



La Biblioteca digital

Curso "SIDRA: La Biblioteca Digital"

SIDRA. Sistema de Información Documental en Red de Asturias

Julio Alonso Arévalo
Universidad de Salamanca
alar@usal.es





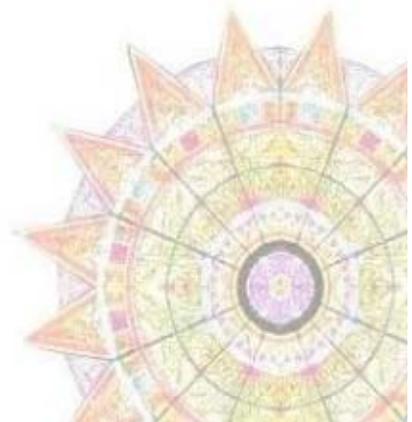
"¿Como va la informatización de la biblioteca y la hemeroteca?.

-Fenómeno...acá es todo digital.

-- No diga.

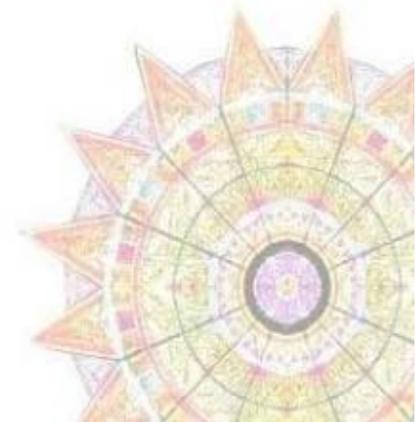
-Sí... uno se moja el dedito y va pasando la página."

(Daniel Paz)

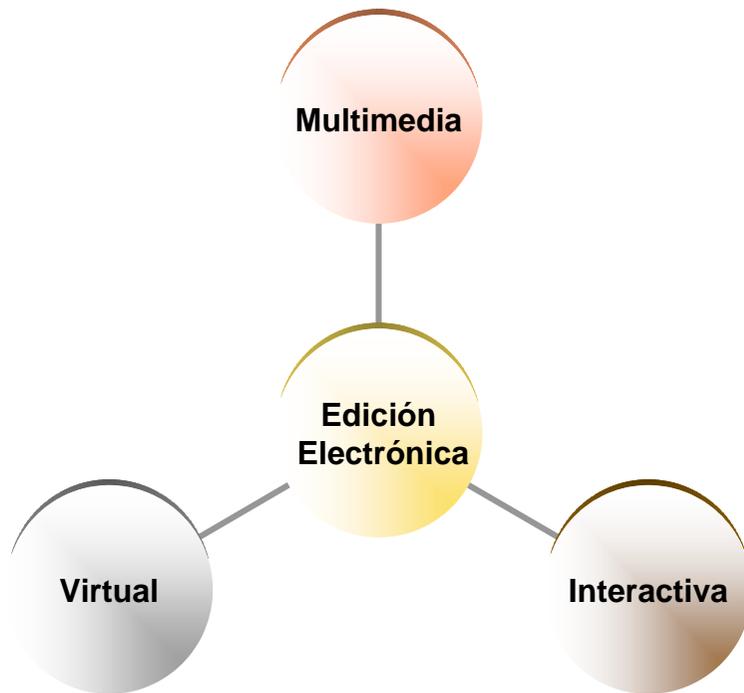


Impacto de las TIC

- Durante los últimos años las innovaciones tecnológicas han **transformado la manera en que se procesa, se guarda, se accede, se comparte y se analiza la información.**
- **Los papeles tradicionales** asignados al productor, al procesador y al usuario de la información están experimentando importantes transformaciones.
- El hecho de que la tecnología de la información haya dotado de competencias informáticas a los investigadores ha **disminuido la especialización entre las habilidades del autor y el editor**



Información electrónica



Fácil de elaborar

Económica

Fácil de almacenar

Rápida distribución

Fácil de recuperar

Uso multidimensional

Material adicional



Contexto

Internet

Sobreinformación

Brecha digital

Intermediación

**Nuevos retos
posibilidades**

**Nuevos
Servicios**

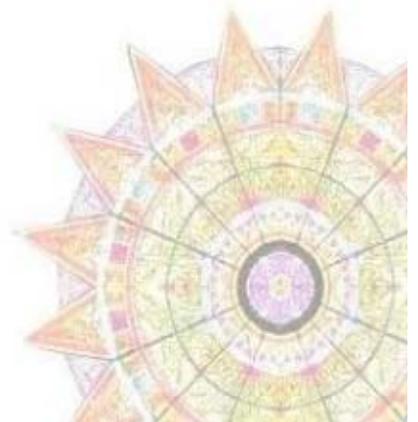
Nuevos medios

**Formación
continuada**



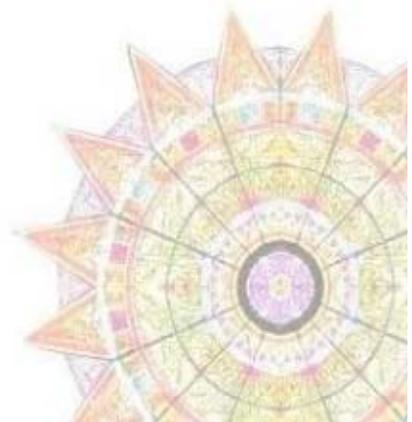
Concepto y Definición

- El concepto se está utilizando con una **amplia acepción** para referirse a cuestiones muy diversas, con diferentes puntos de vista y objetivos
- En los años 90, se conceptualiza en una línea **centrada en los aspectos tecnológicos**
- Será en los primeros años de este siglo cuando se empieza a abordar la consideración de **BD como un SI**
- En la literatura profesional se utiliza como **sinónimo biblioteca electrónica, biblioteca virtual y biblioteca digital**

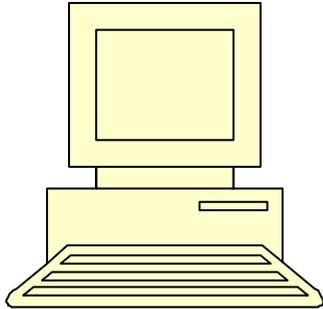


biblioteca electrónica, biblioteca virtual y biblioteca digital

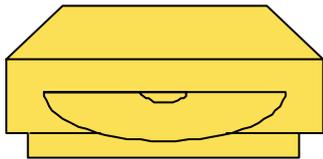
- Una **biblioteca electrónica** estaría formada por objetos físicos que necesitan de medios electrónicos para el acceso a la información contenida en los mismos.
- Una **biblioteca digital** consistiría en un conjunto de materiales y servicios almacenados, procesados y accedidos mediante la utilización de herramientas y redes de comunicaciones digitales
- Una **biblioteca virtual** sería aquella que sólo existe organizada en un espacio informativo virtual.
- El término **más comúnmente aceptado es el de Biblioteca Digital**
- La **evolución** en el tiempo ha sido :
 - la biblioteca automatizada
 - la biblioteca electrónica
 - la biblioteca digital



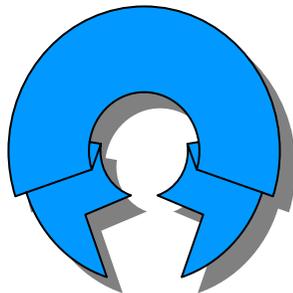
Evolución **histórica**



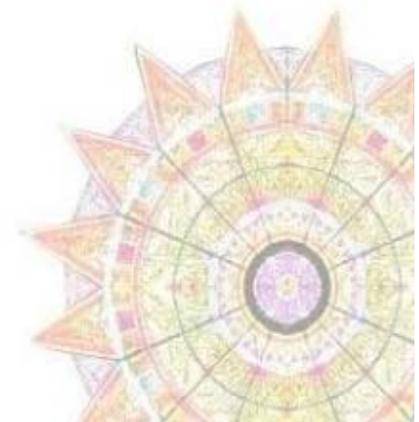
BIBLIOTECA AUTOMATIZADA



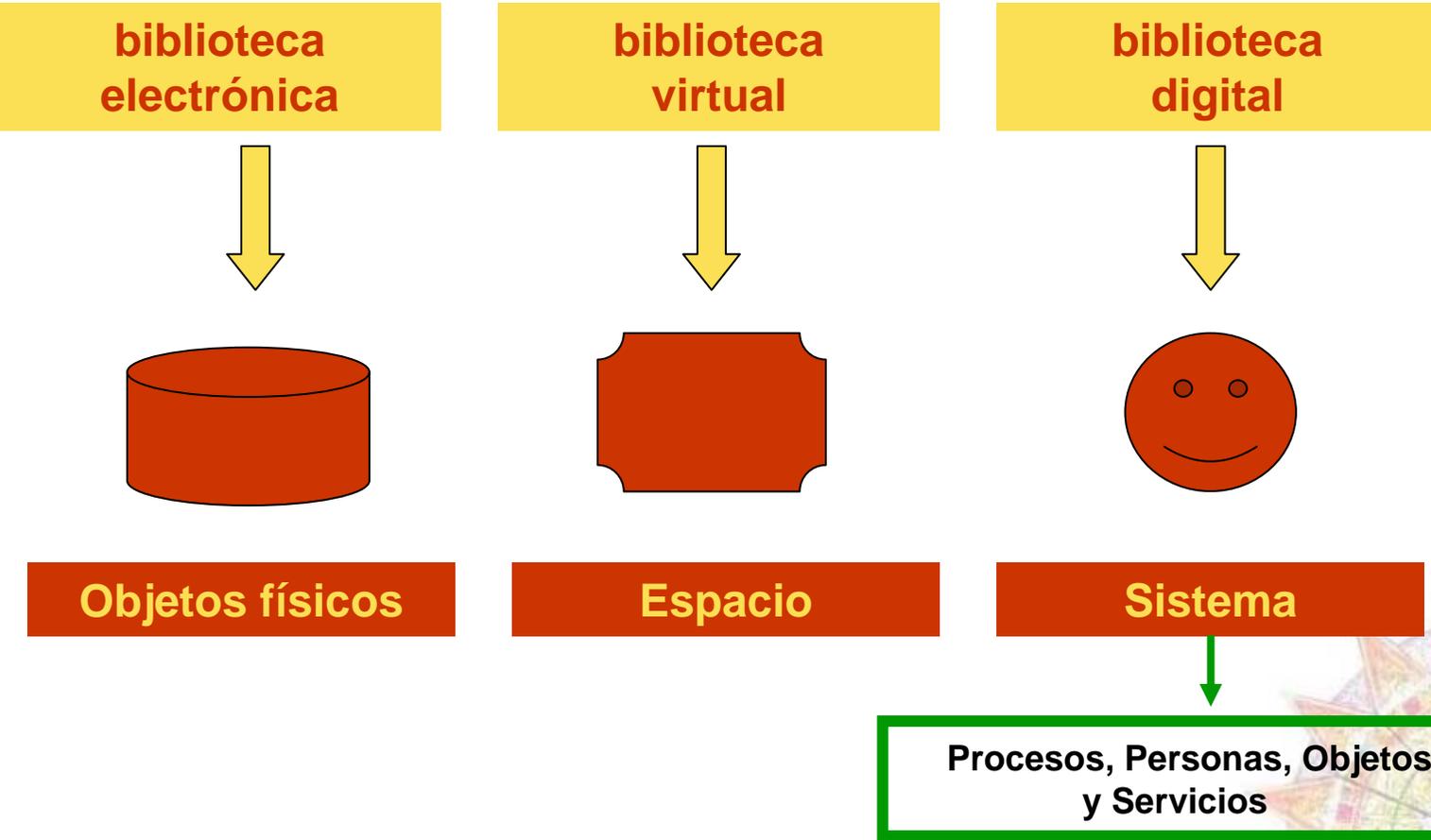
BIBLIOTECA ELECTRONICA



BIBLIOTECA DIGITAL

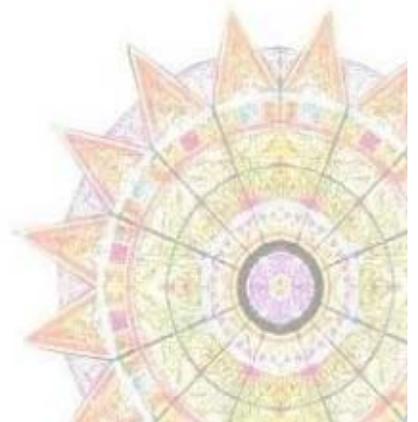
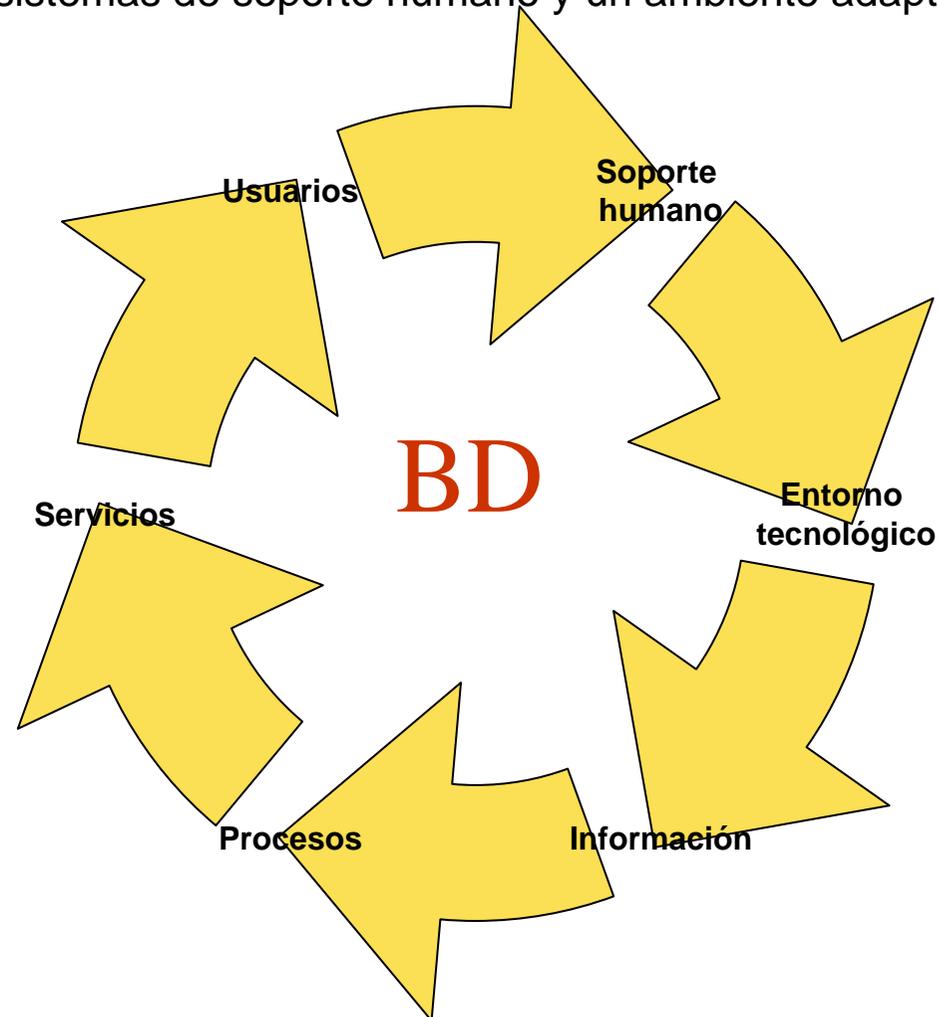


biblioteca electrónica, biblioteca virtual y biblioteca digital



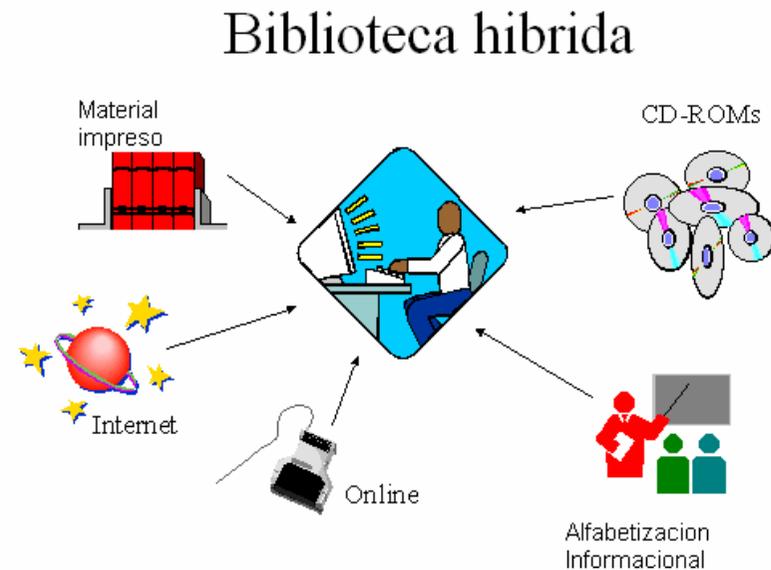
Biblioteca Digital como SI

Seria un espacio virtual en el que interactuarían no sólo fuentes de información, sino también herramientas tecnológicas, sistemas de metadatos, sistema de recuperación, sistemas de soporte humano y un ambiente adaptado



La Biblioteca **híbrida**

- El término fue acuñado por primera vez por programa “eLib” del Reino Unido, para referirse a aquellas bibliotecas que **ofrecen colecciones y funciones de la biblioteca tradicional con aquellos servicios de la biblioteca electrónica**, digital o virtual.
- “La biblioteca híbrida que puede definirse como una biblioteca **diseñada** “para armonizar todo tipo de tecnologías en el **contexto de una biblioteca en funcionamiento**, al mismo tiempo que para explorar la integración de sistemas y servicios en contextos electrónicos e impresos”.



TENDENCIAS

Integración de todos recursos en un único interfaz (Web – OPAC)

Integración con entornos virtuales de aprendizaje (Plataformas de E-learning. CRAI)



Definición de BD

Investigadores

Contenidos
Colecciones
Comunidades de usuarios
Edición digital
Adquisición y Selección
Comunicación Científica

Profesional

Organización
Gestión
Servicios de valor añadido
Recursos de información digital.
Problemática de acceso
Perduración de las colecciones

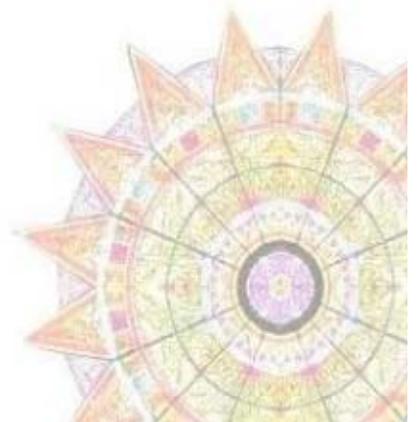
una **biblioteca digital es un sistema de tratamiento técnico, acceso y transferencia de información digital, estructurado alrededor del ciclo de vida de una colección de documentos digitales, sobre los cuales se ofrecen servicios interactivos de valor añadido para el usuario final.**

Tramullas, J., "Propuestas de concepto y definición de la biblioteca digital," *Jornadas de Bibliotecas Digitales*, Vol. 3, 2002. <http://mariachi.dsic.upv.es/jbidi/jbidi2002/Camera-ready/Sesion1/S1-1.pdf> [Consultado 23-octubre 2006]



Tipos de BD

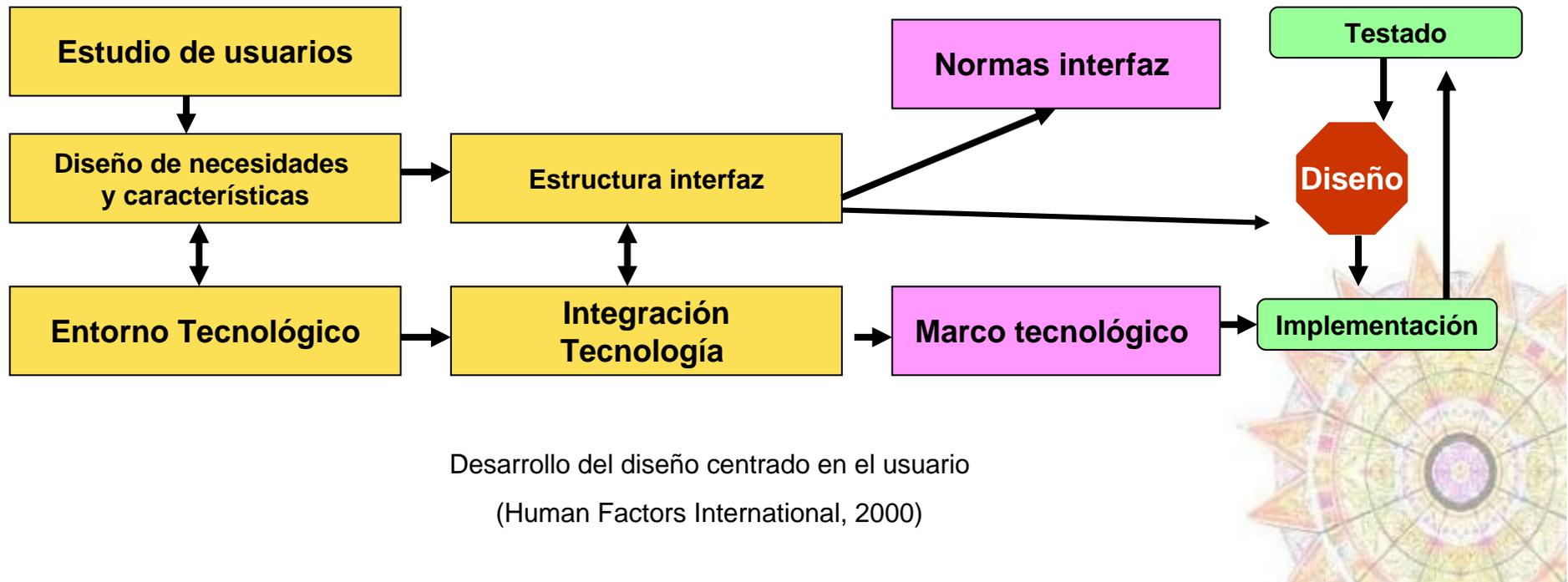
- **Según el formato** de los documentos: texto, gráfico, pdf, audio, vídeo, html.
- **Según el almacenamiento** de la información: un único servidor, distribuida.
- **Según la temática**: multidisciplinarias, especializadas.
- **Según el ámbito geográfico**: generales, locales.
- **Según la institución**: privadas, públicas, mixtas.
- **Según el acceso**: público, restringido, semipúblico



Planificación y Diseño de la BD

En el mundo de la Información ha habido una oscilación de puntos de vista. Los últimos años se han consolidado lo que se ha denominado **orientación a servicios o usuarios**.

- Necesidades de información
- Usabilidad
- Estudio de usuarios
- Flujos de información



Planificación y Diseño de la BD

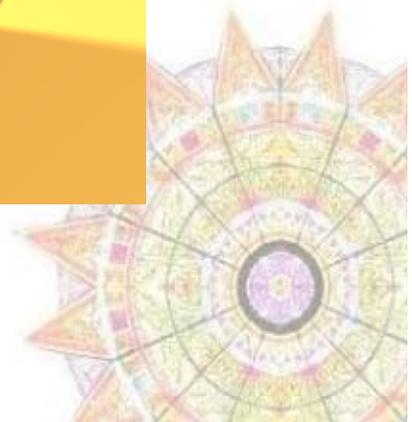
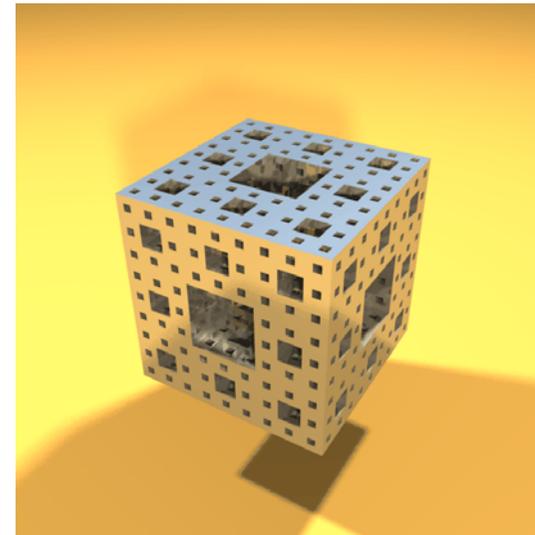
El gran desafío es trasladar el esquema de la biblioteca tradicional a los servicios en línea, aprovechando las posibilidades que ofrecen las tecnologías de la información para poder dimensionar el servicio

Ventajas

- Rapidez
- Comodidad
- Rentabilizar recursos
- Superación de barreras

Inconvenientes

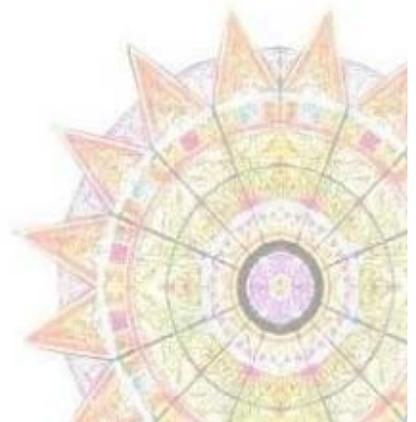
- Problemas legales de acceso (intranets)
- Personal especializado
- Adquisición de software



Planificación y Diseño de la BD

Aspectos a considerar en la planificación

- **Entorno externo:** contexto político, organizativo, social y profesional en el que tiene lugar la planificación.
- **Entorno interno:** objetivos de la organización
- **Necesidades de los usuarios** relacionadas con sus objetivos.
- **Procesos:** acciones de localización, recuperación, adquisición, uso, almacenamiento y evaluación de información.
- **Servicios:** Uso de la biblioteca digital



El Espacio

Biblioteca presencial vs Biblioteca Digital

Espacio real

Conocido

Táctil

Delimitado

Usuario presencial

Servicio presencial



Espacio virtual

Perceptivo

Intangible

Ilimitado

Usuario virtual

Servicio virtual



El bibliotecario en el entorno digital

El empleo de las **TICs influye en el perfil del bibliotecario** ¿cómo?

Los bibliotecarios perciben la aplicación de **las TICs en la prestación de los servicios**. ¿cómo?

Necesita habilidades y conocimientos específicos en el desarrollo de las colecciones.

Afrontar una constante variación en la **definición de su responsabilidad profesional**



La colección **Digital**

Tipología

Tipo de documento

Bases de datos
Revistas electrónicas
Libros electrónicos

Propiedad

Recursos de producción propia
Recursos por suscripción

Tipo de acceso

Libre
Restringido

El grueso de la oferta se basa en recursos comerciales contratados por las bibliotecas a grandes distribuidores (Ebsco, Sweets, Emerald)



Procesos en la BD

Bibliotecas tradicional

Adquisición
Catalogación
Indización
Difusión



Biblioteca Digital

Acceso
Conversión
Presentación
Comunicación

A los procesos tradicionales se añadirían un segundo grupo de operaciones



Gestión de la colección **Digital**

Colección de objetos digitales

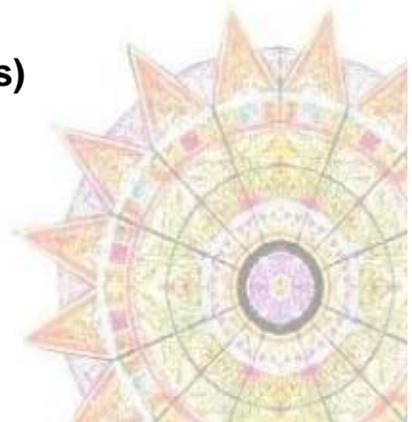
-Gestionado y creados según los **principios de gestión de la colección**

-- Tendrán una **disponibilidad y accesibilidad coherente** como el resto de los materiales de la colección

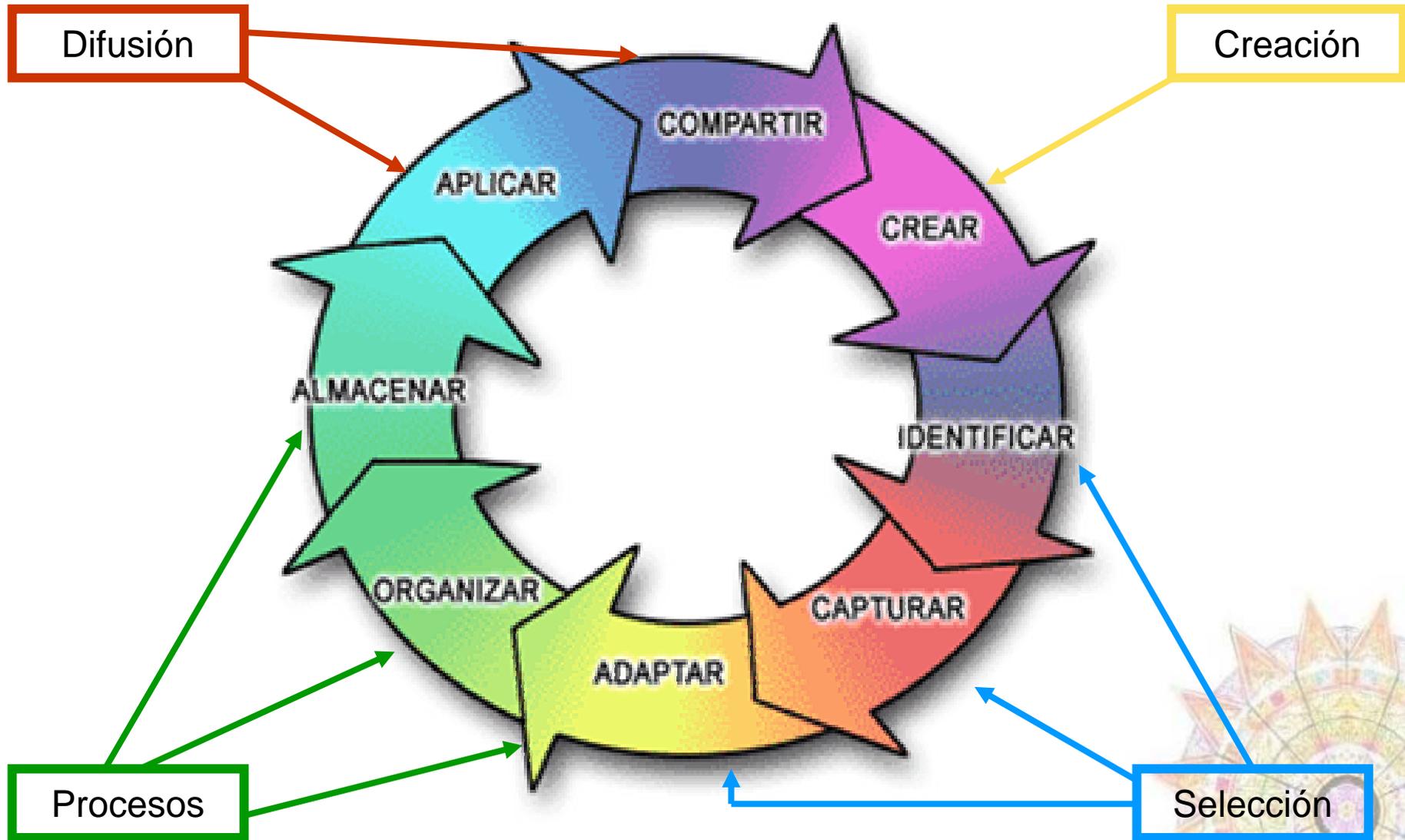
-- Su tratamiento será como recursos a largo plazo para garantizar su **calidad y sostenibilidad**

TAREAS

- **Determinar los recursos existen y la selección de aquellos que son relevantes para las necesidades de los usuarios.**
- **Describir** los recursos seleccionados. (Metadatos)
- **Proporcionar los puntos de acceso necesarios, incluyendo el control de autoridad de los mismos.**
- **Analizar el contenido** de los recursos seleccionados.
- **Proporcionar la información necesaria para la localización y recuperación**
- **Disponer de datos de uso (Indicadores)**

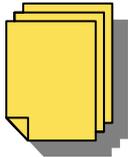


Gestión de la Colección en BD

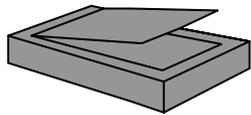


Selección y adquisición de recursos en la BD

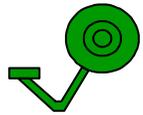
RECURSOS



Materiales. No digitales. Gestión, Localización y Acceso (PI)



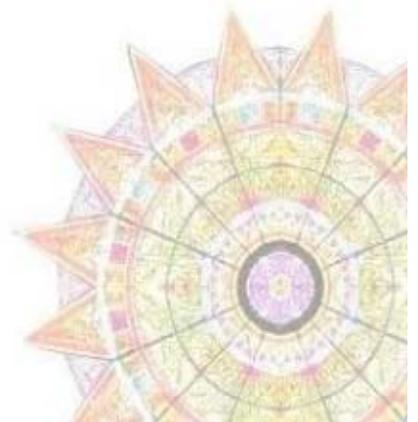
Transición. Digitalizados. Problemas de calidad y de Derechos de Propiedad Intelectual



Digitales. Creados y diseñados para entornos digitales



Accesibles. Materiales en entorno digital que están disponibles libremente



Selección y adquisición de recursos en la BD

- **Bajo licencia de acceso**

- Bases de datos
- Revistas electrónicas



Acceso limitado

- **Recursos propios**

- Material digitalizado (?)
- Seleccionar Información de Internet
- Generar recursos propios
 - Guías temáticas
 - Bases de datos
 - Repositorios
 - Listas de correo



Acceso libre



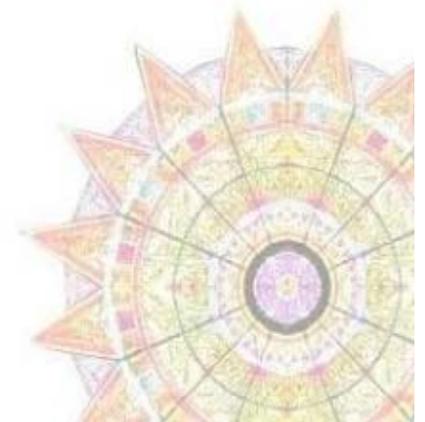
Se trata de un producto nuestro, generado por nosotros y que puede ser susceptible de crear una demanda externa. UNICO y EXCLUSIVO



Mantenimiento y Procesos

Información accesible

- **Identificar y actualizar continuamente** esta información accesible en línea (caen un 14 % de los enlaces)
- **Política coherente de desarrollo** de colecciones
- **Incremento** del volumen de la colección
- **Descripción analítico-sintético** del contenido (Metadatos, XML)
- **Interoperatividad**
- **Usabilidad**
- **Visibilidad**



Metadatos

Datos asociados a un objeto sistema de información con la finalidad de descripción, administración, señalar los requisitos legales, la funcionalidad técnica, uso, utilización y preservación

- La definición más simple de metadatos es "datos sobre los datos."
- Los metadatos pueden tanto generarse a "a mano", como obtenerse automáticamente a través de un software
- Usamos **ficheros .rdf** separados que contienen la información de metadatos en lugar de incluir éstos en el HTML
- Los metadatos pueden **almacenarse en cualquier tipo de base de datos**, y proporcionan un enlace al recurso descrito, más que incluirse dentro de él. Este método es probablemente el más práctico y frecuente
- Aunque también algunas aplicaciones permiten incluir sus metadatos dentro del propio recurso HTML



Dublin Core

- El "**Dublin Core Simple**" son los metadatos Dublin Core que no usan cualificadores ; sólo se expresan los 15 elementos principales del Conjunto de Elementos Dublin Core, como pares simples de atributo-valor sin ningún "cualificador"
- El "**Dublin Core Qualified**" emplea cualificadores adicionales para matizar o clarificar [*refine*] el significado de un recurso

| Contenido | Propiedad Intelectual | Versión |
|--------------------|-----------------------|----------------------|
| <u>Cobertura</u> | <u>Patrocinador</u> | <u>Fecha</u> |
| <u>Descripción</u> | <u>Creador</u> | <u>Formato</u> |
| <u>Tipo</u> | <u>Editor</u> | <u>Identificador</u> |
| <u>Relación</u> | <u>Derechos</u> | <u>Lengua</u> |
| <u>Fuente</u> | | |
| <u>Materia</u> | | |
| <u>Título</u> | | |
| <u>Audiencia</u> | | |



Elementos DC

| | | | |
|---|---|--|---|
| <p>Cobertura</p> <p>Coverage="1995-1996" Coverage="Boston, MA" Coverage="17th century"</p> | <p>Descripción</p> <p>Description="Illustrated guide to airport markings and lighting signals, with particular reference to SMGCS (Surface Movement Guidance and Control System) for airports with low visibility conditions."</p> | <p>Tipo</p> <p>Type="Image" Type="Sound" Type="Text"</p> | <p>Relación</p> <p>Relation="IsVersionOf Elton John's 1976 song Candle in the Wind"</p> |
| <p>Fuente</p> <p>Source="Image from page 54 of the 1922 edition of Romeo and Juliet"</p> | <p>Materia</p> <p>Subject="Aircraft leasing and renting" Subject="Dogs" Subject="Olympic skiing" Subject="Street, Picabo"</p> | <p>Título</p> <p>Title="A Pilot's Guide to Aircraft Insurance"</p> | <p>Audiencia</p> <p>Audience="elementary school students" Audience="ESL teachers" Audience="deaf adults"</p> |
| <p>Patrocinador</p> <p>Contributor="Universidad de Salamanca"</p> | <p>Autor</p> <p>Creator="Shakespeare, William" Creator="Wen Lee" Creator="Hubble Telescope" Creator="Internal Revenue"</p> | <p>Editor</p> <p>Publisher="University of South Where" Publisher="Funky Websites, Inc." Publisher="Carmen Miranda"</p> | <p>Derechos</p> <p>Rights="Access limited to members" Rights="http://cs-tr.cs.cornell.edu/Dienst/Repository/2.0/Terms"</p> |
| <p>Fecha</p> <p>Date="1998-02-16" Date="1998-02" Date="1998"</p> | <p>Formato</p> <p>Format="image/gif 6" Format="40 x 512 pixels"</p> | <p>Identificador</p> <p>Identifier="http://purl.oclc.org/metadata/dublin_core/" Identifier="ISBN:0385424728" Identifier="H-A-X 5690B" [publisher number]</p> | <p>Lenguaje</p> <p>Language="en" Language="fr" Language="Primarily English, with some abstracts also in French." Language="en-US"</p> |

Contenido

Propiedad

Versión

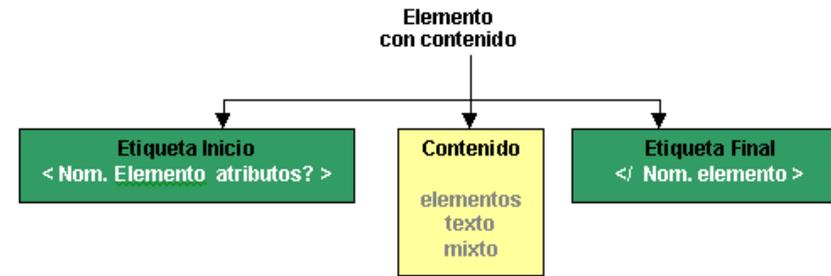
Dublin Core

```
<?xml version="1.0" ?>
- <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
- <rdf:Description rdf:about="http://es.dublincore.org/resources/faq/">
  <dc:title>Preguntas frecuentes(FAQ) sobre la Iniciativa de Metadatos Dublin Core</dc:title>
  <dc:description>Respuestas a las preguntas frecuentes sobre la Iniciativa de Metadatos
    Dublin Core y sobre los metadatos Dublin Core.</dc:description>
  <dc:date>2003-05-16</dc:date>
  <dc:format>text/html</dc:format>
  <dc:language>es</dc:language>
  <dc:publisher>Instituto Universitario Agustin Millares. Universidad Carlos III de
    Madrid</dc:publisher>
  <dc:contributor>Iniciativa de Metadatos Dublin Core</dc:contributor>
  <dc:contributor>Eva Mendez</dc:contributor>
  <dc:relation>http://www.dublincore.org/resources/faq/index.shtml.rdf</dc:relation>
</rdf:Description>
</rdf:RDF>
```

Guía de uso del Dublin Core

<http://es.dublincore.org/documents/usageguide/index.shtml#whatismetadata>





Pertenece a los denominados **lenguajes de marcas o marcado**, llamados así porque **describen el contenido de lo que etiquetan**.

A primera vista, un documento **XML es similar a los documentos HTML**, pero realmente no son iguales, hay una diferencia principal: un documento XML contiene datos que se autodefinen exclusivamente. En XML se separa el contenido de la presentación de forma total.

Se trata **Metaetiquetas** que pueden ver e interpretar las máquinas, pero que son invisibles para el usuario

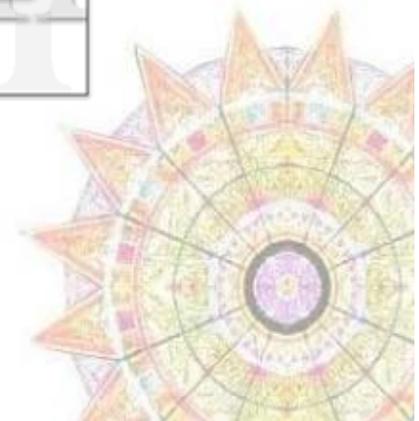
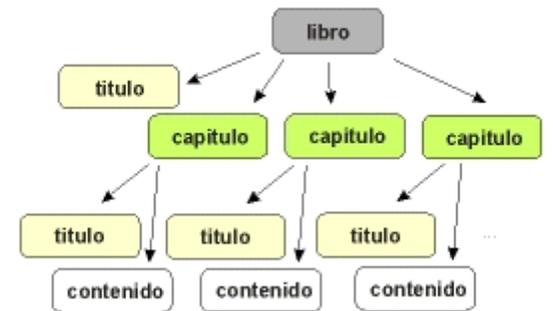
Un editor XML es una aplicación que nos ofrece facilidades para crear y editar documentos XML. (DTD, XML, Notepad ...)

• **Utilizo mis propias etiquetas**. Y es que en XML no estamos trabajando con etiquetas predefinidas. Nosotros podemos crearnos nuestro propio lenguaje de etiquetas en función de nuestras necesidades. .



Objetivos y usos de XML

- Que fuera idéntico a HTML para aprovechar toda la tecnología implantada de este.
- Que fuera extensible, para que lo puedan utilizar en todos los campos del conocimiento.
- Que fuese fácil de leer y editar e implementar. (Se trata de una versión de SGML reducida)
- Que fuese fácil de implantar, programar y aplicar a los distintos sistemas.
 - Permite estructurar la información (capítulos, índices, Resumen, apartado, notas, párrafos, secciones, imágenes)
 - Permite definir contenido (Palabras clave ...)
 - Facilita las tareas de gestión, edición y recuperación



HTML

Define Formato

Cursiva, negritas, Time Roman

```
<html>
<head>
<title>Registro nº1</title>
</head>
<body>
<p>Título: Organización de las bases de datos</p>
<p>Autor: Martin, James</p>
<p>Traductor: Marco, Alfredo di</p>
<p>Editorial: Londres: Prentice Hall PTR, 1975c</p>
<p>Depósito legal: M-5487-1975</p>
<p>ISBN: 0-25-25987-4</p>
<p>Notas: 528 páginas, 56 il. en blanco y negro,
índices de contenido y acrónimos</p>
<p>Resumen: Esta obra pretende sistematizar, de
alguna manera, el trabajo de diseño y gestión de bases
de datos. Para ello se analizan los principales modelos,
describiendo con especial énfasis el modelo entidad-
relación</p>
</body>
</html>
```

XML

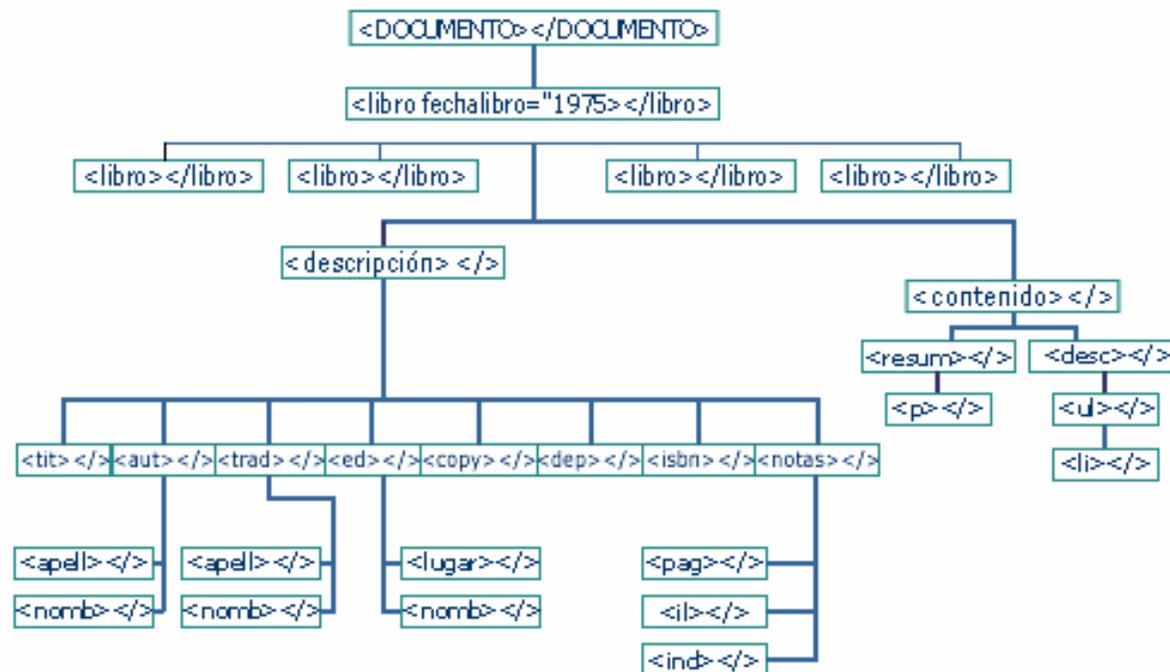
Define Estructura y Contenido

Título, Resumen, capítulos, párrafos

```
<?XML version=» 1 .O»?>
<LIBRO>
<!--La primera edición en 1970 -->
<título>Organización de las bases de datos</título>
<capítulo l>
<título>Introducción</título>
<párrafo>El desarrollo de las bases de datos será
sin duda una de las actividades.. .</párrafo>
<párrafo>A medida que aumenta...</párrafo>
<nota>Las estructuras de datos...</nota>
</capítulo l>
...
</LIBRO>
<ARTÍCULO>
<!--Rev. Esp. de Doc. Cient., v. 21, n. 4, 1998 -->
<título>Norma z39.50, actitud, posibilidades</título>
</resumen>
<párrafo>La tecnología de los sistemas...</párrafo>
</resumen>
<palabras clave>
<párrafo>z39.50, SQL, recuperación...</párrafo>
</palabras clave>
<apartado l>
<título>Introducción: el contexto...</párrafo>
<párrafo>Pensamos que en...</párrafo>
</título l>
</apartado 1>
<ARTÍCULO>to
```



Árbol de etiquetas XML



XML estructura la información de manera que pueda ser tratada como un árbol con raíz (libro), ramas primarias (capítulos), ramas secundarias (secciones) y hojas (párrafos)

```
<?xml version="1.0" encoding="UTF-8" ?>
- <!--
  These data are only offered for use pursuant to the license agreement
  posted at http://webservices.rhapsody.com/rws-license.html.
  Any use of these data indicates your agreement to the terms and conditions
  set forth therein.
-->
- <rss version="2.0">
- <channel>
  <title>Verve Jazz Masters 39 by Cal Tjader on Rhapsody Online</title>
  <link>http://play.rhapsody.com/caltjader/vervezazzmasters39</link>
  <description>Verve Jazz Masters 39 by Cal Tjader on Rhapsody Online</description>
  <category>Afro-Cuban Jazz</category>
  <language>en</language>
  <ttl>720</ttl>
  <pubdate>Tue, 7 Nov 2006 13:33:13 -0800</pubdate>
- <image>
  <url>http://static.realone.com/rotw/images/logo_rhapsody_225x47.gif</url>
  <title>Verve Jazz Masters 39 by Cal Tjader on Rhapsody Online</title>
  <link>http://play.rhapsody.com/caltjader/vervezazzmasters39</link>
  <description>Verve Jazz Masters 39 by Cal Tjader on Rhapsody Online</description>
</image>
- <item>
  <title>Soul Sauce (AKA "Guachi Guoro") - Cal Tjader</title>
  - <description>
    - <![CDATA[
      <script type="text/javascript" src="http://www.rhapsody.com/simple.js"></script>
      <strong>Verve Jazz Masters 39</strong><br />
      <br />
      ¶#169; 1994 The Verve Music Group, a Division of UMG Recordings, Inc.<br />
      <a href="http://play.rhapsody.com/caltjader/vervezazzmasters39/soulsauceakaguachiguoro" onclick="RhapsodyPlayer.playRhapsody( { id
      <a href="http://www.rhapsody.com/caltjader/vervezazzmasters39">Verve Jazz Masters 39 Page on Rhapsody Online</a><br />
    ]]>
  </description>
  <link>http://play.rhapsody.com/caltjader/vervezazzmasters39/soulsauceakaguachiguoro</link>
  <category>Afro-Cuban Jazz</category>
  <pubDate>Sat, 1 Jan 1994 00:00:00 -0800</pubDate>
  <source url="http://play.rhapsody.com/caltjader/vervezazzmasters39">Verve Jazz Masters 39 by Cal Tjader on Rhapsody
  Online</source>
  <guid isPermaLink="false">2176431</guid>
  <rhap:rcid xmlns:rhap="rhap">tra.2176431</rhap:rcid>
  <rhap:artist xmlns:rhap="rhap">Cal Tjader</rhap:artist>
  <rhap:artist-rcid xmlns:rhap="rhap">art.5766</rhap:artist-rcid>
  <rhap:album xmlns:rhap="rhap">Verve Jazz Masters 39</rhap:album>
  <rhap:album-rcid xmlns:rhap="rhap">alb.202069</rhap:album-rcid>
  <rhap:album-art xmlns:rhap="rhap">http://image.listen.com/img/170x170/9/8/8/5/415889_170x170.jpg</rhap:album-art>
</item>
</channel>
</rss>
```

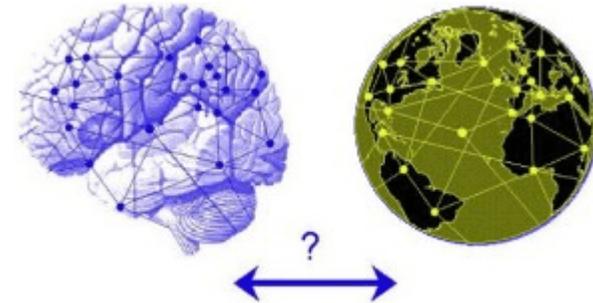


La **WEB** Semántica

- La Web Semántica ha sido impulsada por **Tim Berners-Lee**, creador de la **WWW**, y otras personas relacionados con el **W3C (World Wide Web Consortium)**.

Se trata de una aspiración, más que de una realidad

- Tal como se codifican las **páginas web** actualmente tienen poco sentido para los robots. Se trata de que estas páginas dejen de ser simples cadenas de caracteres para convertirse en textos con sentido que puedan ser interpretados por las máquinas para tomar decisiones adaptadas al contexto en el que se desarrollan



La **Web Semántica** es una extensión de la Web actual, dotada de mayor significado, en la que cualquier usuario en Internet podrá encontrar respuestas a sus preguntas de forma más rápida y sencilla gracias a una información mejor definida

La **WEB** Semántica

Web para humanos

```
<h1>Homogenic</h1>
<p>Album de música electrónica,
compuesto por la artista islandesa
<a href="/artista/Bjork">Björk</a>
en el año 2001.</p>
<p>Incluye canciones como Hunter,
Pluto, entre otras.</p>
```

Web para programas

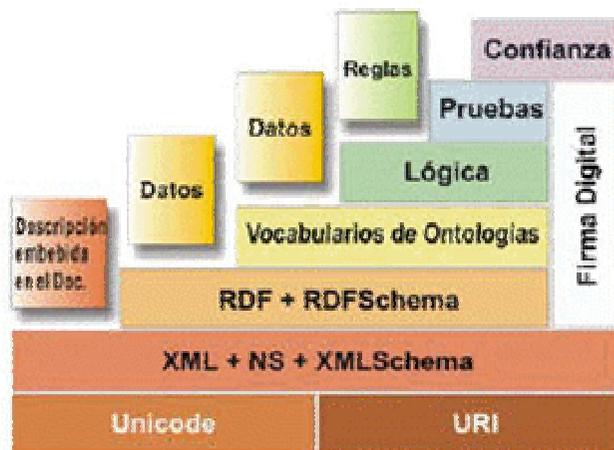
```
<Album>
<name>Homogenic</name>
<year>2001</year>
...
<artist resource='...' />
<genre type='...#Electronica' />
<tracks>
  <track resource='...' />
  ...
</tracks>
</Album>
```

Base
de
Datos

La Web Semántica debería ser capaz de **procesar contenido**, **razonarlo y hacer deducciones lógicas** a partir de éste, y realizar, cuando un **usuario quiera delegar ciertas tareas en el software**, todas estas acciones de forma automática.



Medios de la **Web Semántica**



XML Conjuntos de etiquetas específicos para cada necesidad

Metadatos. Datos sobre los datos

Esquemas Clasificaciones, tesauros...

RDF Sistema de Normalización de todos los metadatos a los que proporciona una sintaxis

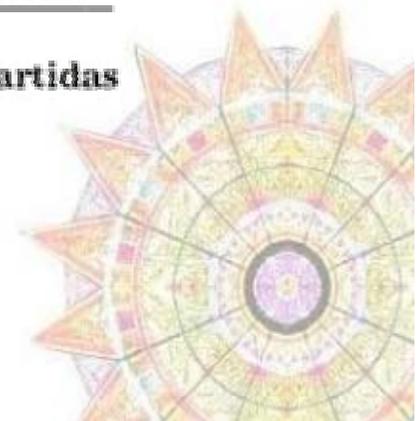
Ontologías. Representación del conocimiento complejo

XML – Hace documentos con más significado

Metadatos – Facilitan la representación, indización y búsqueda-recuperación

Programas de nueva generación –que pueden interpretar y gestionar preguntas y contenidos

ReD del Conocimiento



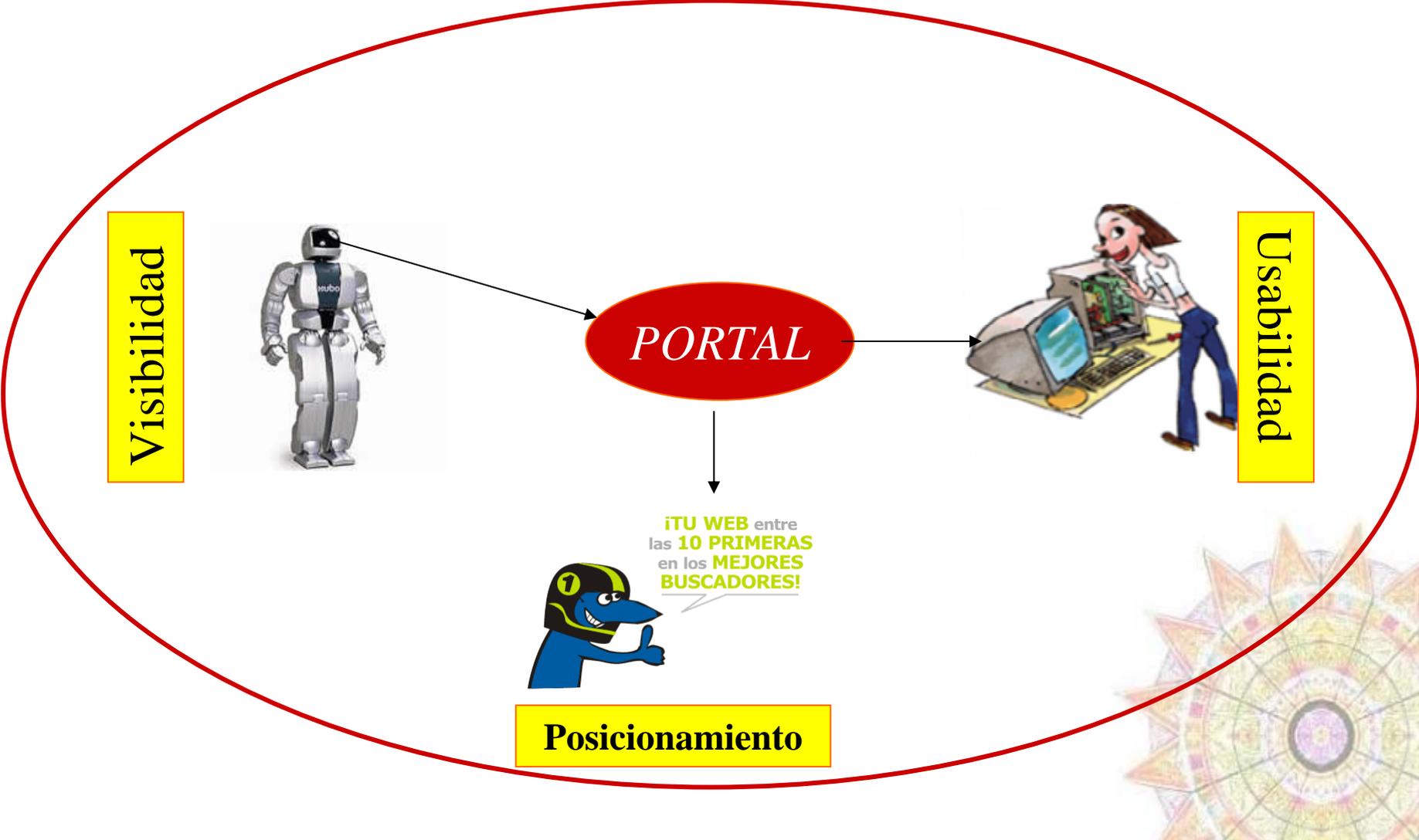
Visibilidad, **posicionamiento** y usabilidad



- En internet el producto último son **los portales**, su visibilidad y aceptación depende de su posicionamiento, usabilidad, contenidos, diseño y rapidez.
- **Visibilidad** implica estar en la Web y poder ser encontrado y visitado.
- **Posicionamiento**. Colocar un sitio Web en los primeros puestos dentro de los motores de búsqueda cuando se realice una determinada búsqueda en Internet. Se realiza mediante la optimización de las páginas y de complicados análisis de las mismas.
- **Usabilidad** supone la aplicación de una serie de métodos que buscan que el sistema sea fácil de usar y de aprender.
- Visibilidad, posicionamiento y usabilidad son conceptos que **actúan interrelacionados y se complementan**



Visibilidad, posicionamiento y usabilidad



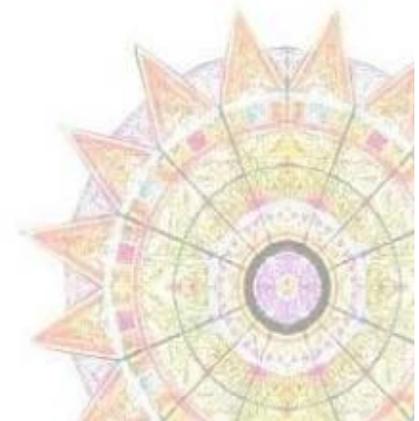
“Internet Invisible”



- Según algunas estimaciones **estos motores solamente indexan un 25 % de todos los contenidos de la red**
- A esta porción de **información que es accesible, pero no está indexada por los buscadores** se le denomina “Internet Invisible”
- Las **páginas webs dinámicas**, es decir, las que se generan como resultado de una búsqueda, no existen hasta que no se realiza dicha consulta y, por tanto, no son detectadas ni incorporadas a la base de datos del buscador
- El problema es importante, no sólo por el volumen estimado, más aún por la **calidad de estos recursos que aparecen en la cara oculta** y que precisamente son los más relevantes para investigación.
- Por ello es **importante generar servicios de información especializada que organicen y faciliten el acceso a los mismos** en función de los intereses de sus más inmediatos usuarios.



“Internet Invisible”



Como **ser Visible?**



- **No existe un acuerdo unánime**, ni una formula infalible que determine que aspectos son determinantes para ser altamente visible a los buscadores, si se apuntan una serie de aspectos que ayudarán a la mejora de su visibilidad:
 - **Registrando la página en el directorio de recursos** de los principales buscadores
 - **Profundidad** del sitio web.
 - **Estructura interna de la Web**, tener un código ESTANDARIZADO de lenguaje (HTML)
 - **Lenguaje e idioma**. Utilizar un idioma de uso común para la ciencia: Inglés
 - Utilización de **metadatos y lenguajes de marcado RDF, XML**
 - **Volumen, actualización y permanencia**



Como podemos medir la visibilidad

- **Numero de visitas**. Estos datos siempre hay que utilizarlos con mucha **prudencia** y tener en cuenta que número de visitas no se corresponde con número de usuarios, ni que todas las visitas son reales.
- **Enlaces recibidos**. Cantidad, y calidad de los mismos
 - **Informe popularidad Exit**
<http://www.kronosdoc.com/gtbib/directorios/exit.php>
 - **TouchGraph GoogleBrowser**
<http://www.touchgraph.com/TGGoogleBrowser.html>
- **Rango de página** “PageRank” Algoritmo basado en los vínculos existentes entre páginas, de manera que cuanto más referenciada está una página por otras páginas de calidad más calidad tendrá.

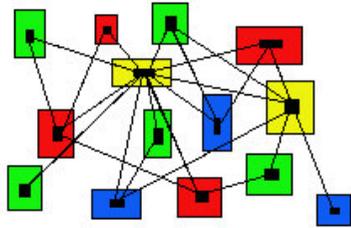


“PageRank”



- Google mide el rango de **1/10 puntos**. Se trata de un algoritmo secreto, en el que influyen diferentes factores, pero parece ser que el **principal argumento que utiliza es que a esa página sea enlazada por páginas de calidad**
- Se trata de un **secreto comercial** que los propios gestores guardan con mucho celo, solo mediante algunos análisis es posible entrever alguna de sus características de funcionamiento, pero además los propios técnicos **cambian algunos de los criterios, para mejorar y castigar prácticas abusivas**
- Los enlaces emitidos por **páginas que son en sí mismas "importantes"** pesan más y ayudan a **convertir a otras páginas también en "importantes"**
- Los sitios importantes y de alta calidad reciben un PageRank más alto, que **Google recuerda el rango de página cada vez que realiza una búsqueda y emite unos resultados**





Page Rank

- Básicamente, el valor del Page Rank de un enlace se calcula tomando el **PageRank de la página que te enlaza y multiplicándola por 0.85 y dividiéndolo por el número total de enlaces en esa página.**

– **Cuantos enlaces me deben hacer para crecer hasta PageRank 6?**

Desde páginas con PR3 necesitas 2750 enlaces.

Desde páginas con PR4 necesitas 550 enlaces.

Desde páginas con PR5 necesitas 110 enlaces.

Desde páginas con PR6 necesitas 22 enlaces.

Desde páginas con PR7 necesitas 5 enlaces.

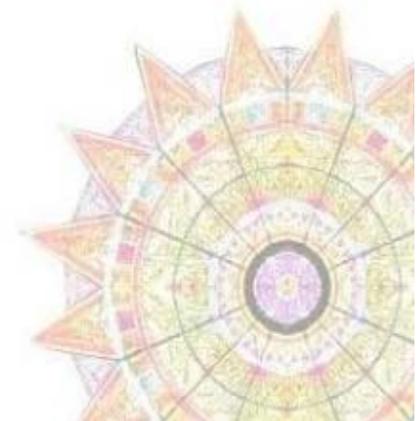
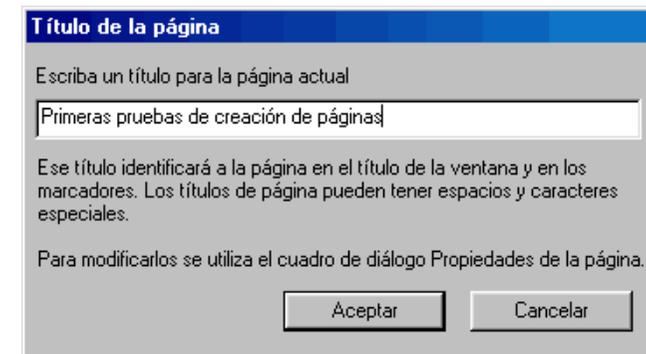
Desde páginas con PR8 necesitas 1 enlace.

Desde páginas con PR9 necesitas 1 enlace.

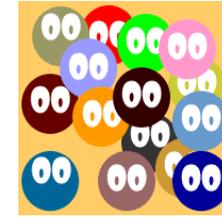
Desde páginas con PR10 necesitas 1 enlace.

Etiquetas META

- Además de PageRank, Google combina en sus búsquedas diversas técnicas que rastrean **coincidencias de las palabras buscadas entre las páginas de su base de datos.**
- Las búsquedas de coincidencias de textos, como ya se sabrá, abarcan gran cantidad de lugares, como el título, **etiquetas META**, cuerpo de la página y además valoran cada aparición según donde se produzca y en que condiciones
- Las búsquedas de coincidencias también se extienden a las páginas que enlazan con la página que Google pretende posicionar, es decir, **también busca coincidencias en las páginas que enlazan con otra** para valorar si ese voto otorgado es adecuado
- Por lo cual se puede decir que **el rango de pagina no es el único factor que influye en el posicionamiento**, y, por tanto, no es nada raro encontrar webs con un Pagerank 2 por delante de páginas que tienen un valor de 4



Posicionamiento



- Otro aspecto relacionado con la visibilidad es el posicionamiento, o sea, **no es sólo importante que la página sea encontrada por el buscador, sino también que la página aparezca en los lugares preferentes** en los resultados que arroja la operación de búsqueda
- En este aspecto parece ser que tiene una influencia fundamental que la **Web este registrada en el directorio de recursos del buscador**, pues el motor busca en primer lugar en los recursos que tiene directamente indexados.
- Cuando realizamos una consulta en un buscador, los resultados suelen figurar de diez en diez. **Rara vez los usuarios pasan del resultado 25**, por lo que si tu web no se encuentra en las primeras posiciones, puede considerarse invisible
- Actualmente muchos **profesionales del marketing** están concienciados de la importancia de que las páginas web estén optimizadas técnicamente para la indexación natural en los principales buscadores



Posicionamiento



- Un estudio de usuarios desarrollado recientemente por la empresa **Eyetoools** e buscador **Google**, se ha determinado que casi el **100% de los mismos revisa los 3 primeros enlaces** que aparecen en la búsqueda. La visión de los resultados de búsqueda **cae en picado por debajo del 50% en los siguientes** resultados y hasta un **20% en los del final de la página**

Organic Ranking Visibility Eyetoools Research and Reports

(shown in a percentage of participants looking at a listing in this location)



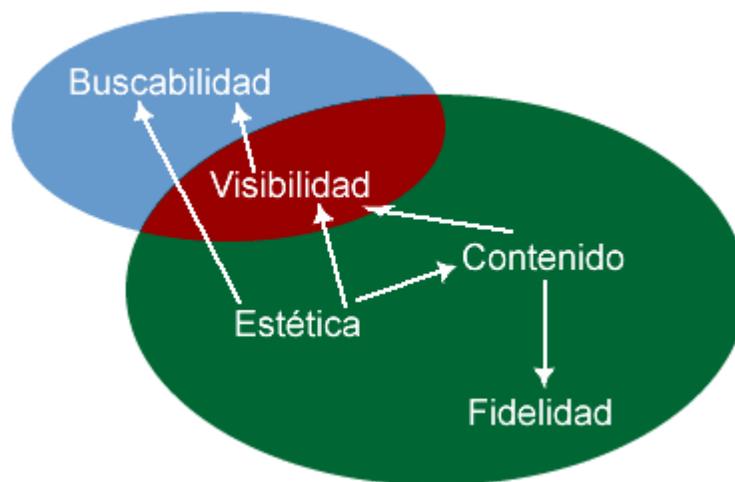
- Rank 1 – 100%
- Rank 2 – 100%
- Rank 3 – 100%
- Rank 4 – 85%
- Rank 5 – 60%
- Rank 6 – 50%
- Rank 7 – 50%
- Rank 8 – 30%
- Rank 9 – 30%
- Rank 10 – 20%



USABILIDAD



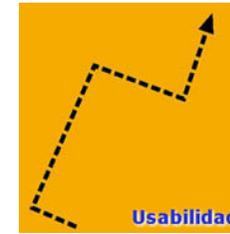
Si encontrar una sede web y que sea visible es importante, no lo es menos que está **sea adecuadamente utilizable**, es decir que **el usuario encuentre lo que busca en la web de forma eficiente, rápida y segura**; aspecto que va a tener su incidencia en la frecuencia de uso de la página – **fidelización del cliente**-.



- Aproximación al usuario
- Amplio conocimiento del contexto del usuario
- Satisfacer las necesidades



Usabilidad



- Cuando hablamos de usabilidad hay dos cuestiones que son fundamentales: **Diseño y contenido**
- La usabilidad **es un método de ingeniería de sistemas que se orienta a usuarios**, y que por lo tanto estos deben definirse en función de los intereses de los mismos, de manera que respondan de la manera más adecuada a sus necesidades
- **Satisfacción subjetiva**. Una cuestión fundamental respecto a la usabilidad es **que el usuario sienta que controla el sistema, y que este complementa sus habilidades de manejo**, a esto se lo denomina satisfacción subjetiva, y se debe basar en un conocimiento de los usuarios y sus potencialidades.
- **La comunicación entre el sistema de información y el usuarios** a través de esa interfaz que es la WEB sea válida.



Los usuarios en el entorno de las BD

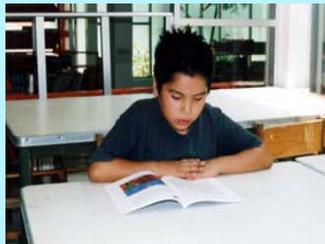
Con acceso pleno



Con Acceso limitado



Presencial



Virtual



Servicios de la BD

Cualquier unidad de información que quiera ser competitiva deberá proporcionar servicios de información digital

- **Información general:** horarios, directorios, Normativa FAQs, Guías, Noticias, Buzón de sugerencias, Carta de Servicios
- **Acceso:** Identificación de usuarios, área de trabajo, consulta (browser o base de datos), Descargas, Peticiones de préstamo, reserva, renovación, desideratas, boletín de novedades,
- **Servicios de información:** Servicios de alerta, Formación interactiva, DSI, Información de referencia, bibliografía recomendada, Guías temáticas, Listas de correo & empleo, Actividades de formación)
- **Otros servicios:** Exposiciones, recursos temáticos, repositorio, listas de discusión, espacio web

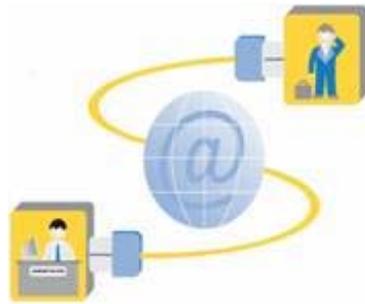




Gobierno electrónico

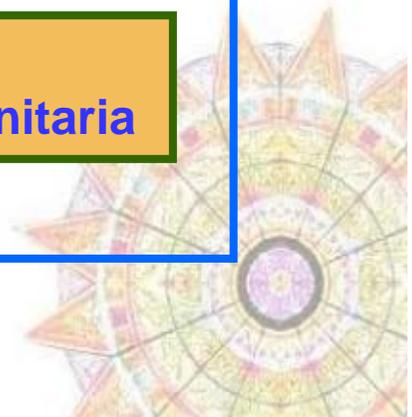
E-GOVERN

Formas de relación entre los ciudadanos y las Administraciones Públicas realizadas mediante las **TIC's** desde donde el público pueda **encontrar información pública, realizar trámites y contactar representantes del gobierno.**



Servicios de Apoyo

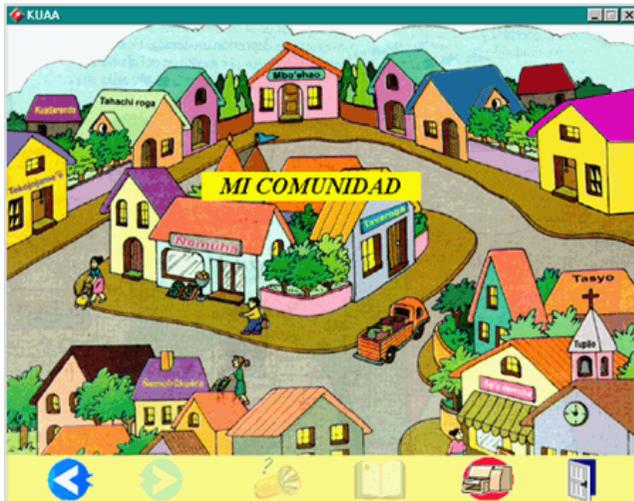
Servicios de información a la Comunidad
integración e implicación de la biblioteca en la vida comunitaria



Servicios de información a la Comunidad

integración e implicación de la biblioteca en la vida comunitaria

- El **nuevo papel de la biblioteca en la sociedad de la información** exige una integración cada vez mayor de la biblioteca en la comunidad de usuarios a la que atiende
- La biblioteca desarrolla un **papel activo en la vida comunitaria** y colabora con las instituciones y grupos existentes en la localidad.
- La biblioteca no debe entenderse únicamente como un agente cultural, debe **ser un elemento potenciador del desarrollo social** de su comunidad



Atender las necesidades informativas de toda índole

Acciones y actividades que las bibliotecas realizan en colaboración con los centros y colectivos de la localidad

Afianzar su presencia en la vida ciudadana

Marcado aspecto social



Servicios de información a la Comunidad

- **Información Ciudadana:** Autobuses, trenes, taxis. Carreteras. Servicios de urgencia. Teléfonos de aviso.
- **Agendas y páginas amarillas:** Agenda de actividades
- **Instituciones, directorios.** Asociaciones, consumo
- **Datos sobre la localidad:** Plano de la ciudad. Industrias. Estadísticas. Datos y precisiones del tiempo.
- **Servicios sociales:** Servicios sociales a minorías, Asistencia médica , Asistencia social, Empleo, Estudios y becas Jóvenes, Medio ambiente, Mujer, ONGs, Alcoholismo
- **Actividades locales.** Aficciones Caza, Pesca. Folklore, Patrimonio Etnográfico
- **Apoyo a los centros escolares**

Socialización de acceso a las TIC

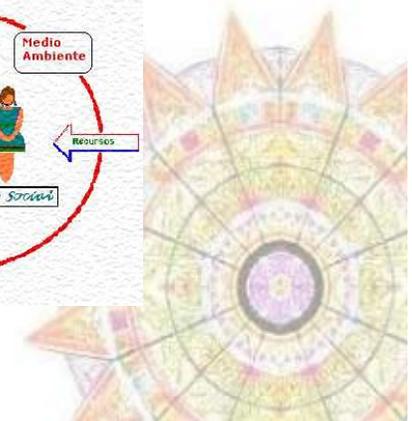
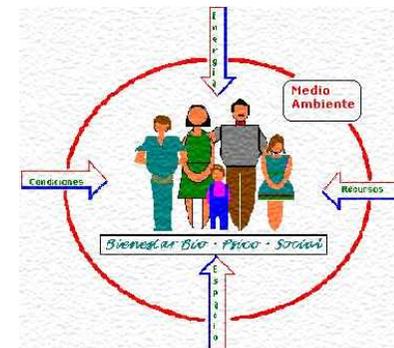
Cursos a mayores, aulas de informática

•Apoyo a la investigación local

•Iniciativas de carácter editorial y literario

Espacio Web, Blogs, Listas

•Alfabetización Informativa



Derechos de autor y copyright

Derechos Morales

Reconocimiento

Integridad

inalienables



No se pueden ceder

Derechos explotación

Reproducción

Distribución

Comunicación Pública

Exclusivos

Se pueden ceder

En conjunto

En parte

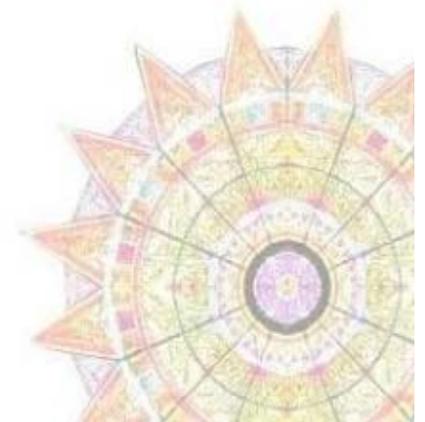
En Exclusiva



Preservación en las BD

El documento digital tiene unas características propias, tales como la facilidad de copia y el uso interactivo, que hacen que su conservación y preservación sea distinta a otro tipo de materiales

- No presenta deterioro de ningún tipo con el paso del tiempo
- Se envían reproducciones idénticas entre copia y copia
- Se pueden manipular mucho más fácilmente
- Importancia de la salvaguarda del patrimonio por el dinamismo de la propia Internet



Evaluación y calidad en las BD

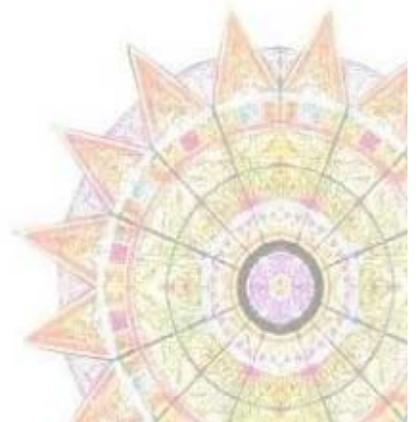
Todo sistema requiere unos mecanismos de medición, que nos orienten en que áreas de trabajo y servicios están cubriendo las expectativas de los usuarios y cuales son susceptibles de mejora.

E-metrica o Cibermetría se plantea establecer una serie de parámetros que sean válidos para la especificidad del impacto de los recursos electrónicos

La medición tiene una enorme importancia de cara a **evaluar si la selección de los recursos electrónicos**, sobre todo aquellos que suponen un alto coste, **se rentabilizan con el uso**

Entre las **cuestiones** que nos orientaran para la toma de decisiones estarían;

- Relación visitas/descargas
- Visibilidad
- Usabilidad del sistema



Association of Research Libraries

"Measures for Electronic Resources"

Accesos a recursos electrónicos:

- R1: Número de revistas electrónicas a texto completo.
- R2: Número de fuentes electrónicas de referencia.
- R3: Número de libros electrónicos.

Uso de recursos y servicios en red:

- U1: Número de transacciones de referencias electrónicas.
- U2: Número de sesiones en bases de datos electrónicas.
- U3: Número de búsquedas realizadas en bases de datos.
- U4: Número de registros consultados en bases de datos.
- U5: Visitas virtuales al sitio web de la biblioteca y catálogo.

Gastos relacionados con los recursos disponibles en red y en infraestructura relacionada:

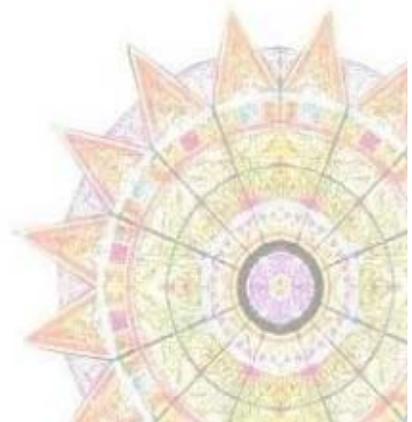
- C1: Costo de revistas electrónicas a texto completo.
- C2: Costo de fuentes electrónicas de referencia.
- C3: Costo de libros electrónicos.
- C4: Gastos de la biblioteca para utilitarios bibliográficos, redes y consorcios.
- C5: Gastos externos de para utilitarios bibliográficos, redes y consorcios.

Actividades de la digitalización de la biblioteca:

- D1: Tamaño de la colección digital de la biblioteca.
- D2: Uso de la colección digital de la biblioteca.
- D3: Costo de la construcción y mantenimiento de la colección digital.

A partir de este conjunto de datos, se elaboran las medidas de ejecución, esta son:

- P1: Por ciento de transacciones de referencias electrónicas del total de referencias.
- P2: Por ciento de visitas a la biblioteca virtual del total de las visitas a las bibliotecas.
- P3: Por ciento de libros electrónicos de todas las monografías.



La Biblioteca digital

Curso "SIDRA: La Biblioteca Digital"

SIDRA. Sistema de Información Documental en Red de Asturias

Julio Alonso Arévalo
Universidad de Salamanca
alar@usal.es

