SWAD–Europe deliverable 4.2: Semantic Web and Web Services: RDF/XML and SOAP for Web Data Encoding

Project name:
W3C Semantic Web Advanced Development for Europe (SWAD-Europe)

Project Number:
IST-2001-34732

Workpackage name:
4. Semantic Web and Web Services

Workpackage description:

Deliverable title:
SWAD-Europe deliverable 4.2: Semantic Web and Web Services: RDF/XML and SOAP for Web Data Encoding

URI:
http://www.w3.org/2001/sw/Europe/reports/sw_soap_design_report/

Authors:
Nikki Rogers

Abstract:
This document is in support of the work conducted for SWAD-Europe deliverable 4.2. It is a Public design document describing the architecture of a working SOAP-based "Semantic Web Service". For example we aim to show the use of W3C's XML Protocol and Web Service specifications to describe, discover and communicate with remote RDF/Semantic Web query or inference services. This design document is illustrated throughout with working examples.

Contents
- Introduction
- Section 1: Overview of an example web service architecture
- Section 2: Architecture for an example Semantic Web Service
- Section 3: Guide to integrating Axis and Sesame
- Section 4: The Soap data model and the RDF data model - transforming from the former to the latter

Introduction

In order to explore object marshalling and unmarshalling possibilities in the context of distributed semantic web services in which various RDF-based and non-RDF SOAP solutions may proliferate, an example web services architecture is shown in this report.

The architecture aims to demonstrate how an RDF-based web service (in this case we happen to use Sesame for the RDF service 'backend') might be combined with SOAP (in this case we happen to use Apache Axis for the RDF service's SOAP 'frontend') in order to serialise object data. In the examples, the service's WSDL is used in order to generate client-side stubs so that we may demonstrate the ease with which some client on the network (whether it be "RDF-savvy" or otherwise) may consume such a service.

We also show how soap-encoded data may be transformed (via an XSL stylesheet) to RDF data and suggest scenarios where this may be useful.

The example dataset we use is very small and chosen merely to demonstrate how object data may be serialised as SOAP encoded data, for exchange on the network. We example how SOAP encoded data may be transformed 'back' to RDF data (the "round trip") in order to show the relationship between the two data models. However, we do not attempt a full comparison of the data models: we do not provide a transformation from RDF to SOAP for example, hence we do not consider the serialisation of arbitrary RDF data via the soap data encoding syntax. Instead we aim to demonstrate some potential for using the SOAP data encoding for standardised data exchange in a network environment in which both semantic web services and more traditional web services reside.

We note that at the time of writing this report, Document/Literal message styled SOAP Web Services have gained industry support over RPC/Encoded services (or any other combination). We give a brief outline of the impact of this document/literal versus RPC/Encoded debate for SOAP services in Section 1.

In Section 1 of this report we give an overview of our chosen hypothetical illustration of a semantic web service scenario. We present a diagram of this example web service architecture. Then, in Section 2 we explain our example architecture for a Semantic Web Service at a more technical level. In Section 3 we explain how we have opted for the most generic approach with which to integrate Apache Axis and Sesame software, in order to run trials. We give our example data and configurations used and we hope to make the given sample implementation downloadable via a link on this page in the near future. In Section 4 we briefly discuss the Soap data model and the RDF data model - transforming from the former to the latter.

Section 1: Overview of an example web service architecture

Overview diagram of the exampled architecture -
The main advantage of using RDF to store the data in the backend database is that, whilst a range of content providers may publish data all related to a similar domain, their choice of data modelling actually varies from provider to provider. The RDF standard (plus OWL for example) makes it easier to aggregate such related but heterogeneous data. It provides the basis for a simple inferencing layer, meaning that for example my aggregated data store can be simply designed to 'know' that a "semi-detached bungalow" and a "detached bungalow" are both sub-classes of a type called "residential property" and so it may continue.

Relevant Scenarios - We give here an example high-level scenario to illustrate the potential of using the SOAP data encoding syntax for a web services architecture as outlined above.

Suppose I wish to purchase a house in the South West of the UK. I may be interested in various data sets: houses for sale in the area (obviously), according to price band and location, plus environmental information (such as pollution factors, tendency for landslides, flooding and so on), plus also schools information on which I can base decisions regarding my children's education (government Offsted reports, or schools' own web sites for example). It currently tends to be the case that no single online service will provide such information. So these data repositories exist in the closed worlds sense.

However, if that information is available via the network via SOAP services then I can use a SOAP client to query these services. I should be able to aggregate the results using the unique identifiers these sets of information have in common - postcodes. I may well use RDF-based tools to assist aggregation at my end, with inferencing capabilities to support querying over the aggregated data (perhaps using OWL for example in order to link equivalent terms for the cases where various data providers have used different terms for "number of bedrooms" and so on).

In a perhaps more globally "highly evolved" - semantically - version of the same context, a data provider on the web - an estate agent's online service for example - might in fact choose to use a "intelligent RDF backend" to aggregate objects in order to centralise access to some of these resources. By doing this, a data provider thereby enhances its business services by allowing end users (or other services - maybe an online jobs recruitment service) a central point to access distributed information, rather than having to develop their own aggregator clients.

The main advantage of using RDF to store the data in the backend database is that, whilst a range of content providers may publish data all related to a similar domain, their choice of data modelling will naturally vary from provider to provider. The RDF standard (plus OWL for example) makes it simpler to aggregate such heterogeneous data. It provides the basis for a simple inferencing layer, meaning that for example my aggregated data store can be simply designed to "know" that a "semi-detached bungalow" and a "detached bungalow" are both sub-classes of a type called "bungalow", which itself is a sub-class of a type called "residential property" and so it may continue.

But that resources which one content provider may classify as "flats/apartments" are the same as what another content provider classifies as "apartments or flats". And so on, this system being fully extensible and flexible in RDF, meaning that not only is it quite 'clever' in terms of the sorts of questions that can be asked of and answered by it, but it also may be maintained with less overheads in terms of human effort than when using, say, a traditional relational database approach. It may also be the case that even if a data provider should add new attributes to the objects that they are modelling, when the objects are unmarshalled at the client end they will not break the service's model used for
aggregating the data (RDF resources are allowed arbitrary, non-reserved properties).

The main advantage of using SOAP as the transportation protocol for such objects is that the SOAP specification supports the SOAP encoding data model which promotes the standardised serialisation of arbitrarily complex objects (allowing multi-recursive and cyclic data structures). And this remains consistent in either an RDF-aware or RDF-non-aware context.

The document/literal versus RPC/Encoded debate for SOAP services - Industry support for SOAP messaging styles has recently - over the lifetime of the SOAP project - swung in favour of Document/Literal. In this section we take a brief look at the place of SOAP encoding in the SOAP 1.2 specification and also its absence in the WS-I Basic Profile. We consider some of the arguments against using SOAP encoding for Web Services and note the potential for using document/literal styled SOAP Web Services for the Semantic Web.

The place of SOAP Encoding in SOAP 1.2

Having been in section 5 of the SOAP 1.1 specification, the SOAP encoding is revised and moved to an adjacent section in the SOAP 1.2 specification (see http://www.w3.org/TR/soap12-part2#soapenc). Here it states "This encoding MAY be used to transmit data in SOAP header blocks and/or SOAP bodies. Other data models, alternate encodings of the SOAP Data Model as well as unencoded data MAY also be used in SOAP messages".

The WS-I Basic Profile

The WS-I Web Services Interoperability Organization, formed in February 2002 (and led by Microsoft and IBM), published the WS-I Basic Profile in August 2003. The specifications covered by the Basic Profile include SOAP 1.1, WSDL 1.1, UDDI 2.0, XML 1.0 and XML Schema, with the aim of promoting Web services interoperability. With regards to soap encoding, the specification says "The soap:encodingStyle attribute is used to indicate the use of a particular scheme in the encoding of data into XML. However, this introduces complexity, as this function can also be served by the use of XML Namespaces. As a result, the Profile prefers the use of literal, non-encoded XML. And it goes on to specify that this attribute must not be used. It appears that the argument against using soap encoding is "why bring another sort of encoding into the arena when we already have XML and XML namespaces?". However, semantic web initiatives frequently point to the clear gap left by XML in a context where it is important to maximise semantic coherence at machine-to-machine level across distributed web services. In contrast to XML, the SOAP data encoding allows for multi-referencing on a grand scale and, when augmented with RDF, RDF's and OWL opens up the space occupied by web services such that the links between vast repositories of related but heterogeneous data may be exploited in a largely automated and arguably efficient fashion. It is hard to see how such links can be managed sensibly using the XML tree structure. A directed graph structure such as specified by the SOAP data encoding offers plenty in terms of managing the way information may be mapped to the network - with RDF-based layers making it usefully 'meaningful'.

None-the-less, RDF has its own RDF/XML syntax and parallel SWADE work is reviewing the relationship of RDF/XML to XML, for example looking at the development of a normalised RDF/XML. This means that, should a SOAP-based web service require to be document/literal based, it might then be configured to offer SOAP message responses where the SOAP message content is encoded in RDF/XML. Schemas - both RDF and XML schemas - might then be associated with this same content (via namespaces) to permit either or both of the XML processing and RDF processing alternatives on the client side. Again, this promotes a relaxed situation in which both RDF-aware and non-RDF-aware services may share out data to RDF-aware and non-RDF-aware consumer applications equally.

It is recommended that work that follows on from this workpackage look more closely at just this sort Document/Literal option. In the example architecture described below, scenarios are limited by the sort of data structures that can be handed over from Sesame (the RDF database) to the Axis server software (for creating SOAP data encoded SOAP body content). Only Sesame's tabulated data structures for result sets can be very easily mapped to Java beans for straight forward consumption by Axis' bean and array serialisers. RDF result sets encoded in RDF/XML would require a custom serialiser to be written for Axis - one that would handle a complete RDF to SOAP mapping. We have not conducted full investigations in this respect, but such a feature would be required to handle arbitrary RDF structures. At this point the utility of a Document/Literal options becomes apparent again.

Section 2: Architecture for an example Semantic Web Service

This section takes a quick look at the data being exchanged over the network in the specific scenario we have produced for this report: how that data originates, how objects are "preserved" during transportation, and what can be done with the data on the receiving end as it were.

The exampled web service architecture is obviously not intended to be representative of the scale and capabilities of current or future web services, but simply to illustrate the extent to which SOAP-to-RDF mappings are possible, and indeed useful in the context of web services. In other words, this is merely a proof-of-concept demonstration.

We begin with another diagram, similar to the more general overview above, but specifically labelled with the java classes and stylesheets used in our example scenario.

In our example scenario the SOAP, RPC-Encoded style service that we have called "rubyinfo" happens to hold a small amount of data about Ruby software products. These software products are encapsulated as what we have named "rubyinfo" objects which essentially are objects that have "product", "owner", "update" and "category" properties - themselves objects, with further properties such as "category_major", "category_minor" etc. This is best explained quickly with the diagram from this workpackage's first report:
WSDL for the Web Service - Please take a look at the WSDL we have for our example (localhost) Ruby Soap Service. You will see that the "request" message name is rubyQueryRequest and that it contains one part, (not named - we could have more sensibly given this a generic name such as "query_expression", say) of type "xsd:string".

```xml
<wsdl:message name="rubyQueryRequest">
  <wsdl:part name="in0" type="xsd:string" />
</wsdl:message>
```

REQUEST MESSAGE - Here is an example request message in which we see that the single string part to the RPC message body is in fact an RQL query (for information on the RQL language as interpreted for Sesame, see http://sesame.administrator.nl/publications/rql-tutorial.html). This soap message specifies the encoding style to be used (soapenv:encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"). And the RQL query effectively says:

"give me all objects that have PRODUCT_NAME, VERSION etc properties, using the namespace ruby = http://example.ruby-language.org/rda-vocab/test2.rdf#.

```
POST /swadWP4/services/RubySoapServiceTest HTTP/1.0
Content-Type: text/xml; charset=utf-8
Accept: application/soap+xml,
application/dime, multipart/related, text/*
User-Agent: Axis/1.1
Host: 127.0.0.1
Cache-Control: no-cache
Pragma: no-cache
SOAPAction: ''
Content-Length: 1271

<?xml version="1.0" encoding="UTF-8"?>
<soapenv:Envelope xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/"
xmlns:xsd="http://www.w3.org/2001/XMLSchema"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <soapenv:Body>
    <ns1:rubyQuery soapenv:encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
xmlns:ns1="urn:rubysoapservices" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
      <arg1 xsi:type="xsd:string">select PRODUCT_NAME, VERSION, STATUS, HOMEPAGE, DOWNLOAD, LICENSE, DESC, EMAIL, OWNER_NAME, OWNER_ID, CATEGORY_MAJOR, CATEGORY_MINOR from
        (X) ruby:ruby_product . ruby:product_name {PRODUCT_NAME},
        (X) ruby:ruby_product . ruby:version {VERSION},
        (X) ruby:ruby_product . ruby:status {STATUS},
        (X) ruby:ruby_product . ruby:homepage {HOMEPAGE},
        (X) ruby:ruby_product . ruby:download {DOWNLOAD},
        (X) ruby:ruby_product . ruby:license {LICENSE},
        (X) ruby:ruby_product . ruby:desc {DESC},
        (X) ruby:ruby_owner . ruby:email {EMAIL},
        (X) ruby:ruby_owner . ruby:owner_name {OWNER_NAME},
        (X) ruby:ruby_owner . ruby:owner_id {OWNER_ID},
        (X) ruby:ruby_category . ruby:category_major {CATEGORY_MAJOR},
        (X) ruby:ruby_category . ruby:category_minor {CATEGORY_MINOR}
        using namespace ruby = http://example.ruby-language.org/rda-vocab/test2.rdf#</arg1>
    </ns1:rubyQuery>
  </soapenv:Body>
</soapenv:Envelope>
```

We have permitted the request message to contain a single string parameter deliberately so as to allow the testing of different query expressions on the remote service. We could of course specify several separate queries (for e.g. asking all ruby information objects with product name value is so-and-so) in the WSDL. Normally we would deploy the service so as the parameter would be more informatively named than simply "arg1" (the Axis default).
<?xml version="1.0" encoding="UTF-8"?>
<soapenv:Envelope xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/
xmlns:xsd="http://www.w3.org/2001/XMLSchema"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <soapenv:Body>
    <ns1:rubyQueryResponse
soapenv:encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
xmlns:ns1="urn:rubysoapservices">
      <rubyQueryReturn xsi:type="soapenc:Array"
soapenc:arrayType="ns1:RubyInfo[3]"
xmlns:soapenc="http://schemas.xmlsoap.org/soap/encoding/">
        <item href="#id0"/>
        <item href="#id1"/>
        <item href="#id2"/>
      </rubyQueryReturn>
    </ns1:rubyQueryResponse>
  </soapenv:Body>
</soapenv:Envelope>
We provide a stylesheet for transforming soap encodings from the SOAP 1.1 specification to that of the 1.2 specification (Author: Max Froumentin). After transformation, the body of the above response message looks like this:

```xml
<Body xmlns="http://www.w3.org/2003/05/soap-envelope">
  <soapenv:Body>
    <soapenv:Fault>
      <faultcode>soapenv:Client</faultcode>
      <faultstring>Unknown Fault</faultstring>
    </soapenv:Fault>
  </soapenv:Body>
</Body>
```
<rubyQueryResponse xmlns:soapenc="http://schemas.xmlsoap.org/soap/encoding/"
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 ns0:itemType="ns1:RubyInfo" ns0:arraySize="3">
 <item ns0:ref="id0"/>
 <item ns0:ref="id1"/>
 <item ns0:ref="id2"/>
 <rubyQueryReturn>
 <multiRef xmlns="http://schemas.xmlsoap.org/soap/encoding/"
 xmlns:soapenc="http://schemas.xmlsoap.org/soap/encoding/"
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xmlns:ns2="urn:rubysoapservices"
 xmlns:ns1="http://schemas.xmlsoap.org/soap/envelope/
 xmlns:ns0="http://www.w3.org/2003/05/soap-encoding"
 ns0:id="id0"
 ns0:encodingStyle="http://www.w3.org/2003/05/soap-encoding"
 xsi:type="ns2:RubyInfo">
 <ruby_category ns0:ref="id3"/>
 <ruby_owner ns0:ref="id4"/>
 <ruby_product ns0:ref="id5"/>
 <update xsi:type="xsd:string" xsi:nil="true"/>
 </multiRef>
 <multiRef xmlns="http://schemas.xmlsoap.org/soap/encoding/"
 xmlns:soapenc="http://schemas.xmlsoap.org/soap/encoding/"
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xmlns:ns1="http://schemas.xmlsoap.org/soap/envelope/
 xmlns:ns0="http://www.w3.org/2003/05/soap-encoding"
 ns0:id="id1"
 ns0:encodingStyle="http://www.w3.org/2003/05/soap-encoding"
 xsi:type="ns2:RubyInfo">
 <ruby_category ns0:ref="id6"/>
 <ruby_owner ns0:ref="id7"/>
 <ruby_product ns0:ref="id8"/>
 <update xsi:type="xsd:string" xsi:nil="true"/>
 </multiRef>
 <multiRef xmlns="http://schemas.xmlsoap.org/soap/encoding/"
 xmlns:soapenc="http://schemas.xmlsoap.org/soap/encoding/"
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xmlns:ns1="http://schemas.xmlsoap.org/soap/envelope/
 xmlns:ns0="http://www.w3.org/2003/05/soap-encoding"
 ns0:id="id2"
 ns0:encodingStyle="http://www.w3.org/2003/05/soap-encoding"
 xsi:type="ns2:RubyInfo">
 <ruby_category ns0:ref="id9"/>
 <ruby_owner ns0:ref="id10"/>
 <ruby_product ns0:ref="id11"/>
 <update xsi:type="xsd:string" xsi:nil="true"/>
 </multiRef>
 <multiRef xmlns="http://schemas.xmlsoap.org/soap/encoding/"
 xmlns:soapenc="http://schemas.xmlsoap.org/soap/encoding/"
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xmlns:ns1="http://schemas.xmlsoap.org/soap/envelope/
 xmlns:ns0="http://www.w3.org/2003/05/soap-encoding"
 ns0:id="id7"
 ns0:encodingStyle="http://www.w3.org/2003/05/soap-encoding"
 xsi:type="ns2:Owner">
 <email xsi:type="xsd:string">mailto:neoneye@diku.dk</email>
 <owner_id xsi:type="xsd:string">neoneye@diku.dk-Simon-neoneye-Strandgaard</owner_id>
 <owner_name xsi:type="xsd:string">Simon-neoneye-Strandgaard</owner_name>
 </multiRef>
 <multiRef xmlns="http://schemas.xmlsoap.org/soap/encoding/"
 xmlns:soapenc="http://schemas.xmlsoap.org/soap/encoding/"
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xmlns:ns1="http://schemas.xmlsoap.org/soap/envelope/
 xmlns:ns0="http://www.w3.org/2003/05/soap-encoding"
 ns0:id="id4"
 ns0:encodingStyle="http://www.w3.org/2003/05/soap-encoding"
 xsi:type="ns2:Owner">
 <email xsi:type="xsd:string">mailto:hd@rubylinks.de</email>
 <owner_id xsi:type="xsd:string">hd@rubylinks.de-Hans-Dieter-Stich</owner_id>
 <owner_name xsi:type="xsd:string">Hans-Dieter-Stich</owner_name>
 </multiRef>
 <multiRef xmlns="http://schemas.xmlsoap.org/soap/encoding/"
 xmlns:soapenc="http://schemas.xmlsoap.org/soap/encoding/"
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xmlns:ns1="http://schemas.xmlsoap.org/soap/envelope/
 xmlns:ns0="http://www.w3.org/2003/05/soap-encoding"
 ns0:id="id3"
 ns0:encodingStyle="http://www.w3.org/2003/05/soap-encoding"
 xsi:type="ns2:Category">
 <category_major xsi:type="xsd:string">Library</category_major>
 <category_minor xsi:type="xsd:string">Distributed Communication</category_minor>
 </multiRef>
 <multiRef xmlns="http://schemas.xmlsoap.org/soap/encoding/"
 xmlns:soapenc="http://schemas.xmlsoap.org/soap/encoding/"
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xmlns:ns1="http://schemas.xmlsoap.org/soap/envelope/
 xmlns:ns0="http://www.w3.org/2003/05/soap-encoding"
 ns0:id="id10"
 ns0:encodingStyle="http://www.w3.org/2003/05/soap-encoding"
 xsi:type="ns2:Owner">
 <email xsi:type="xsd:string">mailto:cilibrar@ofb.net</email>
 <owner_id xsi:type="xsd:string">cilibrar@ofb.net-Rudi Cilibrasi</owner_id>
 <owner_name xsi:type="xsd:string">Rudi Cilibrasi</owner_name>
 </multiRef>
 <multiRef xmlns="http://schemas.xmlsoap.org/soap/encoding/"
 xmlns:soapenc="http://schemas.xmlsoap.org/soap/encoding/"
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xmlns:ns1="http://schemas.xmlsoap.org/soap/envelope/
 xmlns:ns0="http://www.w3.org/2003/05/soap-encoding"
 ns0:id="id9"
 ns0:encodingStyle="http://www.w3.org/2003/05/soap-encoding"
 xsi:type="ns2:Category">
 <category_major xsi:type="xsd:string">Library</category_major>
 <category_minor xsi:type="xsd:string">Parallel</category_minor>
 </multiRef>
</rubyQueryResponse>
Transforming from SOAP 1.2 encoding to RDF/XML - the “round trip” - The stylesheet we have developed (Max Froumentin and Nikki Rogers) for this transformation is 'soap_to_rdf.xsl'. It is not completely generic - the first element in the soap body to be treated as an ‘edge’ has to be identified as noted in the stylesheet. Also the choice of name for SOAP accessors that have ‘id’ attributes is arbitrary and for Axis is "multiRef". Therefore these multiref accessors are matched as part of the transformation in order to be mapped to RDF typed elements.

When applied to the above response body data, it produces:

```xml
<?xml version="1.0" encoding="utf-8"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:env="http://www.w3.org/2003/05/soap-envelope"
xmlns:enc="http://www.w3.org/2003/05/soap-encoding"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:ns1="urn:rubysoapservices"
xmlns:ruby="urn:rubysoapservices">
  <rdf:Description>
    <rubyQueryResponse xmlns="urn:rdf2soap">
      <rubyQueryReturn>
        <item rdf:nodeID="id0"/>
        ...
      </rubyQueryReturn>
    </rubyQueryResponse>
  </rdf:Description>
</rdf:RDF>
```
<owner_name xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Rudi Cilibrasi</owner_name>
</rdf:Description>
</rdf:Description>
</rdf:Description>

<category_major xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Library</category_major>
<category_minor xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Parallel</category_minor>
</rdf:Description>
</rdf:Description>
</rdf:Description>
<rdf:Description rdf:nodeID="id8">
<rdf:type rdf:resource="ruby:Product"/>
<desc xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">pure ruby editor for pair-programming</desc>
<download xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">http://prdownloads.sourceforge.net/metaeditor/aeditor-ruby-0.3.tar.gz?download</download>
<homepage xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">http://metaeditor.sf.net/</homepage>
<license xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">BSD</license>
<product_name xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">aeditor</product_name>
<status xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">alpha</status>
<version xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">0.3</version>
</rdf:Description>
</rdf:Description>
</rdf:Description>
<rdf:Description rdf:nodeID="id11">
<rdf:type rdf:resource="ruby:Product"/>
<desc xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">MPI Ruby</desc>
<download xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">http://prdownloads.sourceforge.net/mpiruby/mpi_ruby-0.4.tar.bz2?download</download>
<homepage xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">http://mpiruby.sourceforge.net/</homepage>
<license xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">OpenSource</license>
<product_name xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">mpi_ruby</product_name>
<status xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Beta - seems to work with some testing</status>
<version xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">0.4</version>
</rdf:Description>
</rdf:Description>
</rdf:Description>
<rdf:Description rdf:nodeID="id5">
<rdf:type rdf:resource="ruby:Product"/>
<desc xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Ruby/Spread is an interface to the Spread Library at http://www.spread.org/. Spread is a toolkit and daemon that provides multicast and group communications support to applications across local and wide area networks. Spread is designed to make it easy to write groupware, networked multimedia, reliable server, and collaborative work applications.</desc>
<homepage xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">http://www.rubylinks.de/</homepage>
<license xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">LGPL</license>
<product_name xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">ruby-spread</product_name>
<status xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">0.1</status>
</rdf:Description>
</rdf:Description>
</rdf:Description>
<rdf:Description rdf:nodeID="id10">
<rdf:type rdf:resource="ruby:Product"/>
<desc xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Ruby/Spread is an interface to the Spread Library at http://www.spread.org/. Spread is a toolkit and daemon that provides multicast and group communications support to applications across local and wide area networks. Spread is designed to make it easy to write groupware, networked multimedia, reliable server, and collaborative work applications.</desc>
<homepage xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">http://www.rubylinks.de/</homepage>
<license xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">LGPL</license>
<product_name xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">ruby-spread</product_name>
<status xmlns="urn:rdf2soap" rdf:datatype="http://www.w3.org/2001/XMLSchema#string">0.1</status>
</rdf:Description>
</rdf:Description>
</rdf:Description>
You may run both the original data uploaded to sesame and the above RDF (obtained from a transformation of the SOAP response body's data) through the online RDF validator at 
@http://www.w3.org/RDF/Validator/ to compare triples that both graphs produce. They are similar with a couple of differences [TODO : 1. leading edge of soap body response 2. datatyping added in stylesheet - our version of sesame could not cope with this, more recent version might be able to]

Section 3: Guide to integrating Axis and Sesame

Installation of axis and sesame is quite a lengthy process, and we hope to wrap up what we have put together for this report as a downloadable war. file with an ant build to assist installation and configuration. In the meantime, we provide a summary and some more detailed instructions to supplement those available from the Axis and Sesame websites.

Summary instructions - Obtaining Sesame

See the @Sesame server installation instructions@ which describe how the downloads for a sesame installation. Briefly they are:

Overall: you will need to download and install Sesame, and relational database software (MySQL for example) plus a suitable JDBC driver. You will also need to download Axis and also have Ant installed so as to use the Axis build targets. Here are some summary instructions (more detailed instructions follow):

- For Sesame, you will need to create a new user account (e.g. called 'sesame') in MySQL, a new database (e.g. called 'rubyinfodb'), and then grant the user all privileges on the database you created. You can upload our @sample 'Ruby Info' data@ to the database as described below.

Assuming tomcat is your servlet container: In your tomcat web apps directory you will need 2 directories - one for sesame & one for axis. We have used:

[path to tomcat]/webapps/ -- [sesame_dir] -- [swadwp4_dir]

In the[/sesame directory you will need to extract the [sesame.war file and edit this file as discussed below:

[Sesame_dir]/WEB-INF / system.conf

And also add the appropriate jars (those found with the sesame distribution, plus the jdbc driver) in

[Sesame_dir]/WEB-INF/lib/

Then you will need to edit

[Sesame_dir]/WEB-INF/web.xml/

We example a test java class below, for testing out suitable queries on the 'Ruby Info' data, from the commandline, over http.

For Apache Axis, you will need to download and install Axis. There are several jar files that need to be placed in Axis' lib/ directory. When you build Axis, the compiled classes produced need to be placed in the appropriate [SWADWP4_DIR] in the tomcat servlet container.

Add the appropriate jars to

[SWADWP4_DIR]/WEB-INF/lib/

You will need to check you can get the axis test page and also validate Axis with "happyaxis" before you proceed.

Next you can compile the sample Axis provider class we provide (RubySoapServiceTest.java) and copy the class to a directory you create in tomcat: [SWADWP4_DIR]/WEB-INF/classes/rubysoap/services. In addition you will need the deploy.wdd & undeploy.wdd files we provide, and the classes for each of the beans required for this demo (named @Category, @Product, @RubyInfo, @Owner).

You can now use the AdminClient tool that comes with Axis to deploy the Ruby service (this should update your server-config.wdd file in the [SWADWP4_DIR]/WEB-INF directory with the information for your new service).

Once successfully deployed .... you can test the service via the Axis client we provide (Client.java), which is called from the commandline and can take various query strings. This client will do two things 1) It will unmarshall returned data back into javabean (using Axis' built in array and bean deserialisers) 2) Together with soap_1_1_to_1_2.xsl and soap_to_rdf.xsl, it will output a file named 'transformed_to_rdf_latest.xml' (the results of a soap encoded data reponse from the service, the data now transformed and serialised in the RDF/XML syntax). 1) and 2) represent prerequisite stages in how the
client may want to handle data it consumes from the service. 1) is for handling data in an (java) object-oriented, non RDF-aware environment. 2) is for handling data in an RDF environment. 1) and 2) correspond to e.g. B and e.g. A respectively in the diagram that represents this service in Section 2.

Sesame itself,
A DBMS with a JDBC-driver (we have used MySQL version 3.23, installed on the same machine as the sesame software, with the mysql-connector-j-2.0.14-bin.jar for database connectivity) and
a Java servlet container for running the Sesame servlets (we've used jakarta-tomcat-4.1.12 with j2sdk1.4.0_02)

Sesame and the servlet container
As described in the sesame installation instructions, you will need to place the extracted sesame.war file in a directory (called, say 'sesame') in your [TOMCAT_DIR]/webapps/ directory. And you will need to place the appropriate jar files (the jdbc jar and xerces.jar, soap.jar, icu4j.jar and jena.jar that are found in the sesame distribution) in [SESAME_DIR]/WEB-INF/lib/ (see http://sesame.aidministrator.nl/publications/users/ch01s05.html)

To configure a fresh installation of Sesame you can copy the example sesame configuration file at [SESAME_DIR]/WEB-INF/system.conf.example to [SESAME_DIR]/WEB-INF/system.conf. Then you will need to edit this file for the correct path to the directory where Sesame is installed on your machine, and for a new SESAME_PASSWORD. Again, this is all detailed in the Sesame installation instructions.

Uploading the Ruby test data to Sesame
Similarly to as described in the Sesame installation instructions, you can:
- create a new user account (with all privileges granted) - called 'sesame' and then
- create a new MySQL database called 'rubyinfodb' (something like this:

```bash
C:\mysql\bin>mysql -u sesame -p
Enter password: ******
Welcome to the MySQL monitor. Commands end with ; or \g.
Your MySQL connection id is 74 to server version: 3.23.54-max-nt
Type ‘help;’ or ‘\h’ for help. Type ‘\c’ to clear the buffer.
mysql> CREATE DATABASE rubyinfodb
->
Query OK, 1 row affected (0.01 sec)
mysql> GRANT ALL ON rubyinfodb.* TO sesame;
Query OK, 0 rows affected (0.01 sec)
mysql> QUIT
Bye
C:\mysql\bin>
```
- With tomcat stopped, now open the file [SESAME_DIR]/WEB-INF/system.conf and add a new entry for the new repository, copying the format of the test one that comes with sesame e.g.:

```xml
<repository id="mysql-rubyInfo">
  <title>MySQL Ruby Products Information DB</title>
  <sailstack>
    <!-- SyncRdfSchemaRepository prevents multiple concurrent transactions and reads during a transaction -->
    <sail class="nl.aidministrator.rdf.sail.sync.SyncRdfSchemaRepository"/>
    <sail class="nl.aidministrator.rdf.sail.sql92.MySQLSail">
      <param name="jdbcDriver" value="com.mysql.jdbc.Driver"/>
      <param name="jdbcUrl" value="jdbc:mysql://localhost:3306/rubyinfodb"/>
      <param name="user" value="sesame"/>
      <param name="password" value="sesame"/>
    </sail>
  </sailstack>
  <acl worldReadable="true">
    <user login="testuser" readAccess="true" writeAccess="true"/>
  </acl>
</repository>
```

- startup tomcat
- Bring up http://localhost:8080/[SESAME_DIR] and you should see the new database in the drop down list
- Select this database (rubydb) and choose (from the toolbar) to Add (copy-paste) data (if not url). You can copy the data (representing "Rubyinfo" objects) that we have used for this example: from here: upload data. You should now get all the data parsed and uploaded (Sesame will tell you the results
of the transaction) and you can choose to extract it - have a look at it.

A standalone test of Sesame over http

It is possible to do command-line testing of Sesame (over http) via the SesameClient.java file (SESAME_SOURCE_DIRECTORY/ai/administrator/rdf/client/SesameClient.java). We have adapted this client to form the "bridge" between Axis and Sesame as described below. The adapted file is {PATH TO SESAME}/java/rubysoap/services/SesameTestStandalone.

Call this class from the commandline

```java
java rubysoap/services/SesameTestStandalone
```

It issues a hardcoded query against the sesame database "RubyInfodb" in which we uploaded data in order to test we can soap-->rdf roundtrip. All matching RubyInfo "Objects" are returned using Sesame's tabulated form for result sets. The result set is then mapped to javabeans in a proprietary way.


The query string can be adapted according to RQL conventions. The beans required are

```java
RubyInfo.java, Category.java, Owner.java, Product.java
```

Obtaining Axis

We installed Axis as described in Axis' installation instructions and also in the README that comes with the axis downlaod. REQUIRE: We used Ant Version 1.5.1 in order to use the build for Axis. We placed the Axis classes from the build in TOMCAT_DIR/webapps/ in a directory we have called 'swadWP4'. We performed the "validate" and testing tasks as required for an Axis installation.

### Axis Happiness Page

**Examining webapp configuration**

<table>
<thead>
<tr>
<th>Needed Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tomcat SAAF API (javax.xml.soap.SOAPMessage) at C:\Program Files\tomcat-4.1.2\common\lib\soap.jar</td>
</tr>
<tr>
<td>Tomcat JAX-WS API (javax.xml.soap.JAX-WS) at C:\Program Files\tomcat-4.1.2\common\lib\jws.jar</td>
</tr>
<tr>
<td>Apache-Ant (org.apache.antlr) at <a href="http://AntServer@C:%5CProgram">http://AntServer@C:\Program</a> Files\tomcat-4.1.2\webapps\swadWP4\WEB-INF\lib\ant.jar</td>
</tr>
<tr>
<td>Jakarta-Commons Discovery (org.apache.commons.discovery) at C:\Program Files\tomcat-4.1.2\webapps\swadWP4\WEB-INF\lib\commons-discovery.jar</td>
</tr>
<tr>
<td>Jakarta-Commons Logging (org.apache.commons.logging) at C:\Program Files\tomcat-4.1.2\webapps\swadWP4\WEB-INF\lib\commons-logging.jar</td>
</tr>
<tr>
<td>Log4j (org.apache.log4j) at C:\Program Files\tomcat-4.1.2\common\lib\log4j-1.2.14.jar</td>
</tr>
<tr>
<td>IBM's WSDA4Java (com.ibm.wsdlfactory.WSDLFactoryImpl) at C:\Program Files\tomcat-4.1.2\webapps\swadWP4\WEB-INF\lib\wss4j.jar</td>
</tr>
<tr>
<td>JAX-WS implementation (javax.xml.soap.IntrospectFactory)</td>
</tr>
<tr>
<td>Activemq AAR (javax.activemq.ConsumerFactory) at C:\Program Files\tomcat-4.1.2\common\lib\activemq.jar</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Optional Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tomcat Mail API (javax.mail.internet.Message) at C:\Program Files\tomcat-4.1.2\common\lib\mail.jar</td>
</tr>
<tr>
<td>Tomcat XML Security API (org.apache.xml.security.Init) at C:\Program Files\tomcat-4.1.2\webapps\swadWP4\WEB-INF\lib\xalan.jar</td>
</tr>
</tbody>
</table>

**The core axis libraries are present. The optional components are present.**

**WSDD**

The appropriate classes for the service @RubySoapServiceTest.java, @RubyInfo.java, @Category.java, @Owner.java, @Product.java are placed in the TOMCAT_DIR/webapps/swadWP4/WEB-INF/classes directory, and the appropriate jar files (example for sesame and the sesame client) in the lib directory for this webapp.

```java
java org.apache.axis.wsd1.Java2WSDL
-o somewebservice.wsd1
-l "http://localhost:8080/axis/services/SesameServiceTest"
-n "urn:seweservices"
-p somewebservice="urn:seweservices" somewebservice.SesameServiceTest
```

Where:

- `o` indicates the name of the output WSDL file
- `l` indicates the location of the service
- `n` is the target namespace of the WSDL file (which should more correctly be http:/.../not urn)
- `p` indicates a mapping from the package to a namespace. There may be multiple mappings, the class specified contains the interface of the webservice.

In order to deploy a service via Axis it is necessary to create a WSDL file as mentioned earlier. This can be authored "by hand", or from a temporary folder we can now run the Axis wsd1java tool - with the server-side switch - on the service's classes. This will generate the deploy.wsd and undeploy.wsd files (and also server-side skeletons which in fact we do not use in this example). The wsd1 generated will look something like this:

```xml
  <wsdl:service name="SesameServiceTest">
    <wsdl:port type="wsdl:SoapService" operation="SoapOperation"/>
  </wsdl:service>
</wsdl:definitions>
```

```java
java org.apache.axis.axis.client AdminClient deploy.wsd1
-o somewebservice.wsd1
```

Where:

- `o` indicates use of the server-side WSDL file
- `l` indicates the location of the service
- `n` is the target namespace of the WSDL file (which should more correctly be http:/.../not urn)
- `p` indicates a mapping from the package to a namespace. There may be multiple mappings, the class specified contains the interface of the webservice.

```xml
  <wsdl:service name="SesameServiceTest">
    <wsdl:port type="wsdl:SoapService" operation="SoapOperation"/>
  </wsdl:service>
</wsdl:definitions>
```

```java
java org.apache.axis.axis.client AdminClient deploy.wsd1
-o somewebservice.wsd1
```

Where:

- `o` indicates use of the server-side WSDL file
- `l` indicates the location of the service
- `n` is the target namespace of the WSDL file (which should more correctly be http:/.../not urn)
- `p` indicates a mapping from the package to a namespace. There may be multiple mappings, the class specified contains the interface of the webservice.

```xml
  <wsdl:service name="SesameServiceTest">
    <wsdl:port type="wsdl:SoapService" operation="SoapOperation"/>
  </wsdl:service>
</wsdl:definitions>
```
As you can see from this file, we use the Axis built in RPC provider, with a SOAP data-encoded format for the response message body. The deploy and undeploy.wsdd files in the servlet container (i.e. tomcat in our case) with the service's java class files. In the same directory it is now possible to deploy the service using Axis' AdminClient tool from the command line, like this:

```
```

Accessing the service

Having deployed the service it should now be possible to access the WSDL for the service from the localhost machine, say, thus:

```
```

You can access the service:
- via a custom 'client', the required beans, and the two stylesheets: `soap_1_1_to_1_2.xsl`, `soap_to_rdf` (note we used `Saxon` for the XSLT to RDF)
- watch the soap messages go back and forth via the Axis tcpmon (TCP Monitor) tool (this is a very handy tool).

---

**Section 4: The Soap data model and the RDF data model - transforming from the former to the latter**

[TODO] Rough thinks:
- SOAP data model - has no equivalent way of defining a schema as with RDF schema, but you can do some typing in your instance data (optional typing of nodes). It's not as "rich" as RDF & certainly, new, undeveloped - RDF - has RDFS & also layers s.a. Owl & also rdf query etc, whereas there's no SOAP query. However SOAP in conjunction with RDF schema allows "meaningful" data aggregation
opportunities?

- SOAP nodes & edges roughly correspond to RDF resources & properties & In RDF 'everything is a resource' - and in SOAP syntax ELEMENT INFORMATION ITEM elements are used for nodes/edges

- SOAP simple values are like rdf literals - in the syntax SOAP uses CHARACTER INFORMATION ITEM elements for these

- Where RDF has b-nodes, SOAP has ref/id pairings - using ATTRIBUTE INFORMATION ITEM attributes to reference them

- Where RDF kinda has support for collections, SOAP supports ARRAYS

More research required .. e.g. re trying to serialise more arbitrary structures in the SOAP data encoding syntax. For e.g. stuff described in OWL - how would that map to the soap data encoding model easily via XSLT?