
El rol de las herramientas y tecnologías de la web semántica en las bibliotecas digitales.

María Rosa Mostaccio

II Encuentro Nacional de Catalogadores “La cooperación y las normas para la organización y tratamiento de la información en las bibliotecas argentinas”

25, 26 y 27 de noviembre de 2009 ,

Biblioteca Nacional, Buenos Aires, Argentina

Estructura de la presentación

- ¿Qué es la web semántica?
- ¿Cómo es la arquitectura de la web semántica?
- ¿Qué rol juegan las ontologías?
- ¿Por qué la aplicación de herramientas y tecnologías de la web semántica en las bibliotecas digitales?
- ¿Qué ejemplos existen de aplicaciones de tecnologías web en bibliotecas digitales?
- ¿Por qué es necesario que los bibliotecarios se interesen y utilicen estas tecnologías?

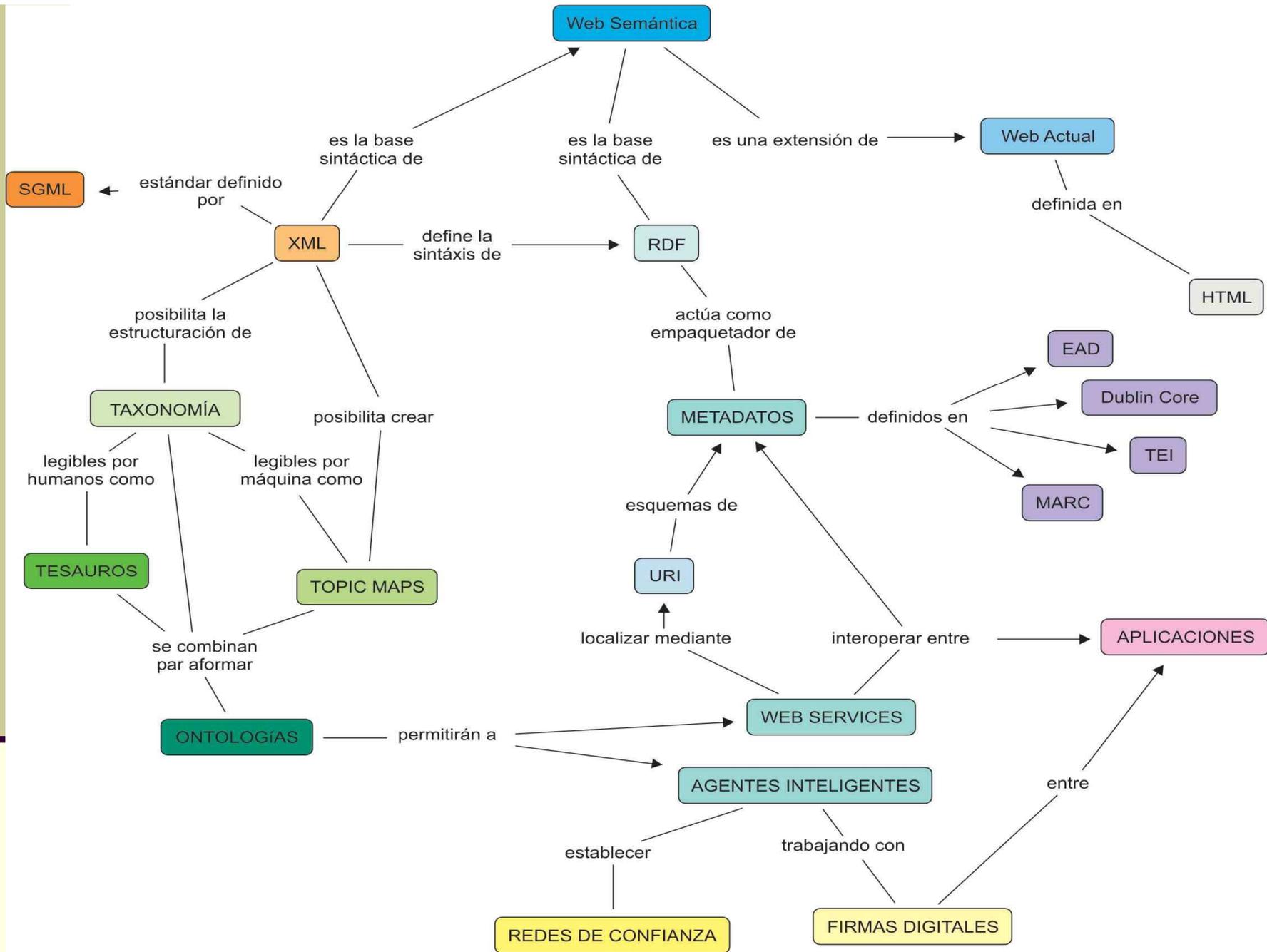
¿Qué es la web semántica?

“La **Web Semántica** es una **Web extendida**, 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**.

Al dotar a la Web de más **significado** y, por lo tanto, de **más semántica**, se pueden obtener soluciones a problemas habituales en la búsqueda de información gracias a la utilización de una infraestructura común, mediante la cual, es posible *compartir*, *procesar* y *transferir* información de forma sencilla.

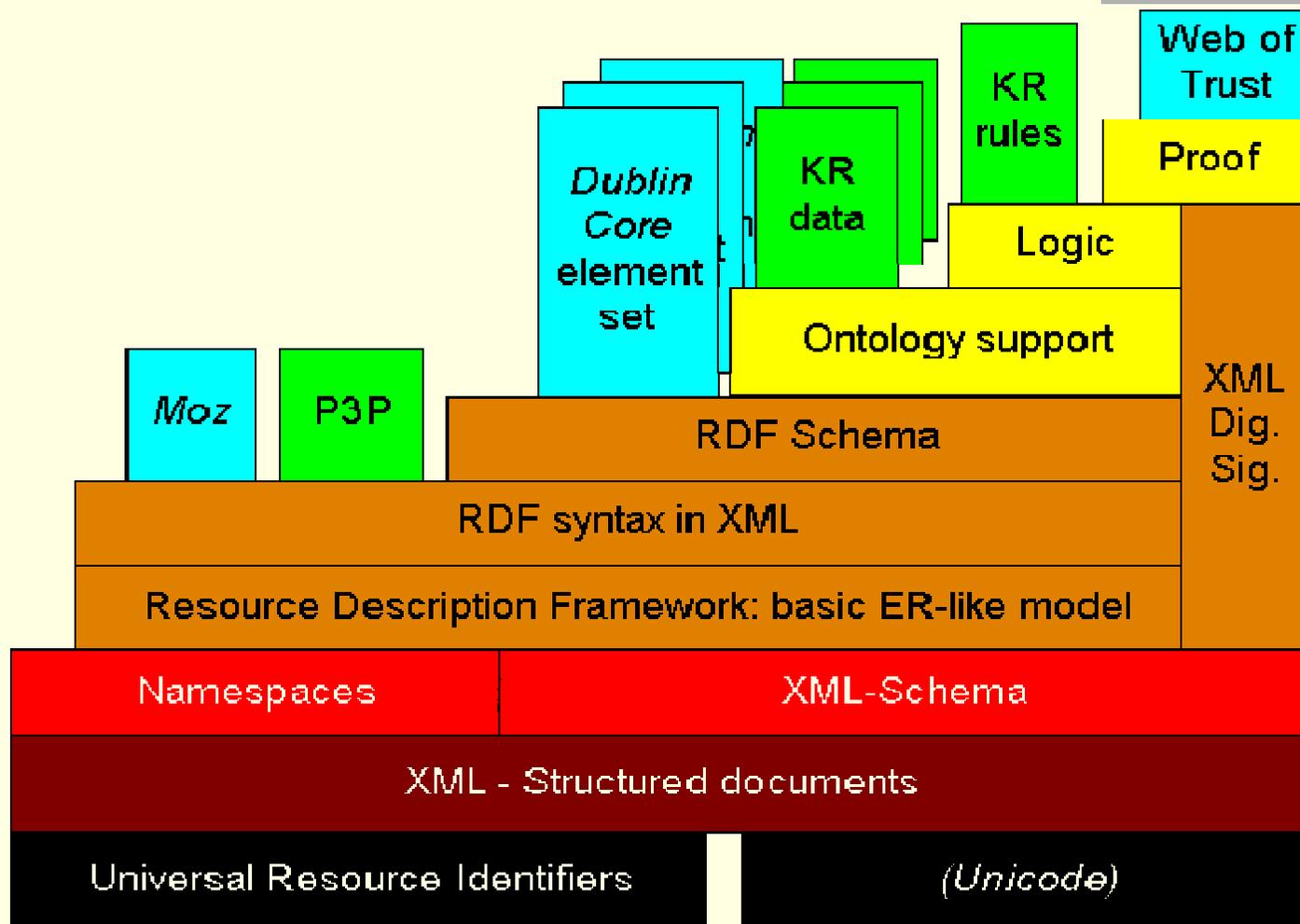
Esta Web extendida y basada en el significado, se apoya en **lenguajes universales** que resuelven los problemas ocasionados por una Web carente de semántica en la que, en ocasiones, el acceso a la información se convierte en una tarea difícil y frustrante.”

<http://www.w3c.es/divulgaci3n/guiasbreves/WebSemantica>



Fuente: "Web Semántica: un nuevo enfoque para la organización y recuperación de información en la web". http://bvs.sld.cu/revistas/aci/vol13_6_05/aci030605.htm

Arquitectura de la Web Semántica



Ontologías

“es una especificación explícita y formal de una conceptualización ”Gruber (1993)

- es un sistema de representación del conocimiento en un dominio o ámbito del conocimiento.
- está formada por una taxonomía relacional de conceptos y por un conjunto de axiomas o reglas de inferencia mediante los cuales se podrá inferir un nuevo conocimiento.
- comparte la comprensión común de la estructura de información entre personas o agentes de software

¿Qué rol juegan las ontologías ?

- Permiten la descripción semántica de recursos para una recuperación más eficiente.
- Permiten una comunicación e interacción con los usuarios.
- Permiten navegar y descubrir nuevos conocimientos.
- Permiten interconexión con otras bases de datos heterogeneas.

¿Por qué la aplicación de herramientas y tecnologías de la web semántica en las bibliotecas digitales?

- Porque potencian las posibilidades de búsquedas mediante el descubrimiento y la navegación por los datos de los recursos,
- Porque permiten utilizar conceptos bien definidos e inferir datos implícitos a partir de descripciones bibliográficas,
- Porque posibilitan la interoperabilidad entre distintos sistemas y usuarios, describiendo y organizando los recursos en un lenguaje comprensible por las computadoras o agentes.

¿Qué ejemplos existen de aplicaciones de tecnologías web en bibliotecas digitales?

- Iniciativa MarcOnt prototipo utilizado en la Biblioteca Digital Jerome, del Digital Enterprise Research Institute (DERI) y la Gdansk University of Technology (GUT).
- Modelo FRBR utilizado en Greenstone 3 , New Zealand Digital Library Project en la Universidad de Waikato.(desarrollado y distribuido en cooperación con la Unesco y la ONG Human Info, en Bélgica)

Iniciativa MarcOnt prototipo utilizado en la Biblioteca Digital Jerome

The screenshot shows a web browser window titled "JeromeDL - Iceweasel". The address bar displays "http://bleedingedge.jeromedl.org/?lang=es_ES". The browser's menu bar includes "Archivo", "Editar", "Ver", "Historial", "Marcadores", "Herramientas", and "Ayuda". The page content features a header with "v2.0" logos and a search bar with "Idioma" set to "es_ES" and a "Buscar" button. The main heading reads "JEROMEDL E-BIBLIOTECA CON SEMANTICA". Below this is a dark banner with the text "Haciendo Bibliotecas Digital real". The main content area is divided into four columns: "Navegar en la biblioteca" (listing categories like "Informes técnicos(22)", "En libros(4)", "Tesis(18)", "Diversos(236)", "Artículos(34)", "Libros(9)", "Memorias de conferencias(8)", "En colecciones(1)", "En memorias de conferencias(245)", "Remitibles(61)"); "Browse Collections" (listing "undefined (0)", "publications from corrib (0)", "blabla (0)"); "Búsqueda de recursos" (listing "Búsqueda simple", "Búsqueda avanzada", "Búsqueda semántica", "SQE search"); and "Investigación en DL" (listing "Página del proyecto JeromeDL", "Página de la iniciativa MarcOnt", "Página de DERI.Galway", "Página de la universidad Politécnica de Gdansk, Polonia"). A right sidebar contains "JeromeDL abreviatura", "Perfil del lector" (with "Login" and "Registro" links), and "Sindicar" options for "RSS 1.0", "RSS 2.0", and "ATOM FEED". The footer contains copyright information: "(C) Copyright 2003-2006 by Digital Enterprise Research Institute (DERI) and WET1 & Main Library Gdansk University of Technology Poland and Sebastian Ryszard Kruk. Todos los derechos reservados." The Windows taskbar at the bottom shows the system tray with "Listo" and the time "10:09".

Jerome DL : interfaz de búsqueda en lenguaje natural con plantillas

The screenshot displays the JeromeDL search interface in a web browser. The browser's address bar shows the URL: `http://bleedingedge.jeromedl.org/search?lang=es_ES&type=semantic&donotexecute=false&rdfqltype=nlq&nqlc`. The page title is "JeromeDL - Búsqueda de recursos : Búsqueda semántica - Iceweasel".

The main content area is titled "Resultados de búsqueda de recursos RDF y interfaz NL". It features a search input field with the text "Show me all publications written by Sebastian Kruk...". Above the input field, there are tabs for "Consulta de Lengua Natural" and "Consulta RDF". Below the input field, there are buttons for "Buscar" and "Limpiar el formulario".

The search results are displayed below the input field, showing a list of publications. The first result is by Sebastian Ryszard Kruk, Sławomir Grzonkowski, Adam Gzella, and Mariusz Cygan, titled "DigiMe - Ubiquitous Search and Browsing for Digital Libraries". Other results include publications by Sebastian Ryszard Kruk, Bernhard Haslhofer, Piotr Piotrowski, Adam Westerski, Tomasz Woroniecki, Sebastian Ryszard Kruk, Adam Gzella, Filip Czaja, Władysław Bultrowicz, Ewelina Kruk, Wit Novacek, Siegfried Handschuh, Diana Maynard, Loredana Laera, Sebastian Ryszard Kruk, Max Völkel, Tudor Groza, and Valentina Tamma, and Sebastian Ryszard Kruk, Kerstin Zimmermann, Brahmananda Sapkota, Hee-Chul Choi, Sebastian Ryszard Kruk, Sławomir Grzonkowski, and Stankiewicz.

The left sidebar contains navigation options such as "Navegar en la biblioteca", "Browse Collections", and "Búsqueda de recursos". The right sidebar contains options for "JeromeDL abreviatura", "Login", "Registro", "Mostrar metadata sin formatear", and "Sindicar".

Jerome DL: búsqueda semántica

The screenshot shows a web browser window with the URL `http://bleedingedge.jeromedl.org/exhibit/index.jsp?state=search&ref=http%3A%2F%2Fbleedingedge.jeromedl.org`. The page title is "JeromeDL - JeromeDL with Exhibit {Select only marcont in facet Palabras claves:} - Iceweasel".

The main content area displays "Resultados de búsqueda de recursos RDF y interfaz NL" (Results of search for RDF resources and NL interface) under the heading "(Browsing with Exhibit)". It shows 2 books filtered from 60 originally. The search results are sorted by "bookType: luego por..." and grouped by order.

The search results list two items:

- 1. *The Role of Ontologies in Semantic Digital Libraries***
Author: Sebastian Ryszard Kruk, Bernhard Haslhofer, Piotr Plotowski, Adam Westerski, and Tomasz Woroniecki
Keywords: ontology, foaf:realm, marcont, digital libraries, and semantic web
Event: Related to: The Role of Ontologies in Semantic Digital Libraries
Conference: The 10th European Conference on Research and Advanced Technology for Digital Libraries
- 2. *Collaborative Ontology Development with MarcOnt Portal***
Author: Sebastian Ryszard Kruk and Maciej Dąbrowski
Keywords: marcont portal, corrib, marcont, ontology, and collaboration
Event: Related to: Collaborative Ontology Development with MarcOnt Portal
Conference: Semantic Technology Conference 2007

The left sidebar contains navigation options like "Navegar en la biblioteca" and "Browse Collections". The right sidebar shows the "Exhibit filter panel" with "Palabras claves:" (learning objects, library, marc21 bibtext, marcont) and "Categories:" (missing this field).

Jerome DL: resultado de una búsqueda

The screenshot shows a web browser window with the address bar displaying `http://bleedingedge.jeromedl.org/resource/zjrwqqa`. The page title is "JeromeDL - SSCF Evaluation Summary - Iceweasel". The browser's menu bar includes "Archivo", "Editar", "Ver", "Historial", "Marcadores", "Herramientas", and "Ayuda".

The website interface features a navigation sidebar on the left with sections: "Browse e-Library" (listing various document types like reports, books, and articles), "Browse Collections" (with sub-sections for "publications from corrib" and "biabla"), "Semantic browse" (with options for navigation and exhibit), "Search for Resources" (with simple, advanced, semantic, and SQE search options), and "DL Research" (with a link to the project web page).

The main content area displays search results for the query "Łukasz Porwoł , Sebastian Ryszard Kruk , Adam Gzella :SSCF Evaluation Summary". It includes a thumbnail of a globe with the text "JeromeDL e-Library with Semantics", a detailed abstract of the document, a list of keywords ("collaborative filtering, social networks, bookmarking, tagging, semantic web"), and an additional description with metadata such as date (September 27, 2007), projectName (eLITE), and deliverableCode (D1.5.300). A note indicates the resource has been viewed 136 times.

Below the main result is a "Recommendations" section titled "Overall recommendations" with a table:

| Resource | Percentage |
|---|------------|
| JeromeDL and FOAFRealm - Taking Advantage of Semantic Social Collaborative Filtering in Digital Libraries | 100.00% |
| Semantic Web enabled Web Services | 67.50% |
| Marcont Initiative Research Roadmap Report | 50.00% |
| JeromeDL 2.0.1 User Guide | 50.00% |
| Adaptive Hypermedia and Semantic Web 2.0 | 50.00% |

Keywords for recommendations: "recommendations". Annotations for recommendations: "recommendations".

On the right side, there is a "JeromeDL Shorthand" sidebar with a "Reader's profile" section (Login, Register), a "Show raw metadata" section with various RDF format buttons (RDF, RDF/XML, N3, Turtle, N-Triples, JSON, SIOC, etc.), and a "New resources feeds (global)" section with RSS and ATOM feed buttons.

Jerome DL : Servicio de Mediación MarcOnt

The screenshot shows a web browser window titled "JeromeDL - Iceweasel". The address bar contains the URL: `http://bleedingedge.jeromedl.org/servlet/mms?uri=http://bleedingedge.jeromedl.org/resource/SQFvJ3uu`. The browser tabs include "JeromeDL - Search for Resou...", "JeromeDL", and "IRM - Recursos Infor...".

The main content area displays "JeromeDL: MarcOnt Mediation Services". Below this title is a navigation bar with buttons for "BibTeX", "BibTeXXML", "BibTeXRDF", "DublinCore", "MARC21", "MARC-XML", "MARC-RDF", and "MarcOnt" (which is highlighted in green).

The main content area shows an RDF/XML snippet:

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:structure="http://www.jeromedl.org/structure#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:ns19="http://jonto.corrib.org/ontology#"
  xmlns:xmarcont="http://www.marcont.org/xontology#"
  xmlns:xfoaf="http://www.foafrealm.org/xfoaf/0.1/"
  xmlns:marcont="http://www.marcont.org/ontology#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  <rdf:Description rdf:about="http://jonto.corrib.org/i18n_wordnet#d9b0c031e08c4b8de9ee511fde46109425cb7c48">
    <rdf:type rdf:resource="http://www.marcont.org/ontology#Keyword" />
    <ns19:hasWord xml:lang="en">semantic web</ns19:hasWord>
    <rdfs:label xml:lang="en">semantic web</rdfs:label>
  </rdf:Description>
  <rdf:Description rdf:nodeID="node134ktq6m4x1572">
    <rdf:subject rdf:resource="http://bleedingedge.jeromedl.org/resource/SQFvJ3uu" />
    <rdf:predicate rdf:resource="http://www.marcont.org/ontology#hasCreator" />
    <rdf:object rdf:resource="mailto:tomasz.woroniecki@deri.org" />
    <rdf:type rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#Statement" />
    <marcont:order rdf:datatype="http://www.w3.org/2001/XMLSchema#integer">5</marcont:order>
  </rdf:Description>
  <rdf:Description rdf:about="http://jonto.corrib.org/i18n_wordnet#db1bb442d5ad458ba49bd81d354e6973b1682a65">
    <rdf:type rdf:resource="http://www.marcont.org/ontology#Keyword" />
    <ns19:hasWord xml:lang="en">digital libraries</ns19:hasWord>
    <rdfs:label xml:lang="en">digital libraries</rdfs:label>
  </rdf:Description>
  <rdf:Description rdf:about="mailto:bernhard.haslhofer@univie.ac.at">
    <rdf:type rdf:resource="http://xmlns.com/foaf/0.1/Person" />
    <xfoaf:timestamp>1180862057238</xfoaf:timestamp>
    <foaf:mbox rdf:resource="mailto:bernhard.haslhofer@univie.ac.at" />
    <foaf:mbox_sha1sum>0bbdab671f03f52ff81ba058c09da405be437f24</foaf:mbox_sha1sum>
    <foaf:name>Bernhard Haslhofer</foaf:name>
    <foaf:family_name />
  </rdf:Description>
</rdf:RDF>
```

At the bottom of the browser window, the taskbar shows the system tray with the time "10:23" and the text "Listo" on the left.

Greenstone 3 y FRBR

- Otra aplicación es el modelo FRBR en **Greenstone 3**, en etapa de investigación está desarrollando una nueva plataforma para bibliotecas digitales semánticas.
- Este cambio tendrá profundos efectos en la recuperación del contenido para los usuarios finales e introducirá también un impacto significativo en la arquitectura de la biblioteca digital actual.
- Se establece una comparación para mostrar que a partir de las jerarquías semánticas del modelo FRBR es posible establecer un perfil de servicio de alerta de acuerdo a necesidades de los usuarios.

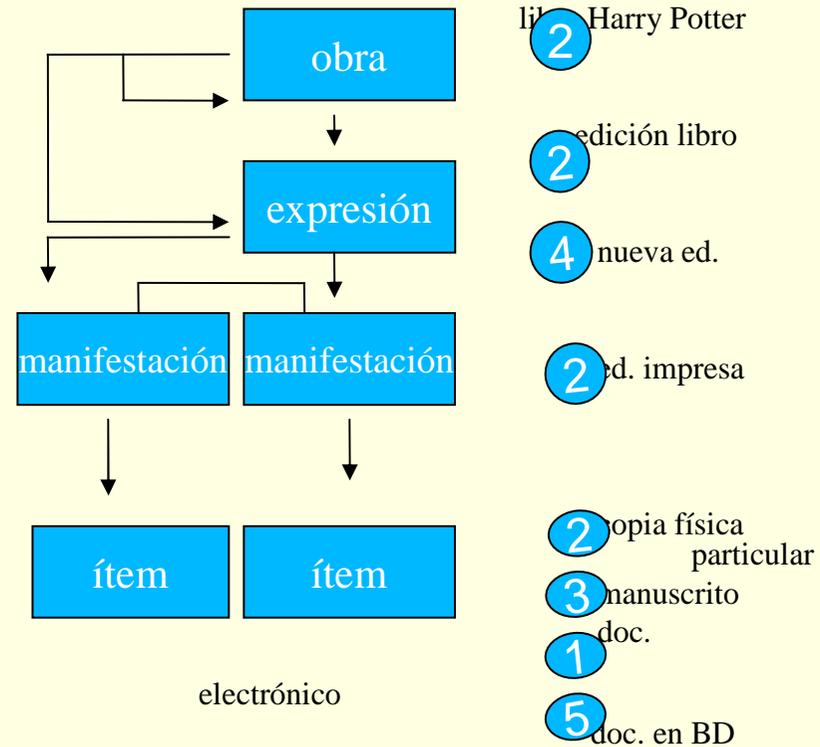
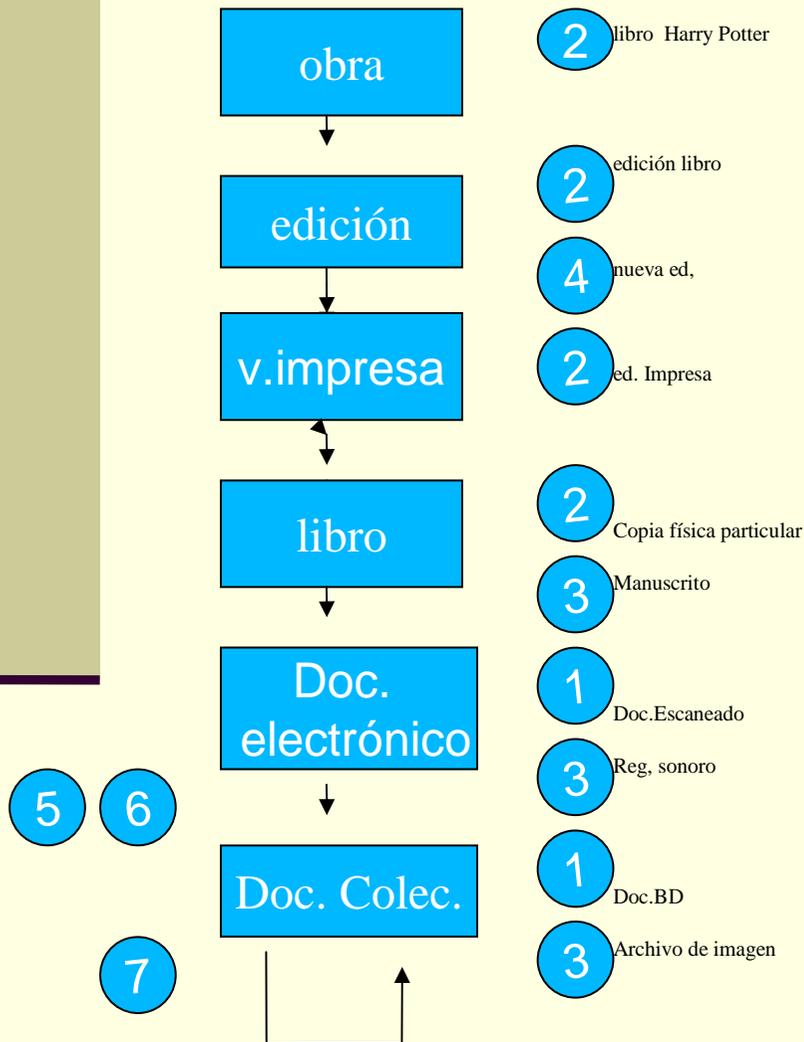
Greenstone 3 : Visualización perfil servicio de alerta

Acciones básicas consideradas para la creación de perfiles de usuarios fueron las siguientes:

Un usuario puede querer ser notificado cuando :

1. Una vez que un nuevo documento electrónico se ha incorporado y se encuentra disponible en la colección de la biblioteca digital.
2. Una nueva obra ha sido publicada.
3. Un manuscrito ha sido nuevamente digitalizado , en una resolución mayor, o en un formato diferente.
4. Otra edición del mismo libro se publica.
5. Un documento electrónico ha sido recientemente publicado en la biblioteca digital.
6. Un documento se elimina de la biblioteca digital.
7. Un documento electrónico ha cambiado

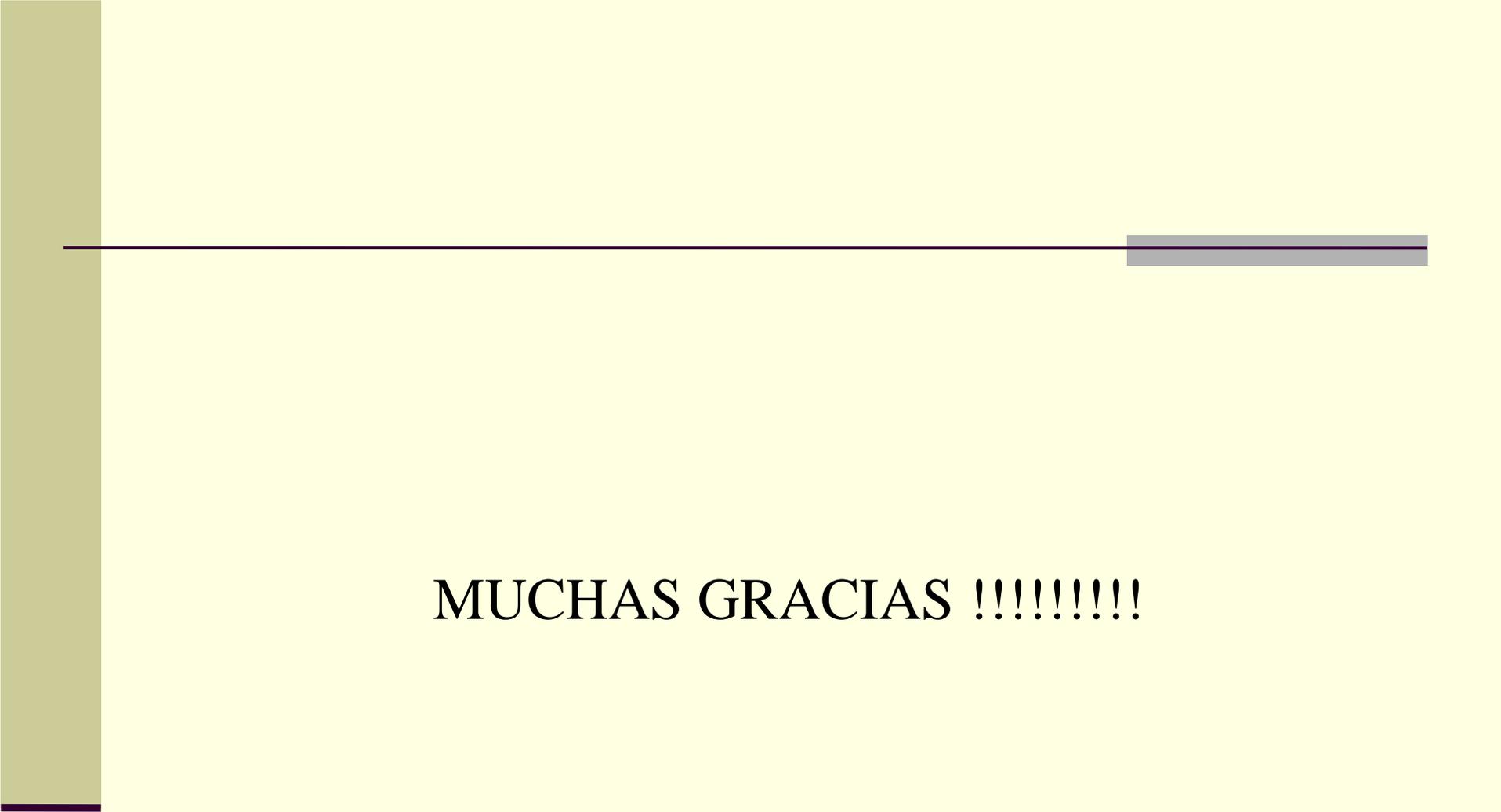
Visualización Servicio de Alerta en Greenstone 3 y FRBR



Por qué es necesario que los bibliotecarios se interesen y utilicen estas tecnologías?

Porque estas tecnologías permitirán :

- normalizar los atributos de los metadatos aplicables a los documentos
- crear una red de relaciones que aporte especificación y fiabilidad
- trabajo cooperativo al funcionar como soporte de conocimiento entre las organizaciones
- integración de diferentes perspectivas de usuarios ,y la interoperatividad entre sistemas distintos
- construcción automatizada de mapas conceptuales y temáticos , tratamiento ponderado del conocimiento para recuperar información de forma automatizada ,
- establecer modelos normativos que permitan la creación de la semántica de un sistema y un modelo para poder extenderlo y transformarlo entre diferentes contextos
- servir de base para la construcción de lenguajes de representación del conocimiento ,
- mejorar la recuperación y organización de la información, como así también compartir conocimiento con la comunidad de usuarios.



MUCHAS GRACIAS !!!!!!!!