

SWAD-Europe Deliverable 3.18: RDF Query Standardisation

Project name:

Semantic Web Advanced Development for Europe (SWAD-Europe)

Project Number:

IST-2001-34732

Workpackage name:

3 Dissemination and Exploitation

Workpackage description:

<http://www.w3.org/2001/sw/Europe/plan/workpackages/live/esw-wp-3.html>

Deliverable title:

3.18 RDF Query Standardisation

URI:

http://www.w3.org/2001/sw/Europe/reports/rdf_std_query_report/

Author:

Dave Beckett

Abstract:

This report describes the work done on supporting RDF Query Standardisation (March 2004-September 2004), based on earlier project work (reports, workshops, meetings, software) in the form of participation in the W3C RDF DAWG activity, development of query languages and software to evaluate query language designs and implementation strategies.

STATUS:

Completed 2004-09-30.

Contents

1. [Introduction](#)
2. [RDF Query Languages](#)
3. [RDF Data Access Working Group](#)
4. [Rasqal Software](#)
5. [Demonstrations](#)
6. [Future work](#)
7. [References](#)

1. Introduction

The background to this work is earlier software development and reporting in the area of RDF and query languages. This was reported in Workpackage 7 [\[SWADE-WP7\]](#) deliverable [\[SWADE-D7-2\]](#) *Databases, Query, API, Interfaces: report on Query languages* by Libby Miller, ILRT along with *Strawman query language implementation (Squish QL)* reported on in D7.4 [\[SWADE-D7-4\]](#) by Libby Miller, ILRT.

In addition to reports and software development the project has supported community building for RDF query via the use of several meetings discussing query use cases and testcases (see the [SWAD-Europe events page](#)) to provide evidence and concrete information.

This work in the project was one driver into showing the timeliness for standardisation activity in this area, and during the SWAD-Europe project the W3C formed an activity to progress this, the *RDF Data Access Working Group* (DAWG) [\[DAWG\]](#) beginning work in February 2004. This opportunity allowed SWAD-Europe to progress and support the query work developed, backed with practical implementation knowledge as well as using the team experience with other W3C standardisation work ([\[RDF Core WG\]](#), [\[Web Ontology WG\]](#)).

This report describes an additional deliverable supporting the RDF query standardisation work by the means of participating in the standards activity along with supporting software development.

2. RDF Query Languages

There has been a common thread to one SQL-like set of RDF query languages going from one of the earliest, *RDFDB QL* [\[RDFDB-QL\]](#) onwards to several related, derived or similar languages such as *Tinkling* [\[TINKLING\]](#), *SquishQL* [\[SWADE-D7-4\]](#), [\[SQUISH\]](#) and finally to the *RDF Data Query Language* (RDQL) [\[RDQL\]](#). Several of the earlier pieces of work were reported on or influenced by the W3C *QL'98 - Query Languages 1998* workshop [\[QL98\]](#) in November 1998.

RDQL has been widely implemented language with at least 5 different and very complete implementations in different programming languages known and systems. RDQL was co-developed by Libby Miller (ILRT, SWAD-Europe) along with Andy Seaborne (HP Labs Europe) and Alberto Reggiori (@semantics)

SWAD-Europe has supported this work with development of test cases [\[TESTCASES\]](#), implementations and discussions of the issues it raised [\[TESTCASES-REPO\]](#). It was found that after some time the core of the language was becoming very stable, there was a drift in some of the detail of the RDQL features which was an indicator that more form standardisation would be beneficial. A version of RDQL was [submitted to the W3C](#) in October 2003 by HP Labs as input to the future standards work.

3. RDF Data Access Working Group

The DAWG [\[DAWG\]](#) was chartered to start in February 2004, with the first meeting in March 2004. The main work was to create a query language and protocol and to substantially base that on existing work, rather than design from scratch. Dave Beckett from ILRT joined the DAWG to participate on behalf of SWAD-Europe.

The initial work was to form use cases for the query language and protocol which was the first part of the activity performed in the first few months of the working group including the first face-to-face meeting in Leiden, Netherlands. This resulted in the *Use Cases and Requirements W3C Working Draft* [\[DAWG-UC\]](#) first published August 2004.

In July 2004 several RDQL implementors including SWADE team members met to discuss a strawman language based on RDQL to meet the DAWG use cases and requirements. This resulted in *BRQL: A Query Language for DAWG* [\[BRQL-1\]](#) which was shortly afterwards implemented in the most part by Andy Seaborne (not under SWAD-Europe). This language work was accepted as the basis of the strawman query language for DAWG at the second face-to-face meeting in July 2004. However there remained and remain many open issues

about the query language design in the large and in detail.

4. Rasqal Software

A new implementation of an RDF query system was developed, looking forward to the DAWG work. This was initially based on the W3C RDQL member submission [\[RDQL\]](#). The implementation was designed to integrate with the *Redland RDF Application Framework* [\[REDLAND\]](#) already used in SWAD-Europe for other work (see [\[SWADE-WP10\]](#), [\[SWADE-D10-1\]](#), [\[SWADE-D10-2\]](#)). This would be used to allow the query language support to be added and exported to all the systems and languages that Redland supports - including C#, Java, Perl, PHP, Python, Ruby and Tcl as well as crossing Linux, Unix, OSX and possibly Win32.

The software developed was the *Rasqal RDF Query Library* [\[RASQAL\]](#), a C library that is now part of the Redland set of libraries. It uses the Redland *Raptor RDF Parser Toolkit* [\[RAPTOR\]](#) to perform parsing and deal with web features (URIs, WWW retrieval).

Rasqal was designed to be targeted at the kinds of query language similar to RDQL, as that seemed the likely best goal for the standardisation work under the DAWG. The library as designed to separate the detail of query language syntaxes (RDQL, BRQL, SeRQL etc.) from the model of the query which was at core, a conjunctive set of triple patterns over a graph. The query engine was designed to be flexible in terms of being capable to adapt to changing designs of query processing that would likely evolve over the DAWG activity.

Rasqal has allowed experimentation with issues that have come up in the ongoing DAWG standardisation work including the important concept of data provenance (`SOURCE` in BRQL) and data aggregation tracking, previously reported on in earlier deliverables D10.1 *Scalability and Storage: Survey of Free Software / Open Source RDF storage systems* [\[SWADE-D10-1\]](#), D12.4.1 *Large Scale Resource Discovery and Presentation Demonstrator* [\[SWADE-D12-3-1\]](#) and D3.11 *Workshop on Semantic Web Storage and Retrieval* [\[SWADE-D3-11\]](#) and at this time (Sep 2004) still under discussion in the DAWG.

At this point there have been three major releases of Rasqal [\[RASQAL-RELEASE\]](#) over the period of this work (March-September 2004) that completely implements RDQL and has been already used successfully in several tools external to SWAD-Europe and been a driver for Redland, encouraging two more ports to Objective-C (OSX) and more work on porting to Win32.

Summary of the Rasqal technical work:

- Open source licensed software
- Abstract model of RDF query (sources, triples, constraints ...)
- Query execution engine implementation
- API in C to access the query, engine and results
- Internal SPIs for triple stores with Raptor or Redland
- Language bindings in Perl, Python, C# via Redland Bindings [\[REDLAND-BINDINGS\]](#) (Objective C for OSX done externally)
- Ported to Linux, Unix, OSX, Win32 (in progress)
- RDQL fully supported
- Draft support for DAWG QL based on BRQL

Rasqal is available from the *Rasqal web site* [\[RASQAL\]](#) as source and binaries for Redhat and Debian Linux. Binaries for other platforms may be available separately. Rasqal requires the open source *Raptor library* [\[RAPTOR\]](#) Bindings

to other languages are available from the *Redland Bindings web site* [\[REDLAND-BINDINGS\]](#) which requires Redland additionally.

5. Demonstrations

Rasqal has been used to enhance the earlier demonstrator for large scale data (provenance, Redland Contexts) to allow user queries. This is available as one of the demonstrations under the *Large Scale Resource Discovery and Presentation Demonstrator* [\[LARGE-SCALE-DEMO\]](#) as demonstration *Perform an RDQL query* over the previously crawled FOAF data. It does not at present allow use of BRQL or the developing provenance support in Rasqal as that is at this support is still being worked on in the DAWG activity.

A separate *Rasqal RDF Query demonstration* [\[RASQAL-DEMO\]](#) service allows application of a query in RDQL (soon BRQL) to be given against any RDF source of data on the web, reading the data into a Redland in-memory data store and then executing it. This is implemented with Redland's Perl API and will ship with a future release of Redland Bindings.

6. Future work

Rasqal will likely be the basis for ongoing work to support the DAWG standardisation process, and as an open source application, can be enhanced and receive patches from users, as well as being used in commercial applications (since the license is flexible).

References

[SWADE-WP7]

[\[SWAD-Europe WP 7: Databases, Query, API, Interfaces\]](#)

[SWADE-D7-2]

[\[SWAD-Europe D7.2 Databases, Query, API, Interfaces: report on Query languages\]](#), Libby Miller, ILRT, University of Bristol, 2003-04-01

[SWADE-D7-4]

[\[SWAD-Europe D7.4 Public release of a "strawman" query language implementation incorporating current best practices\]](#), Libby Miller, ILRT, University of Bristol, 2003-11-05

[DAWG]

[\[RDF Data Access Working Group \(DAWG\)\]](#)

[RDFDB-QL]

[\[rdfDB query language\]](#), R.V. Guha, 2002

[TINKLING]

[\[Tinkling, a small RDF API and Query Language implementation\]](#), Libby Miller, ILRT, University of Bristol, July 2002

[SQUISH]

[\[Three Implementations of SquishQL, a Simple RDF Query Language\]](#), Libby Miller, Andy Seaborne, Alberto Reggiori, Proceedings of First International Semantic Web Conference, Sardinia, Italy, June 9-12 2002

[RDQL]

[\[RDQL - A Query Language for RDF\]](#), W3C Member Submission, Andy Seaborne, HP Labs Bristol, 9 January 2004

[QL98]

- QL'98 - Query Languages 1998, W3C Workshop, 15 November 1998
- [TESTCASES]**
- Summary of RDF query tests work, February-May 2003, Libby Miller, Dan Brickley and others, ILRT, University of Bristol, March 2004.
- [TESTCASES-REPO]**
- RDF Query (and Rule) Testcase Repository, Dan Brickley, W3C, March 2003
- [DAWG-UC]**
- RDF Data Access Use Cases and Requirements, W3C Working Draft, Kendall Grant Clark, 2 August 2004
- [BRQL-1]**
- BROL: A Query Language for DAWG, by Dave Beckett, Chris Dollin, Nicholas Gibbins, Steve Harris, Andy Seaborne, Damian Steer, July 2004
- [REDLAND]**
- Redland RDF Application Framework, Dave Beckett, ILRT, University of Bristol. URL <– <http://librdf.org/>>
- [SWADE-WP10]**
- SWAD-Europe WP10: Tools for Semantic Web Scalability and Storage
- [SWADE-D10-1]**
- SWAD-Europe D10.1 Scalability and Storage: Survey of Free Software / Open Source RDF storage systems, Dave Beckett, ILRT, University of Bristol, 2003-02-17.
- [SWADE-D10-2]**
- SWAD-Europe D10.2 Mapping data from RDBMS, Dave Beckett and Jan Grant, ILRT, University of Bristol, 2003-02-18.
- [RASQAL]**
- Rasqal RDF Query, Dave Beckett, ILRT, University of Bristol. URL <– <http://librdf.org/rasqal/>>
- [RAPTOR]**
- Raptor RDF Application Framework, Dave Beckett, ILRT, University of Bristol. URL <– <http://librdf.org/raptor/>>
- [SWADE-D12-4-1]**
- SWAD-Europe Deliverable 12.4.1: Large Scale Resource Discovery and Presentation Demonstrator, Dave Beckett, ILRT, University of Bristol, June 2004
- [SWADE-D3-11]**
- SWAD-Europe Deliverable 3.11: Developer Workshop Report 4 - Workshop on Semantic Web Storage and Retrieval, Dave Beckett, ILRT, University of Bristol, January 2004
- [RASQAL-RELEASE]**
- Rasqal Release Notes, Dave Beckett, ILRT, University of Bristol, 2004
- [REDLAND-BINDINGS]**
- Redland Language Bindings, Dave Beckett, ILRT, University of Bristol, 2004
- [LARGE-SCALE-DEMO]**
- Large Scale Resource Discovery and Presentation Demonstrator, Dave Beckett, ILRT, University of Bristol, 2004
- [RASQAL-DEMO]**
- Rasqal RDF Query demonstration, Dave Beckett, ILRT, University of Bristol, 2004