Setting up SISS Services

Vocabulary Services
Overview

• Why Vocabularies?

• What constitutes a vocabulary

• Vocabulary service deployment and setup
  • SISSVoc
  • OpenRDF
## Rationale

### Names of beer glasses in various Australian cities

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>115 ml (4 fl oz)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>small beer</td>
<td>–</td>
<td>shetland</td>
</tr>
<tr>
<td>140 ml (5 fl oz)</td>
<td>pony</td>
<td>pony</td>
<td>pony</td>
<td>–</td>
<td>horse/pony</td>
<td>pony</td>
<td>pony</td>
</tr>
<tr>
<td>170 ml (6 fl oz)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>six (ounce)</td>
<td>small glass</td>
<td>bobbie/six</td>
</tr>
<tr>
<td>200 ml (7 fl oz)</td>
<td>seven</td>
<td>seven</td>
<td>boer[^5]</td>
<td>butcher</td>
<td>seven (ounces)</td>
<td>glass</td>
<td>glass</td>
</tr>
<tr>
<td>225 ml (8 fl oz)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>eight (ounce)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>255 ml (9 fl oz)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>schooner</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>350 ml (12 fl oz)</td>
<td>schmiddy</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>570 ml (20 fl oz)</td>
<td>pint</td>
<td>pint</td>
<td>pint</td>
<td>imperial pint</td>
<td>pint</td>
<td>pint</td>
<td>pint</td>
</tr>
</tbody>
</table>

**Notes:**

1. Entries in **bold** are common.
2. Entries in *italics* are old-fashioned and/or rare.
3. Entries marked with a dash are not applicable.
4. The “fl oz” referred to here is the *imperial fluid ounce*.
5. The term “beor” is ambiguous. Historically it could be a 170 ml or a 285 ml glass in Tasmania, 200 ml or 285 ml in Far North Queensland, etc. These days ordering a “beor” will get the standard glass size for the particular pub you are in, which typically could be 285 ml, 425 ml, 570 ml, etc.
6. “Half Pint” is much more common in Canberra than Sydney.
7. “Pot” or also known as Pot glass
8. Traditionally 425 ml is a size not found in Western Australia or Victoria.

**References:**

1. The Aussie Beer Baron
2. Buying Beer in Australia
3. Guidelines at a glance
4. Ordering Beer
5. Liquor Merchants Association of Australia
6. Which Size Beer Do Ya Want, Mate?
e.g. Data record

• You just received 37 spreadsheets from a colleague containing all the data you need…

<table>
<thead>
<tr>
<th>t</th>
<th>T</th>
<th>v</th>
<th>l</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>1221</td>
<td>0322</td>
<td>17.89</td>
<td>0.783</td>
<td>2300</td>
</tr>
</tbody>
</table>

+ t, T = temperature or time or tonne..
- if temperature, is it in K? °C?
+ Is v a velocity or a volume? In what unit?

This could certainly be improved!
Ok, I’m convinced!
I need some sort of controlled vocabulary.

But, what does it look like?
How do we develop a Vocabulary Service?

Resource Data Framework (RDF) is an abstract data model based upon the idea of making statements about resources in the form of triples:

Subject → Predicate → Object

Example:

shiraz hasColour red
SISSVoc Overview

• SISSVoc encompasses:
  • Formalised concepts and relationships in the Earth & Environmental Science domain
  • HTTP API on top of Semantic Web Technologies
  • Query and Navigation of ontologies
Formalised in RDF using SKOS

Validated online using w3c validator tool

Imported into Triple Store
SISSVoc Architecture

User

Internet

SISSVoc – HTTP Rest Interface

HTTP

Sesame OpenRDF Triple Store

SKOS/RDF Vocabularies

Persistent Storage
SQL/Flat File etc

SPARQL Queries
SISSVoc RESTful Interface

• RESTful SISSVoc query to lookup a vocabulary concept via a human readable label:

```
http://localhost:8080/SISSVoc
/getConceptByLabel?/Lithology/Andesite
```

(SISSVoc API)
(Repository)
(Label to lookup)

• Entire API is documented on our wiki
So how do I deploy this?

- Surprisingly easily …
  - …if you are familiar with Java web servlet deployments
- OpenRDF Sesame + Web Interface
  - [http://www.openrdf.org/](http://www.openrdf.org/)
  - Simple WAR deployment
- SISSVoc
  - Same as above
  - Need to configure the location of OpenRDF Sesame.
- Process is documented on our wiki
Example

• So I have an RDF Document... Now what?
  • For this example we’ll be jumping into the realm of ERML

• We firstly want to load this RDF into a triple store
  • OpenRDF will be our store of choice

• We can use the OpenRDF instance on your SISS in a box VM
  • However if you want to deploy a new OpenRDF instance:
    https://twiki.auscope.org/wiki/Trash/GridSESAMERepositoryManagement

• To test our triple store we use some example SPARQL queries
  • They are representative of what SISSVoc will be using

• The final test is to query our vocabulary through SISSVoc
Conclusions

• Vocabulary services are crucial for data interoperability
  • Just because you’ve standardised the model, doesn’t mean the content of the model will be understandable.

• RDF provides us with a means for expressing a vocab
  • There are numerous tools for creating RDF documents

• The RDF can be queried once put in a triple store

• SISSVoc provides a ‘restful’ interface into a triple store