

# SWAML

## Semantic Web Archive of Mailing Lists

Sergio Fernández

`sergio.fernandez@fundacionctic.org`  
Departamento I+D+i (Fundación CTIC)

E.U.I.T.I.O (Universidad de Oviedo)

**<http://swaml.berlios.de/>**

11 de Mayo de 2007

# Introducción

SWAML

Sergio Fernández

Introducción

La Web semántica

Herramientas de software

Conclusiones y trabajo futuro

## ▶ **Situación actual:**

- ▶ miles de listas de correo
- ▶ publicación (sintáctica) en (X)HTML

## ▶ Situación actual:

- ▶ miles de listas de correo
- ▶ publicación (sintáctica) en (X)HTML

## ▶ Problemas:

- ▶ pérdida de información
- ▶ marcado sin ningún valor semántico
- ▶ problemas usando motores de búsqueda convencionales

# Archivos de una lista de correo en HTML

SWAML

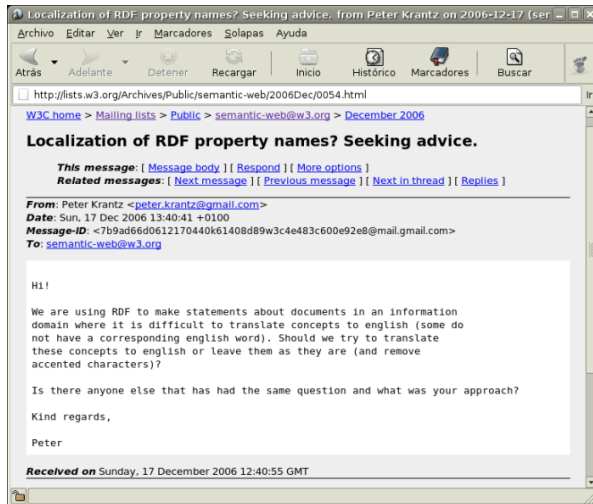
Sergio Fernández

Introducción

La Web semántica

Herramientas de software

Conclusiones y trabajo futuro



The screenshot shows a web browser window with the title "Localization of RDF property names? Seeking advice, from Peter Krantz on 2006-12-17 (ser)". The address bar contains the URL "http://lists.w3.org/Archives/Public/semantic-web/2006Dec/0054.html". The browser's menu bar includes "Archivo", "Editar", "Ver", "Ir", "Marcadores", "Solapas", and "Ayuda". The toolbar contains icons for "Atrás", "Adelante", "Detener", "Recargar", "Inicio", "Histórico", "Marcadores", and "Buscar".

The main content area displays the following text:

[W3C home](#) > [Mailing lists](#) > [Public](#) > [semantic-web@w3.org](#) > [December 2006](#)

## Localization of RDF property names? Seeking advice.

**This message:** [ [Message body](#) ] [ [Respond](#) ] [ [More options](#) ]  
**Related messages:** [ [Next message](#) ] [ [Previous message](#) ] [ [Next in thread](#) ] [ [Replies](#) ]

---

**From:** Peter Krantz <[peter.krantz@gmail.com](mailto:peter.krantz@gmail.com)>  
**Date:** Sun, 17 Dec 2006 13:40:41 +0100  
**Message-ID:** <7b9ad66d0612170440k61408d89w3c4e483c600e92e8@mail.gmail.com>  
**To:** [semantic-web@w3.org](mailto:semantic-web@w3.org)

Hi!

We are using RDF to make statements about documents in an information domain where it is difficult to translate concepts to english (some do not have a corresponding english word). Should we try to translate these concepts to english or leave them as they are (and remove accented characters)?

Is there anyone else that has had the same question and what was your approach?

Kind regards,

Peter

---

**Received on** Sunday, 17 December 2006 12:40:55 GMT

# Búsquedas caóticas

SWAML

Sergio Fernández

Introducción

La Web semántica

Herramientas de software

Conclusiones y trabajo futuro

The screenshot shows a web browser window with the title "recursive dns servers ddos as a growing ddos problem - Buscar con Google - Iceweas". The address bar contains "http://www.google.com/search" and the search query "growing ddos:". The Google logo is visible, along with navigation buttons and a search box containing the query. Below the search box, there are search options: "Buscar en la Web" and "Buscar sólo páginas en español". The search results are displayed under the heading "La Web Resultados 1 - 10 de aproximadamente 28,200 de recursive dns servers ddos as a growing ddos problem. (0.2)". The first result is "recursive DNS servers DDoS as a growing DDoS problem" with a link to "Traduzca esta página". The snippet for this result reads: "We discussed recursive DNS servers before (servers which allow to query anything - including what they are not authoritative for, . www.derkeiler.com/Mailing-Lists/securityfocus/bugtraq/2006-03/msg00003.html - 16k - En caché - Páginas similares)". The second result is "Full Disclosure: recursive DNS servers DDoS as a growing DDoS problem" with a link to "Traduzca esta página". The snippet reads: "recursive DNS servers DDoS as a growing DDoS problem. This message : [ Message body ] [ More options ]; Related messages : [ Next message ] [ Previous ... seclists.org/fulldisclosure/2006/Feb/0697.html - Páginas similares)". The third result is "SecurTeam Blogs » Recursive DNS servers as a growing DDoS problem" with a link to "Traduzca esta página". The snippet reads: "[2] DNS-operations - http://lists.oarci.net/mailman/listinfo/dns-operations. [Changed title from: Recursive DNS servers DDoS as a growing DDoS problem] ... blogs.securiteam.com/index.php/archives/332 - 42k - En caché - Páginas similares)". The fourth result is "recursive DNS servers DDoS as a growing DDoS problem" with a link to "Traduzca esta página". The snippet reads: "Re: recursive DNS servers DDoS as a growing DDoS problem. From: Ventsislav Genchev. Re: recursive DNS servers DDoS as a growing DDoS problem ... archive.cert.uni-stuttgart.de/bugtraq/2006/03/msg00003.html - 10k - En caché - Páginas similares)". The fifth result is "Re: recursive DNS servers DDoS as a growing DDoS problem" with a link to "Traduzca esta página". The snippet reads: "Re: recursive DNS servers DDoS as a growing DDoS problem ... Follow-Ups.: Re: recursive DNS servers DDoS as a growing DDoS problem. From: Geo ...". The browser status bar at the bottom shows "Terminado" and "PR:n/a".

La Web semántica es una **extensión** sobre la Web actual donde el contenido Web debe ser expresado de forma que sea entendido, interpretado y usado por máquinas.

Proviene de la visión de Tim Berners-Lee (Director del W3C) de la Web como un medio **universal** para el intercambio de **datos**, **información** y **conocimiento**.

La siguiente versión de la Web, la **Web 3.0**

# RSS, un caso de uso muy conocido

SWAML

Sergio Fernández

Introducción

La Web semántica

Herramientas de software

Conclusiones y trabajo futuro

## News



[30-04-2007] [SAW2007 slides](#)

Published the slides used in talk of last friday in the workshop on Social Aspects of the Web (aka SAW2007). You could find it at: <http://pdownload.berlios.de/swaml/SAW2007-Mailing-Lists-Meet-The-Semantic-Web.pdf>

[25-04-2007] [SWAML finalist](#)

Yesterday we have received the notification that SWAML has entered the final phase of the "I Concurso Universitario de Software Libre", a competition between free projects developed into the Spanish universities. We will have to wait until the final phase , that will be celebrated in May in Seville, to know the final result. More information at: <http://concurso-softwarelibre.us.es/noticias.html#finalistas>

[16-04-2007] [SAW2007's programme published](#)

The programme of Workshop on Social Aspects of the Web (SAW 2007), where we'll present a paper about SWAML, has been published. See it at: [http://bis.lj.ie.poznan.pl/10th\\_bis/programme.shtml](http://bis.lj.ie.poznan.pl/10th_bis/programme.shtml)

[23-03-2007] [SWAML files to Poland](#)

A paper about SWAML, titled "Mailing lists meet the Semantic Web", had celebrate on april in Poznan (Poland). SAW 2007 is organised in conjunct in co-operation with ACM SIGMIS. More information at the conference's Web

[09-02-2007] [Buxon also accepted in Debian](#)

Buxon 0.0.3 also has been accepted in Debian GNU/Linux.

[28-01-2007] [SWAML accepted in Debian](#)

SWAML 0.0.5 has been accepted in Debian GNU/Linux. In a few hours it hadho for the excellent work that he has done <https://lists.berlios.de/pipermail/swaml-devel/2007-January/000045.html>

[28-12-2006] [New releases: SWAML 0.0.5 and Buxon 0.0.3](#)

Today we released two new versions: SWAML 0.0.5 and Buxon 0.0.3, final packages from <http://swaml.berlios.de/#files>

[27-12-2006] [Closed phase](#)

Last week Sergio successfully presented his Degree Project. This project phase doesn't mean that project development finish here, it is left much to

[04-12-2006] [SWAML on BibSonomy](#)

All entries of SWAML's BibTeX files has been imported into BibSonomy: <http://www.bibsonomy.org/>

[02-12-2006] [PFC is on the print shop](#)

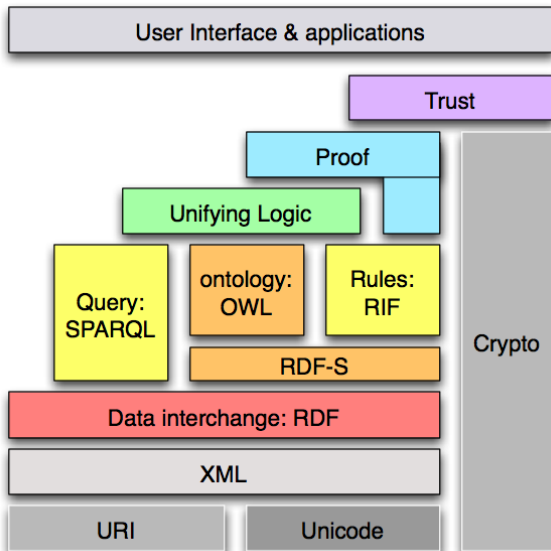
Yesterday I took the final version of PFC document to the print shop. The book (in Spanish) about the SWAML project for my degree dissertation.

Código fuente de [http://developer.berlios.de/export/rss20\\_bsnnews.php?](http://developer.berlios.de/export/rss20_bsnnews.php?)

Archivo Editor Ver Ayuda

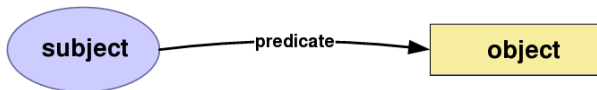
```
<?xml version="1.0" encoding="ISO-8859-1" ?>
<rss version="2.0">
<channel>
<title>Berlios Developer Project: SWAML - News</title>
<link>http://developer.berlios.de/news/?group\_id=4806</link>
<description>Berlios Developer Project News of SWAML</description>
<language>en-us</language>
<copyright>Copyright 2006-2007 Fraunhofer FOKUS</copyright>
<webMaster>wikiusers.berlios.de (Sergio Fdez)</webMaster>
<lastBuildDate>Wed, 09 May 2007 15:54:00 GMT</lastBuildDate>
<docs>http://blogs.faw.harvard.edu/tech/rss/</docs>
<generator>Berlios RSS generator</generator>
<image>
<url>http://developer.berlios.de/images/berlios\_small\_logo.png</url>
<title>Berlios Developer</title>
<link>http://developer.berlios.de/</link>
<width>124</width>
<height>32</height>
</image>
<item>
<title>SAW2007 slides</title>
<link>http://developer.berlios.de/forum/forum.php?forum\_id=25310</link>
<description>Published the slides used in talk of last friday in the workshop on Social Aspects of the Web (aka SAW2007). You could find it at: http://pdownload.berlios.de/swaml/SAW2007-Mailing-Lists-Meet-The-Semantic-Web.pdf
```

# La pila de la Web semántica



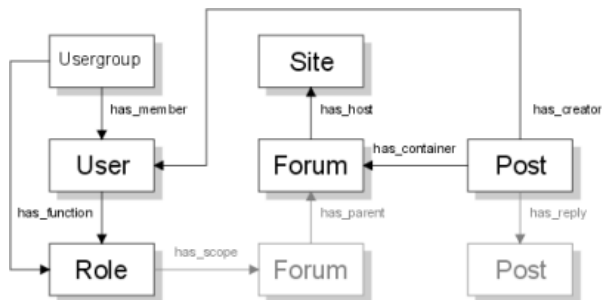
**R**esource **D**escription **F**ramework es una familia de tecnologías (RDF, SPARQL, etc...) de W3C para modelar conocimiento.

RDF se basa en un modelo de tripletas:



El **sujeto** es un recurso, identificado por un URI (*Uniform Resource Identifier*), que se relaciona mediante un **predicado** con un **objeto**, que puede ser a su vez otro recurso o un literal.

**Semantically-Interlinked Online Communities** es una ontología que provee un vocabulario para interconectar diferentes sitios de discusión, como blogs, wikis, foros y/o listas de correo.



# Herramientas de software

Principalmente hemos desarrollado dos herramientas de software como parte de este proyecto:

Principalmente hemos desarrollado dos herramientas de software como parte de este proyecto:

- ▶ **SWAML** es un aplicación no interactiva en línea de comandos cuyo principal propósito es *traducir* mailboxes en instancias de `sioc:Forum` en RDF.

Principalmente hemos desarrollado dos herramientas de software como parte de este proyecto:

- ▶ **SWAML** es un aplicación no interactiva en línea de comandos cuyo principal propósito es *traducir* mailboxes en instancias de `sioc:Forum` en RDF.
- ▶ **Buxon** es un navegador gráfico para visualizar instancias de `sioc:Forum`.

# Herramientas de software

SWAML

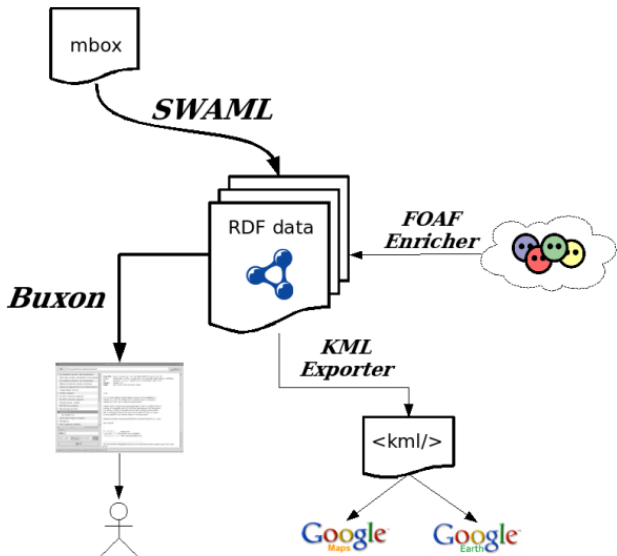
Sergio Fernández

Introducción

La Web semántica

Herramientas de software

Conclusiones y trabajo futuro



proceso batch:



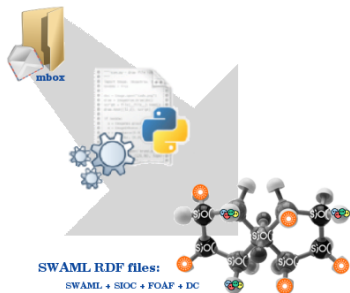
proceso batch:

1. mbox



proceso batch:

1. mbox
2. parser



proceso batch:

1. mbox
2. parser
3. serializar a RDF/XML

# SWAML y KML

SWAML

Sergio Fernández

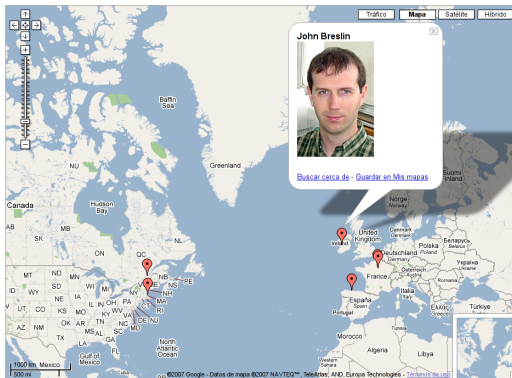
SWAML también exporta en KML las coordenadas geográficas de los suscriptores de una lista de correo.

Introducción

La Web semántica

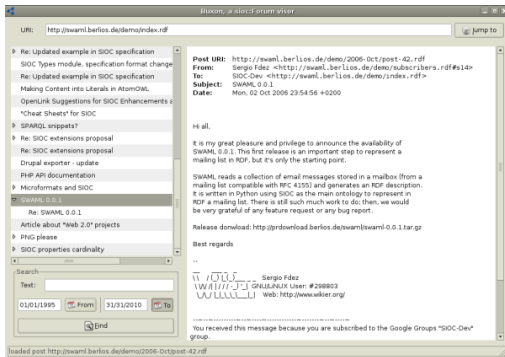
Herramientas de software

Conclusiones y trabajo futuro



SWAML + SPARQL + FOAF = mash-up!

- ▶ un lector de `sioc:Forum`
- ▶ recompone los hilos de los mensajes
- ▶ quizás la implementación más madura de SIOC



## Conclusiones técnicas

- ▶ Conseguir que la inmensa cantidad de conocimiento que almacenan las distintas listas de correo pueda llegar a ser **útil** de verdad.
- ▶ Apoyo a la submission de SIOC al W3C.
- ▶ Presencia internacional (paquetes en Debian GNU/Linux) con una relativa comunidad ya.

# Conclusiones personales

- ▶ Mayor involucración en la comunidad del Software Libre
- ▶ El haber llegado a esta fase final
- ▶ Y conseguir un título :-)

- ▶ Generar RDFa, incluyendo las tripletas en el marcado XHTML.
- ▶ Acceso a cuentas POP/IMAP para no depender de quien aloja la lista de correo.
- ▶ Publicar los datos a través de un SPARQL endpoint.

*La tarea de la universidad no es ofrecer lo que la sociedad demanda, sino lo que la sociedad necesita.*

*E. W. Dijkstra*

Y como el Software Libre es algo que ayuda a que una sociedad sea mejor, **la Universidad debe apostar por el Software Libre** para que estos beneficios lleguen a la sociedad.

¿preguntas...?

En primer lugar a todas las personas que han puesto su granito de arena en el proyecto, especialmente a **Diego Berrueta** y **Jose E. Labra**, pero también a John Breslin y Uldis Bojars de DERI Galway, y a Nacho Barrientos.

Y también a los organizadores, patrocinadores y colaboradores del concurso.

Más información en la página Web del proyecto:

<http://swaml.berlios.de/>

# SWAML

## Semantic Web Archive of Mailing Lists

Esta charla se distribuye bajo los términos de la licencia:

**CreativeCommons**

