Public Final Report (D1.13)

IST-2001-34732, SWAD-Europe

W3C Semantic Web Advanced Development for Europe

Libby Miller, Dan Brickley, Kate Sharp, Dave Reynolds, Stephen Buswell, Martin Pike, Michael Wilson, Brian Matthews

Report version 1
Report preparation date: 2004-12-23
Classification: Public report
Contract Start Date: 2002-05-01
Project Co-ordinator: University of Bristol (ILRT)
Partners: ERCIM (W3C), HP Laboratories, CCLRC, Stilo

SWAD-Europe Consortium

European Research Consortium for Informatics and Mathematics (ERCIM/W3C)
2004, route des Lucioles
BP 93
06902 Sophia Antipolis Cedex
France
⊣ http://www.w3.org/
Dan Brickley (danbri@w3.org), SWAD-Europe Project Director

Institute for Learning and Research Technology (ILRT)
University of Bristol
8-10 Berkeley Square
Bristol BS8 1HH, UK
⊣ http://www.ilrt.bris.ac.uk/
Libby Miller (libby.miller@bristol.ac.uk), Coordinator

HP Laboratories Bristol
Filton Road
Stoke Gifford
Bristol BS34 8QZ
U.K.
⊣ http://www.hpl.hp.com/bristol/
Dave Reynolds (dave.reynolds at hp.com)

CCLRC Rutherford Appleton Laboratory
Chilton
Didcot
Oxfordshire
OX11 0QX
UK
⊣ http://www.cclrc.ac.uk/
Michael Wilson (M.D.Wilson at rl.ac.uk)
Stilo (UK) Head Office
Windmill Hill Business Park
Whitehill Way
Swindon
SN5 6QR
United Kingdom
⊣ http://www.stilo.com/
contact@stilo.com

Contents

Acknowledgements

SWAD-Europe has had two project officers over the duration of the project, Brian Macklin, and Hans-Georg Stork. We would like to thank them for their patience and kind help during the project.

Introduction

This report summarises the activities, results and achievements of the SWAD-Europe project, which was funded for 30 months between May 2002 and October 2004. SWAD-Europe was a Combined RTD and Demonstration project, contract number IST-2001-34732, funded by the European Commission under the 7th call of the Semantic Web technologies area of Framework Programme 5. The project website is ⊣ http://www.w3.org/2001/sw/Europe/.

The aim of the project was to support the World Wide Web Consortium (W3C) in its Semantic Web activities in Europe through a mixture of research, demonstrations, outreach and dissemination. The W3C (via its European host, ERCIM) was the lead technical partner in the project and the University of Bristol (via the ILRT department) provided administrative coordination.

The W3C is a membership organisation created in 1994 to lead the Web to its full potential. Its work includes the creation of many Web standards, including HTML, XML (eXtensible Markup Language), RDF (Resource Description Framework) and SVG (Scaleable Vector Graphics). SWAD-Europe's focus was on the Semantic Web area within W3C which includes the RDF and OWL (Web

Ontology Language) standards. A main aim of SWAD-Europe was to advance understanding of the implementation of these standards through conversations with Web developers, particularly through the Semantic Web Interest Group, an open community of those interested in the Semantic Web. SWAD-Europe also aimed to connect these Semantic Web activities to other Web technologies such as XML and SVG.

The sections below discuss the project's activities grouped into the following sections:

- Community Building (WP 3)
- XML and RDF Mapping and Modelling (WP 4, 5, 6)
- Information Integration Demonstrators (WP 12.1)
- Semantic Web Thesauri (WP 8)
- Semantic Web Tools Support (WP 7, 10, 12.4)
- Trust and the Semantic Web (WP 11)

1. Project Overview

The SWAD-Europe project ran from May 2002 to October 2004, and aimed to support W3C's Semantic Web initiative in Europe, providing targeted research, demonstrations and outreach to ensure Semantic Web technologies move into the mainstream of networked computing.

SWAD-Europe originated within W3C's Semantic Web Activity, extending the approach taken in previous SWAD work at MIT to focus on collaborative work undertaken alongside the World Wide Web Consortium's RDF (now Semantic Web) Interest Group. While the project was about and in support of W3C's Semantic Web Activity, it was managed independently. The project was created to complement and feed into 'standards track' work within the World Wide Web Consortium (W3C). SWAD-Europe was a collaboration between the World Wide Web Consortium in Europe, hosted by ERCIM, research institutions at the Central Laboratory of the Research Councils (CLRC) and the University of Bristol (ILRT), and the companies Stilo International Plc and Hewlett-Packard Ltd.

A brief overview of the achievements of the project is provided immediately below, more detailed discussion is provided in section 2, and a table listing all the project deliverables is in the appendix. All project deliverables are available from the reports page of the project website: ⊣ http://www.w3.org/2001/sw/Europe/reports/intro.html.

Community Building (WP 3)

The aim of this part of the project was to increase participation in and knowledge of the Semantic Web among the target groups of:

- Internet, Web and Open Source developer communities
- Academic and Research Community
- Content and Tool Producers
- Industry and Commerce

Work included online presence and scheduled IRC discussions; t-shirt and postcard distribution; answering frequently asked questions on mailing lists and maintaining the project wiki and weblog. The work focussed around the Semantic Web Interest Group at W3C, which is chaired by Dan Brickley, also the Director of SWAD-Europe.

The main deliverables in this area were eleven workshops in eight countries. The workshops were themed in response to topics of interest in the community

and were held on the following topics:

- Dublin Core and RDF, in Florence, Italy (D3.4)
- EARL / Image Annotation, in Bristol, UK (D3.6)
- Semantic Web Calendaring, in Bristol, UK (D3.7)
- Geographic Information Management Virtual workshop (D3.9)
- Semantic Web Storage and Retrieval, in Amsterdam, Netherlands (D3.11)
- Image Description and Annotation, in Madrid, Spain (D3.12)
- Introduction to the Semantic Web (held in Spanish), in Madrid, Spain (D3.13)
- Metadata in a multilingual world, in Copenhagen, Denmark (D3.14)
- Semantic Web in Latin America (held in Spanish), in Buenos Aires, Argentina (3.19)
- Geospatial information, in Budapest, Hungary (D3.20)
- Final workshop: Friend of a Friend, Social Networking and the Semantic Web, in Galway, Ireland (D3.16)

XML and RDF Mapping and Modelling (WP 4, 5, 6)

Three workpackages were aimed at integrating Semantic Web and XML technologies including SOAP and Web Services. The main deliverables were:

- A comparison of RDF and SOAP approaches to encoding graphs in XML (D4.1)
- A report on the architecture of a working SOAP-based "Semantic Web Service" (D4.2)
- A survey of schema annotation and mapping technology (D5.1)
- A demonstrator showing the extraction of RDF from annotated XML (D5.2)
- Testcases and demonstrator comparing XSLT and N3 as rules languages for MathML data (D5.3)
- A report describing an approach to mapping ontologies and XML schemas (D6.1)
- A report describing three use cases where RDF/OWL has a role to play together with other XML syntaxes in encoding data for information capture, manipulation and dissemination (D6.2)
- Software and documentation showing an approach to the problem of unifying related but different types of information encoded in different formats and residing in potentially different repositories (D6.3)

Information Integration Demonstrators (WP 12.1)

Two high specification demonstrators were produced in this workpackage. The demonstrator topics were chosen during the project for maximum topicality and effectiveness in the aim of showing realistic application-driven tests of Semantic Web technology. The demonstrator topics chosen were a Semantic Web weblogging tool and a Semantic Web portal, and a report on the reasons for these decisions is also available. The main deliverables were:

- Open demonstrators: Semantic web applications - analysis and selection (D12.1.1)
- Semantic Blogging demonstrator (D12.1.4)
- Semantic Portals Demonstrator (D12.1.6)
- Report on lesson learned in creating the demonstrators (D12.1.8)

Semantic Web thesauri (WP 8)

This workpackage aimed to reconnect the Semantic Web and Digital Libraries communities by specifying an RDF model for thesaurus data, based on discussions within these communities and suggesting approaches to mapping thesauri to OWL ontologies. The work included the development of SKOS-Core, an RDF schema for thesauri, and associated explanatory documents:

- A guide to using the SKOS-Core 1.0 RDF schema for thesauri (D8.1)
- A review of existing work (seven schemas) on RDF-based systems for encoding thesaurus data (D8.2)
- An introduction and guide to using SKOS-Core and SKOS-Mapping RDF schemas for encoding of multilingual thesauri (D8.3)
- An introduction and guide to the RDF thesaurus vocabulary extension for defining inter-thesaurus mappings and equivalence relationships (D8.4)
- A report describing an RDF encoding of the Physics and Astronomy Classification Scheme (PACS). (D8.5)
- A report describing the features of OWL that are used in the SKOS-Core 1.0 schema for thesaurus data to express additional semantics and constraints that may not be expressed with RDF schema alone (D8.6)
- A thesaurus research prototype demonstrating the SKOS schema by means of the SKOS API web service (D8.7)
- Guidelines, methods and examples for migrating current thesaurus systems to RDF based thesaurus systems (D8.8)

Semantic Web tools Support (WP 7, 10, 12.4)

The aim of this area of work in SWAD-Europe was to build on existing work and experience in the Semantic Web community and within the project to help progress data access methods for RDF (APIs and query languages) to the stage where they could be standardized. To do this, the project undertook surveys of existing APIs and query languages, implementations of RDF query languages, including testcases; and demonstrators showing scalability of the technologies. The focus was on near-term practical goals and grounding in running code. This work has fed directly into the W3C's Data Access working group.

Deliverables in this area:

- A report comparing existing RDF APIs and setting out the requirements for RDF APIs (D7.1)
- A report gathering together frequently asked questions about RDF query from developer mailing lists and answering them (D7.2)
- Implementations of an RDF API (D7.3) and a an RDF query language (D7.4) and associated work within the Semantic Web developer community on cross-implementation testcases for RDF query.
- A report surveying the state of Semantic Web storage for RDF / triple data using existing free software tools (D10.1)
- A report surveying the schemas used in mapping between Semantic Web data and relational databases, and mapping approaches to and from relational schemas (D10.2)
- Work on the Redland Free Software/Open Source RDF storage system written in C and documentation on implementation issues, including RDF query implementation (D10.3)
- Report and demonstrator showing access to a large scale RDF database (D12.4.1)

Trust and the Semantic Web (WP 11)

The aim of this area of work in the project was to investigate the multiple ways in which trust is relevant to the Semantic Web. There were two deliverables in this

area, but each of these has been split into multiple documents in order to do justice to this large area of research.

- A report examining some current standards for data exchange formats and protocols related to security, authorization policy and trust management (D11.1)
- A report detailing some scenarios in which RDF might be used to model trust and access control in networked systems (D11.1)
- A report analysing aspects of trust in service-oriented computing for virtual organisations and the Grid (D11.1)
- A report describing a framework for deploying trust policies on the Semantic Web (D11.1)
- An analysis of the role of recommendation systems as a basis for trust on the Semantic Web (D11.1)
- A description of an ontological description of a trust assertion, and the implementation in N3 of a trust system for an advanced electronic house scenario (D11.2)
- An RDF schema and an OWL ontology for trust statements (D11.2)
- An RDF Schema for Recommendations and Subjective Logic (D11.2)
- An RDF Schema for Trust Metrics (D11.2)
- An RDF Schema for Agent Types (D11.2)
- A report describing the use of RDF metadata in configuring Internet access from a home network (D11.2)

2. Project objectives, results and achievements

This section summarises the original objectives of SWAD-Europe and then provides a discursive overview of the project results and achievements.

The original objectives of SWAD-Europe were to

- Implement scenario-led examples showing the integration of multiple Semantic Web technologies drawing practical use cases from industry, consumer, developer perspectives.
- Develop a Semantic Web technology integration strategy that emphasises the utility of XML dialects (such as SVG, HTML, MathML, XLink) as complementary rather than competing components of the Web.
- Ensure that the European developers, citizens and content creators are kept aware of Semantic Web technology for supporting universal accessibility, device independence and internationalisation.
- Ensure that European Community is kept aware of international best practice, and that best practice within Europe is recognised internationally.
- Undertake targeted research and development in support of these objectives, and in collaboration with the wider European developer community, W3C member organisations, and related Open Source initiatives.

The overarching aim of the project was to provide, through all appropriate means, a body of answers to questions that had previously gone unanswered, and to foster grassroots communities within which such concerns are addressed. Amongst its many themes, SWAD-Europe provided detailed answers to developer questions about RDF query and storage (analysis of scalability issues; query languages; APIs), and human-oriented classification (SKOS for thesauri, bookmark sharing etc.; semantic blogging). The project's final workshop was on the theme of FOAF, Social Networking and the Semantic Web, and illustrated some of the strengths of the project, combining presentations from academic, commercial and Open Source perspectives with active collaborative work on tools

and applications.

As always with such projects, some areas proved more rewarding than others, and the emphasis of the project evolved in response to successes and opportunities. A very successful part of the project has been in the area of outreach and community-building. The SKOS work has helped to re-engage the digital library / thesaurus community with the Semantic Web initiative. The eleven workshops held as part of the project have attracted diverse participants from multiple countries and specialities, and from the research, Open Source and business communities. Other successful work in the project has included software development, in particular the leading Open Source C implementation of RDF in Redland/Raptor; and well-crafted and appealing demonstrators in the areas of Semantic Blogging and Semantic Portals, showing that Semantic Web applications can be simple, practical and easy. The pragmatic 'walk before we run' focus of the project was appreciated both by Semantic Web sceptics and by enthusiasts. Project members have also made substantial contributions to the editing and chairing of the RDF Core standards, and later helped to establish 'Semantic Web phase two' groups at W3C: the Data Access working group and the Semantic Web Best Practices and Deployment working group.

In summary, the main achievements of SWAD-Europe were:

- Semantic blogging and Semantic Web portal demonstrators created and made available online, with downloadable code and documentation
- Eleven European workshops held on aspects of Semantic Web technologies, including calendaring, image annotation, Semantic Web tools, introduction to the Semantic Web in Spanish, multilingual metadata, and social networking
- A diverse community of Semantic Web developers was developed maintained via weblog, wiki, mailing list and IRC discussions
- Significant technical input went into the RDFCore, Data Access and Semantic Web Best Practices W3C working groups, and the Semantic Web interest group
- Pre-standardisation surveys, reviews and code in the area on RDF Query and APIs were produced
- A Pre-standardisation thesaurus vocabulary and documentation was created, with community participation
- Solid, Open Source code written in the C programming language for the storage and query of RDF documents
- Open Source software demonstrating the integration of RDF and Semantic Web tools with a range of XML and Web standards
- Research on the important medium-term Semantic Web priority of trust

More detailed discussion under each project area is provided below.

Project Achievements: Community Building (WP 3)

The aim of this part of the project was to increase participation in and knowledge of the Semantic Web among the target groups: Internet, Web and Open Source developers; academics and researchers; content and tool producers; and industry and commerce. There was considerable overlap between these groups, with many academics also creating tools and content and companies creating tools.

Internet, Web and Open Source developer communities

The project reached this community principally by using the appropriate communication tools. Much of the dissemination effort was focussed around the Semantic Web Interest Group, a mailing list and IRC channel for those who already have an interest in the Semantic Web, but dissemination also occurred

more broadly in related areas such as XML developer mailing lists, SVG lists and IRC channels and Web Accessibility lists. The use of IRC (Internet Relay Chat) was particularly helpful for gathering remote developers together to provide real-time feedback and discussion on a variety of topics. 21 scheduled discussions were held on IRC on topics as diverse as image annotation, calendaring, RDF query testcases and geographical information, with most being initiated and chaired by project members. These discussions attracted participants from Europe, Japan and the United States. Agendas, logs and summaries were created so that others could use the information after the event.

The project also created a weblog and a wiki for use within the project and beyond. Weblogs (sites created using short news items with syndicated content) are a good way to communicate with a substantial section of the online community, and enable outreach well outside the Semantic Web community via syndication and inclusion into aggregators such as PlanetRDF. A big advantage of weblogs is that they enable direct feedback and participation from outside the project either via comments or on other weblogs via the trackback mechanism. SWAD-Europe used the weblog to report on results within the project, to highlight items of interest and to answer FAQs. After the end of the project the weblog is being used by the Applications and Demos taskforce of the W3C Semantic Web Best Practices working group, to describe Semantic Web applications.

Wikis are systems in which anyone can create and link new pages and add to existing pages about topics of interest in a community. A wiki was created on the W3C site for the project to use and has grown into a body of knowledge which continues to be created, augmented and used by the community.

Academic and Research Community

As well as interacting with this community in their capacities as Open Source developers, SWAD-Europe members attended all the major Semantic Web conferences, and conferences in related areas such as XML and Dublin Core. Project presentations were given on various topics, including thesauri, trust, semantic blogging, and RDF and SVG, and also introductions to the project and the Semantic Web. Promotional postcards were also distributed at conferences.

A major source of interaction with both academics and industry has been through participation in W3C working groups, including RDF Core, Data Access and Semantic Web Best Practices. Through this mechanism and through workshops, the project has also been able to coordinate with other EU projects such as Knowledge Web and the SDK Cluster. In particular, the project's final workshop on "Friend of a Friend, Social Networking and the Semantic Web" held in conjunction with DERI Galway, was run as an academic workshop with anonymously reviewed papers, and approximately 50% of the 100+ attendees were from academia.

Content and Tool Producers

SWAD-Europe encouraged and rewarded the creation of open tools, data, documentation and participation by distributing 500 promotional t-shirts to content and tool producers. The t-shirts were carefully designed to appeal to young developers while presenting an informative representation of the Semantic Web to help people explain the topic to those who were unfamiliar with it. The t-shirts were very popular and helped enhance the strength of the community. The images used on the shirt have been made into image components for others to use for promotion, under an non-restrictive Creative Commons license.

Industry and Commerce

Workshops were a major plank of the SWAD-Europe dissemination strategy, and reached many developers in multiple countries, including developers working in industry. Feedback suggests that workshops can be a very useful way for companies to gain knowledge and train their employees, and also acquire an overview of a technical topic.

SWAD-Europe also helped encourage wider uptake of Semantic Web technologies by creating and promoting the development of Open Source tools and documents. At this relatively early stage of the Semantic Web, wide adoption depends on the availability of tools and documentation for companies to experiment with without heavy investment.

Project Achievements: XML and RDF Mapping and Modelling (WP 4, 5, 6)

The work packages on XML and RDF integration developed generic tools and techniques for mapping between XML Schemas, and RDF Schemas and OWL Ontologies. A mapping language was developed to support this mapping activity, and a demonstrator built to assist the user in the development of a mapping language. This mapping could then guide the construction of XSLT transforms for practical mapping.

A further unexpected result of this work was the development of ideas, techniques and demonstrations on integrating Semantic Web technologies into the development process. Established development processes for software engineering (e.g. UML, B) and knowledge engineering (e.g. CommonKADS) were studied as methods which could be adapted for the development of distributed end user applications which could be deployed over the web using RDF and OWL for their data model, and Web Services for process deployment and control. This was explored in more detail using CommonKADS, developing Problem Solving Methods as OWL-S components for refinement onto web services. This results in a pair of well-received publications.

Project Achievements: Information Integration Demonstrators (WP 12.1)

The information integration demonstrators comprise two working, online demonstrations of Semantic Web applications which together illustrate Semantic Web usage in the publishing, collecting, integrating and organizing information.

The first demonstrator, semantic blogging, shows how to enhance a web-logging or "blogging" platform with the ability to embedded RDF description of structured and classified information items. The demonstrator shows this capability in the context of publishing bibliographic records where the blog entry itself might be a comment on some publication and the embedded structured item data gives a full bibliographic reference including classification information. This extends the blogging paradigm beyond chronological diaries. Through this approach it becomes possible to use a blog as an informal information repository and offers enhanced navigation, search and view of the structured items. The information items are embedded within the newsfeed output from the blog so that this becomes a way of publishing structured information items onto the Semantic Web. It maintains the lightweight, personal publishing paradigm of blogging while moving beyond purely text-based content.

The second demonstrator, semantic portals, shows how the Semantic Web can be used to construct rich, distributed information portals. Information is published in RDF by a community of suppliers. A portal harvests this information, integrates it (by applying optional inference rules and disambiguating references based on inverse-functional properties) and publishes it via a faceted-browsing user interface. The demonstrator, SWED (Semantic Web Environmental Directory), applies this approach to information on environmental and wildlife organizations. The key benefits of this approach to information portals are richness, sustainability and extensibility.

The "richness" arises from the use of multiple domain ontologies and thesauri to classify the published information and support a convenient browsing interface. It is also possible to publish relational links between items (between organizations in the case of SWED). The "sustainability" arises from the distributed design. Rather than having information stored in a central location in a proprietary format or database each supplier publishes the information in RDF format on their own web site. Thus the information remains under the control of the suppliers who can easily maintain and update their data and alternative portals can easily be created based on the same harvestable data. Thus there is not a single point of failure and the published information is reusable for related purposes - it is easy to create value added portals which extend the published information by linking it to new information sources. The "extensibility" arises out of the nature of the RDF data model in particular the fact that it is statement-based rather than object-based and that it makes an open world assumption. This makes it easy for third parties to publish additional relational links, classifications and properties of existing items as additional RDF statements, referring the same items. This allows the community to organize and enrich the published information without need to control or rewrite the base data. Similarly the information providers themselves, as they find a need to publish additional properties for items, can do so without breaking the existing data and tools.

We can view these demonstrators at several levels - as illustrations of the Semantic Web, as separate interesting applications in their own right and as early components of a more general informal information management toolkit.

The primary purpose of the work was to illustrate the nature of the Semantic Web and to help developers get started working with it. This goal has been met. Between them the two demonstrators successfully illustrate the key features of the Semantic Web - an extensible semi-structure data model plus the ability to publish and reuse vocabularies combined with a web-compatible decentralized model. Both demonstrators manage semi-structured data, index it using thesauri and ontologies and do so in a distributed web setting. We have evidence of groups, such as the Natural History Museum, for whom the demonstrators have enabled them to better understand the Semantic Web, see it as a practical technology and begin explore applications of it that would not have happened without the demonstrators.

The two applications were chosen primarily to be good Semantic Web demonstrators - to illustrate the key features of the Semantic Web approach in simple, practical settings that are fairly easy to communicate. However, the demonstrators are also useful applications in their own right. The semantic blogging demonstrator itself is a pilot demonstration rather than a routinely used service but several groups, including ourselves, are exploring developing real services based on the ideas (and in some cases based on the software developed as part of the project). The semantic portals demonstrator, SWED, is a live service though only has modest data volumes in it at the time of writing. However, the capability is seen as valuable enough that the UK Environment Council have expressed interest in helping to create a consortium to maintain the service beyond the end of the SWAD-E project and potentially expand it to wider coverage.

Finally the two demonstrators together provide an interesting toolkit for development of information management applications. A semantic blog can be used to quickly capture information such as references to web pages, news items, papers. A semantic portal can harvest information from a set of known blogs and provide a faceted browsing view over the collection of blogs. We have built a knowledge management demonstrator which illustrates how this approach allows the combination of informal personal information captured in blogs with classified document information published in the portal. In addition to combining the overall functionality in this way the software used to create the two demonstrators has all been made available open source as part of the SWAD-E outputs. Components of the software can be reused in other ways. In particular the semantic portal toolkit uses a template-driven approach to rendering RDF data into web pages. The same template engine can be reused for applications which require a different

information from the faceted-browsing paradigm used for the SWED demonstrator.

Project Achievements: Semantic Web Thesauri (WP 8)

The work on Semantic Web Thesaurus representation and Methods resulted in a set of deliverables presenting a proposed format for representing thesaurus information as a standard in the Semantic Web. Further an online web service was produced as a demonstrator of the tools which such a format could support.

The core of the work has been the development an RDF vocabulary for representing thesaurus structures, called SKOS-Core (the Simple Knowledge Organisation System). This is now well-established, and forms the basis of further work in other projects. Further work develops this core vocabulary into more sophisticated use of thesauri, with a proposal for a vocabulary to support mappings between thesauri, and also a discussion on methods to represent multi-lingual thesauri. Further consideration has been given to the relationship between thesauri and OWL; although progress has been made, this is still an ongoing issue. This core work was supported by the development of an API and web-service based thesaurus browsing tool to allow programmatic and human access to thesauri on the Web, and the development of several demonstration thesauri, notably GEMET.

Finally it is notable how the thesaurus workpackage in the SWAD-Europe project brought together and motivated a community. Thesauri and other classification schemes are well established in the digital library world, but is has been recognised by this community that there is an urgent need for common standards for representation exchange and access to terminology data. The SWAD-Europe thesaurus discussion group has therefore caused much interest in that community and has lead to the participation of members of that group. This includes those revising British Standards in the area, which ultimately will result in a revised ISO standard.

The primary purpose of the work was to show existing information scientists how to migrate from their current systems to RDF based Semantic Web ones, and motivate this with examples and documentation. The level of ongoing interest and participation in the work of this WP is evidence that that this was achieved to the full.

Project Achievements: Semantic Web Tools Support (WP 7, 10, 12.4)

The goal of this area of work was to build on existing work and experience in the Semantic Web community and within the project to help get various data access methods for RDF (APIs and query languages) to the stage where they could be standardized. This was achieved through a mixture of research, implementation and community interaction.

Project members spent a considerable amount of time assessing and reporting on the state of the art in RDF APIs and query. Reports were produced comparing and evaluating the APIs used in existing RDF tools including Jena, Sesame and Redland and evaluating methods of RDF storage in relational databases, contributing to the body of information available for the community to use for implementations in this area.

The main success in this area of the project has been input into the W3C's Data Access Working Group which started towards the end of the project. A series of IRC discussions was held with developers from Jena, Sesame, RDF Store and SWAD-Europe to create a series of testcases for evaluating RDF query implementations, including a resultset format and a schema for describing queries in a language-neutral way. This work fed directly into the Data Access group and so helped achieved the aim of moving RDF query to the stage where it could be standardised. In addition, SWAD-Europe members participated directly in the

group, taking this work forward and providing an implementation of the query language in the process of standardisation, SPARQL.

SWAD-Europe also contributed to the development of Redland, a mature open Source/Free Software RDF storage and application framework written in C. Work on Redland began before the start of the project and will continue afterwards: the work funded by SWAD-Europe included major improvements in scalability, the addition of query language support and many bindings to different programming languages, and substantial documentation and online help for users. Redland is a good example of a case where SWAD-Europe was able to contribute to the longevity of a popular Open Source project to the benefit of the Semantic Web community.

Project Achievements: Trust and the Semantic Web (WP 11)

The goal of this deliverable was to explore the interaction between various security and trust-related mechanisms, and to identify enough of the purpose and content of these existing standards to indicate how they might be integrated in a Semantic Web trust and policy management framework. As such, the scope of this deliverable was extremely wide and at the beginning of the project relatively under-defined. Consequently, as a major part of the early work of the workpackage we undertook a major survey of the state of the art and standard used to support security within the web; this has been disseminated and used as a significant resource in the security community.

To further refine and identify problems in this field, we also undertook to identify a set of scenarios which form requirements for trust and security on the Web, including access control, e-commerce, and network management.

It soon became apparent that since the beginning of the project, large amount of interest was emerging on the use of trust in the Semantic Web, partly stimulated by the dissemination of our early work on some websites. Two major themes emerged: Community based trust-networks, which are looking at the transfer of trust in communities which operate over the Semantic Web, using community vocabularies such as FOAF; and trust to support virtual collaborations are set up between entities who are participating in some work towards a common goal across organisational and national barriers.

The first theme became the subject of much research outside the project, so we provided links to it and concentrated on the other theme. Closely related to web and grid services, in this theme virtual organisations wish to set up a "virtual firewall" to ensure that they can control the information flows between members of the virtual organisation. We presented use scenarios for the use of trust in virtual organisations across the Semantic Web, and began to develop a framework, laying out the issues facing the establishment and use of trust in VOs.

As demonstrators, we have developed the use of trust in a home network configuration management scenario, and developed an RDF vocabulary supporting the use subjective logic to support reputation management in the Semantic Web.

3. Methodologies

The methodology used throughout the project drew upon approaches that exploit the potential of the Web, and of Web technologies. In particular, Open Source-style collaboration, and test-case based development practices (related to those from the Extreme Programming community) were combined with the use of W3C standards to provide a lightweight approach to coordinating a highly diverse set of project deliverables.

The project itself was set up to help develop and promote W3C's Resource Description Framework (RDF) and Web Ontology (OWL) technologies. These

technologies formed the basis for all the technical work conducted within the project. In many ways, RDF proved central to the technical methodology of the project. RDF itself was designed by W3C as a common, neutral platform for information exchange amongst "loosely coupled" parties. It supports independent extension, incremental development practices (allowing for rapid prototyping and review cycles) and allows user interface design to be decoupled from database and storage systems. The project helped to progress the design of new RDF technologies, showcase RDF-based applications, and improve the stability and scalability of RDF tools. By actually using as well as talking about RDF technology, RDF was shown to be a valuable tool for project coordination methodology. It allows independent workpackage managers to share tools and technology without creating an overly-rigid, highly-interdependent project plan.

A similar approach was adopted for the project's relationship with the outside world. Developers from the project partners were actively engaged with the communities surrounding W3C's Semantic Web initiative. Through Workshops and online presence, SWAD-Europe's own development activities were highly visible within the wider community. This methodology, coupled with test-cases and Open Source-licensed code, allowed for a very rapid turnaround of ideas and prototypes.

4. European Added Value

The European dimension of the project has been very important. Our commitment was that we would hold eight workshops in European countries over the course of the project, and we have exceeded that, holding ten within Europe (including Spain, Denmark, Netherlands, Ireland and Hungary) and one in South America (Argentina). Two of these workshops were held entirely in Spanish, and this has contributed to a healthy Spanish-speaking Semantic Web community. The aim was to improve knowledge of the Semantic Web within Europe and to improve the availability of Semantic Web information in languages other than English, and the project has enabled some steps to be made in this direction.

SWAD-Europe has been committed to a major European role in the creation of the Semantic Web: it has held presentations and workshops in multiple languages, physical outreach in eight European countries, and virtual dissemination reaching many more. SWAD-Europe has increased the profile of Europe within the Semantic Web community through online presence and the creation and maintenance of community fora in the worldwide Semantic Web community. The project has also worked in the area of vocabularies to support multilingual thesauri and image annotation. The use of a multilingual thesaurus can provide a very cost effective means to re-purpose content to new audiences across national and linguistic boundaries.

5. Outlook

Outputs from SWAD-Europe will be taken forward in a number of different ways by several different groups. The website and reports will be held indefinitely on the W3C site and will provide a valuable resource for developers after the end of the project. Various outputs including the work on thesauri (SKOS) and RDF query testcases will feed into W3C Working Groups (Semantic Web Best Practices and Data Access Groups respectively). The demonstrators are being taken forward by a number of groups: For semantic blogging, a number of related projects are known to be underway and one SME is planning a product release in this area. The semantic portals demonstrator and related applications are likely to be pursed by consortia led by the UK Environment Council and by the Natural History Museum. HP is planning internal pilot applications of the technology to

knowledge management. Finally, developer workshops, online presence and t-shirt distribution have helped to reach young European developers and thereby contributed to better knowledge of the Semantic Web in Europe.

Outlook: W3C

ERCIM is the World Wide Web Consortium (W3C) host site in Europe. The SWAD-Europe project was initially proposed to support and promote W3C's Semantic Web initiative in Europe, and to help ensure European input into the evolving Semantic Web. The goals of the project of ERCIM/W3C as project partner are consequently very closely aligned, and the project's objectives (→ http://www.w3.org/2001/sw/Europe/plan/proj_objectives.html) closely reflect those of ERCIM regarding the development of Web technology in general, and W3C-led standards work in particular. Specifically, SWAD-Europe's support of the Semantic Web developer community in Europe has been of particular value. A public article elaborating on the key achievements of the project was circulated in ERCIM's newsletter in Oct 2004 (→ http://www.ercim.org/publication/Ercim_News/enw59/brickley.html).

Outlook: ILRT

ILRT and the University of Bristol have benefited from participation in SWAD-Europe primarily because of the opportunities the project provided to apply Semantic Web technologies in other sectors such as digital libraries and web development. The Semantic Web group has been able to apply knowledge gained of Semantic Web, Web Services and thesauri applications to these very practical domains. Two new projects at ILRT use the knowledge gained in SWAD-Europe. The PARIP project (Practice as Research in Performance) is an Arts and Humanities Research Board funded project visualizes the connections between researchers, using Web Services and community vocabularies. IUGO, funded by JISC (UK's Joint Information Funding Council) is a project to connect conference materials together to maximise the value of investment in conference participation; again, it will use community-driven vocabularies and a Semantic Web framework for this very practical purpose.

The Semantic Web group at ILRT has greatly improved its knowledge and contacts in the Semantic Web area both through interaction with the partners in the project and through broad participation in workshops, conferences and online discussions. This will enable further projects to be developed in this area, both research and in more practical areas.

Outlook: HP Labs

The information integration demonstrators have benefited HP already in several ways. Firstly, they have met their original goals of helping to illustrate the nature of the Semantic Web to developers. HP has used them in discussions with both customers and with internal HP groups. They have been effective at both illustrating that interesting applications are possible and at making the Semantic Web concepts more tangible. An example of the resulting impact is the National Collections UK (NCUK) initiative begun by the Natural History Museum. This is an attempt to provide rich integrated descriptions of natural world collections. As a result of seeing the demonstrators they now wish to experiment with the use of Semantic Web technologies to enhance the extensibility and sustainability of NCUK - something that would not have happened without the demonstrators.

Secondly, HP has found customers interested in the ontology-driven browsing interface used in the semantic portals demonstrator. In particular, at the time of writing we have already won a competitive tender for a pilot educational

information portal on the basis of the demonstrators and will be using the software developed under SWAD-E in this contract.

In terms of future exploitation HP will continue to use them as demonstrators and communication aids in explaining the nature and potential of the Semantic Web and expect that to result in further internal and external collaborations leading to new applications or product enhancements. HP also intends to exploit the demonstrators directly. Internally to HP we expect to pilot the use of both demonstrators in combination to support information knowledge management practices within and across work groups. We also anticipate other groups both within and outside the consortium exploiting the semantic portals software for building structure information portals.

Outlook: CCLRC

The benefit to CCLRC of the SWAD-Europe project is firstly the development of tools and techniques which are of direct benefit to CCLRC's prime line of business, providing support to the UK Science and Engineering base, and secondly developing experience and perspectives on areas which are of application of the Semantic Web which are foreseen as being vital to the future IT infrastructure of CCLRC. For example, in the first category, the work on thesaurus formats is of direct relevance to the IT infrastructure of UK Science, providing vital tools for digital libraries, and for the long-term preservation and utility of scientific data. In the second category, the more long-range development of ideas in trust and virtual organisations across the Semantic Web, will be important in the overall infrastructure supporting science in the UK and Europe, especially in the development of the Next Generation GRID. The number of projects which CCLRC is now involved in which is taking these technologies forward is indicative of the benefit and reputation which CCLRC has garnered through being involved in the SWAD-Europe project.

XML and RDF Mapping and Modelling

The work on mapping and modelling begun in this workpackage has possibilities for further research and practical use. The problem of mapping between different information representations, such as XML formats, or Database Schemas is a very general one, which causes problems across industry and commerce. The ideas explored in this workpackage of using OWL as a high-level representation of meaning to control mapping are thus very exploitable. For example, CCLRC is using these ideas in the context of oceanographic data in the EC project Marine XML.

The problem of integrating existing development methods into the Semantic Web is also of wide interest, especially to the knowledge engineering community, and the papers presented from this workpackage did excite a great deal of interest. CCLRC will continue to research in this area, and will use the techniques developed in further projects.

Semantic Web Thesaurus Representation and Methods

The work on setting standards for formats and APIs for Knowledge Organisation Systems within the Semantic Web developed in this workpackage is now being continued within the W3C Semantic Web Best Practices and Deployment Working Group. Within this working group, the Task Force on Porting Thesauri to RDF and OWL is coordinated by CCLRC. SWAD-Europe's thesaurus webpage and mailing list will be taken over and maintained by this group. The continued participation of member of the NKOS community (→ http://nkos.slis.kent.edu/) and members of the BSi committee IDT/2/2 working on revised national and international standards should ensure that the work of this

work package continues to have an impact.

The SKOS format itself is being used in other thesauri and projects. It will continue to be used in systems developed under SWAD-Europe, notably the Semantic Portals developed by HP Labs. Further, it has been adopted by the Simile project (⊣ http://simile.mit.edu/, the GEMET thesaurus ⊣ http://www.eionet.eu.int/GEMET and by the W3C itself for its glossary (⊣ http://www.w3.org/2003/glossary/).

In terms of future exploitation CCLRC will continue to develop the format, and intends to use it within external projects, for example to support its own intranet, and electronic publications system. It will continue to seek further opportunities to develop and exploit the work within this workpackage in further research and development projects and in production.

Distributed trust systems

Trust is and continues to be a theme of vital interest to the Semantic Web community which has now attracted a great deal of attention from researchers and developer, partly as a result of the SWAD-Europe project.

The work on distributed trust systems in communities is continuing within the wider Semantic Web community, building vocabularies for trust and systems for analysing it in the context of the ongoing FOAF project, partly supported by SWAD-Europe.

CCLRC is continuing its work on using and developing Semantic Web tools, vocabularies and techniques within the context of virtual organisations. CCLRC is promoting the work in the context of the iTrust WG, and intends to present the next conference in the series. Further work is being undertaken in two EC FP6 integrated projects: TrustCoM, and AkoGriMo. In both of these projects, trust and security considerations are being analysed and supported in the context of (different aspects of) virtual organisations. CCLRC is also leading the trust and security workpackage in the EC FP6 NoE CoreGRID, which will continue to support the website from SWAD-Europe.

Outlook: Stilo

The results and achievements of the project have benefited Stilo in two related ways. In general, the experience gained in the various tools and technologies improves the quality of product design and implementation and also supports Stilo's professional services and consultancy activities. More specifically, inclusion of results and their architectural principles in product offerings can improve time to market .

Working on the tools and techniques for exchanging information between the RDF and XML paradigms has given us valuable insight into the design of collaborating multi-namespace schema structures and ontologies. In particular, the combination of RDF and XML, together with a mapping framework allowing semantics-preserving transformations between formats is a powerful technique which we will use to enhance the functionality of our Knowledge Engineering platform, SophX.

Experience gained and technologies utilized in developing the dynamic forms-based user interface for the SWAD-E ontology editing web application will form the basis of an ongoing forms-based model-editing development.

6. Conclusions

The SWAD-Europe project has shown that it is both important and rewarding to provide an environment where members of the larger community that surrounds

W3C can interact and collaboratively explore the practical issues around Web technology. The formal work of the W3C is based on small, highly focussed Working Groups where individuals commit a lot of time to creating new Web standards. SWAD-Europe's primary contribution was to help create a supportive background environment for such work, by allowing a much larger number of geographically-dispersed individuals to participate (through email, IRC, workshops and the Web) in the Semantic Web initiative. The project was, in the Semantic Web scene, unique in its emphasis on the practical and Web aspects of "Semantic Web" for a Web developer audience. The support that SWAD-Europe provided to the RDF and Semantic Web Interest Group was an important exploratory step towards a model for wider participation in Web standardisation work, showing that W3C's successful Working Group-led approach can be complimented by a broader participation model which allows individual researchers and implementers to make real contributions to the deployment of existing standards and to the creation of new ones. The challenge for the future is to work towards a Web in which all European research efforts contribute to the communities which underpin the evolution of Web standards.

Annex: Deliverables (tabular)

All public deliverables are available in full from ⊣ http://www.w3.org/2001/sw/Europe/reports/intro.html.

Deliverables Table

Project Number: IST-2001-34732
Project Acronym: SWAD-Europe
Title: W3C Semantic Web Advanced Development for Europe

Number	Title	Type	Classification	Date Due	Issue Date
1.12	Annual Report May 2003-May 2004	R	Pub	Jan 04	May 04
2.1	Project Technical Plan	R	Pub	Aug 02	Jan 03
2.2	Project Findings Input	R	Pub	Apr 03	Nov 03
2.3	Advisory meeting report	R	Pub	Oct 03	Feb 04
3.1	Project fact sheet	R	Pub	May 02	May 03
3.2	Project Website	R	Pub	May 02	May 02
3.3	Project mailing list	R	Pub	May 02	May 02
3.4	Initial workshop report	W	Pub	Aug 02	Dec 02
3.5	Dissemination and Use Plan	R	Pub	Jul 02	Jul 02
3.6	First Developer Workshop Report (EARL/Image Annotations)	W	Pub	Sept 02	Sept 02
3.7	Second Developer Workshop Report (Semantic Web Calendaring)	W	Pub	Dec 02	Dec 02
3.8	Project Presentation	R	Pub	Feb 03	July 03
3.9	Third Developer Workshop Report (Geographic Information Management)	W	Pub	Apr 03	Apr 03
3.10	Dissemination and Use Plan	R	Pub	Jul 03	Jul 03

3.11	Fourth Developer Workshop report: (Semantic Web Storage and Retrieval)	W	Pub	Aug 03	Dec 03
3.12	Fifth Developer Workshop report: Image description and annotation	W	Pub	Dec 03	June 04
3.13	Sixth Developer Workshop report: Introduction to the Semantic Web	W	Pub	Apr 04	June 04
3.14	Seventh Developer Workshop report: Metadata in a multilingual world	W	Pub	Aug 04	Sep 04
3.15	Technological Implementation Plan	R	Pub	Sep 04	Oct 04
3.16	Final Workshop Report	W	Pub	Oct 04	Oct 04
3.17	Web Showcase	R	Pub	Oct 04	Oct 04
3.18	RDF Query Standardisation	R	Pub	N/A	Sep 04
3.19	Eighth Developer Workshop report: Semantic Web in Latin America	W	Pub	N/A	Oct 04
3.20	Ninth Developer Workshop report: Geospatial information on the Semantic Web	W	Pub	N/A	Oct 04
4.1	RDF/XML and SOAP for Web Data Encoding	R	Pub	Aug 02	Jan 03
4.2	Semantic Web and Web Services: RDF/XML and SOAP for Web Data	R	Pub	Dec 02	Feb 03
5.1	Schema Technology Survey	R	Pub	Oct 02	Jul 03
5.2	Extracting Semantics from XML Structure	R	Pub	Feb 03	Oct 03
5.3	RDF/XML Test cases for RDF Logic, Web Ontology and Maths content	S	Pub	Jul 03	Oct 03
6.1	XML and Semantic Web: Ontology-Schema Mapping Tool	S	Pub	Dec 03	June 04
6.2	Service illustrating use cases for combining MathML, RDF and Web Ontology	S	Pub	Jan 04	June 04
6.3	Documentation for both prototypes, with pointers into relevant literature	R	Pub	Feb 04	June 04
7.1	RDF API requirements and comparison	R	Pub	Oct 02	Feb 03
7.2	Report on Query languages	R	Pub	Dec 02	March 03
7.3	Public release of reference implementation of an RDF API	S	Pub	Apr 03	Oct 03
7.4	Public release of a "strawman" query language implementation	S	Pub	Aug 03	Oct 03
8.1	Core RDF Vocabularies for Thesauri	R	Pub	Aug 03	Oct 03
8.2	Thesaurus Ontologies Report	R	Pub	Jun 02	Oct 03
8.3	RDF Encoding of Multilingual Thesauri	R	Pub	Oct 03	Jan 04
8.4	Inter-Thesaurus Mapping	R	Pub	Dec 03	Feb 04
8.5	RDF Encoding of Classification Schemes	R	Pub	Mar 04	June 04
8.6	OWL Ontology for Thesaurus Data	R	Pub	Apr 04	June 04
8.7	RDF Thesaurus Prototype	S	Pub	May 04	Sept 04
8.8	Migrating Thesauri to the Semantic Web	R	Pub	Jul 04	Sept 04
9.1	Visualisation and Accessibility - Initial Analysis	R	Pub	Apr 03	Jul 03
9.2	Accessibility Adoption Report	R	Pub	Dec 03	Dec 03
9.3	Semantic Web tools to help authoring: A Semantic Web image annotation tool	S	Pub	Apr 04	Sept 04

10.1	Scalability and Storage: Survey of Free Software / Open Source RDF	R	Pub	Jul 02	Jul 02
10.2	Mapping data from RDBMS	R	Pub	Sept 02	Jan 03
10.3	Implementation report on scalable Free Software/Open Source RDF storage	S	Pub	Feb 03	Feb 03
11.1	Recommendation systems as a basis for trust on the Semantic Web	R	Pub	Mar 04	June 04
11.2	Trust implementation	S	Pub	June 04	Oct 04
12.1.1	Open demonstrators: Semantic web applications - analysis and selection	R	Pub	Oct 02	Oct 02
12.1.2	Semantic blogging and bibliographies - requirements specification	R	Pub	Dec 02	Dec 02
12.1.3	Demo 1 Implementation	S	Pub	June 03	Sept 03
12.1.4	Semantic Blogging - Lessons Learnt	R	Pub	Aug 03	Dec 03
12.1.5	Semantic Portals - Requirements Specification	R	Pub	Oct 03	Oct 03
12.1.6	Semantic Portals Demonstrator	S	Pub	Apr 04	June 04
12.1.7	Semantic Portals Demonstrator - Lessons Learnt	R	Pub	June 04	Aug 04
12.1.8	Demonstrators - Lessons Learnt	R	Pub	Aug 04	Aug 04
12.2.1	Annotation Demo Single Schema	R	Pub	Sept 02	Sept 02
12.2.2	Annotation Demonstrator Server Report	R	Pub	Oct 02	Jan 03
12.3.1	Transformation libraries for RDF	S	Pub	June 03	Jan 04
12.3.2	Transformation libraries report	R	Pub	Oct 03	Jan 04
12.4.1	Large Scale Resource Discovery and Presentation Demonstrator	S	Pub	Aug 03	Sept 04