldo



El reto de la WS es proveer un lenguaje que exprese tanto los datos como las reglas para el razonamiento de los datos y que facilite las reglas para la existencia de un sistema de representación del conocimiento que será exportado a la Web.



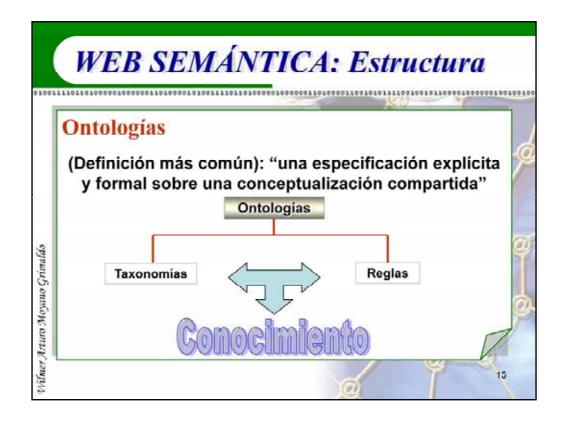
- a) XML (*Extensible Markup Language*). Es el metalenguaje que actualmente se usa para intercambiar datos en la red.
- b) RDF (*Resource Description Framework*). Es un lenguaje que describe metadatos y fuentes de información (documentos, imágenes, etc.).
- c) Ontologías. Son conceptualizaciones que definen el significado de un conjunto de conceptos para un determinado dominio. Las ontologías se expresan en lenguajes formales como OWL.
- **d) Lógica.** El razonamiento lógico permite determinar si los datos son correctos e inferir conclusiones a partir de ellos.
- e) Prueba. Las pruebas explican y verifican los pasos de los razonamientos lógicos.
- **f) Confianza.** Se refiere a técnicas que aseguran la identidad y fiabilidad de los datos y servicios.



El RDF es la base de la Web Semántica, porque permite que las máquinas y los seres humanos hablen en el mismo sentido gramatical.

Antes del RDF, no era lo mismo para una maquina la expresión Cervantes es el autor del Quijote, a el autor del Quijote es Cervantes. Ahora si, porque RDF permite que por su modelo de entidades y relaciones se la palabra Cervantes siempre sea el sujeto que actúa en la oración sin importar en que parte de la misma se encuentre. Esto permite que las máquinas nos puedan interpretar hoy en día.

Los dos ejemplos mas conocidos del uso de las RDF son las RSS 2.0 (para sindicación en Blogs) y el FOAF para redes sociales como E-Conozco o Linked in donde se usa para la descripción de perfiles.



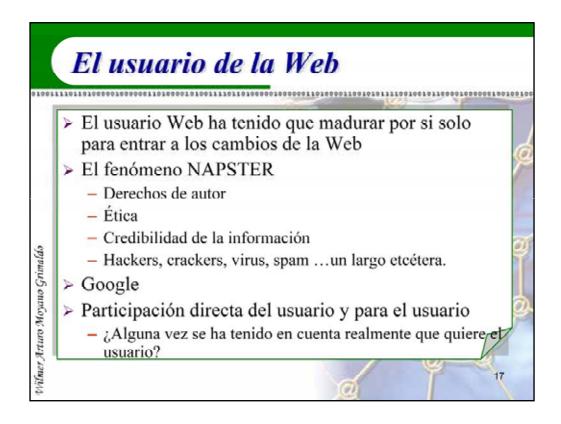
Las ontologías proporcionan la vía para representar el conocimiento y son las que explican como conectar un dato con otro.

Una ontología define un vocabulario común para investigadores que necesitan compartir información en un dominio. Ella contiene definiciones de conceptos básicos y sus relaciones que pueden ser interpretadas por una maquina.

Una ontología palabras mas palabras menos sirve para recuperar información marcándole el camino a seguir a los buscadores, de una forma casi como la que hizo Pulgarcito cuando ponía trocitos de pan para recordar el camino a casa.



Si todo esto se logra desarrollar, lo que se espera es que cuando el usuario haga una consulta, los resultados sean relevantes para lo que necesita.

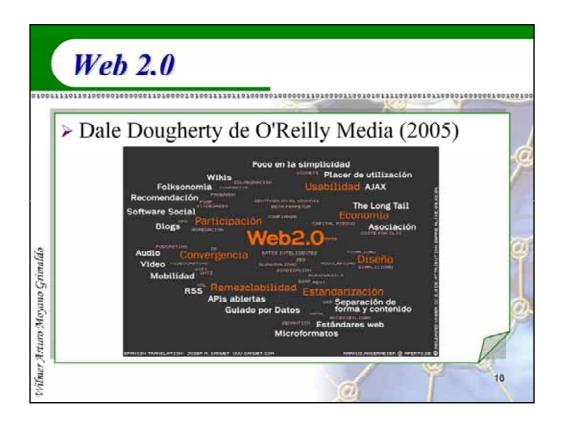


Esta es la esencia de la presentación, donde se podría indicar que aun estamos en una Web adolescente y que debe madurar para conseguir ser una verdadera Web con compromiso social.

Es una Web donde el usuario ha tenido que adaptarse casi a la fuerza a los cambios que suceden, como sucede con la mayoría de las tecnologías. Sin embargo la actual Web es una internet social donde al usuario le ha tocado abrirse camino para generarse verdaderos espacios, uno de estos es lo sucedido con Napster.

Por otra parte el futuro de los negocios ya no está en las tecnologías de la información sino en la gestión que se tenga de estas, por eso empresas como Google han conseguido posicionarse en el mercado y empresas como Microsoft, empiezan a competir mas en lo que tiene que ver con este tema que en Hardware.

Todo esto es el comienzo de lo que se ha denominado Web 2.0, pero que prefiero llamar Web Social, porque lo explica mejor, el darle versiones a la Web lo que hace es indicar que todo es un proyecto tecnológico y en realidad no es asi.



La Web 2.0 es un nombre que surge de la idea de Dale Dougherty quien en una conferencia en el año 2005 decidió llamarla asi para mostrar que la Web como se conoce ha evolucionado en su plataforma desde su creación y que se ha convertido en una Web mucho mas participativa.

En cierta forma si se observa, contiene muchos de los postulados de lo que es la idea de una Web Semántica, pero se diferencia en que la WS es mas hacia la organización de información con significado y la Web 2.0 esta mas orientada hacia las plataformas y el accesibilidad por parte del usuario.

Nuestra idea aca no es usar los dos conceptos por aparte sino complementarlos, algo que podría llamarse Web Semantica 2.0, aunque ya muchos empiezan a hablar de una Web 3.0 cuando estos dos convergen.

Es una evolución natural de las tecnologías que se manejan en la Web, pero no es un cambio novedoso de lo que ya existía. Lo mas importante aca es que se empieza a hacer consciente que el usuario es parte del desarrollo de la misma. Sin embargo, el punto en nuestra profesión es que tenemos que trabajar mas en esa interacción y adaptación del usuario, además de velar por los contenidos y su calidad, mas no ser los responsables de éstos. Mi propuesta es hacia la creación de bibliotecas digitales semánticas, con la aplicación de estas tecnologías. Desde mediados de la década pasada se hablaba ya de el usuario como parte de las bibliotecas digitales y que eso hacía que estos proyectos no fueran bases de datos en texto completo, por lo tanto una vez mas se está mostrando la necesidad de hacer esto una realidad.



Los dos ejemplos mas representativos de lo que ha sido la Web 2.0, lo demuestran los Blogs y las Wikis.

En el caso de los primeros es muy interesante, porque muestran claramente que la Web es accesible para cualquier tipo de usuario con conocimientos mínimos de tecnologías Web (html, css, etc...), algo que en cierta forma se podría comparar con lo que en medicina se llama "medicina basada en la evidencia" o incluso es un ejemplo muy básico de lo que se ha denominado "Gestión del Conocimiento", donde yo como persona natural, usuario común, tengo también cosas que aportar al mundo y en verdad lo hago y soy escuchado.

Este es un Blog sobre diseño creado por una diseñadora industrial, ella lo está desarrollando tanto por satisfacción personal y profesional como por dar su punto de vista, si lo hubiera querido hacer de esta forma hace unos 10 años, quizás lo que se vería sería "La Pagina Personal de Carolina Ayerbe" y tendría libros de visitas y habría tenido que usar obligatoriamente FrontPage o Dreamweaver o cualquier editor de html del tipo WYSIWYG, y si hubiera sido unos 5 años antes quizás habría tenido que conocer html y desarrollarla con un editor de html.

Con lo que tiene hoy en día este Web llega a muchas mas personas, obtiene la posibilidad de interactuar mas rápidamente con los lectores y a la par generar discusiones mas dinámicas en un diseño muy sencillo, algo que no existe en la mayoría de las bibliotecas digitales. Por otra parte este blog contiene tecnologías para sindicación de contenidos, en este caso particular, usa ATOM. Con lo cual yo como usuario puedo enterarme del contenido y las actualizaciones de la misma sin necesidad de estar mirando la URL todo el tiempo.

Ahora la tendencia se ha movido hacia generar Blogs mas semánticos, pero estos ya son una aproximación porque contienen organizaciones de contenidos temáticas aunque no de la forma como se vería en una ontología o un tesauro.



El otro componente son los llamados Wikis (en hawaiano quiere decir: rápido) que son unas herramientas muy útiles para poder compartir conocimiento, están basados en el motor llamado MediaWiki. El caso mas conocido y funcional es el de la famosa Wikipedia, la cual contiene información de todos los temas compuesta por personas comunes lo que la hace una fuente muy rica de conocimiento. Su principal inconveniente sin embargo es el hecho de que por mas que se asegura que sus contenidos son confiables, igual el grado de confiabilidad es variable.

Asi como los Blogs son herramientas propias de un usuario que quiere darse a conocer, la importancia de las Wikis radica en que la interacción del usuario es muy directa en la creación de conocimiento, y es una herramienta que ya varios han mirado como importante para sus contenidos, como lo es la página del software Dspace.

Aquí entraría a hacer parte uno de los problemas mas grande que tiene la Web actual, diferente a la de la cantidad de información que pueda contener, y es la veracidad en los contenidos, y en cierta forma lo que se podría denominar la "honestidad intelectual"

En este caso por ejemplo, quise mostrar que incluso en la Wikipedia podemos conseguir información muy completa sobre la Universidad de Salamanca. Herramientas de este tipo pueden ser realmente valiosas, pero no se pueden manejar sin una supervisión.

De esta herramienta también surge la idea de unificarla a los proyectos de Web Semantica, en lo que se ha querido denominar: Semantic MediaWiki, con la cual se busca que los datos que los usuarios ingresen a la Wiki sean mas estructurados y organizados de acuerdo a los preceptos de la WS. El proyecto Wikipedia está actualmente basado en este SMW y su estructura de organización está basada en RDF.

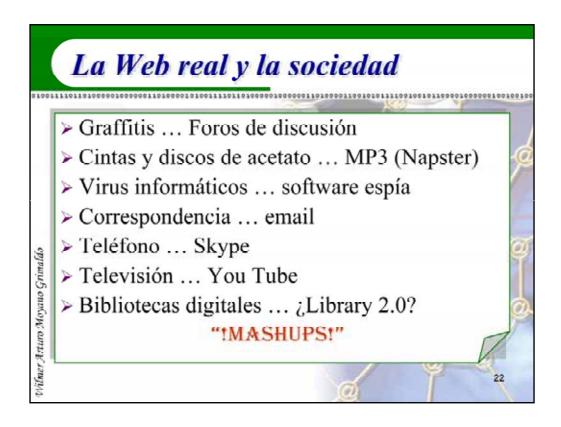


Y no se puede dejar por fuera el tema de Google, porque es el mejor ejemplo de cómo aprovechar la Web 2.0 y posicionarse en el mercado de Internet y conseguir dinero.

Muchos autores incluso hablan de la Googleización de los catálogos, que sería algo que a nosotros como documentalistas nos interesa del fenómeno Google directamente. De Google se pueden aprender muchas cosas, una quizás es a ser ambiciosos en lo que se quiere ofrecer, antes de Gmail a nadie se le ocurría que pudiera existir un correo tan grande, ahora abundan hasta correos que dan 45 gb de espacio.

Igualmente algo que nos debe importar mucho es la convergencia de tecnologías y de diseño que maneja Google, que permiten que el usuario desde su propia cuenta de correo pueda manejar un calendario, un procesador de texto, fotos, entre muchas mas aplicaciones, lo que para mi es fantastico ya que lo tengo todo en un solo lugar e incluso si estoy en la Antartida y me puedo conectar a Internet, sin necesidad de tener mucho software instalado puedo por ejemplo escribir un documento.

El ejemplo que tenemos en pantalla es para ver en lo que vamos, desde cuando a principios de los ochenta se usaban procesadores de texto como Wordstar a lo que ahora nos ofrece Google.

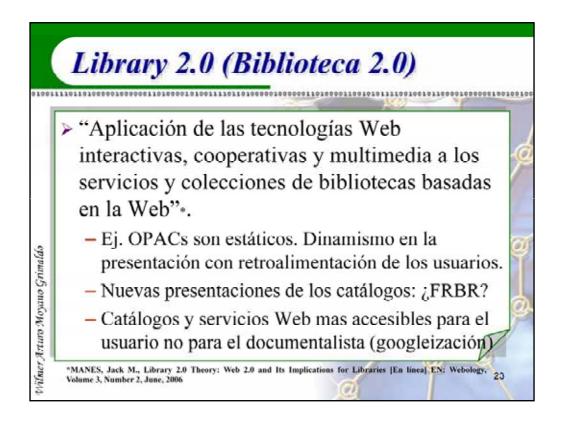


En mi opinión, yo considero que la evolución de la Web llevará a que se tenga incluso que reevaluar el término "usuario Web" a quizás uno que sea "habitante Web", cuando Berners-Lee propuso su idea de que la WWW fuera una Web inmediata y casi calcada de la vida real, no estaba muy alejado de esto.

Hoy en día la Web es realmente el reflejo de lo que sucede en la sociedad "no Web", ya las personas no tienen que salir a pintar graffitis en las calles, expresan su opinión en los foros de discusión, ya no se usan los disco antiguos, ahora se usan los MP3, y no necesariamente tengo que comprarlos, se han aumentado otro tipo de delitos, ahora se habla de delitos electrónicos.

La vida está girando hacia la virtualidad, donde incluso ya no necesito prácticamente salir de mi casa para conocer a una persona, la puedo conocer por intermedio de servicios como Hi5 o e-conozco. Se habla incluso de la adicción a Internet, o se empieza a llamar la atención por cosas tan sencillas como que hoy en día los niños están empezando a perder la habilidad de la escritura manuscrita.

En nuestro caso es obvio que nos ha llegado lo mismo, pero no se debe ver como una amenaza sino como la posibilidad de extender lo que ya hemos venido haciendo, lo importante aca es que tenemos que salirnos de la idea de una profesión puramente técnica, no somos catalogadores únicamente, el fuerte del bibliotecario está en las ciencias humanas y en el servicio, y es lo que mas hemos descuidado en la Web. En el estudio realizado por Hong (Iris) Xie, sobre como los usuarios evaluarían las bibliotecas digitales, uno es precisamente este, que abandonamos al usuario y se siente perdido y abrumado.



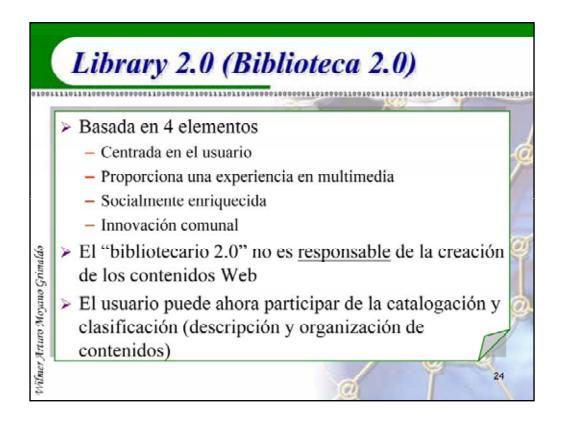
Ahora ha surgido a la par de la Web 2.0 lo que se ha llamado Library 2.0, algunos han considerado que es la evolución también de las bibliotecas digitales, pero el término solo indica la aplicación de todo esto que se ha visto en un ambiente virtual.

Realmente es una etiqueta que yo no comparto, porque es lo que siempre se ha buscado con los proyectos de biblioteca digital y no se ha podido, a finales de la década pasada en Colombia desarrollamos el primer proyecto funcional de una biblioteca digital en Latinoamérica, y fue un proyecto enriquecedor para todos los que participamos en él, ya que nos abrió la mente a la posibilidad de ver que había mas trabajo en la profesión de lo que se creía. Y en los informes que presentábamos hacia el año 2001 en que finalizó el proyecto, ya indicaban cosas de esto que se está hablando ahora.

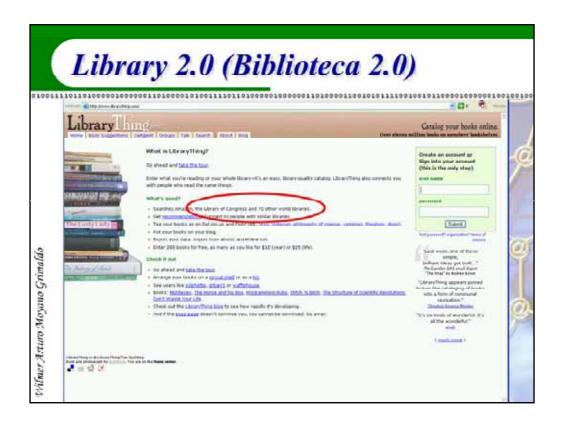
Yo creo que nuestra labor aca es mas de tipo realmente documental y social que tecnológica y técnica, pero nuestra visión sigue llevándonos a que lo importante es la creación de servicios Web 2.0, y nos sentimos a la moda porque hablamos de esto, como es el caso de las listas de discusión de biblioteconomía como lwetel. La experiencia nos va enseñando que nos enfrascamos en discusiones innecesarias y muy pocos se atreven a pasar a la práctica.

Aun en muchos sitios se discute si es mejor crear una biblioteca digital, una electrónica, o una virtual, y mas recientemente una hibrida y ahora quizás en si es mejor crear una biblioteca 2.0, y mientras estamos discutiendo esto, los informáticos no se han quedado esperando y estan desarrollando nuestro trabajo.

En el año 2004 tuve la fortuna de poder conocer al Dr. Edward Fox en persona, quien es el director de la NDLTD en el Virginia Tech Institute, uno de los gestores y participantes de aquel ya famoso Dlib-1, y el nos decía algo que realmente es muy cierto: "no importa como se llamen estas cosas, simplemente hay que hacerlas, evaluarlas y seguir adelante" En nuestro caso, sin olvidar que nuestro fin ultimo es la satisfacción de las necesidades de información de un usuario.



- •Centrada en el usuario: participación de los usuarios en la creación de la biblioteca y los servicios
- •Proporciona una experiencia en multimedia: porque el usuario se beneficiará especialmente de contenidos en audio y video
- •Socialmente enriquecida: La presencia de la biblioteca en la Web incluye obligatoriamente la presencia activa del usuario y su comunicación permanente con el bibliotecario (bibliotecólogo)
- •Innovación comunal: fundamentado en el principio de las bibliotecas de servicios a la comunidad, pero con la tendencia a los cambios de la comunidad; es decir, la biblioteca debe cambiar hacia los usuarios pero la actitud de los usuarios también debe cambiar hacia la biblioteca.



Estos son dos ejemplos de lo que se ha comentado, por un lado está Del.icio.us un proyecto de organización de imágenes donde los usuarios empiezan a crear sus propias etiquetas de descripción de sus recursos, y los comparten entre si, de tal forma que si yo hablo de mi viaje a Madrid y otro usuario hace lo mismo, nuestras imágenes quedarán agrupadas y normalizadas de acuerdo a esto. Esto es a lo que se le ha denominado Folksonomías y que es algo para mirar, porque si bien es algo que puede servir también es algo que puede generar problemas, además porque tendencias como Dublin Core o Marc iniciaron en cierta forma con esta misma visión de que el usuario lo pudiera manejar por si mismo y finalmente no se pudo.

Library Thing es otro ejemplo muy particular de cómo lo que nosotros, en general, no hacemos como bibliotecarios o documentalistas, otros lo van a hacer. Es un proyecto que busca que los usuarios puedan catalogar sus materiales por si mismos, e intercambiar registros, es una especie de consorcio que cada vez toma mas fuerza, ya que los participantes tienen acceso directo a otros participantes y pueden generar incluso intercambio de información sobre obras o temáticas de su preferencia.

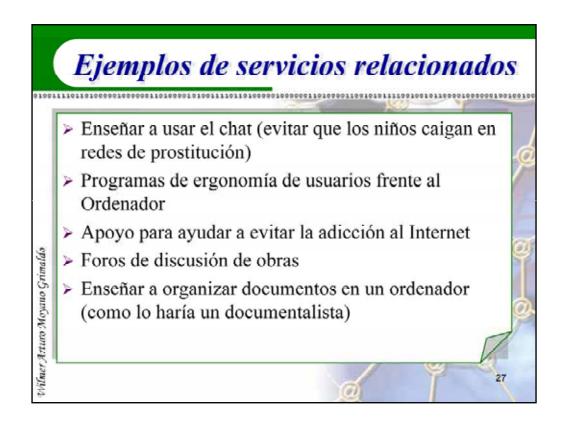
Algo que es muy interesante en este proyecto es que el usuario puede acceder incluso directamente a los catálogos y la catalogación de la Biblioteca del Congreso a través del protocolo Z39-50 y descargar sus registros para su propia catalogación, o incluso organizar sus colecciones con el uso del Sistema de Clasificación Dewey. No se que tan legal sea esto, pero es realmente funcional y crece porque es pensado en exactamente lo que los usuarios quieren y esperan.

# Web: desde nuestra profesión

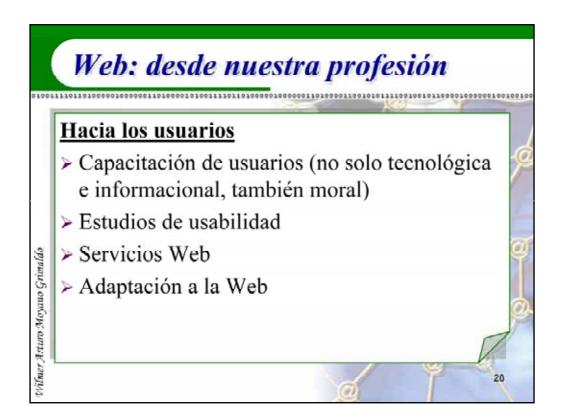
### Hacia la concepción de la Web

- Creación de servicios Web (Herramientas Web pero no solo éstas) \*
- > Estudios de Accesibilidad (en el diseño)
- > Conformación de objetos digitales
- > Diseño de bibliotecas digitales
- Creación y gestión de contenidos en Bibliotecas digitales, museos digitales, archivos digitales
- "Web of Trust"

Wilner Arturo Moyano Grimaldo



Estos son unos ejemplos de servicios que yo llamaría como Servicios Paraweb, es decir, servicios paralelos a la Web y que si bien no son exclusivos de nuestro quehacer, si tienen que ver con todo lo que se ha construido de la Web, y consecuencias de lo mismo. En ultimo lugar, una biblioteca debería poder ofrecer apoyo en este tipo de cosas.



# WS: desde nuestra profesión

### Organización y recuperación del conocimiento

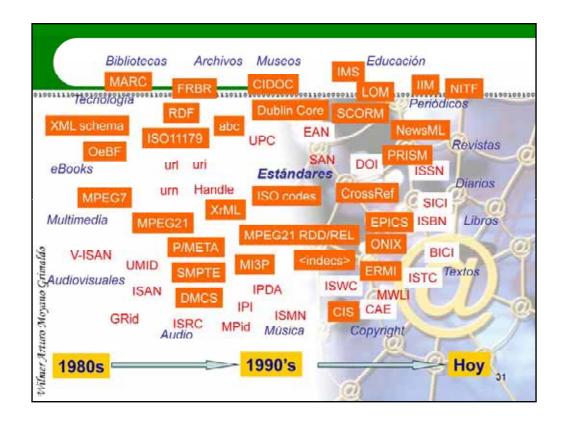
- Ontologías:
  - Nuevas y evaluación de las existentes
  - Adaptaciones (mapping):
    - · Sistemas de Clasificación + tesauros
    - FRBR + RDF + OWL
- Unidades de información Digitales "Bibliotecas 2.0"
  - Orientadores de usuarios

# ¿Cuál es el reto para la Web?

- > Intelectual findability Honestidad Intelectual
  - Ética de la información (Digital)

Wilmer Arturo Moyano Grimaldo

- Ya se está buscando una Web 3.0 sin haber consolidado una Web 2.0 (Web Semántica 2.0)
- Dejar de perder el tiempo en usar términos raros y mas bien trabajar.



La evolución de los estándares tiende a la información digital, pero la mayoría de los estándares no son de carácter documental, ni descriptivo, la pregunta es: ¿Qué papel tiene nuestra profesión en la actual sociedad de la información digital y la globalización? ¿qué estamos haciendo para ayudar en esta evolución? ¿O será mas bien que estamos siendo arrastrados hacia una evolución forzada?



Para termina, mi llamado es aquí a ser un poco como Nicolás Copernico, quien a raíz de sus observaciones y mediciones se dio cuenta que la tierra giraba alrededor del sol y no al contrario, aunque la tendencia en su tiempo era a que todos consideraran como verdad lo primero. De la misma forma, tenemos que salirnos de la idea que nuestro trabajo es solo en torno a la información y a la documentación, ya que ahora todo gira en torno al usuario.

Ya no somos solo guardianes de conocimiento (el universo hace tiempo que dejó de girar en torno al bibliotecario) ahora somos gestores de conocimiento, con todo lo que esto implica.



Un ejemplo de un proyecto colaborativo entre bibliotecas y museos

# **CONCLUSIONES**

- > La Web Semántica no es una nueva Web es la potenciación de la ya existente hacia una forma mas semántica de la misma.
- > Es también un proyecto político que implica apoyo de los gobiernos.
- ➤ La Web 2.0 y la Biblioteca 2.0 son solo unos nombres mas, lo que hay es que hacerlos prácticos.

Wilmer Arturo Moyano Grimaldo

# **CONCLUSIONES**

Wilmer Arturo Moyano Grimaldo

- Es un campo de mucho trabajo e investigación por parte de nosotros, pero si no lo hacemos otros lo harán (...ya lo están haciendo)
- No es un cambio en la forma de ejercer la profesión, es una evolución de lo que se ha venido haciendo desde siempre.
- El cliente (usuario) tiene la razón, pero hay que orientarlo y no dejarlo actuar solo y sobre todo ESCUCHARLO.

