

# A Survey of the Effectiveness of Current Interoperability Protocols Technical Report

Jorgina K. do R. Paihama, Hussein Suleman  
Digital Libraries Laboratory, Department of Computer Science  
University of Cape Town, South Africa  
{jpaihama, hussein}@cs.uct.ac.za

## Introduction

Interoperability is the capability of different systems to communicate and exchange data with one another, using a set of predefined formats and protocols that will allow the systems to use one another's services successfully [2][12].

This report presents the results of the "Effectiveness of current interoperability protocols" survey.

The purpose of the survey was to investigate the level of understanding and use of the current interoperability protocols, and also to find out what are the best features and/or the shortcomings of each protocol from experts or people who have some knowledge/experience in working with them.

This information will be used to assess what makes a good interoperability protocol, and then design an experimental set of interoperability specifications as well as a framework to evaluate interoperability protocols in general.

## Background

There are a number of different protocols that provide interoperability between heterogeneous systems at various levels of abstraction. The protocols assessed for this particular survey are:

### *The Atom Publishing Protocol*

The Atom Publishing Protocol or APP [10] is an XML based application level protocol used to generate and edit Web resources of constantly updated websites, like news websites.

### *The Atom Syndication Format*

The Atom Syndication Format or ATOM [1] is an XML based feed format for representing lists of related Web resources known as feeds.

### *The Open Archives Initiative - Object Reuse and Exchange*

The Open Archives Initiative - Object Reuse and Exchange or OAI-ORE [6] defines a set of standards for the description and exchange of compound digital objects also called aggregations, of Web resources. These digital objects can be anything from text, images and data to video.

### *The Open Archives Initiative - Protocol for Metadata Harvesting*

The Open Archives Initiative - Protocol for Metadata Harvesting or OAI-PMH [5] makes it possible for Web-accessible repositories to interoperate by

providing a platform that allows repositories to share, publish and archive one another's metadata records.

### *RSS*

RSS, known as Really Simple Syndication or Rich Site Summary [13], is an XML based Web feed format used to publish frequently updated Web resources, such as blogs and news sites. An RSS document/feed contains full or summarised text and metadata.

### *The Search/Retrieval Web Service*

Search/Retrieval is a service for search and retrieval of Web resources across the Internet [14]. The aim is to promote interoperability between distributed databases by providing a common utilisation framework [7]. This protocol works in two ways: as SRU [11], which is Search/Retrieval via the URL, (Uniform Resource Locator) and as SRW [8], which is Search/Retrieval via SOAP (Simple Object Access Protocol). Both SRU and SRW use CQL (common/Contextual Query Language) for representing queries to information retrieval systems [11].

### *The Simple Web-service Offering Repository Deposit*

The Simple Web-service Offering Repository Deposit or SWORD (a profile of APP) is a lightweight protocol for depositing content from one location to another [9]. The aim of this protocol is to lower barriers to deposit content into repositories that support SWORD.

### *Z39.50*

The Z39.50 protocol is a NISO (National Information standards Organization) application layer protocol that supports distributed search and retrieval between structured network services by stipulating data structures and interchange rules that allow a client machine to search and retrieve records from databases on a server machine, across different platforms [3][4].

## **The survey**

The method chosen for this survey was the administration of a questionnaire which was conducted online<sup>1</sup>. The participants had to answer 6 questions (Appendix A), on the various protocols.

In order to get a fair result, people with interests in different areas of interoperability were invited to take part in this survey. An invitation to participate was sent to the following mailing lists: SRW, eprints-tech, dspac-tech, owner-atom-protocol list, ore-implementers, oai-implementers, rss-public and Computer Science postgraduate students at UCT (see email addresses in Appendix B).

A total of 23 people took part on the survey from 12/06/2009 to 17/07/2009. Below is a summary of the findings obtained from the survey.

---

<sup>1</sup> <http://banzai.cs.uct.ac.za/survey/index.php?sid=97137&lang=en>

## Most Popular Protocols

Participants were asked to rate their level of expertise with each of the protocols according to the following criteria:

Level 1: Expert implementer

Level 2: Implemented (have written code for an implementation of the protocol)

Level 3: Read and understood

Level 4: Heard about it, but do not know the details

Level 5: Never heard about it

The table below shows each of the above mentioned protocols and the corresponding number of people (in bold) as well as the percentage that they represent at the different levels of expertise.

Protocols	Number of people at each level of expertise					
	1	2	3	4	5	No answer
<b>RSS</b>	<b>3</b> (13%)	<b>6</b> (26.1%)	<b>9</b> (39.1%)	<b>3</b> (13%)	<b>1</b> (4.3%)	<b>1</b> (4.3%)
<b>ATOM</b>	<b>1</b> (4.3%)	<b>4</b> (17.4%)	<b>7</b> (30.4%)	<b>9</b> (39.1%)	<b>1</b> (4.3%)	<b>1</b> (4.3%)
<b>APP</b>	<b>0</b> (0%)	<b>1</b> (4.3%)	<b>1</b> (4.3%)	<b>3</b> (13%)	<b>17</b> (73.9%)	<b>1</b> (4.3%)
<b>Z39.50</b>	<b>0</b> (0%)	<b>0</b> (0%)	<b>6</b> (26.1%)	<b>12</b> (52.2%)	<b>4</b> (17.4%)	<b>1</b> (4.3%)
<b>OAI-PMH</b>	<b>7</b> (30.4%)	<b>7</b> (30.4%)	<b>5</b> (21.7%)	<b>0</b> (0%)	<b>3</b> (13%)	<b>1</b> (4.3%)
<b>OAI-ORE</b>	<b>1</b> (4.3%)	<b>1</b> (4.3%)	<b>5</b> (21.7%)	<b>9</b> (39.1%)	<b>6</b> (26.1%)	<b>1</b> (4.3%)
<b>SRU/W</b>	<b>0</b> (0%)	<b>0</b> (0%)	<b>7</b> (30.4%)	<b>5</b> (21.7%)	<b>10</b> (43.5%)	<b>1</b> (4.3%)
<b>SWORD</b>	<b>0</b> (0%)	<b>0</b> (0%)	<b>7</b> (30.4%)	<b>5</b> (21.7%)	<b>10</b> (43.5%)	<b>1</b> (4.3%)

Table 1: Number/percentage of people and their levels of expertise

### *OAI-PMH*

OAI-PMH is most popular protocol, with 60.9% of participants (14 people) as experts/implementers (levels 1 & 2).

### *RSS*

RSS is the second most popular protocol, with 39.1% of participants (9 people) in levels 1 & 2.

### *ATOM*

ATOM is in third place with 21.7% of participants (5 people) in levels 1 & 2.

### *OAI-ORE*

8.7% of the participants (2 people) are OAI-ORE experts/implementers.

### *APP*

None of the participants are APP experts and 4.3% of the participants (1 person) are in level 2.

### *Z39.50*

There were no experts/implementers of the Z39.50 protocol, and only 26.1% of the participants have read and understood this protocol.

### *SRU/W*

There were no experts/implementers of the SRU/W protocol, and only 30.4% of the users have read and understood the protocol.

## *SWORD*

Just like SRU/W, there were no experts/implementers of the SWORD protocol, and similarly there were only 7 people who have read and understood the protocol, and 43.5% of participants never even heard of the protocol.

## **Most Unpopular Protocols**

From the data in the above section, it can be deduced that:

APP is the least known protocol, with 73.9% (17 people) who have never heard of it. This is followed by 43.5% ((10 people) per protocol) who never heard of SRU/W and/or SWORD. 26.1% (6 people) of participants never heard of OAI-ORE, 17.4% (4 people) never heard of Z39.50, 13% (3 people) never heard of OAI-PMH, and 4.3% ((1 person) per protocol) never heard of RSS and/or ATOM.

## **Most Useful Features**

To the participants the features that make the protocols the best in their categories are:

### *RSS*

- Aggregation
- Flexible namespace use (dc, PRISM)
- Flexible use areas (not restricted to only one area of use, e.g. can be used by non-librarians)
- Link association with date
- Popularity (Its popularity attracts more people, trust)
- Sharing of quick text information and links
- Simplicity (easy to learn and implement)
- The use of popular standards (XML)
- Time saving (only get information of your interest)
- Value added services (podcasts)

### *ATOM*

- Aggregation
- Clear data model
- Flexible use areas (not restricted to only one area of use, e.g. can be used by non-librarians)
- Link association with date
- More standardised (i.e. fewer versions) and has better specifications than RSS
- Multiple formatting (dc, PRISM)
- Powerful
- Simplicity
- The use of popular standards (XML)

### *Z39.50*

- Allows federated search
- Provides library data interchange
- Standard for a particular community (i.e. libraries)

### *OAI-PMH*

- Allows community based aggregation of metadata
- Flexible metadata formats
- Integration into software packages (Dspace, CONTENTdm)
- Harvesting of vast amounts of data at once (saves time)
- Low server load
- Platform independent metadata retrieval
- Resumption tokens allows users to download on their own schedule
- Simplicity in implementation
- Specifically designed for digital repositories
- Strings open archives together
- Very good at generically providing db data.

### *OAI-ORE*

- A solution to the aggregation problem
- Platform independent metadata retrieval

### *SWORD*

- Interoperability with other systems (e.g. CRIS) possible upload to repository through MS word
- Multiple deposit methods
- Platform independent repository deposit
- Reduces work load (multiple deposits at once)

### **Least Useful Features**

The participants have also indicated which features they find least useful and/or the gaps in each protocol.

### *RSS*

- Lacks features (most features are extended modules)
- Lacks semantics
- Poor documentation (often leads to bad implementations)
- Poor specs of content type (programmers use it for different things)
- Too many versions

### *ATOM*

- Lacks semantics
- Too heavily geared towards blog posts, therefore it is misappropriated in systems interoperability

### *APP*

- Overly zealous in its requirements

### *Z39.50*

- Complex/arcane
- Heavy server load
- Outdated
- Use restricted to a specific community only (librarians)

### *OAI-PMH*

- Assumption of item-level description
- Cumbersome syntax and operation
- Lacks variables (e.g. searching function)
- ListIdentifiers (no further details were provided on this issue)
- Multiple metadata formats make the protocol harder than it needs to be
- Sets are chaotic (no further details were provided on this issue)
- Underutilized (can be used for much more than transporting metadata records, it can transport an entire archive including the digital signatures)

### *OAI-ORE*

- Complex
- Sets are chaotic (no further details were provided on this issue)

### *SRU/W*

- Complex to implement
- It is difficult to translate CQL to backend search engines

### *SWORD*

- Does not allow withdrawals (restricted to deposits only)
- Relies exclusively on packaging for content passing
- Unfinished (still evolving)

One user said that RSS does not have any feature which is not useful, while 2 users said the same about OAI-PMH.

## **Suggested Improvements**

### *RSS*

- A standard interpreter with graphical user interface
- The ability to query specific time frames (a layer above the protocol)
- Adding new tag (element) by the user upon need

### *ATOM*

- The ability to query specific time frames (a layer above the protocol)
- Adopting some of the semantics of OAI-PMH

### *Z39.50*

- Allow full harvest
- Combined with SRU, Z39.50 becomes simpler

### *OAI-PMH*

- Provide services other than metadata extraction (e.g. negotiate the legal issues and allow access to the actual data)
- Consider persistent resumptionTokens
- Definition of sets restriction (nomenclature, hierarchy)
- Although interoperability is the key issue, other aspects of the protocol are also important (i.e. efficiency)

### *OAI-ORE*

- Consider persistent resumptionTokens
- Definition of sets restriction (nomenclature, hierarchy)

### *SRU/W*

- Allow full harvest and simplify

### *SWORD*

- Better integration with OAI-ORE
- Extension for non-packaged material
- Implementation of replace and add functions

### *Improvements for all the protocols*

- Although interoperability is the key issue, other aspects of the protocol are also important (i.e. efficiency)
- Better documentation
- Create a protocol for the WWW rather than one for a specific community
- Do not require data that is not essential
- More standardised namespaces, to better tag data
- Simplicity equals a protocol that is easier to be widely adopted
- There should be a protocol with a standardised interface that allows harvesting, syndication and searching with a very small set of mandatory operations/parameters (for the sake of simplicity)

### **General comments**

- *RSS & ATOM* - Do not perform well for interoperability on bibliographic applications
- Standards are the only way to make interoperability work. Using these standards it is possible to upgrade systems, change systems completely (Dspace to Fedora) without losing expensive work.

- There is much work going into developing protocols for interoperability, but little work on the interoperability of the content. All of these protocols deal with the transport of some kind of data, but most do not concern themselves with the data itself. But data interoperability is the core of the interoperability problem.

## Conclusions

The number of participants in this survey represents a very small sample of the “interoperability world population”, but the answers were very useful for the purpose of this research. And from them we can draw the following conclusions:

Although OAI-PMH is classified as the best-known protocol in this survey, it has been observed that this result is based on the fact that most of the respondents are OAI-PMH implementers, which means that OAI-PMH is mostly known by its implementers only. RSS and ATOM on the other hand are well known to both implementers (levels 1 & 2) and the general Web users (levels 3 & 4). There was only one person (level 5) who never heard of either RSS or ATOM, compared to 3 people who never heard of OAI-PMH. There was some disagreement on the issue of whether multiple metadata formats for OAI-PMH was a good or bad feature - overall most participants said that multiple metadata formats makes things harder.

The results of this survey lead us to conclude that a good interoperability protocol should be: simple enough to allow programmers to implement, explore, and experiment while requiring only operations that are crucial to the performance of the protocol, but robust enough to provide value added services for the data extracted. The perfect solution, if perfection can ever be achieved, is some combination of the simplicity and lightness of RSS, with the structure and semantics of OAI-PMH and the efficiency of SWORD.

## References

- [1] Atom Enabled: Atom Syndication Format.  
<http://www.atomenabled.org/developers/syndication> Last accessed on the 05<sup>th</sup> of May 2009.
- [2] ISO/IEC 2382-01, *Information Technology Vocabulary, Fundamental Terms*
- [3] JISC Information Environment Architecture Glossary. At:  
<http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/glossary/> Last accessed on the 03<sup>rd</sup> of May 2009.
- [4] Lynch, C.A., The Z39.50 Information Retrieval Standard. Part I: A Strategic View of its Past, Present and Future. April 1997. D-Lib Magazine.
- [5] Open Archives forum: OAI for beginners – The Open Archives Forum Online Tutorial. <http://www.oaforum.org/tutorial/> Last accessed on the 03<sup>rd</sup> of May 2009.
- [6] Open Archives Initiative, Object Reuse and Exchange. Aggregations of Web Resources. <http://www.openarchives.org/ore/> Last accessed on the 16<sup>th</sup> of July 2009.

- [7] SRW: Search and Retrieve Web Service <http://srw.cheshire3.org/docs/>  
Accessed on the 06<sup>th</sup> of May 2009.
- [8] SRW Editorial Board. Search/Retrieve Website Version 1.1. May 2004.  
<http://srw.cheshire3.org/SRW-1.1.pdf> Last accessed on the 06<sup>th</sup> of May 2009.
- [9] SWORD Protocol. <http://www.swordapp.org/> Last accessed on the 03<sup>rd</sup> of  
May 2009.
- [10] The Atom Publishing Protocol: draft-eitf-atompub-protocol-04.txt.  
<http://bitworking.org/projects/atom/draft-eitf-atompub-protocol-04.html> Last  
accessed on the 08<sup>th</sup> of May 2009.
- [11] The Library of Congress. Standards: SRU.  
<http://www.loc.gov/standards/sru/specs/cql.html> Last accessed on the 06<sup>th</sup> of  
May 2009.
- [12] Wikipedia, The free encyclopedia. *Interoperability*. April 2009  
<http://en.wikipedia.org/wiki/Interoperability> Last accessed on the 03<sup>rd</sup> of May  
2009.
- [13] Wikipedia: The Free Encyclopedia. RSS. <http://en.wikipedia.org/wiki/RSS>  
Last accessed on the 16<sup>th</sup> of July 2009.
- [14] Wikipedia, The free encyclopedia. Search/Retrieve Web Service.  
[http://en.wikipedia.org/wiki/Search/Retrieval\\_Web\\_Service](http://en.wikipedia.org/wiki/Search/Retrieval_Web_Service) Last accessed on  
the 06<sup>th</sup> of May 2009.

## Appendix A (Questionnaire)

1. **What is your level of confidence with each of the following protocols? Choose your answer from 1 to 5 according to the values explained below.**

Choose your answer according to the following values:

1 - Expert implementer

2 - Implemented (have written code for an implementation of the protocol)

3 - Read and understood

4 - Heard about it, but do not know the details

5 - Never heard about it

Please choose the appropriate response for each item:

RSS      1    2    3    4    5

ATOM    1    2    3    4    5

APP      1    2    3    4    5

Z39.50    1    2    3    4    5

OAI-PMH    1    2    3    4    5

OAI-ORE    1    2    3    4    5

SRU/W     1    2    3    4    5

SWORD     1    2    3    4    5

Other(s)    1    2    3    4    5

2. **If in question one you choose other(s), please name the other protocol/s that you are familiar with.**

Write the name of the protocol followed by the number that indicates your level of confidence with the protocol. If there is more than one protocol separate the list with commas (,).

3. **What in your opinion is the most useful feature(s) of each of the protocols?**

RSS:

ATOM:

APP:

Z39.50:

OAI-PMH:

OAI-ORE:

SRU/W:

SWORD:

Other(s):

**4. What in your opinion is the least useful feature(s) of each of the protocol?**

RSS:

ATOM:

APP:

Z39.50:

OAI-PMH:

OAI-ORE:

SRU/W:

SWORD:

Other(s):

**5. How do you believe these protocols could be improved?**

Your opinion on improvements for interoperability protocols in general. If you are not making a generalized comment please name the protocol before making the actual comment.

**6. General comments about the protocols. If you are not making a general comment please name the protocol before making the actual comment.**

**Appendix B (Mailing Lists email addresses)**

1. SRW: {srw@mail.dei.unipd.it}
2. eprints-tech: {eprints-tech@ecs.soton.ac.uk}
3. dspace-tech: {dspace-tech@lists.sourceforge.net}
4. owner-atom-pro: {owner-atom-protocol@vpnc.org}
5. ore-implementers: {ore-implementers@openarchives.org}
6. oai-implementers: {oai-implementers@openarchives.org}
7. rss-public: {rss-public@yahoogroups.com}
8. Computer Science Postgraduate students at UCT: {grads@cs.uct.ac.za}