



---

# Aplicaciones de la web semántica

---

**Diego Berrueta**  
**[diego.berrueta@fundacionctic.org](mailto:diego.berrueta@fundacionctic.org)**  
**Oviedo, 15/Junio/07**





# Contenidos

- 1) Introducción
- 2) Vocabularios comunes: FOAF, DOAP, RSS, Dublin Core, SIOC, EARL, SKOS
- 3) El huevo y la gallina: D2R, GRDDL, RDFa y microformatos, DBpedia
- 4) Navegadores semánticos: Tabulator, Piggy Bank
- 5) Otras piezas del puzzle: Semantic Web Client Library, PingTheSemanticWeb
- 6) Computación humana y web semántica (ESP Game)



# Contenidos

- 1) Introducción
- 2) Vocabularios comunes: FOAF, DOAP, RSS, Dublin Core, SIOC, EARL, SKOS
- 3) El huevo y la gallina: D2R, GRDDL, RDFa y microformatos, DBpedia
- 4) Navegadores semánticos: Tabulator, Piggy Bank
- 5) Otras piezas del puzzle: Semantic Web Client Library, PingTheSemanticWeb
- 6) Computación humana y web semántica (ESP Game)



# Web semántica, ¿para qué?

- ¿Quién necesita a la web semántica teniendo Google?
- ¿Quién necesita servicios web semánticos teniendo Yahoo! Pipes?
- ¿Dónde estamos y a dónde queremos llegar?



# Web semántica: motivación



Fuente: Amber Mac, <http://www.flickr.com/photos/ambermac/249272580/>



# Web semántica: realidad o ilusión

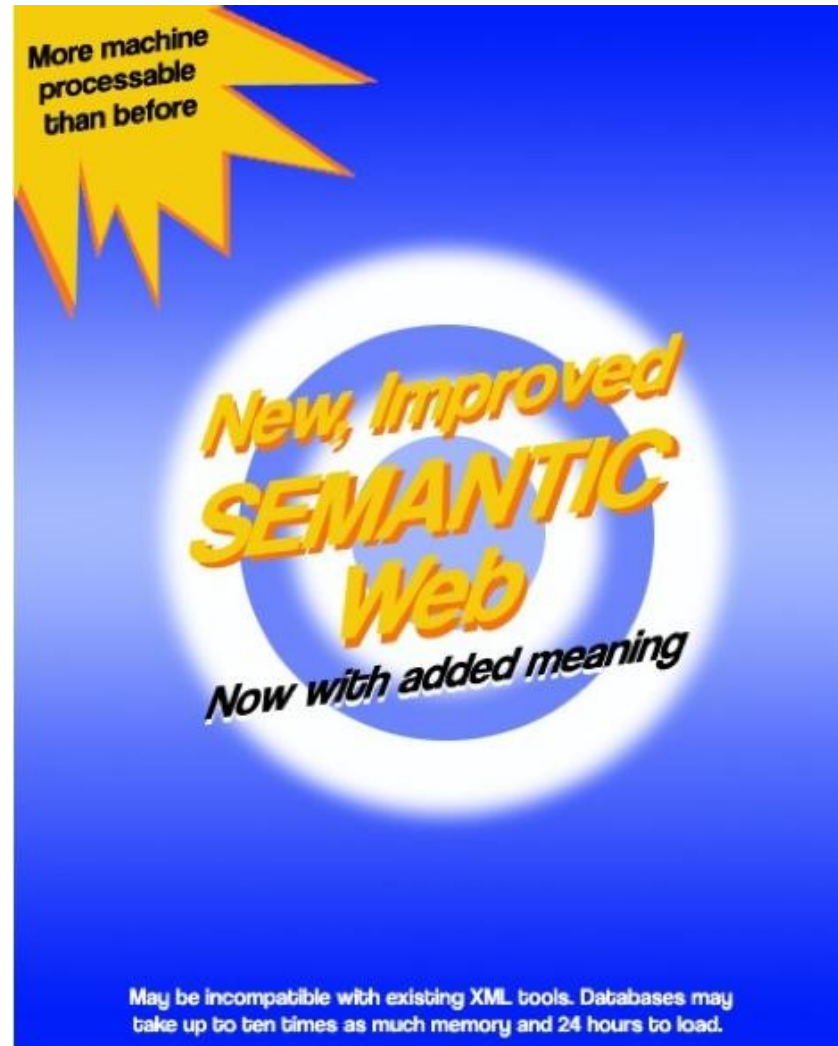
## BEWARE OF THE HYPE!

¿es la web semántica la  
Inteligencia Artificial del siglo XXI?

Fuente: Bernardo Cuenca, curso de verano, Uniovi, 2006



# Web semántica lava más blanco



Fuente: Mark Butler, <http://www.flickr.com/photos/dullhunk/303503677/>



# Contenidos

- 1) Introducción
- 2) Vocabularios comunes: FOAF, DOAP, RSS, Dublin Core, SIOC, EARL, SKOS
- 3) El huevo y la gallina: D2R, GRDDL, RDFa y microformatos, DBpedia
- 4) Navegadores semánticos: Tabulator, Piggy Bank
- 5) Otras piezas del puzzle: Semantic Web Client Library, PingTheSemanticWeb
- 6) Computación humana y web semántica (ESP Game)

(Coautor de la Sección 2: Luis Polo)

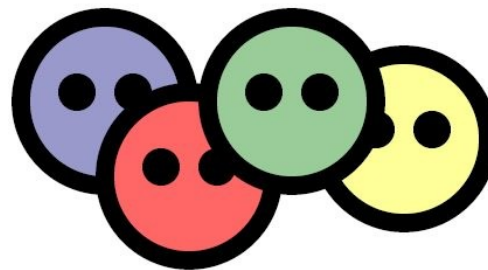


# Vocabularios comunes

- Son vocabularios RDF
- Creados para satisfacer una necesidad concreta:
  - **FOAF**: descripción de personas
  - **DOAP**: descripción de proyectos
  - **RSS**: publicación de noticias
  - **Dublin Core**: catalogación y descripción de recursos (libros, imágenes, etc.)
  - **SIOC**: descripción de comunidades online
  - **EARL**: informes de resultados de pruebas
  - **SKOS**: representación de tesauros
- Sólo EARL y SKOS nacieron en el W3C
- Pueden reutilizarse y extenderse



# FOAF





# FOAF: objetivos

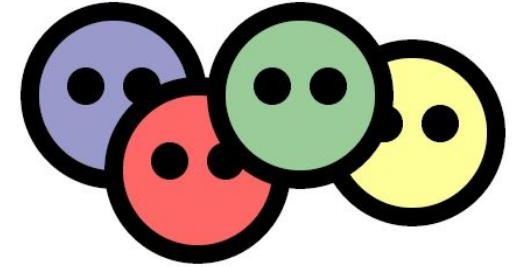
## Objetivos:

- Publicar información “personal” en la web semántica
- Tejer una red de conocidos que sea un reflejo de las relaciones entre las personas en el mundo real

## Utilidad y aplicación:

- Páginas web personales
- Redes de confianza
- Visibilidad de proyectos y actividades
- Conexión con otros recursos de la red (cuentas de correo, cuentas en páginas web, canales de IRC...)

# FOAF: conceptos fundamentales



- Concepto básico: `<foaf:Person>`
- Propiedades simples: *name*, *title*, *givenName*, *familyName*, *nick*
- Enlaces web: *depiction* (foto), *homepage*, *workplaceHomepage*
- Propiedad inv. funcional / identificador de usuario: *mbox* (dirección de correo)
  - Para evitar el spam, se suele usar *mbox\_sha1sum*
- Relaciones entre personas: `<foaf:knows>`
  - Se identifican mediante el *mbox* / *mbox\_sha1sum*
  - Importante: enlazar el fichero FOAF con `<rdfs:seeAlso>`



# FOAF: ejemplo

```
<rdf:RDF [...]>

<foaf:Person rdf:nodeID="me">
  <foaf:name>Diego Berrueta Muñoz</foaf:name>
  <foaf:nick>berrueta</foaf:nick>
  <foaf:mbox_sha1sum>98a99390f2fe9....3f50e59a5ecb</foaf:mbox_sha1sum>
  <foaf:homepage rdf:resource="http://www.berrueta.net"/>

  <foaf:knows>
    <foaf:Person>
      <foaf:name>Jose Emilio Labra</foaf:name>
      <foaf:mbox_sha1sum>8114083efd....d9906080ae89</foaf:mbox_sha1sum>
      <rdfs:seeAlso rdf:resource="http://di002.edv.uniovi.es/~labra/labraFoaf.rdf"/>
    </foaf:Person>
  </foaf:knows>

</foaf:Person>

</rdf:RDF>
```



# FOAF: herramientas

## Visualizadores:

- FOAF Explorer (FOAF → HTML)
- FOAFnaut (FOAF → SVG)
- FOAF-nav (Iván Frade)

## Otros:

- FOAFbot (bot para redes IRC):
  - `<edd> foafbot: picture of edd and yoz`
  - `<foafbot> Picture at http://matthau.yoz.com/cam/20020609/bots_talk4.jpg -- according to Edd Dumbill`
  - `<foafbot> Caption: From left to right: Edd, Yoz, Matt Webb, Jo Walsh, Celia Romaniuk`
- FOAF-o-matic: formulario para crear un fichero FOAF rápidamente
- FOAF map: representación de personas sobre un mapa
- FOAFing the music



# DOAP



# DOAP: descripción

- Vocabulario para describir proyectos [de software]
- Útil para organizar proyectos (ejemplo: Apache)
- Herramientas: navegador, validador, extensión de FOAF, asistentes para crear ficheros DOAP, conversores (SF.net, freshmeat.net), integración con anotaciones Java5



# DOAP: ejemplo

```
<doap:Project rdf:about="http://swaml.berlios.de/doap.rdf">
  <doap:name>Semantic Web Archive of Mailing Lists</doap:name>
  <doap:shortname>SWAML</doap:shortname>
  <doap:homepage rdf:resource="http://swaml.berlios.de/">
  <doap:created>2005-09-24</doap:created>
  <doap:description xml:lang="es">
    SWAML es un proyecto de investigación alrededor de las tecnologías
    de la Web Semántica para publicar los archivos de las listas de correo
    en un formato RDF.
  </doap:description>
  <doap:release>
    <doap:Version>
      <doap:name>SWAML</doap:name>
      <doap:created>2006-12-28</doap:created>
      <doap:revision>0.0.5</doap:revision>
    </doap:Version>
  </doap:release>
  ...
```



**RSS**





# RSS: usos



## Inicialmente:

- Publicar noticias (medios de comunicación)

## Actualmente:

- Blogs, weblogs
- Todo tipo de portales
- Notificaciones activas / alertas
- Modificaciones en contenidos
- Podcasts (multimedia)
- [...]



# RSS: historia

- Creado por Netscape (~1999) para su navegador
- Historial tortuoso de versiones incompatibles:
  - RSS 0.90 (basado en XML)
  - RSS 0.91
  - RSS 0.92
  - RSS 1.0 (basado en RDF)
  - RSS 2.0 (basado en XML)
- Competidor: Atom (basado en XML)
- Acrónimo mutante:
  - Real Simple Syndication
  - Rich Site Summary
  - RDF Site Summary



# RSS: ejemplo (RSS 1.0)

```
<?xml version="1.0" encoding="iso-8859-1"?> <?xml-stylesheet href="/css/rss.css" type="text/css"?>
<rdf:RDF xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-
ns#" xmlns="http://purl.org/rss/1.0/">
  <channel rdf:about="http://www.w3c.es/noticias.rss">
    <title>W3C Oficina Española</title>
    <description>Guiando la Web hacia su máximo potencial...</description>
    <link>http://www.w3c.es</link>
    <dc:date>2005-11-23</dc:date>
    <items> <rdf:Seq>
      <rdf:li rdf:resource="http://www.w3c.es/Noticias/2005/#L2005110774"/>
      <rdf:li rdf:resource="http://www.w3c.es/Noticias/2005/#L2005112382"/>
    </rdf:Seq> </items>
  </channel>
  <item rdf:about="http://www.w3c.es/Noticias/2005/#L2005110774">
    <title>El W3C pone en marcha el Grupo de Trabajo de Formato de Intercambio de Reglas</title>
    <description>2005-11-07:W3C se complace en anunciar la creación del Grupo de Trabajo de
      Formato de Intercambio de Reglas. "Después de [...].</description>
    <link>http://www.w3c.es/Noticias/2005/#L2005110774</link>
    <dc:date>2005-11-07</dc:date>
  </item>
</rdf:RDF>
```



# RSS: herramientas

- Agregadores RSS para el escritorio:
  - Feedreader
  - Liferea
- Agregadores RSS a través de web:
  - <http://www.bloglines.com/>
- Creación de comunidades:
  - PlanetPlanet
- Validador de RSS:
  - <http://feedvalidator.org/>



# Dublin Core





# Dublin Core: introducción

- Objetivo: etiquetar recursos multimedia (bibliotecas)
- Historia: creado en 1995 para estandarizar lenguajes de etiquetado con metadatos
  - Curiosidad: workshop en Dublin, Ohio, USA
- Origen de la iniciativa: bibliotecas convencionales (libros y revistas)
- Desarrollado por la Dublin Core Metadata Initiative (consorcio)

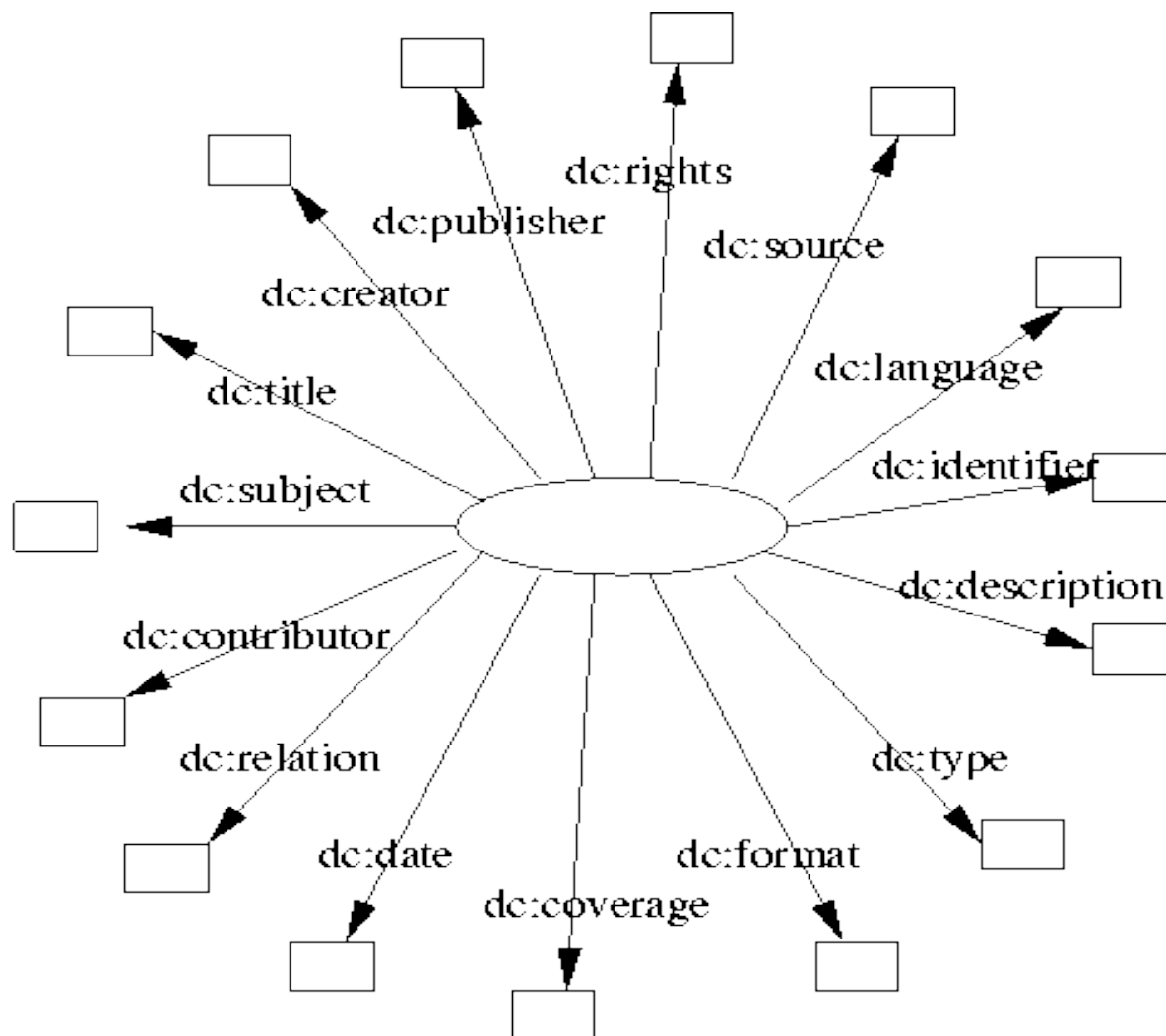




# Definición del Dublin Core

- Estatus del Dublin Core
  - Especificación (2006):  
<http://dublincore.org/documents/dcmi-terms/>
  - ISO Standard 15836-2003 (February 2003):  
<http://www.niso.org/international/SC4/n515.pdf>
  - NISO Standard Z39.85-2001 (September 2001):  
<http://www.niso.org/standards/resources/Z39-85.pdf>
- Formas de expresar Dublin Core
  - HTML/XHTML (elementos *meta*)
  - Sintaxis XML
  - Sintaxis RDF/XML

# Dublin Core: Conceptos Básicos

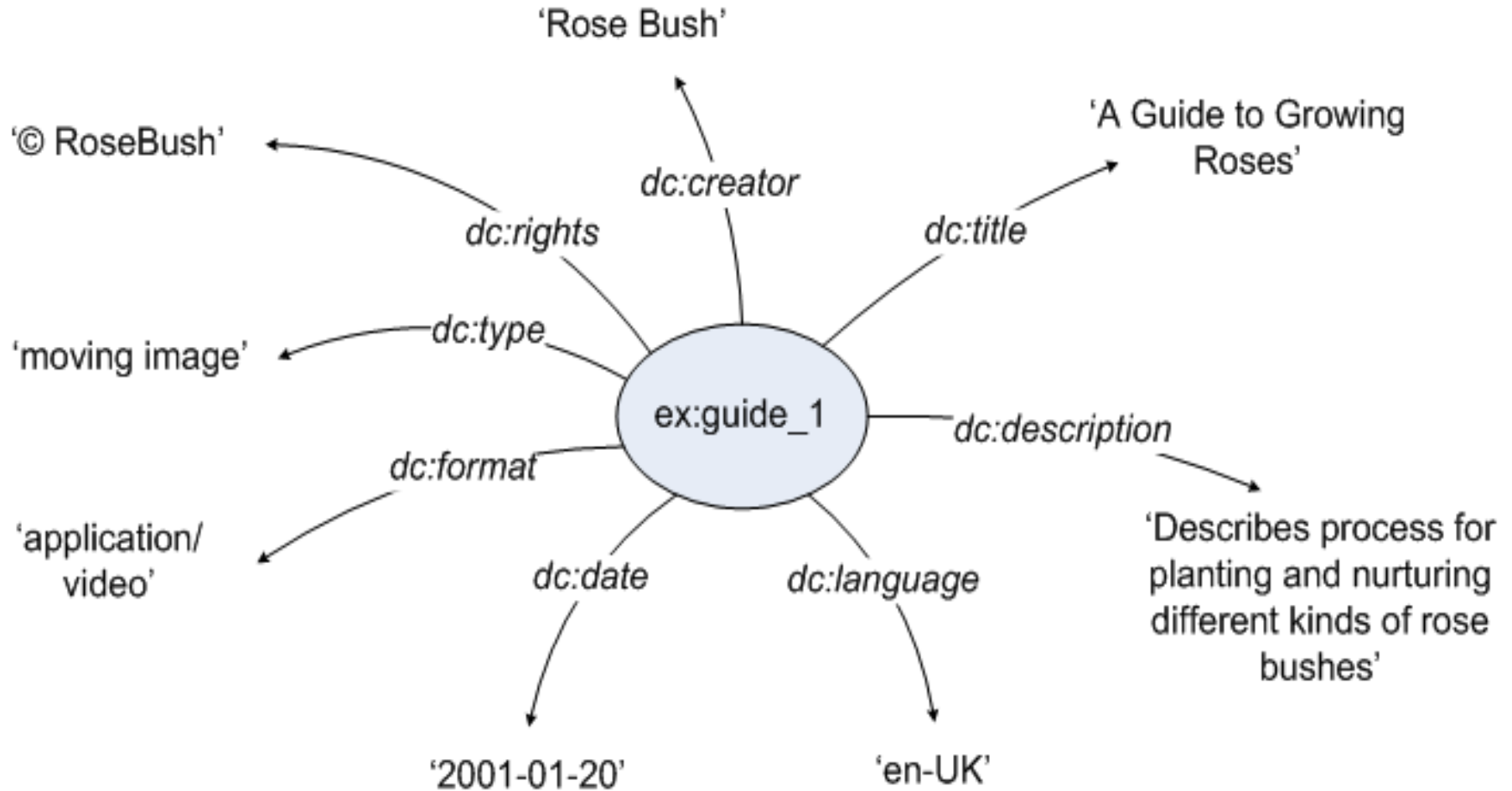




# Dublin Core: ejemplo

```
<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://media.example.com/audio/guide.ra">
    <dc:creator>Rose Bush</dc:creator>
    <dc:title>A Guide to Growing Roses</dc:title>
    <dc:description>Describes process for planting and
      nurturing different kinds of rose bushes.</dc:description>
    <dc:date>2001-01-20</dc:date>
    <dc:language>en-UK</dc:language>
    <dc:rights>(C) Rose Bush</dc:rights>
    <dc:type>moving image</dc:type>
    <dc:format>application/video</dc:format>
  </rdf:Description>
</rdf:RDF>
```

# Dublin Core: ejemplo





# Dublin Core: herramientas

- Asistentes para generar RDF:
  - <http://www.webposible.com/utilidades/dublincore-metadata-gen/>
  - [http://www.webposible.com/utilidades/generador\\_rdf\\_foto.html](http://www.webposible.com/utilidades/generador_rdf_foto.html)
- Licencias Creative Commons (CC):
  - <http://www.creativecommons.org/>
- Firefox Dublin Core Viewer
  - <https://addons.mozilla.org/firefox/528/>
- Herramientas para Dublin Core
  - <http://dublincore.org/tools/>

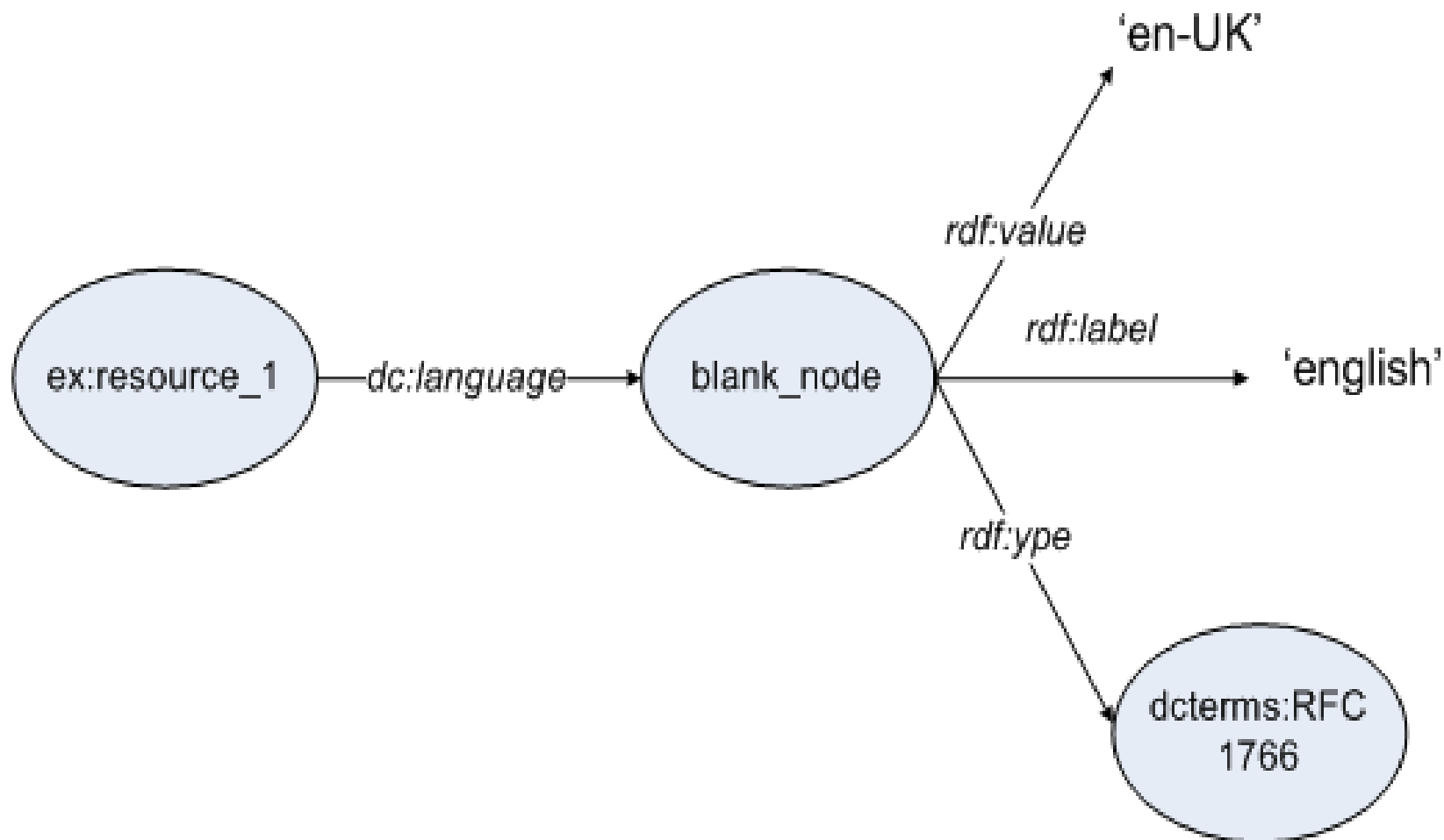




# Refinamiento de Dublin Core

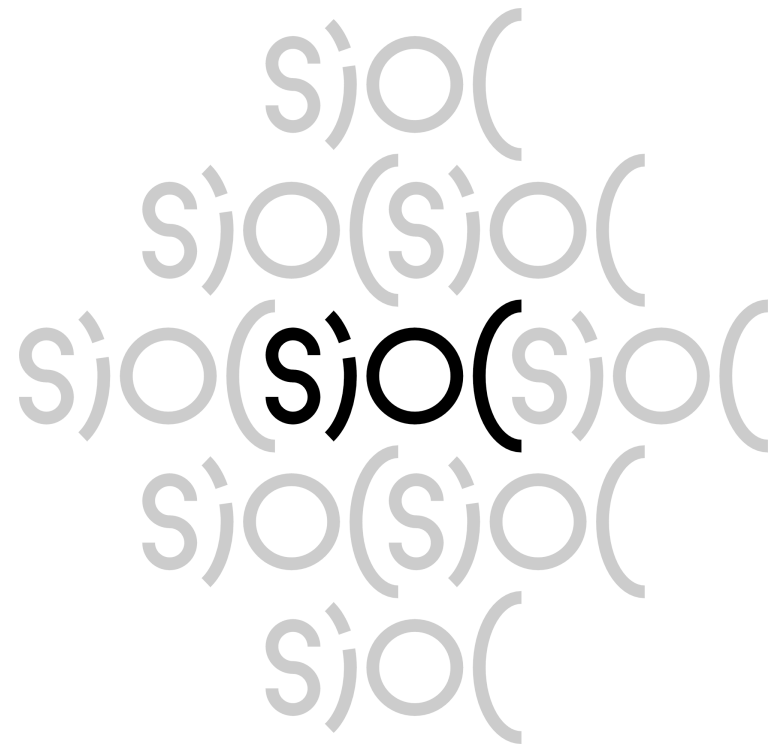
- Elementos de refinamiento de Dublin Core
  - Restringen y especifican el contenido de una etiqueta.
    - `<dc:date>: {<dc:valid>, <dc:created>, <dc:modified>}`
- Tipo de codificación
  - Esquemas que ayudan a la interpretación del valor de un elemento
    - Vocabularios controlados, tesauros, folksonomías, etc.

# Refinamiento DC: ejemplo





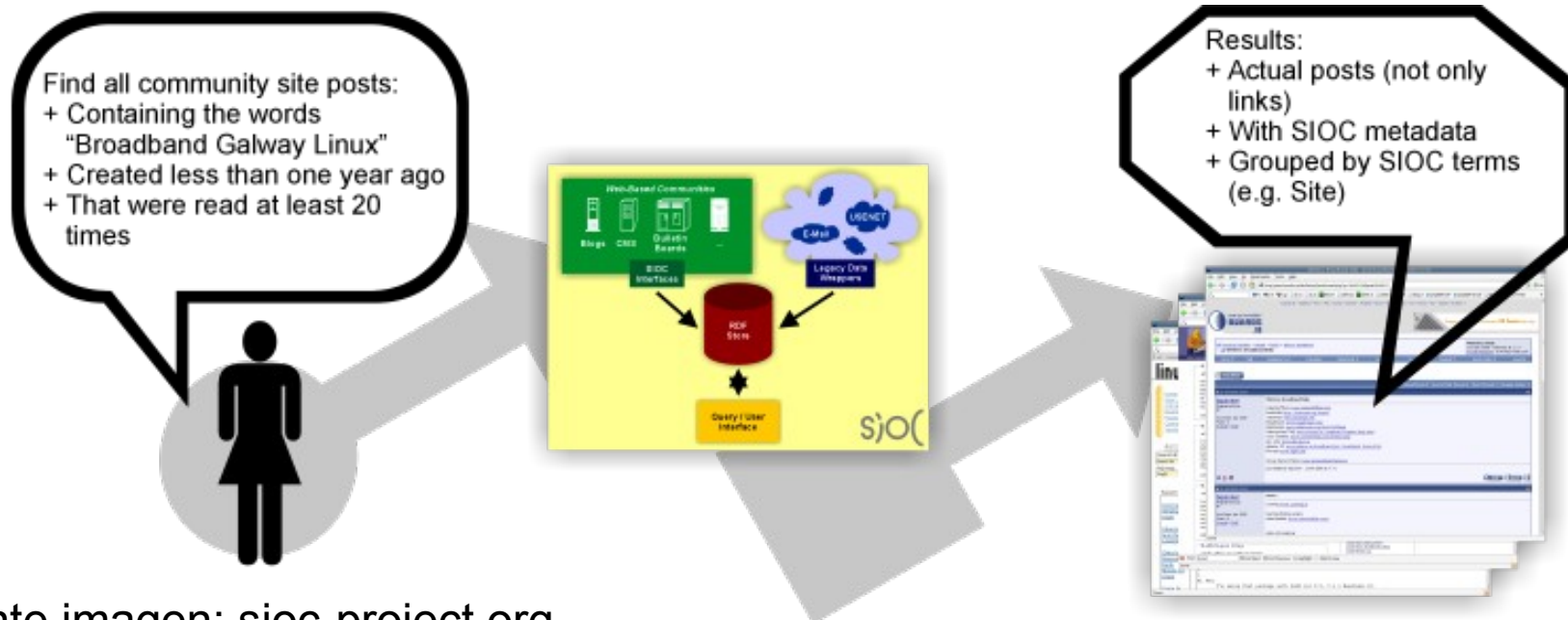
**SIOC**



# SIOC: objetivos

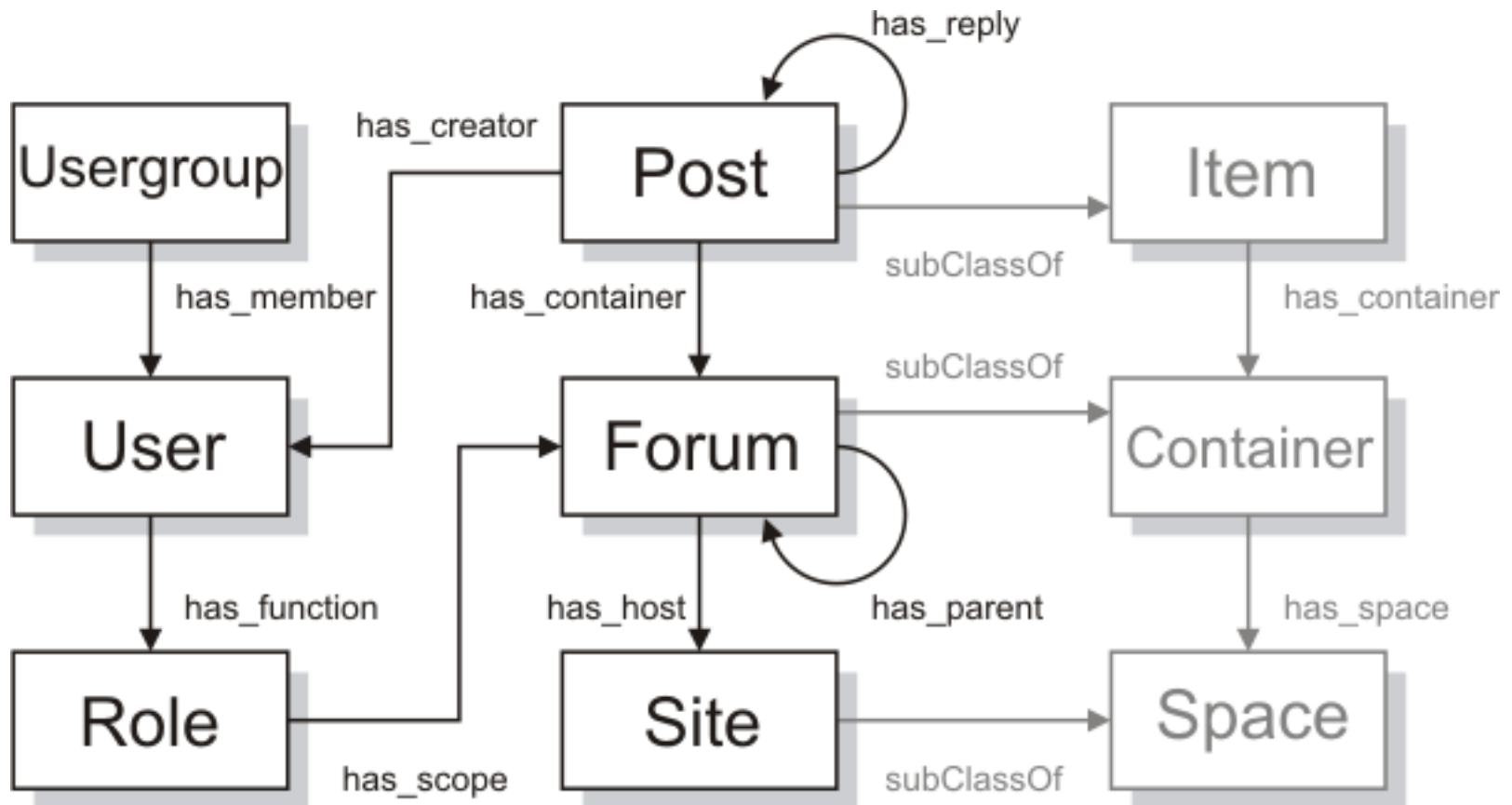
- Describir comunidades virtuales
- Describir la comunicación dentro de estas comunidades (posts en blogs y foros, páginas wiki, listas de correo, grupos de news, etc.)

SIOC  
SIOC(SIOC)  
SIOC(SIOC)  
SIOC(SIOC)  
SIOC(SIOC)  
SIOC(SIOC)



Fuente imagen: [sioc-project.org](http://sioc-project.org)

# SIOC: principales conceptos



Fuente: especificación SIOC, <http://rdfs.org/sioc/spec/>



# SIOC: ejemplo (I)

```
<sioc:Site rdf:about="http://sioc-project.org">
  <sioc:host_of>
    <sioc:Forum rdf:about="http://swaml.berlios.de/demo/index.rdf">
      <sioc:has_subscriber rdf:resource="#s7"/>
    </sioc:Forum>
  </sioc:host_of>
</sioc:Site>

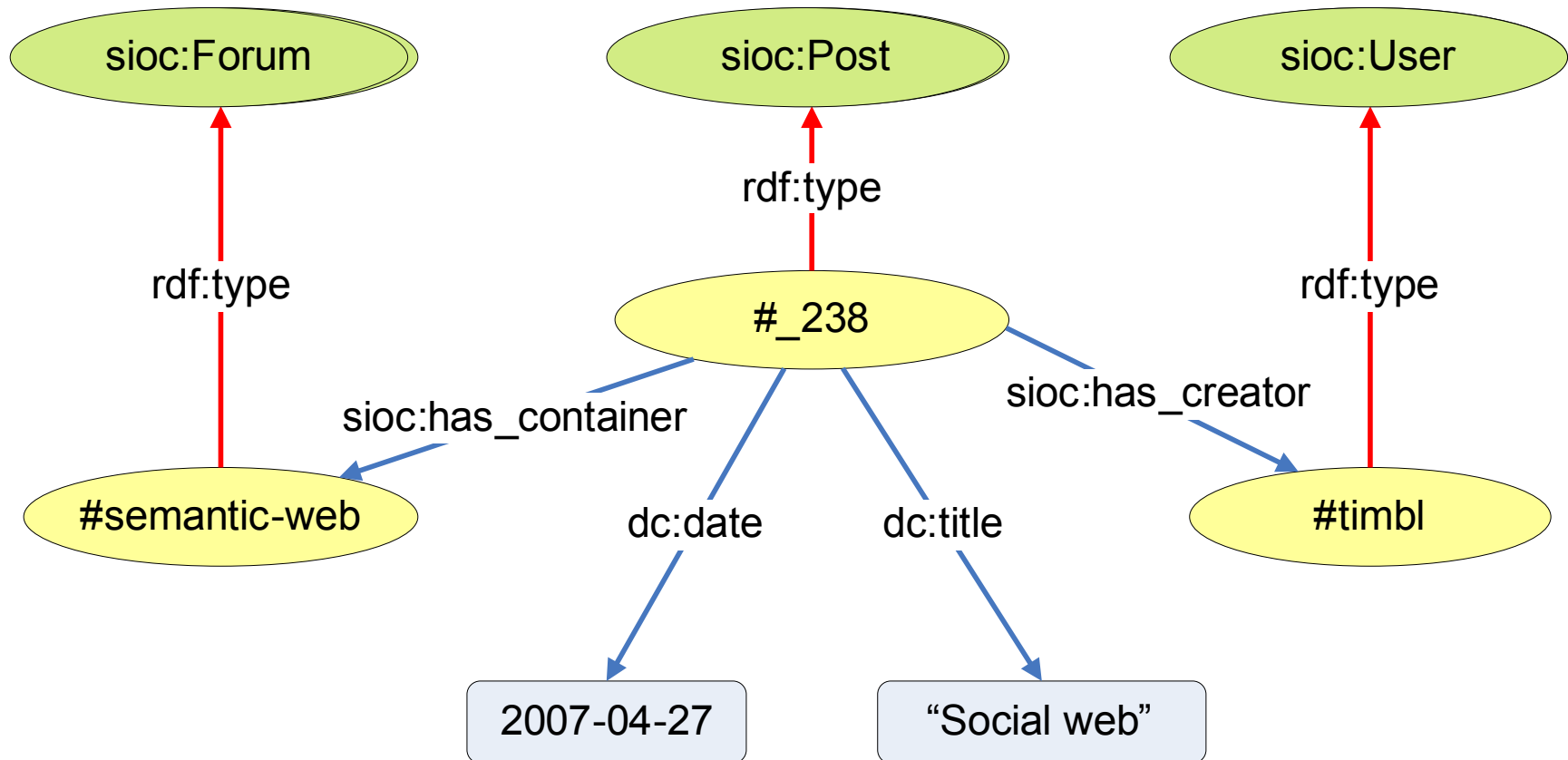
<sioc:User rdf:about="#s7">
  <sioc:name>Axel Polleres</sioc:name>
  <sioc:email_sha1sum>36cf5b9757bdc...02f1eb0</sioc:email_sha1sum>
  <rdfs:seeAlso rdf:resource="http://platon.escet.urjc.es/~axel/foaf.rdf"/>
  <sioc:creator_of rdf:resource="post-27.rdf"/>
</sioc:User>
```



## SIOC: ejemplo (II)

```
<sioc:Post rdf:about="post-7.rdf">
  <sioc:id>b2ea11c9cf8802bc72f82af4c5dbba1d70bb3c86</sioc:id>
  <sioc:has_container rdf:resource="http://swaml.berlios.de/demo/index.rdf"/>
  <dcterms:created>Tue, 19 Sep 2006 16:40:32 +0200</dcterms:created>
  <sioc:reply_of rdf:resource="post-5.rdf"/>
  <dc:title>
    Re: [Fwd: [Expertfinder-dev] Call for ExpertFinder-Telecon]
  </dc:title>
  <sioc:has_creator rdf:resource="#s7"/>
  <sioc:link rdf:resource="http://swaml.berlios.de/demo/2006-Sep/post-7.rdf "/>
</sioc:Post>
```

## SIOC: ejemplo (III)





# SIOC: aplicaciones

## Exportación de contenidos en SIOC:

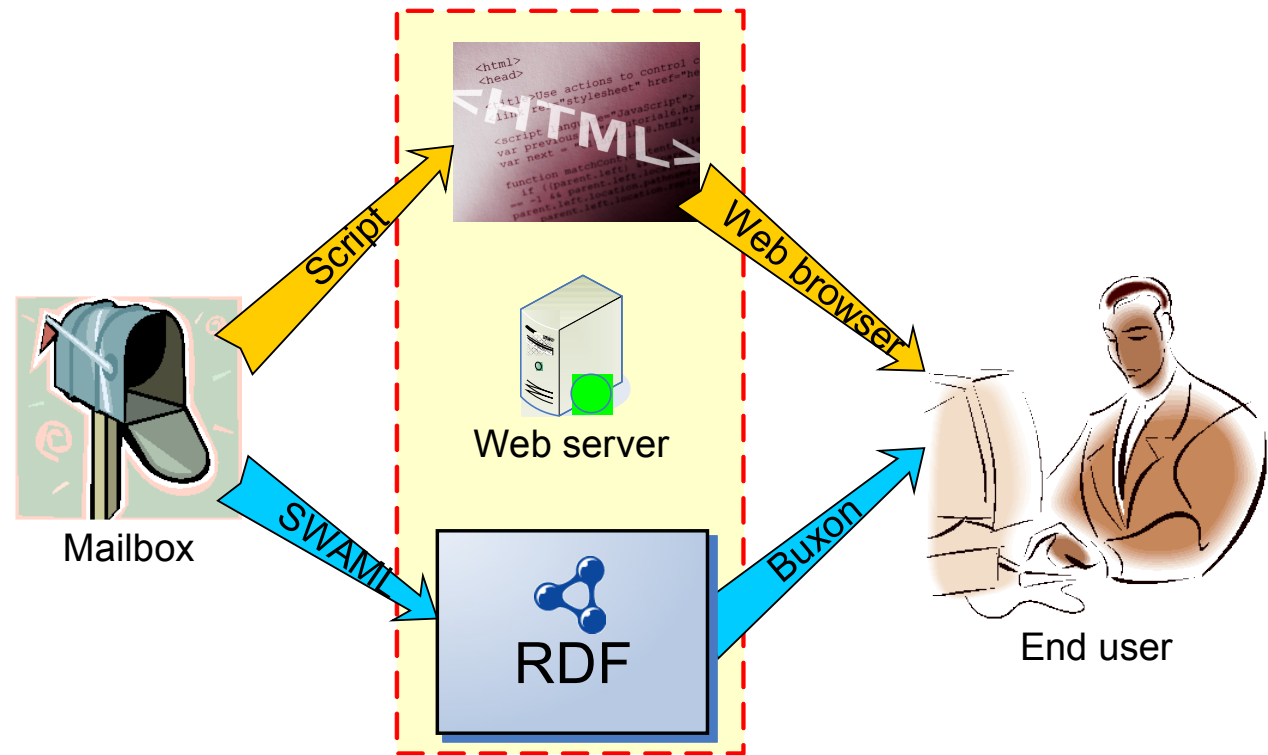
- Gestores de contenidos: Drupal, WordPress, phpBB
- Wikis
- Listas de correo: SWAML (Sergio Fernández, PFC EUITIO)

## Aplicaciones:

- Navegadores: SIOC-explorer, SIOC-browser, Buxon (S. Fernández)
- Extensión para Firefox
- Consultas con SPARQL

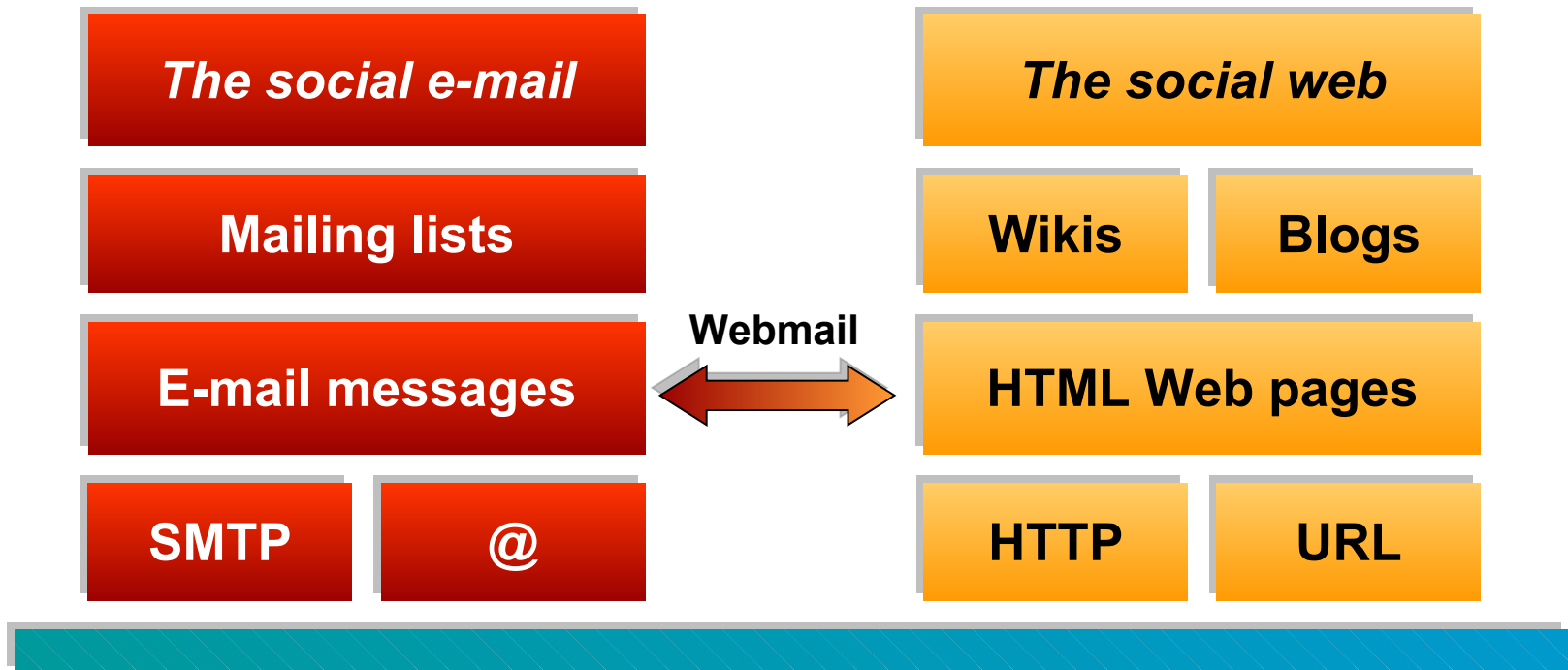
# SWAML/Buxon

- SWAML: convierte archivos de correo (mbox) en documentos RDF usando SIOC
- Buxon: navegador de documentos SIOC



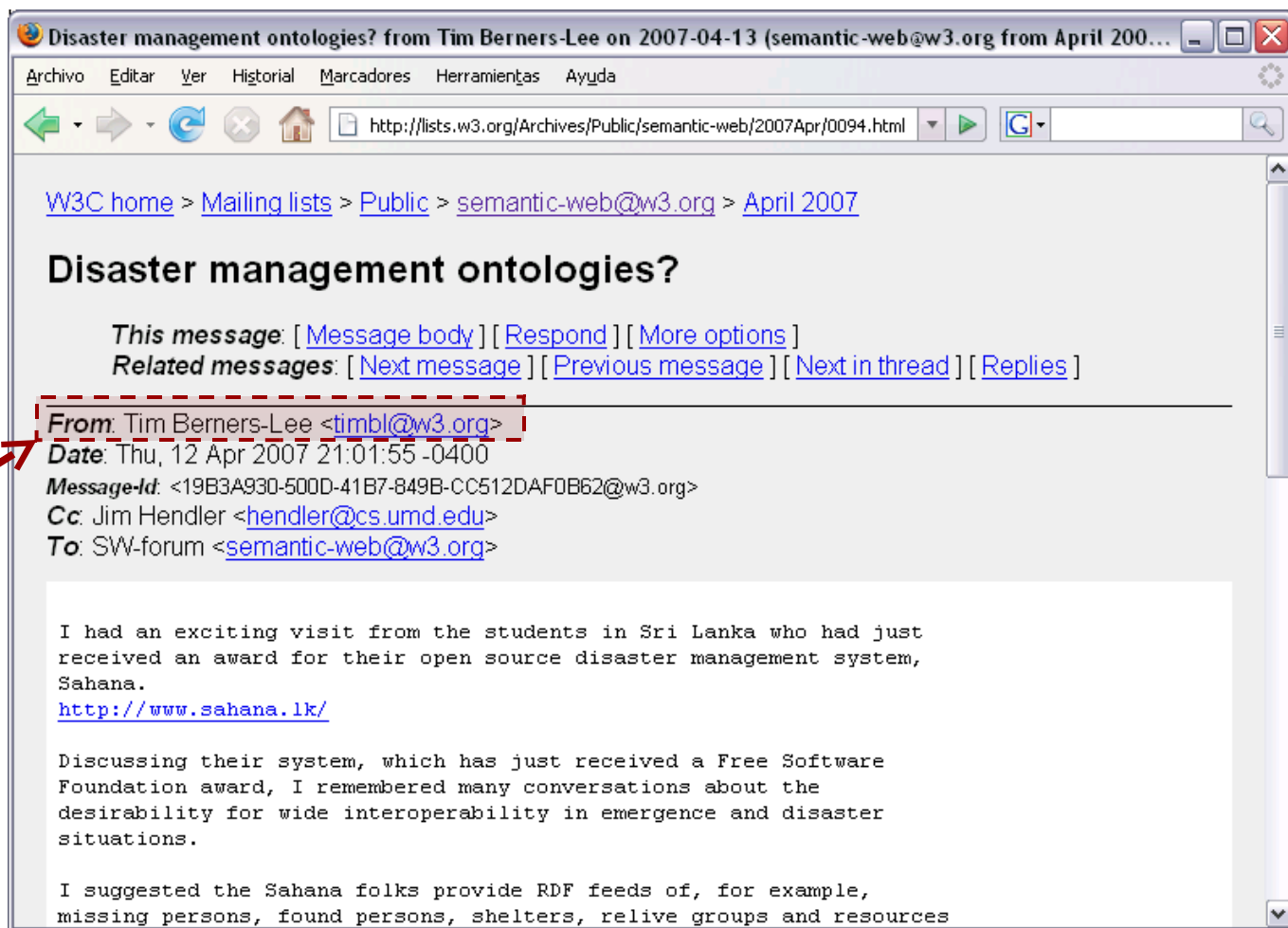


# Web and e-mail: the two towers



Fuente: S. Fernández et al.,  
“Mailing lists meet the Semantic Web”, SAW 2007, Poznan

# Mailing list messages as HTML pages



Author of the  
message

You see it,  
The machine **does not**



# Anatomy of an e-mail

**Received:** from bat.berlios.de (bat.berlios.de [195.37.77.135]) by lagavulin.gandi.net (Postfix) with ESMTP id 8B50B1CCDD for <diego@berrueta.net>; Fri, 13 Oct 2006 19:53:46 +0200 (CEST)  
**From:** Sergio Fdez <sergio@wikier.org>  
**To:** swaml-devel@lists.berlios.de  
**Date:** Fri, 13 Oct 2006 19:51:59 +0200  
**Message-Id:** <1160761919.3787.13.camel@localhost.localdomain>  
**X-Mailer:** Evolution 2.6.3  
**Subject:** [swaml-devel] SWAML 0.0.2 is out  
**Reply-to:** swaml-devel@lists.berlios.de  
**List-Id:** <swaml-devel.lists.berlios.de>

Hi,

I have just finished the work [...]

Best regards



# Buxon

**sioc:Forum  
URI**

**Message  
threads**

**Search/filter  
features**

The screenshot shows the Buxon application window titled "Buxon, a sioc:Forum visor". It features a "URI:" field at the top with the value "http://swaml.berlios.de/demo/index.rdf" and a "Jump to" button. Below this is a list of message threads, with "SWAML 0.0.1" selected. To the right of the list is a detailed view of the selected message, including its URI, from, to, subject, and date. The message body contains text about SWAML 0.0.1 and a release download link. At the bottom of the message details is a signature block for Sergio Fdez. Below the message details is a search bar with a "Find" button and date range filters. The status bar at the bottom indicates the loaded post URI.

URI:

Message threads:

- Re: Updated example in SIOC specification
- SIOC Types module, specification format change
- Re: Updated example in SIOC specification
- Making Content into Literals in AtomOWL
- OpenLink Suggestions for SIOC Enhancements a
- "Cheat Sheets" for SIOC
- SPARQL snippets?
- Re: SIOC extensions proposal
- Re: SIOC extensions proposal
- Drupal exporter - update
- PHP API documentation
- Microformats and SIOC
- SWAML 0.0.1**
- Re: SWAML 0.0.1
- Article about "Web 2.0" projects
- PNG please
- SIOC properties cardinality

Message details:

**Post URI:** http://swaml.berlios.de/demo/2006-Oct/post-42.rdf  
**From:** Sergio Fdez <http://swaml.berlios.de/demo/subscribers.rdf#s14>  
**To:** SIOC-Dev <http://swaml.berlios.de/demo/index.rdf>  
**Subject:** SWAML 0.0.1  
**Date:** Mon, 02 Oct 2006 23:54:56 +0200

Hi all,

It is my great pleasure and privilege to announce the availability of SWAML 0.0.1. This first release is an important step to represent a mailing list in RDF, but it's only the starting point.

SWAML reads a collection of email messages stored in a mailbox (from a mailing list compatible with RFC 4155) and generates an RDF description. It is written in Python using SIOC as the main ontology to represent in RDF a mailing list. There is still such much work to do; then, we would be very grateful of any feature request or any bug report.

Release download: <http://prdownload.berlios.de/swaml/swaml-0.0.1.tar.gz>

Best regards

--

Sergio Fdez  
GNU/LINUX User: #298803  
Web: <http://www.wikier.org/>

You received this message because you are subscribed to the Google Groups "SIOC-Dev" group.

loaded post http://swaml.berlios.de/demo/2006-Oct/post-42.rdf

**Message details**



# SWAML/Buxon: We do mash-ups too!



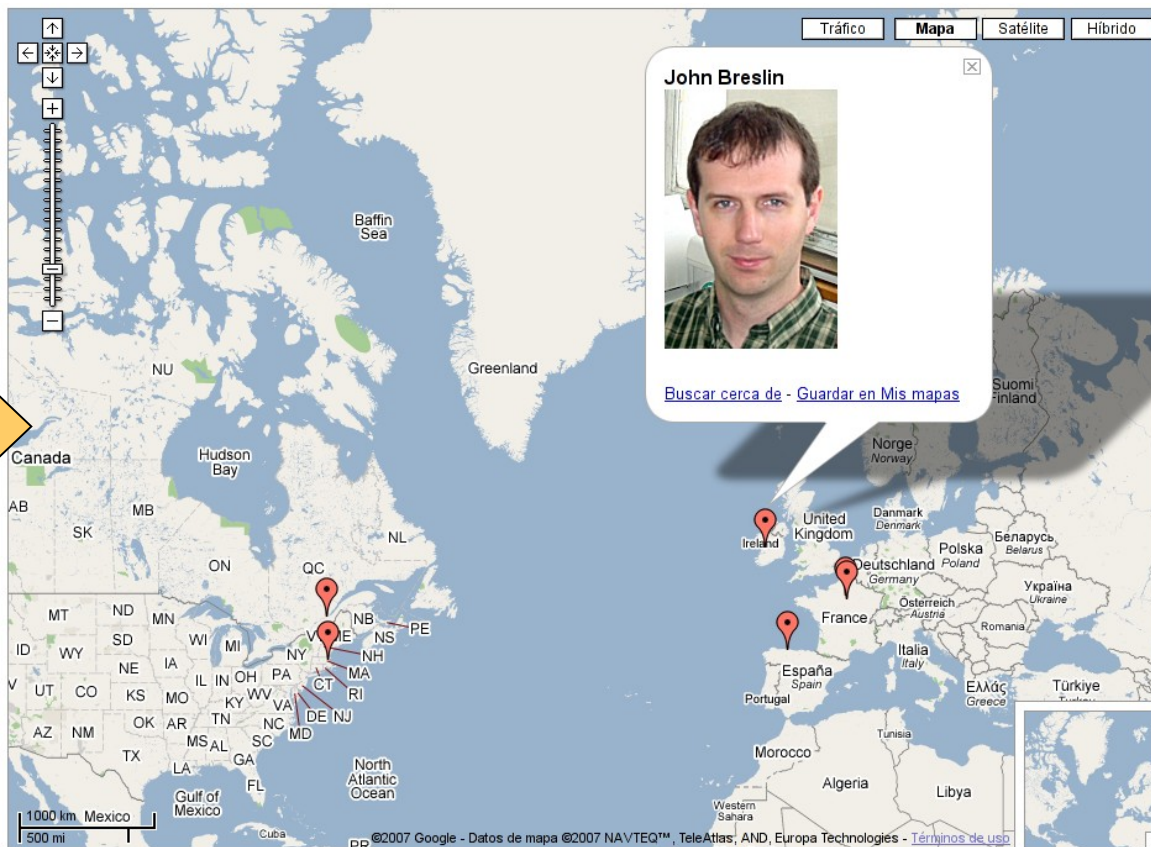
Mailing list  
data in RDF



SPARQL



FOAF  
profiles





# Improving search results quality



Objective: filter out repeated results



# EARL



- **“Evaluation And Report Language”** (origen: Evaluation and Repair tools WG, W3C)
- **Objetivo:** Capturar los resultados de la ejecución de pruebas
- **Conceptos fundamentales:** TestCase, TestSubject, Assertion
- **Aplicaciones:** pruebas unitarias, validadores, etc.
  - Vapour: validador de buenas prácticas en publicación de documentos de web semántica



## Ejemplo (I)

```
<earl:TestCase rdf:about="#TestContentTypeRdf">  
  <dc:title xml:lang="en">  
    Content type should be 'application/rdf+xml'  
  </dc:title>  
  <dc:description xml:lang="en">  
    The Content-Type header of the response should be  
    'application/rdf+xml'  
  </dc:description>  
</earl:TestCase>
```



## Ejemplo (II)

```
<earl:TestSubject rdf:nodeID="subjectRecipe1PropertyUri">
  <dc:title xml:lang="en">Derreferenced property URI</dc:title>
  <dc:date>2007-02-20</dc:date>

  <earl:httpRequest>
    <http:GetRequest>
      <uri:uri>http://vapour.sf.net/recipes-web/example1#propA</uri:uri>
      <http:version>1.1</http:version>
      <http:accept><!-- empty --></http:accept>
      <http:user-agent>Vapour validation tool 1.0</http:user-agent>
    </http:GetRequest>
  </earl:httpRequest>

  <earl:httpResponse>
    <http:Response>
      <http:content-type>application/rdf+xml</http:content-type>
      <http:responseCode>200</http:responseCode>
      <!--http:location-->
    </http:Response>
  </earl:httpResponse>
</earl:TestSubject>
```



## Ejemplo (III)

```
<earl:Assertion>  
  <earl:assertedBy rdf:resource="#vapour1-0"/>  
  <earl:subject rdf:nodeID="subjectRecipe1VocabularyUri"/>  
  <earl:test rdf:resource="&recipes;TestContentTypeRdf"/>  
  
  <earl:mode rdf:resource="&earl;automatic"/>  
  <earl:result>  
    <earl:TestResult>  
      <earl:validity rdf:resource="&earl;pass" />  
    </earl:TestResult>  
  </earl:result>  
</earl:Assertion>
```

## Ejemplo (IV)

### Dereferencing the vocabulary URI



Client

GET  
<http://vapour.sourceforge.net/recipes-web/example1>

200  
Content-type: application/rdf+xml



Web server

### Test results

- ♦ 1st request while dereferencing vocabulary URI without specifying the desired content type (Content type should be 'application/rdf+xml'): **Pass**
- ♦ 1st request while dereferencing vocabulary URI without specifying the desired content type (HTTP response code should be 200): **Pass**

### Further validation

- ♦ [Check the response with the W3C RDF validator](#)

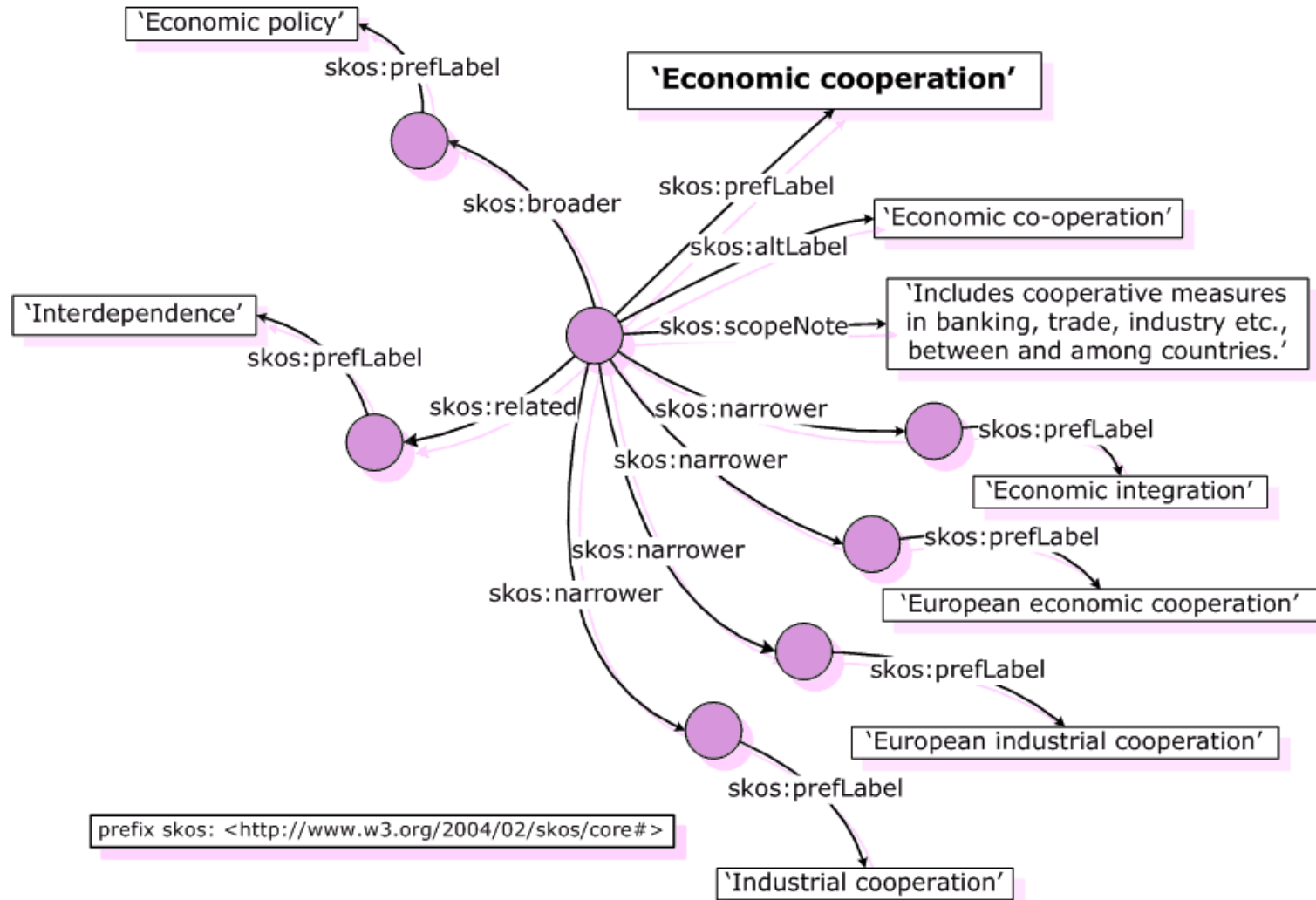
Report generated by [Vapour](#), version 1.0 on 2007-03-11 17:22:16.185252. Please read the [Best Practice Recipes](#) document. You can also [get an RDF version](#) of this report.



# SKOS


- “Simple Knowledge Organization System”
- **Objetivo:** describir esquemas conceptuales (tesauros, clasificaciones, taxonomías, folksonomías, glosarios, terminologías...)
- **Clase fundamental:** *<skos:Concept>*
- **Propiedades:**
  - *skos:prefLabel, skos:altLabel*
  - *skos:prefSymbol, skos:altSymbol, skos:depiction*
  - *skos:definition, skos:scopeNote, skos:example...*
  - *skos:narrower, skos:broader, skos:related*
  - *skos:primarySubject, skos:isSubjectOf*

# Ejemplo (I)



Fuente: SKOS Core Guide, <http://www.w3.org/TR/swbp-skos-core-guide/>

## Ejemplo (II)




semantic web  
environmental directory

☒ all items
 ☐ within current results

[Home](#)
[Adminstration](#)

Recently visited:



### Browse

Organisations and projects are categorised using six characteristics or facets which are shown below. The numbers in the brackets show how many results are in that facet. To start browsing pick an option from one of the facets.

Click  to see full lists of otions for each facet.

☒ **Topic of interest** [\[What is this facet?\]](#)  
[Animal Welfare](#) (12) | [Archaeology](#) (2) | [Arts and Crafts](#) (8) | [Biodiversity](#) (4) | [Built Environment](#) (18) | [Business and Commerce](#) (12) | [Charities](#) (1) [More options](#)

☐ **Project type** [\[What is this facet?\]](#)  
[Campaigning Project](#) (1) | [Conservation Project \(Heritage\)](#) (1) | [Conservation Project \(Wildlife\)](#) (1) | [Economic/Social Development Project](#) (1) | [Educational Project](#) (2) | [Pilot Project](#) (1) | [Research and Development Project](#) (1) [More options](#)

☐ **Operational area** [\[What is this facet?\]](#)  
[International Areas](#) (3) | [Non UK Counties](#) (5) | [United Kingdom](#) (77) | [Worldwide](#) (23)

☒ **Activity** [\[What is this facet?\]](#)  
[Animal Welfare](#) (6) | [Awards Accreditation](#) (6) | [Campaigning and Lobbying](#) (71) | [Conservation Cultural Heritage](#) (2) | [Cultural Activities](#) (7) | [Education and Training](#) (76) | [Employment](#) (2) [More options](#)

☒ **Name** [\[What is this facet?\]](#)  
[A\\*](#) (5) | [B\\*](#) (9) | [C\\*](#) (13) | [D\\*](#) (2) | [E\\*](#) (4) | [F\\*](#) (7) | [G\\*](#) (3) [More options](#)



# Combinación de vocabularios



# Combinación de vocabularios (I)

- **RSS + Dublin Core + módulos RDF (syn):**

```
<rdf:RDF>
```

```
  <channel rdf:about="http://barrapunto.com/">  
    <title>Barrapunto</title>  
    <link>http://barrapunto.com/</link>  
    <description>La información que te interesa</description>  
    <dc:language>es</dc:language>  
    <dc:rights>Copyright &copy; 2003, Barrapunto</dc:rights>  
    <dc:date>2005-11-28T10:12:22+00:00</dc:date>  
    <dc:publisher>Barrapunto S.L.</dc:publisher>  
    <dc:creator>webmaster@barrapunto.com</dc:creator>  
    <dc:subject>Software libre</dc:subject>  
    <syn:updatePeriod>hourly</syn:updatePeriod>  
    [...]
```

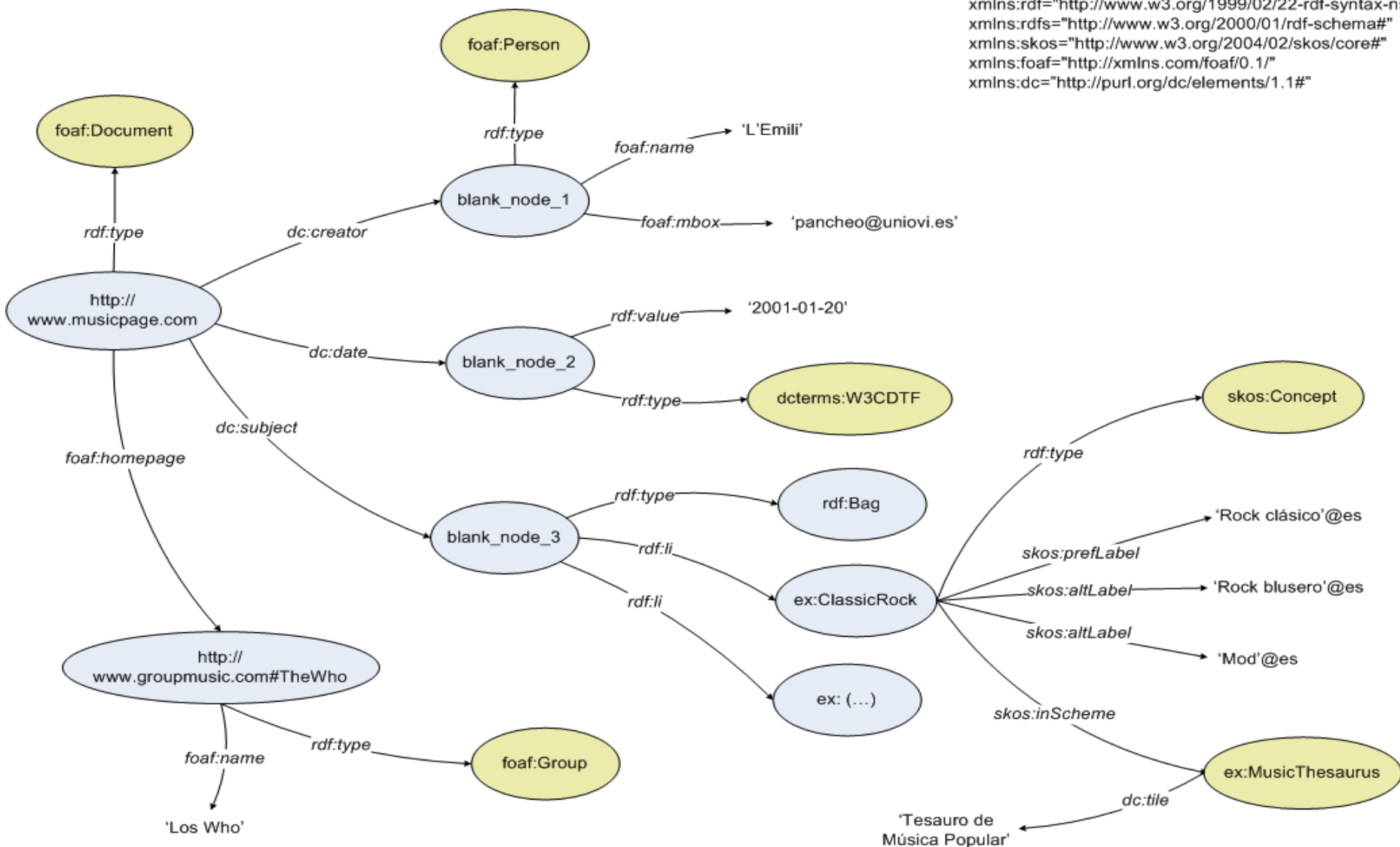


# Combinación de vocabularios RDF (II)

```

xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:skos="http://www.w3.org/2004/02/skos/core#"
xmlns:foaf="http://xmlns.com/foaf/0.1/"
xmlns:dc="http://purl.org/dc/elements/1.1#"

```





# Contenidos

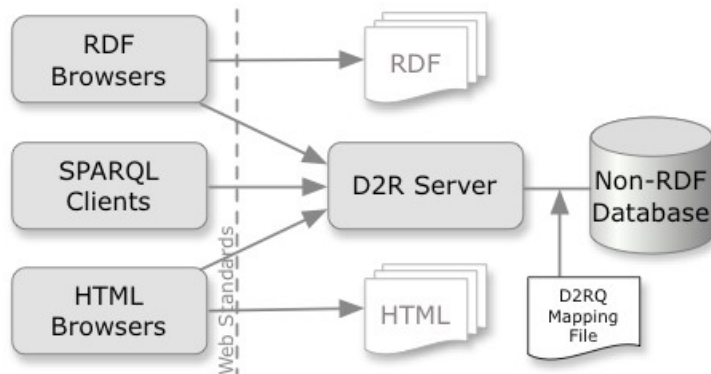
- 1) Introducción
- 2) Vocabularios comunes: FOAF, DOAP, RSS, Dublin Core, SIOC, EARL, SKOS
- 3) El huevo y la gallina: D2R, GRDDL, RDFa y microformatos, DBpedia
- 4) Navegadores semánticos: Tabulator, Piggy Bank
- 5) Otras piezas del puzzle: Semantic Web Client Library, PingTheSemanticWeb
- 6) Computación humana y web semántica (ESP Game)



**D2R**

# D2R: descripción

- Aplicación *wrapper* que publica en RDF los contenidos de una base de datos relacional
- Construcción automática del mapeo (lenguaje D2RQ) examinando la estructura de la BBDD. Posteriormente se puede afinar el mapeo
- Genera documentos RDF y un endpoint SPARQL
- Genera también un interfaz web para “navegar” por los contenidos de la base de datos



Fuente imagen: documentación de D2R



## D2R: ejemplo

```
map:Conference a d2rq:ClassMap;  
  d2rq:dataStorage map:Database1.  
  d2rq:class :Conference;  
  d2rq:uriPattern  
    "http://conferences.org/comp/confno@@Conferences.ConfID  
    @@";  
.
```

```
map:eventTitle a d2rq:PropertyBridge;  
  d2rq:belongsToClassMap map:Conference;  
  d2rq:property :eventTitle;  
  d2rq:column "Conferences.Name";  
  d2rq:datatype xsd:string;  
.
```

Fuente: documentación de D2R y D2RQ



# RDFa y microformatos



# Microformatos



- **Concepto:** formatos empotrados dentro de HTML usando (¿forzando?) elementos y atributos existentes para “describir” semánticamente fragmentos del contenido de la página
- Existen múltiples microformatos ad-hoc: hCard ,XFN, rel-licenses, rel-tag
- ¿Dónde está la semántica en los microformatos?  
¿cómo interpreta una máquina un microformato?



- **Objetivo:** introducir contenidos RDF en documentos XHTML, eliminar la duplicidad
  - ... en otras palabras: los microformatos, pero formalmente bien hechos
- **Funcionamiento:** reutilización de algunos elementos y atributos de XHTML, introducción de otros nuevos
- **Estado actual:** en desarrollo, aún no se puede validar una página XHTML con RDFa (oficialmente)
- **Herramientas:** bookmarklets, validadores, intérpretes
- **Casos de uso:** combinación con los vocabularios comunes, wikis semánticos, etc.



# RDFa: ejemplo

```
<html xmlns:cal="http://www.w3.org/2002/12/cal/ical#">
  <head><title>Jo's Blog</title></head>
  <body>
    <p class="cal:Vevent" about="#xtech_conference_talk">
      I'm giving
      <span property="cal:summary">
        a talk at the XTech Conference about web widgets
      </span>,
      on
      <span property="cal:dtstart"
        content="20070508T1000+0200">
        May 8th at 10am
      </span>.
    </p>
  </body>
</html>
```

Fuente: RDFa Primer, Adida y Birbeck



# RDFA: conceptos

- `xmlns:foaf="http://xmlns.com/foaf/0.1/"`
- `@class="foaf:Person"`
- `@id="xtech_conference_talk"`
- `@about="#xtech_conference_talk"`
- `@property="foaf:name"`
- `@content="20070508T1000+0200"`
- `@rel="foaf:homepage" / @href`
- `@datatype="xsd:date"`



# RDFa frente a microformatos

## RDFa

- Interpretación RDF
- Un único parser/GRDDL
- Respaldo W3C
- Es posible combinar vocabularios
- XHTML válido... **algún día**

## Microformatos

- ¿Cuál es la semántica?
- Parser o GRDDL ad-hoc
- ¿Quién los respalda?
- Difícil combinar vocabularios
- HTML válido **ahora**
- “Designed for humans first and machines second”



# GRDDL





- Objetivo: enlazar un documento XML con una transformación que extraiga su semántica
- Los enlaces pueden hacerse en documentos individuales o para clases enteras de documentos (enlazando el espacio de nombres)
- Caso especial para HTML (sigue siendo válido)
- Habitualmente, las transformaciones son hojas XSLT o funciones JavaScript
- Aplicaciones: “screen scraping”, extraer microformatos, extraer RDFa



# GRDDL: ejemplo (I)

```
<html xmlns="http://www.w3.org/1999/xhtml"
      xmlns:grddl='http://www.w3.org/2003/g/data-view#'
      grddl:transformation=
        "http://www.example.com/glean_title.xsl">
  <head>
    <title>Are You Experienced?</title>
  </head>
  ....
</html>
```



## GRDDL: ejemplo (II)

```
<html xmlns="http://www.w3.org/1999/xhtml">
  <head profile="http://www.w3.org/2003/g/data-view">
    <title>Some Document</title>

    <link rel="transformation"
      href="http://www.w3.org/2000/06/
      dc-extract/dc-extract.xsl" />
    <meta name="DC.Subject"
      content="ADAM; Simple Search; Index+; prototype" />
    ...
  </head>
  ...
</html>
```



# DBpedia



# DBpedia

- (Parte de) La información de la Wikipedia, disponible para la web semántica en RDF
- Extrae la información de las “cajas grises” (Infoboxes) de los artículos de Wikipedia
- Consulta las Wikipedias en varios idiomas
- Enlaza otras fuentes de datos externas a la Wikipedia (recursos geográficos, libros, música, CIA factbook...)
- Tiene un endpoint SPARQL



# Contenidos

- 1) Introducción
- 2) Vocabularios comunes: FOAF, DOAP, RSS, Dublin Core, SIOC, EARL, SKOS
- 3) El huevo y la gallina: D2R, GRDDL, RDFa y microformatos, DBpedia
- 4) Navegadores semánticos: Tabulator, Piggy Bank
- 5) Otras piezas del puzzle: Semantic Web Client Library, PingTheSemanticWeb
- 6) Computación humana y web semántica (ESP Game)



# Tabulator



# Tabulator: descripción

- Navegador de datos basado en árboles y tablas
- Ideado (e inicialmente programado) por Berners-Lee
- Programado en JavaScript/RDF (AJAR). Sólo Firefox y Opera
- No dirigido a usuarios finales, sino a desarrolladores e investigadores
- Sigue enlaces `rdfs:seeAlso`, negociación de contenido, redirecciones
- Soporta GRDDL y `owl:sameAs`
- Vistas de mapas, calendarios y cronogramas
- Consultas SPARQL, asistente para generar consultas



# Tabulator: ejemplo

## ▼ Semantic Web Activity

endingDate 2006-01-31  
startingDate 2001-02-09  
type ▶ ● activity  
homePage ▶ <http://www.w3.org/2001/sw/>  
homePageAddress <http://www.w3.org/2001/sw/>  
part

## ▼ RDF Data Access Working Group

endingDate 2006-01-31  
startingDate 2004-02-25  
type ▶ ● working group  
homePage ▶ <http://www.w3.org/2001/sw/DataAccess/>  
homePageAddress <http://www.w3.org/2001/sw/DataAccess/>  
mailbox ▶ ● <mailto:public-rdf-dawg@w3.org>  
chair ▶ Kendall Clark  
charter ▶ <http://www.w3.org/2003/12/swa/dawg-charter.html>  
name RDF Data Access Working Group  
patent policy ▶ PP  
teamContact ▶ Eric Prud'hommeaux  
is deliverer of ▶ RDF Data Access Use Cases and Requirements  
▶ SPARQL Protocol for RDF Using WSDL 1.1  
▶ SPARQL Query Results XML Format  
▶ SPARQL Protocol for RDF  
▶ SPARQL Query Language for RDF  
is part of ▶ Semantic Web Activity

▶ Rule Interchange Format Working Group

Fuente: The Tabulator: Data browser rewards Semantic Web best practices. Tim Berners-Lee



# Piggy Bank



# Piggy Bank

- Plugin que convierte Firefox en un “navegador semántico”
  - Servidor J2EE empotrado en Firefox!
- Capacidades:
  - Importar y almacenar información (RDF, GRDDL, screen scrapers en JavaScript)
  - Búsqueda y navegación
  - Visualizar la información en mapas, grafos, cronogramas
  - Exportar (compartir) información



Combined Information: restaurants + movies - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://127.0.0.1:1978/piggy-bank/model1116627657351?command=

[My Piggy Bank] | Combined Information

## restaurants + movies

Combined Information

1 filter criterion

- type: Show (remove) Restaurant (remove)

Order Commands

View items as list

1 item(s)

- ♦ [Abacus Cuisine Of China](#)

Map - Satellite New!

©2005 Google Map data ©2005 NAVTEQ™, Tele Atlas

Show Restaurant

Transferring data from mt.google.com...

Type here to search

**cuisine**

Type here to filter

- "Asian" (28)
- "Californian" (1)
- "Canadian" (2)
- "Chinese" (5)
- "Desserts" (1)
- "Dim Sum" (4)

**date**

By hour

Type here to filter

- 10:00 PM (138)
- 11:00 PM (20)

**type**

Type here to filter

- Restaurant (28)
- Show (158)

**address**

Type here to filter



# Contenidos

- 1) Introducción
- 2) Vocabularios comunes: FOAF, DOAP, RSS, Dublin Core, SIOC, EARL, SKOS
- 3) El huevo y la gallina: D2R, GRDDL, RDFa y microformatos, DBpedia
- 4) Navegadores semánticos: Tabulator, Piggy Bank
- 5) Otras piezas del puzzle: Semantic Web Client Library, PingTheSemanticWeb
- 6) Computación humana y web semántica (ESP Game)



# Semantic Web Client Library



# Semantic Web Client Library

- Objetivo: ofrecer a las aplicaciones Java una visión de toda la web semántica en un único grafo RDF
- Sigue automáticamente los enlaces (rdfs:seeAlso)
- Soporte GRDDL
- Usable desde la línea de comandos y como API



# Ping the Semantic Web



# Ping The Semantic Web

- Listado de URLs de documentos RDF
- Servicio para notificar (ping) cada vez que se crea o se modifica un documento
- Modelo push-pull



# Contenidos

- 1) Introducción
- 2) Vocabularios comunes: FOAF, DOAP, RSS, Dublin Core, SIOC, EARL, SKOS
- 3) El huevo y la gallina: D2R, GRDDL, RDFa y microformatos, DBpedia
- 4) Navegadores semánticos: Tabulator, Piggy Bank
- 5) Otras piezas del puzzle: Semantic Web Client Library, PingTheSemanticWeb
- 6) Computación humana y web semántica (ESP Game)



# Computación humana y web semántica

- 9.000 millones de horas-persona jugando al Solitario durante el año 2003
- ¿Y si se dedicase ese tiempo a algo útil?
  - Construcción del Empire State Building: 7 millones de horas-persona (año 1931)
  - Construcción del canal de Panamá: 20 millones de horas-persona (años 1880-1914)
- Mejor, ¿y si pudiésemos divertirnos y hacer algo útil a la vez?
- Human computation: hacer un uso práctico de los “ciclos de CPU” del cerebro humano
  - Ciertos problemas difíciles de resolver para la máquina son fácilmente resueltos por el cerebro humano

Fuente: Luis von Ahn, Human Computation, Google Tech talks, 2006



# ESP Game

- Juego de acertar palabras
- Colaboración entre desconocidos a través de la web
- Recogidas 30 millones de etiquetas de imágenes
  - Atributo “alt” (accesibilidad)
  - Recuperación de imágenes (buscadores)
- ¡Útil y divertido a la vez!
- De los mismos autores: Phetch



# Master Ingeniería Web

# FIN

**diego.berrueta@fundacionctic.org**



Publicado con licencia Creative Commons Reconocimiento NoComercial 3.0

