Land & Water

Summer Internship 2015/16



Developing Environmental Vocabularies

By Xavier Butcher – 29/01/16

Contents

[1 Heading 1 2](#_Toc394073019)

[1.1 Heading 2 2](#_Toc394073020)

[1.2 Formatting [style = Heading 2] 2](#_Toc394073021)

# Keywords

DataBase, Editors, Environment, Excel, Methods, Notepad, RDF, Registry, Semantics, SKOS, Soil, SPARQL, SQL, Taxonomy, TopBraid, URI, Vocabulary, Wiki.

# Introduction

The aim for this project was to create and collate a number of online vocabularies into a standard format and location. Vocabularies are often hosted online, and are used to support environmental applications. They contain labels, definitions and other related information for a particular concept or registry. This project mostly focussed on soil and land operations. Currently, a number of vocabularies are unique to their individual application, and are often in different formats (text, csv, pdf, or lists). Having a definitive vocabulary on a subject is critical, as the same term could have different definitions or meanings in different vocabularies.

It is also necessary to introduce semantic relationships, which essentially add another layer of detail to vocabulary. This means different terms can be placed in a hierarchical order (broader and narrower), as well as relating similar terms. These vocabularies follow the guidelines of SKOS, the Simple Knowledge Organization System, which is an industry standard of rules and guidelines across online vocabularies which aims to make publication and use of vocabularies an easy and standard process.

SPARQL the SPARQL Protocol and RDF Query Language is, as the acronym suggests, an RDF query language. It is tending to be the industry standard query language. SPARQL allows the user to query semantic data in databases, and retrieve and manipulate RDF data.

There were 4 main tasks relating to this project, which were completed successfully:

•          Source relevant existing vocabularies

•          Convert the vocabulary content using semantic web technologies (RDF/SKOS)

• Harmonising vocabulary content with existing vocabularies where applicable

•          Publish vocabularies to SPARQL triple stores where applicable

# Example

For those who have little or no knowledge of vocabularies, SKOS or semantics, this section will give a basic example of how these theories are used.

This example will centre on the ‘Concept’ of a “Margherita Pizza”. As a bare minimum, a Concept must have a Label (Margherita Pizza), a URI (Unique Resource Identifier – a unique web address that the concept can be located) and a Definition (Margherita Pizza is a flatbread generally topped with tomato sauce and cheese and baked in an oven). A Concept may also be part of a hierarchy of other Concepts. In this case, “Margherita Pizza” is a Narrower Concept of “Pizza”, which itself is a Narrower Concept of “Food”. A Concept can also be related to other concepts regardless of hierarchy, for instance “Margherita Pizza” is related to “Mushroom Pizza”. A Concept can also be part of a collection. For example “Margherita Pizza” can belong to the “Vegetarian Pizza” collection.

# Materials

* National Committee on Soil and Terrain (2009), ‘Australian soil and land survey field handbook (3rd edn).’ (CSIRO Publishing: Melbourne)
* ‘Soil Information Transfer and Evaluation System (SITES) – Database design and exchange protocols (version 2.0).’ Jacquier, D; Wilson, P; Griffin, T; Brough, D. (CSIRO Publishing: Canberra)
* The CSIRO environmental informatics Wiki - <https://wiki.csiro.au/display/EI/Environmental+Informatics+Home>
* The CSIRO vocabularies and vocabulary services Wiki - <https://wiki.csiro.au/display/VOCAB/Home>
* The Linked Data Registry - <http://registry.it.csiro.au/sandbox/student/xavier>
* Microsoft office suite
* Notepad++
* TopBraid Composer
* RDF123
* Microsoft SQL Server Management Studio

# Method

Figure 1 illustrates the steps involved in creating a vocabulary.

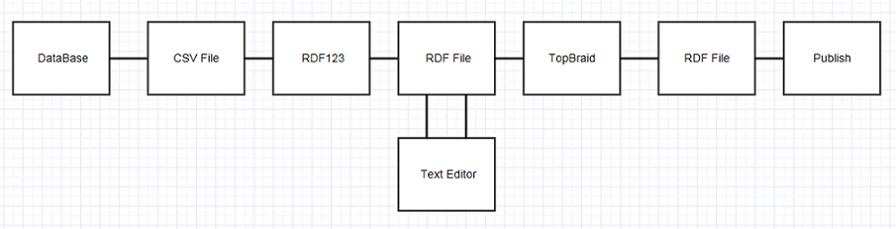


Figure 1: Process diagram

# Results

Blah

# Discussion

Blah

# Further Recommendations

Blah

# Conclusion

Blah

# Acknowledgements

Blah

#### Heading 4

## Formatting [style = Heading 2]

The CSIRO font for all Office programs is Calibri. The styles of the headings vary in font size, weight and colour, and have numbered or not numbered options. Space before and after headings is included in the style, and no blank lines should be inserted between elements.

Main text is 12 pt Calibri, left justified [style = Body text]. Main text must flow in one column spreading over the full width of the page. Space between paragraphs is included in the style, and no blank lines should be inserted between paragraphs.

Two examples of bulleted lists follow. If the dot points are full sentences:

* Each should start with a capital letter. [style = List Bullet]
* Each should end in a full stop.

If the dot points are sentence fragments, follow these conventions:

* no capitals at beginning
  + second level bullet [style = List Bullet 2]
  + second level bullet [style = List Bullet 2]
* third level bullet [style = List Bullet 3]
* no ‘and’ after the second to last dot point
* no full stop at the end of each dot point (except for the last).

An example of a numbered list follows [style = List Number]. These are basic typographic rules for the reports as a whole.

1. In general, all text is aligned on the left, with ragged edges on the right. Do not justify or centre text.
2. second level numbered list [style = List Number 2]
3. second level numbered list [style = List Number 2]

* third level numbered list [style = List Number 3]

1. In general, use sentence case, unless otherwise specified (e.g. subheadings). Capitalise only the first letter of the sentence and proper nouns.
2. Between the full stop and the start of the next sentence, use only one space, not two.
3. Use single quotation marks, not double. Use single quotation marks for quoted speech, and to signal unusual usage.
4. Use an en dash to denote a range of values, e.g. ‘35%–50%’. Do not add spaces on either side.
5. Use an en dash with one space on either side to indicate a break of thought or a side note, e.g. ‘A mix of three of them – myristic, palmitic and palmitoleic acid – triggers a quick upsizing in heart muscle cells.’.
6. Scientific names and titles are set in italics of the same weight, e.g. ‘Homo sapiens first appeared …’. [style = Italics]

Example of a feature box heading [style = Boxed heading]

Example of a feature box body text [style = Boxed text]

### Hyperlinks [style = Heading 3]

All hyperlinks must include a ScreenTip. You can see the ScreenTip text if you hover over the hyperlink below. To enter the ScreenTip, right-click the hyperlink, click ‘Edit Hyperlink’, click the ‘ScreenTip’ button, and then enter ‘[Organisation Name] website’. For example: <[www.australia.gov.au](http://www.australia.gov.au/)> (ScreenTip text is ‘Australian Government website’) and <[www.anu.edu.au](http://www.anu.edu.au/)> (ScreenTip is ‘Australian National University website’). In reference lists, the ScreenTip for hyperlinks should be ‘Hyperlink to: ‘[Title of report]’’. For example: <[publications.csiro.au/rpr/download?pid=csiro:EP132686&dsid=DS5](https://publications.csiro.au/rpr/download?pid=csiro:EP132686&dsid=DS5)> (ScreenTip is ‘Hyperlink to: Water resource assessment for the Great Artesian Basin’).

### Equations

Equations should be numbered consecutively as they appear in the text with Arabic numerals, for example ‘(1)’. Equations must be typed, not hand printed. Go to ‘Insert’ > ‘Equation’ to generate equations within Word. The style includes spacing above and below [style = Equation]. Alternative text must be inserted for equations. This website should be used to generate alternative text for equations: <[www.wiris.net/demo/editor/demo/en/](http://www.wiris.net/demo/editor/demo/en/)>.

 (1)

### Figures and photographs

The resolution of figures should be minimum 300 dpi. The preferred file format is Windows enhanced metafile for figures. This can be achieved by:

* copy the figure in the source file (e.g. Excel)
* in Word, choose ‘paste special’, then click ‘Picture (Enhanced Metafile)’ and OK
* right-click on the image and select Text wrapping and choose ‘In Line with Text’.

Figure numbers and captions appear at the bottom of the figures. Figures should be numbered consecutively with Arabic numerals, in the order in which reference is made to them in the text, e.g. ‘Figure 1’, ‘Figure 2’, etc.

In the figure caption, the figure number is followed by a space. Captions do not end in a full stop, whether they are full sentences or sentence fragments. If required insert a source reference below the figure caption.

Resolution of photographs should be minimum 300 dpi. The preferred file formats are jpg or png for photographs. Do not use any picture styles (rounded corners, drop shadows, frames) on photographs.

All figures require alternative text. To insert a description, right-click the figure, click ‘Size’, and then click the ‘Alt Text’ tab. In instances where it is not possible to provide alternative text, please use the disclaimer: ‘For a description of this image please contact [email address]’.

To add alt text:

* Word 2007 – right-click on your figure, select Size > Alt Text *or* Format Picture > Alt Text
* Word 2010 – right-click on your figure, select Format Picture > Alt Text
* Word 2013 – right-click on your figure, select Format Picture > Layout & Properties tab > Alt Text

Line graph

Figure 2 Example figure, full width (17 cm) [use References/Insert Caption, label = Figure | style = Caption]

Use the caption note to add additional information under the caption, this doesn’t appear in the TOC [style = Caption Note]. To reduce the space between the caption and caption note, select the caption above, go to the Home tab - Paragraph - Line and Paragraph Spacing - in the drop-down click on ‘Remove Space After Paragraph’

1. [Insert text, style = Figure/Table Source]

(a) (b)

Bar graphBar graph

Figure 3 Example figures, half width (8.4 cm) [use References/Insert Caption, label = Figure | style = Caption]

Pie chart

Figure 4 Example figure, full width (17 cm) [use References/Insert Caption, label = Figure | style = Caption]

### Tables

[Table style = Table\_CSIRO]

Table numbers and captions appear at the top of the tables. Tables should be numbered consecutively with Arabic numerals, in the order in which reference is made to them in the text, e.g. ‘Table 1’, ‘Table 2’, etc.

The table itself is left aligned and should fit within the margins. Authors can choose how wide they wish their table(s) to be within the margin: either extending to the margins, or narrower if desired.

For accessibility purposes you must identify the column headers of tables: select the first row of the table, right-click and select Table properties > Row, tick the ‘Repeat as header row at the top of each page’ option. Also try to avoid complex tables, simple data tables without merged/nested cells are preferred and when possible separate the complex tables into simple tables.

In the table caption, the table number is followed by a space. Captions do not end in a full stop, whether they are full sentences or sentence fragments. If required insert a source reference below the table.

Table 1 Example table, full width (17 cm) [use References/Insert Caption, label = Table | style = Caption]

|  |  |  |  |
| --- | --- | --- | --- |
| column heading [style = Columnheading] | column heading | column heading | column heading |
| Text [style = TableText] | Text [style = TableText] | Text | Text |
|  | * Bullet [style = TableBullet] * Bullet | Text | Text |

1. [Insert text, style = Figure/Table Source]

Table 2 Example table, half width (8 cm) [use References/Insert Caption, label = Table | style = Caption]

|  |  |  |  |
| --- | --- | --- | --- |
| column heading [style = Columnheading] | column heading | column heading | column heading |
| Text [style = TableText] | 15.5 | 15.5 | 15.5 |
| Text [style = TableText] | 20.5 | 20.5 | 20.5 |
| Total | 36.0 | 36.0 | 36.0 |

1. [Insert text, style = Figure/Table Source]

Table 3 Example table, full width (17 cm) with coloured row headings [use References/Insert Caption, label = Table | style = Caption]

|  |  |  |
| --- | --- | --- |
| column heading [style = Columnheading] | column heading | column heading |
| Row heading [style = RowHeading] | Text [style = TableText] | Text [style = TableText] |
|  | Text [style = TableText] | Text [style = TableText] |
|  | Text [style = TableText] | Text [style = TableText] |
|  | Text [style = TableText] | Text [style = TableText] |
| Row heading [style = RowHeading] | Text [style = TableText] | Text [style = TableText] |
|  | Text [style = TableText] | Text [style = TableText] |
|  | Text [style = TableText] | Text [style = TableText] |
|  | Text [style = TableText] | Text [style = TableText] |
| Row heading [style = RowHeading] | Text [style = TableText] | Text [style = TableText] |
|  | Text [style = TableText] | Text [style = TableText] |
|  | Text [style = TableText] | Text [style = TableText] |

[Insert blank line if no source]

### Accessibility

Documents that are to be published on the web are required to be accessible to Web Content Accessibility Guidelines (WCAG) 2.0 Level AA. Documents should be provided in both Microsoft Word and Adobe PDF format.

To ensure documents are accessible, these requirements must be followed:

* Use the standard CSIRO report template.
* Use short headings throughout the document, including the title. If necessary, a subtitle could be used.
* Use styles for all text.
  + Use only the recommended styles that are already contained within this template.
  + Table styles are also included.
  + Do not modify the current styles.
* Include alternative text (alt text) for all images, tables, equations and objects.
  + For images, use this statement for the alt text where it is not possible to describe the image: ‘For a description of this image, please contact [email address].’
  + For equations, use this website to generate alt text: <[www.wiris.net/demo/editor/demo/en/](http://www.wiris.net/demo/editor/demo/en/)>.
* All images and objects should be in line with text and not floating.
* Avoid the use of text boxes and columns.
* Tables:
  + Use the recommended table styles contained within the template.
  + Keep tables simple – try to avoid merging and splitting cells and only use simple, single row headers.
  + If complex tables (with merged cells) are needed, add this statement to the table alt text: ‘For a description of this table, please contact [email address].’
  + Specify the table header row by selecting the first row, right-click the table, click ‘Table Properties’, ‘Row tab’ and then tick ‘Repeat as header row at the top of each page’. You must do this for all tables that have a header row, even if they do not flow onto another page. You do not need to do this if you have a table that uses a header column instead.
* All web hyperlinks must include a ScreenTip.
* Do not use empty paragraph returns.
* Do not use double spacing or use multiple spaces for alignment.
* Include a table of contents for large documents.

|  |  |  |
| --- | --- | --- |
|  | | |
| CONTACT US  t 1300 363 400  +61 3 9545 2176  e enquiries@csiro.au  w www.csiro.au  At CSIRO we shape the future  We do this by using science to solve real issues. Our research makes a difference to industry, people and the planet.  As Australia’s national science agency we’ve been pushing the edge of what’s possible for over 85 years. Today we have more than 5,000 talented people working out of 50-plus centres in Australia and internationally. Our people work closely with industry and communities to leave a lasting legacy. Collectively, our innovation and excellence places us in the top ten applied research agencies in the world.  WE ASK, WE SEEK AND WE SOLVE |  | For further information  Insert Flagship/Business Unit name  Insert contact name t +61 0 0000 0000  e first.last@csiro.au  w www.csiro.au/flagship-businessunit  Insert Flagship/Business Unit name  Insert contact name t +61 0 0000 0000  e first.last@csiro.au  w www.csiro.au/flagship-businessunit  Insert Flagship/Business Unit name  Insert contact name t +61 0 0000 0000  e first.last@csiro.au  w www.csiro.au/flagship-businessunit |