SWAD-Europe extra deliverable: Report on developer workshop 8 - The Semantic Web in Latin America.

[⊣ Versión española]

Project name:
> Semantic Web Advanced Development for Europe (SWAD-Europe)

Project Number:
> IST-2001-34732

Workpackage name:
> 3 Dissemination and Implementation

Workpackage description:
> ⊣ http://www.w3.org/2001/sw/Europe/plan/workpackages/live/esw-wp-3

Deliverable title:
> Extra deliverable. Developer Workshop 8 Report

URI:
> ⊣ http://www.w3.org/2001/sw/Europe/reports/dev_workshop_report_8

Author:
> Charles McCathieNevile

Abstract:
> This report summarises the eighth developer workshop, held in Buenos Aires, Argentina on 21 and 22 August 2004. The workshop explored various aspects of the Semantic Web development in Spanish, with a particular focus on work taking place in Europe and in South America

STATUS:
> **Completed** The first draft of this report was published 5 September 2004 in Spanish. The first English draft was published 9 September 2004 This version $Date: Monday 20 September 2004 - 19:21:55$.
>
> Please send any comments on this report to the mailing list ⊣ www-rdf-interest (in English) or ⊣ Web-Semantica-Ayuda (in Spanish).

Contents

Executive summary

This workshop brought together Spanish speakers (the great majority from Latin America, but also from Spain, and the USA) interested in developing Semantic Web technologies. It was held entirely in Spanish..

The workshop was hosted by the Departamento de Diseño de Sistemas of the ⊣ Universidad Argentina John F. Kennedy (in Spanish), in Buenos Aires. Attendees came from the four target communities with about half being from academia, and the rest divided between open-source and commercial developers and the government sector).

In addition the discussion was noted live on IRC, so further participation of a more global character was made possible.

The workshop had the following outcomes:

- Development a range of RDF documentation in Spanish
- Introducing a range of South American developers to the Semantic Web and to each other.
- Demonstrating some work that has been done in Spanish, in particular showing work done in Latin America
- Motivating the use of the Spanish language developers list Web-Semantica-Ayuda
- Motivating closer collaboration between developers in Latin America, and with European (particularly Spanish) developers.
- Introducing the Latin American developer community to some collaboration tools heavily used by the english-speaking developer community.

1 Introduction

This report is part of the ⊣ SWAD-Europe project ⊣ Work package 3: Dissemination and Implementation. It describes the developer workshop "La Web Semántica en América Latina" held in Buenos Aires, Argentina, on 21 and 22 August 2004, in Spanish.

The principal objectives of the workshop were

- Introduce spanish-speaking developers to some of the principal ideas and some tools for the Semantic Web in their own language.
- Present some of the work on the Semantic Web being done in Spanish and in Latin America.
- Motivate more of the spanish-speaking developer community in various areas of work concerning the Semantic Web
- To instigate the development of a working spanish-language Semantic Web developer community

2 Background

There is work done on the Semantic Web in Latin America, but much of it takes place, or is published, only in English. For people interested in the topic it is very difficult to find information available in Spanish. To introduce more members of the Spanish-speaking development community to the Semantic Web, and to introduce developers to each other, some members of the ⊣ Web-Semantica-Ayuda (in Spanish) suggested a workshop in Spanish in Latin America. The success of a previous SWAD-Europe workshop in Spanish, held in Madrid but followed via IRC from Latin America was an additional motivation, and the 30 available places were very quickly filled, with a waiting list established.

The workshop was led by Charles McCathieNevile, of SWAD-Europe, and hosted by the Departamento de Diseño de Sistemas of the ⊣ Universidad

Argentina John F. Kennedy (in Spanish). The organisation committee included:

- Ricardo Baeza-Yates, Universidad de Chile (Chile)
- Diego Ferrayra, Ministerio de Educación, (Argentina)
- Leandro Mariano López, a Semantic Web developer better known as Inkel (Argentina)
- Eva Méndez, Universidad Carlos III Madrid (Spain)
- Luis Olsina, Universidad Nacional de La Pampa (Argentina)
- Rosa Rojas, Universidad Kennedy (Argentina)
- Claudio Segovia, Fundación Sidar (Argentina)
- Gabriel Tiburtini, Universidad Kennedy (Argentina)

Many thanks are due to all the above and their respective organisations for their energy and assistance.

3 The workshop

Presentations -

In the course of the workshop there were a number of formal presentations and discussions on relevant themes. They are described in the order in which they took place.

Introducción a la Web Semántica con CWM (Introduction to the Semantic Web with CWM)

This presentation by Charles McCathieNevile, member of the SWAD-Europe project on behalf of W3C. The presentation gave an introduction to RDF, primarily through the use of practical examples. The presentation covered

- The basic structure of RDF information, its representation in RDF/XML and the use of N3 to test extensions.
- Creating information with Foaf-a-matic and Hera - two tools available in Spanish (the first translated from the english original by Leandro Mariano López and the latter developed originally in spanish and translated to english by Ramiro Benavidez and Charles McCathieNevile) which produce RDF information about different areas but with some overlap.
- The use of FoafNaut as a browser of a specific type of RDF information
- The use of CWM to merge and query RDF information
- The development of vocabularies
- A quick introduction to the family of RDF and OWL specifications

(This presentation and discussion in IRC)

El vocabulario FOAF (Amigo de un Amigo) y su extensión con información sobre qué idiomas alguien lee, escribe o habla (The FOAF Vocabulary and its extension with information about what languages people speak, read or write)

Leandro Mariano Lopez (better known globally as a Semantic Web developer through his nom-de-plume "Inkel) is an independent Buenos Aires-based Semantic Web developer, best-known for his vocabulary " Speaks Reads and Writes".. He has translated a number of important Semantic Web documents, specifications and tools, and maintains a Weblog in Spanish about the Semantic

Web and Web Programming. In his presentation Inkel explained the origins and the basics of the FOAF (Friend Of A Friend) vocabulary, already widely used in the Semantic Web, and the motivation for and implementation of his extensions to the vocabulary describing the languages that people can speakm read or write. (⊣ This presentation and discussion in IRC)

⊣ Tematres

Diego Ferreyra works in the Argentine Ministry of Education, in a group dedicated to inforamtion and communication technologies (ICT). He also participates in an english project to develop a motor for dynamic multimedia presentation, and is the author of TemaTres. He presented and led discussion about this tool, a web-based system developed in Argentina for the manipulation of hierarchical thesauri. His discussion covered how the structure of a thesaurus is easily represented in RDF and how TemaTres can present a thesaurus following the North American standard ⊣ ZThes or the ⊣ SKOS-Core vocabulary developed for the purpose by the SWAD-Europe project, using ⊣ Dublin Core to encode the information used. Finally he offered some perspectives on the future development of TemaTres, leading to a discussion on the use of RDF as a native format for systems such as thesauraus management tools. A ⊣ demonstration of Tematres with some test thesauri is available. (⊣ This presentation and discussion in IRC)

⊣ Búsqueda y Navegación Semántica (Semantic Search and Navigation)

Maria de los Angeles Martín is a professor and researcher in the ⊣ Grupo de I+D en Ingeniería del Software de la Universidad de La Pampa. María de los Angeles presented her project on the development of a bibliometric ontology, and a system of semantic navigation and search for catalogues, using this approach. As part of the discussion she gave a more detailed explanation of the various core techonogies used in the area, such as, XML Schema, RDF and OWL, the relationships between them and the benefits and drawbacks of each. She also covered the methodologies commonly used for the construction of thesauri, including ⊣ Methontology which they have selected for their project, and looked at the different approaches to querying information (syntactic, structural, and semantic) and the architectural requirements they each imply. (⊣ This presentation and discussion in IRC. ⊣ The slides are also available as PDF - 1.5MB)

⊣ Grupo Metadatos, Universidad de Chile

Claudio Gutierrez, profesor in the ⊣ Departamento de Ciencias de la Computación at the Universidad de Chile, one of the principal investigators in the Metadata research group at that University, and member of the ⊣ Centro de Investigaciones sobre la Web presented the work of the group: Its goals, motivation and interests, projects that have been completed or are in progress, applications developed by the group including a ⊣ generator for Dublin Core metadata in Spanish, a ⊣ wiki about metadata (a mix of English and Spanish), and ⊣ DepMark, a project to gather information in RDF about all the departments of computer science in Chilean Universities, their staff, projects and publications. Claudio also presented his reflections on the state of the Semantic Web today, higlighting in the discussion issues that can hold back development including problems of an essentially social nature, difficulties in portability and interoperability of Semantic Web based systems, the problem of the diffuse nature of the Semantic Web, etc. (⊣ This presentation and discussion in IRC)

⊣ Dublin Core en español: Metadatos, DC, etc.

Eva Méndez is a professor and researcher in biblioteconomia at Universidad Carlos III de Madrid in Spain. She also maintains the ⊣ Spanish language mirror of the Dublin Core website and is co-administrator of the Spanish metadata discussion list ⊣ DCMI-ES. She focussed her presentation on metadata schemas, disitnguishing the Dublin Core, as a general model for metadata and an operational infrastructure in the Semantic Web from the many metadata schemes based in particular areas of knowledge and used on the web, highlighting the fact that the current status of the Semantic Web is primirily a set of unlinked "Semantic islands". Eva spoke of the processes of formalisation for Dublin Core and for RDF, and led a discussion on the application of Dublin Core in Spanish-language projects and the utility of metadata, leading to discussion of "web visibility" and the "new visibility" assumed by RSS. Her presentation generally drew attention to the value and importance of working in Spanish, rather than English, for the Spanish-speaking development community. (⊣ This presentation and discussion in IRC)

⊣ Annotea, EARL, y la web fiable (Annotea, EARL and the Web of Trust)

Charles McCathieNevile presented some concrete steps towards a Semantic Web that has stronger management of trust. In this presentation he concentrated on ⊣ Annotea and ⊣ EARL, two applications which for different reasons have directly incorporated some management of trust in their application. Along with some work done in the area of FOAF applications, these projects have provided implementation experience on some possible ways to move towards a more generalised and more standardised "Web of Trust" and the issues that arise in this line of development. (⊣ This presentation and discussion in IRC)

Attendance -

At the workshop itself 30 people were present. Around half of these people represented various universities, mostly from Argentina, and the rest were from software and web companies or organisations, independent developers, with a few people from the government sector.

As well, a number of people participated through the IRC discussion, with the result that some parts of the workshop were translated in real time to english, allowing for a wider disucssion between various developers.

Organisations represented included

- DIGIBIS, S.A. (Spain)
- Dirección de Aduanas, Argentina (Customs agency)
- Fundación Sidar (International)
- Gobierno de la Ciudad Autonoma de Buenos Aires (Argentina)
- Ministerio de Educación de Argentina.
- Universidad de Buenos Aires (Argentina)
- Universidad de Chile
- Universidad Carlos III de Madrid (Spain)
- Universidad John F Kennedy (Argentina)
- Universidad Nacional de la Pampa (Argentina)
- Universidad de Mar del Plata (Argentina)
- University of Hawaii (USA)

4. Outcomes

In the course of the workshop a number of ⊣ Semantic Web tools and projects were presented in detail - see ⊣ Appendix A, as well as many of the important themes in development, and the experiences of developers working in the area. The body of work presented in Spanish provides an important resource for Spanish-speaking developers or Spanish speakers interested in learning more about the topic, in some cases providing the first available documentation in Spanish of a particular area In the course of the workshop a nunmber of mebers of the Spanish-speaking developer community have been introduced to the use of IRC, and tools such as the "chump" as a practical way to facilitate remote collaboration, and in particular its everyday use as an important tool used by the global Semantic Web developer community.

Since the workshop there has been an increase in participation in the discussion list ⊣ Web-Semantica-Ayuda, and in participation in the international IRC channel ⊣ #rdfig from Argentina in particular, and spanish-speakers in general, building further on the similar result from the previous SWAD-Europe workshop held in Spanish.

From this increased level of discussion we hope to see an increase in collaboration among Spanish-speaking developers and between the Spanish-speaking community and the global development community, resulting in increased international recognition of work being done in Spanish and an increase in the quality as well as quantity of that work.

This workshop has demonstrated a concrfete contribution in this direction, motivating such work as a presentation in Spanish of existing steps towards the "Web of Trust" and the development of the Spanish version of DOAP-a-matic

A ⊣ record of the discussion (in spanish) is available, as are the noted highlights (⊣ first day highlights, ⊣ second day highlights) with links which were considered interesting, noted and commented by the participants including those who were following the workshop remotely through IRC.

The evaluations of the workshop by the participants have expressed a very high level of satisfaction with the workshop, the topics presented and the individual presentations.

Appendix A Projects and Tools

⊣ Amaya
A browser/editor for the Web, developed as an open-source project led by W3C and INRIA. It includes an interface for Annotea annotations, and uses the RDF bookmark vocabulary. The interface can be set to a number of of languages including English and Spanish.

⊣ Annotea
An protocol for making annotations about any type of Web content, using RDF. The Annotea project, financed in part by SWAD-E, has invluded development of a server, client library code in several languages, and implementation of a complete user interface in Amaya, as well as the protocol itself.

⊣ Chump
A program that can collect comments about something of interest from an IRC channel, provding a summary page which is more readable than a complete log of the discussion. It is used heavily by the RDF community in the #rdfig channel.

⊣ CWM
An RDF parser developed as an open source project led by Tim Berners-Lee and Dan Connolly. During the workshop an ⊣ Introduction to RDF with CWM was presented.

⊣ DepMark (in Spanish)

A java application for the creation and management of information about departments of computer science, their staff and publications, in Chiiean universities. Developed by the ⊣ Grupo Metadatos at the Universidad de Chile.

⊣ Doap-a-matic
> A tool for describing projects such as software development projects in RDF, by entering data into a form. This tool was developed as a multilingual application by Vincent Tabard and a Spanish version was made for presentation at the workshop by Emmanuelle Gutiérrez y Restrepo of Fundación Sidar.

⊣ EARL
> An RDF vocabulary to describe conformance of a resource to a list of requirements. An introduction to the code, ⊣ EARL by example is available in a number of languages including English and Spanish.

⊣ FOAF
> An RDF vocabulary to describe people, with various attributes including whom they know.

⊣ Foaf-a-matic
> A Web page that permits a user to create a FOAF RDF file by filling in a form. Created by Leigh Dodds, a Spanish translation was made by Leandro Mariano López

⊣ Foafnaut
> A graphic browser (in SVG) for FOAF information. From Jim Ley

⊣ Foaf Explorer
> A browser for FOAF information, presenting a large variety of information in a web page.

⊣ Generador Dublin Core (in Spanish)
> A tool for generating Dublin Core information about documents in Spanish, developed by the ⊣ Grupo Metadatos at the Universidad de Chile.

⊣ Grupo Metadatos (in Spanish)
> Site of the Universidad de Chile's research group investigating metadata and related technology. Created under the auspices of the ⊣ CIW (Centro de Investigación sobre la Web) at the University, the group is dedicated to metadata and other technology which has direct implications for the construction of the Semantic Web. The site also hosts a ⊣ metadata Wiki (mixed English and Spanish), and links to various projects developed or in development within the group.

⊣ Hera
> A tool that permits a user to evaluate the accessibility of a web page according to the Web Content Accessibility Guidelines, and create a report using the RDF vocabulary EARL (supplemented with some specific terms). From the ⊣ Fundación Sidar, it has a web-based interface, available in a number of languages including English and Spanish.

⊣ Jena
> A Java RDF parsing package, developed as an open source project led by Hewlett Packard, and in part through the SWAD-E project.

⊣ Kowari
> An open-source database designed for treating metadata, and offering massive scalability. It uses RDF as its native format and is optimised for storing and querying data as subject-predicate-object triples.

⊣ RDFPic Extended
> A tool developed by Vincent Tabard in colaboration with the ⊣ Fundación Sidar, based on the ⊣ RDFPic tool from W3C. This version uses an infrastructure based on Apache and PHP, a much more common platform than the Jigsaw server that the original version uses..

- Redland
    A framework for RDF programing. Developed as an open source project led by Dave Beckett, in part through the SWAD-E project, it permits a programmer to have access to RDF parsing (normally through Raptor) through APIs available in a number of programming languages (e.g. C, Ruby, PERL, Java)
- SKOS-Core
    An RDF/OWL vocabulary produced as part of the SWAD-E project, for encoding rich thesauri, taxonomies, etc.
- Tematres
    Una aplicación para la gestión de tesaurós que puede presentar sus datos en forma de SKOS (un vocabulario RDF para tesauros) On this topic see also:
    - Automatización de tesauros y su utilización en la Web Semántica (in Spanish), the presentation by José Ramón Pérez Aguëra at the - taller en español sobre la Web semántica (in Spanish) held in Madrid in June 2004).