

org/wp-content/uploads/2015/11/ASI-BestPracticesIndexing2015.pdf

International Association of Business Communicators (n.d.) 'IABC code of ethics for professional communicators'. www.iabc.com/about-us/leaders-and-staff/code-of-ethics/

SI (2007) 'Society of Indexers' code of professional conduct.' www.indexers.org.uk/index.php?id=159

Society for Technical Communication (1998) 'Ethical principles.' www.stc.org/about-stc/the-profession-all-about-technical-communication/ethical-principles

Books

Borko, H. and Bernier, C. L. (1978) 'Ethical considerations', pp. 223–6 in *Indexing concepts and methods*. New York: Academic Press.

Wellisch, H. (1991). 'Narrative indexing', pp. 244–5 in *Indexing from A to Z*. New York: H. W. Wilson.

Articles and items from *The Indexer*

The Indexer contains a great many articles, letters and other items relevant to ethics in indexing. Please see its indexes at www.theindexer.org/contents/indexing-practice.htm and www.theindexer.org/contents/onlineindex.htm. Here are a few I found of particular interest.

ASI Committee on Ethics, Standards and Specifications (1975) 'Ethics and specifications', *The Indexer* 9(4),174–7.

Bell, H. K. (1991) 'Bias in indexing and loaded language', *The Indexer* 17(3),173–7.

Bell, H. K. (1996) 'Whom should we aim to please?' *The Indexer* 20(1), 3–5.

Browne, G. (1996) 'Professional liability of indexers', *The Indexer* 20(2), 70–3.

Eisenschitz, T. (1985) 'Copyright for indexers', *The Indexer* 14(4), 253–4.

Halliday, J. (2007) 'Professionalism and the indexer', *The Indexer* 25(3), 167–8.

Intner, S. S. (1984) 'Censorship in indexing', *The Indexer* 14(2), 105–8.

Jacobs, C. (2007) 'Ethical places, ethical spaces: stopping to listen', *The Indexer* 25(3),161–6.

Mallory, M. and Moran, G. (1994) 'Scholarly search for the truth', *The Indexer* 19(2), 99–101.

Heather Ebbs has been indexing, editing and writing for more than 30 years and has indexed over 700 books, journals and reports in a wide variety of subjects and disciplines. A past president of ISC/SCI, she is currently International Liaison for the society. Email: h1ebbs@gmail.com

Website indexing

Mary Coe

Website indexing may be done using multiple methods and tools, but the basic art of indexing still applies. As websites are often organic and constantly changing, the work of the web indexer is never done. In fact, the need for website indexing may only increase in the near future as new initiatives for collating and linking data appear.

Web indexing includes the use of search engines and metadata, the organization of web links, and the creation of website indexes (Browne and Jermy, 2004: 136). While all of these features can be considered separately, they often work together, with the common aim of helping users to find online information. This article considers indexing within individual websites, not how large search engines, such as Google, trawl the World Wide Web.

There are several methods of website indexing:

- creation of back-of-book (BOB)-style indexes (often called A–Z indexes) within websites
- application of keywords to metadata fields in web pages
- construction of thesauri or taxonomies for various uses in websites.

Back-of-book style indexes

Browne and Jermy (2004: 126) define BOB-style indexing as the 'creation of a website index that looks and functions like a back-of-book index. It will usually be alphabetically organised, give detailed access to information, and contain

index entries with subheadings and cross-references.' They argue that BOB-style indexes on websites offer the user the advantage of direct access to specific subjects of interest and the ability to browse for a quick coverage of the content on a website. They also note that such indexes are in a format that is familiar to most users (Browne and Jermy, 2004: 3–4). In most cases, entries in this type of index are

Web and Electronic Indexing Special Interest Group
Web Dictionary of Cybernetics and Systems
Web indexing
about
Beyond Book Indexing (Rowland and Brenner)
discussion group
Indexing Specialties: Web Sites (Hedden)
and metadata
Web and Electronic Indexing SIG
Website Indexing (Browne and Jermy)

Figure 1 From the ASI website index

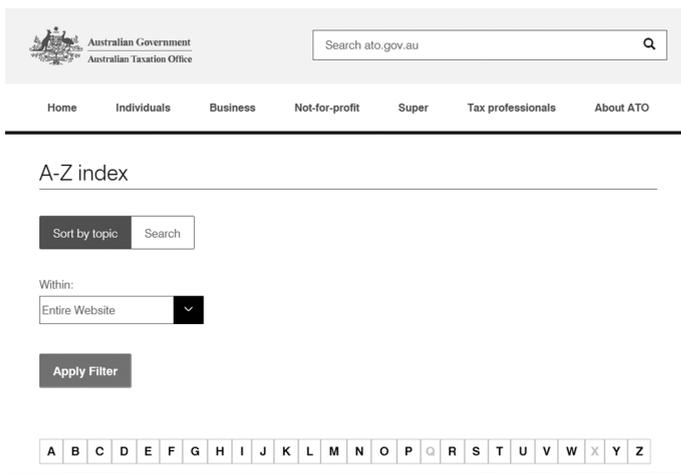


Figure 2 A variety of search tools on the Australian Tax Office website

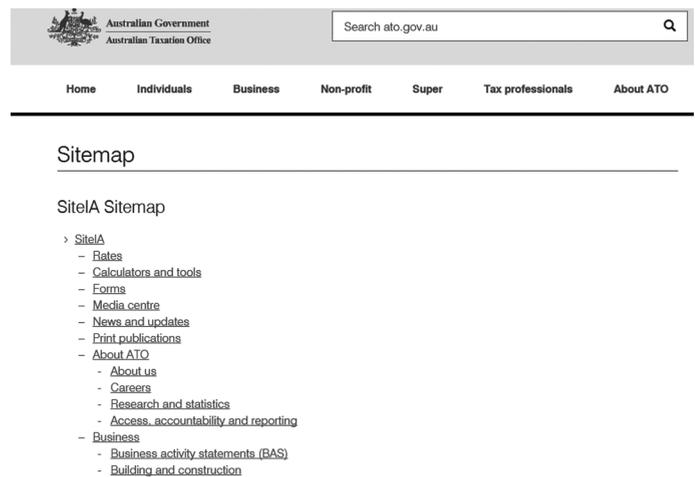
hyperlinked to webpages. An example of a BOB-style index that is presented in a traditional format can be found on the ASI website (www.asindexing.org/a-z-index/). A section of this index is shown in Figure 1.

BOB-style indexes on websites might also include site maps, which are defined as an ‘overview of the navigational structure of a website, acting like a Table of Contents, and used to orient users and show them the scope of the site’ (Browne and Jermy, 2004: 134). These indexes can be used in conjunction with site-specific search engines to enable users to access information in different ways. Hedden (2004) states that ‘A–Z indexes are more accurate than search engines for searching the content of a website’, but it should be noted that users may prefer to use search engines as well as indexes. Offering both choices could provide complementary ways of finding information (Browne and Jermy, 2004: 4). Large sites may offer multiple search tools, such as an A–Z index, a site map and a site-specific search engine. For example, the Australian Tax Office (ATO) website (www.ato.gov.au) offers all three of these options (see Figure 2).

Metadata web indexing

Metadata is defined as ‘structured data about data, which may include information about the author, title and subject of web resources’ (Browne and Jermy, 2004: 131). Metadata is not displayed on a webpage, it is ‘hidden’ in the page code. In some cases, it is used by a site-specific search engine to assist users in finding information. For example, if you were to view the page source for a webpage about fatigue on the NPS MedicineWise site (www.nps.org.au/topics/signs-and-symptoms/fatigue/for-individuals/what-it-is), you would see the metadata shown in Figure 3, which includes information about the title of the page and the subject content. This metadata is applied by human indexers and is then used by the website’s internal search engine, Funnelback, to answer search queries from users.

This type of indexing does not result in a traditionally formatted index, such as the BOB-style index, and users are not often aware of it. However, it is similar to BOB-style indexing in that it groups content, provides synonyms or cross-references, and is created by human indexers who



are considering concepts rather than simply pulling words from text on the page. ‘Keyword’ is a term that is often used to describe subject terms applied as metadata. In the example shown, the keywords can be seen in the ‘Subject’ content section and include the terms ‘fatigue’ and ‘signs and symptoms.’

Thesauri or taxonomies

A thesaurus or taxonomy may also be used with websites. A thesaurus is ‘a structured list of approved subject headings (preferred terms) showing the relationships between them’, and a taxonomy is a ‘controlled vocabulary used primarily for the creation of navigation structures for websites’ (Browne and Jermy, 2004: 135). A thesaurus is usually more hierarchical or structured than a taxonomy. It may also include terms that are not to be used for indexing (non-preferred terms) with cross-references to the preferred terms.

```
<!-- Metadata for Funnelback search -->
-->
<meta name='Title.page' content='What is fatigue?' />
<meta name='Title.HP' content='' />
<meta name='HP.only' content='No' />
<meta name='Description.internal' content='Fatigue is a feeling of constant tiredness or weakness that is not relieved with rest; it can affect anyone. Find out more.' />
<meta name='Stream.keyword' content='' />
<meta name='Content.type' content='Topic' />
<meta name='Date.revised' content='2014-09-02 00:00:00' />
<meta name='Content.section' content='Health topic' />
<meta name='Subject' content='fatigue|signs and symptoms|' />
```

Figure 3 Metadata for the NPS MedicineWise site

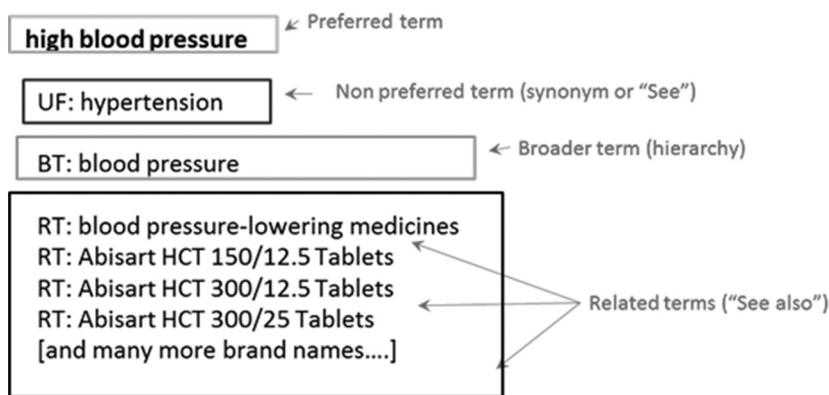


Figure 4 A sample entry from the NPS thesaurus

Figure 4 is an example of an entry for the preferred term ‘high blood pressure’ from the NPS thesaurus. Note that it includes a non-preferred term (synonym or cross-reference) and that its place in the hierarchy of the controlled vocabulary is indicated by broader terms and related terms (which could also be described as cross-references). While an indexer might see the non-preferred term ‘hypertension’ in webpage content, they would need to apply the preferred term ‘high blood pressure’ as a keyword in metadata. In this way, all content on the topic of ‘high blood pressure’ (or ‘hypertension’) will be grouped on the site.

Hedden (2010: 15) states that controlled vocabularies may be used for the following purposes:

- indexing support
- retrieval support
- organization and navigation support.

Indexing support

A thesaurus or taxonomy can assist indexers by ensuring consistency of indexing. This is important for websites that are maintained by multiple indexers, contain large numbers of pages, and/or require indexing over a long period of time (Hedden, 2010: 15).

In some cases, a thesaurus is displayed on a website so that it can be used by others to assist with indexing their own collections. For example, the Australian Thesaurus of Education Descriptors (ATED) is used to index several databases and the ACER library catalogue, and is also displayed online, so that anyone can use it (<http://cunningham.acer.edu.au/multites2007/index.html>). A similar example is the US National Library of Medicine’s Medical Subject Headings (MeSH) thesaurus. The MeSH thesaurus is used by the library to index articles from biomedical journals for the MEDLINE/PubMED database, but it is also displayed online (<https://www.nlm.nih.gov/mesh/MBrowser.html>).

Retrieval support

Use of a thesaurus or taxonomy can support retrieval by providing greater recall and precision. *Precision* is ‘the relevance to the searcher of the items that are retrieved.

If a search retrieves one hundred documents of which ninety-five are very relevant, that search has high precision’ (Browne and Jermeý, 2004: 132). *Recall* is ‘the proportion of relevant information that is retrieved by a search. If a search only retrieves one hundred relevant documents out of three thousand that are available, that search has low recall. If it retrieves all the available documents on the topic, it has high recall’ (Browne and Jermeý, 2004: 133). If indexers apply keywords from a thesaurus as metadata to a website, and the website’s internal search engine uses this metadata to provide search results, a user is more likely to be presented with all of the relevant content on that topic. In this case, the thesaurus is likely to remain ‘behind the

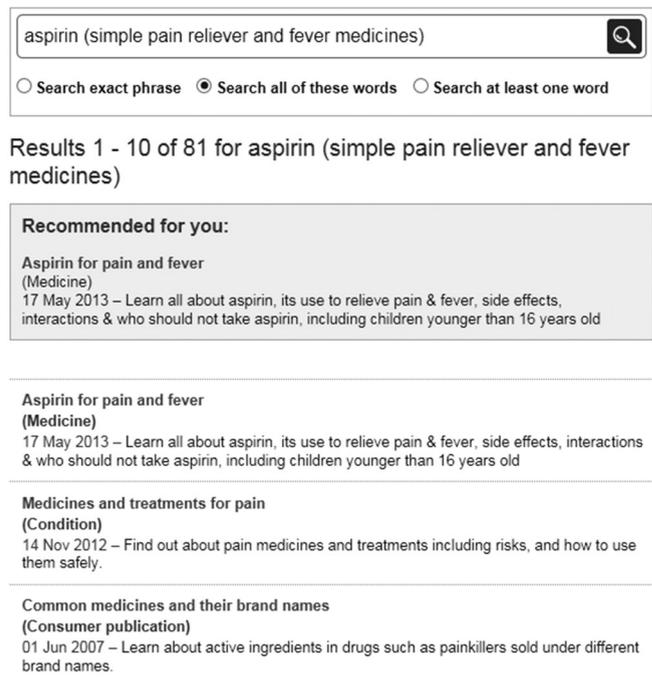
scenes’ and users are not generally aware that it has been used to guide indexing.

For example, keywords applied as metadata on the NPS MedicineWise website (www.nps.org.au/) must be chosen from the NPS thesaurus, which is not visible to external users. Searching for, say, the term ‘GORD’ on the NPS MedicineWise website will reveal content on ‘gastro-oesophageal reflux disease’, ‘GERD’, ‘heartburn’ and ‘reflux’. By grouping all content on this topic in one place using metadata and providing synonyms that users can access from the search box, the indexer has provided the user with better recall and precision than they would achieve from a simple free text search on the term ‘GERD’. A section of the search results page for this query can be seen in Figure 5.

Note that the reader has been advised that their search query has been expanded to ‘heartburn and reflux’ and that they are given the option to conduct a free text search of the website, if they chose to do so. One advantage of the online environment is the ability to give the user a choice of search tools.

Figure 5 Search results on the NPS website

Figure 6 Categories for searching on the NPS website



aspirin (simple pain reliever and fever medicines)

Search exact phrase Search all of these words Search at least one word

Results 1 - 10 of 81 for aspirin (simple pain reliever and fever medicines)

Recommended for you:

Aspirin for pain and fever (Medicine)
17 May 2013 – Learn all about aspirin, its use to relieve pain & fever, side effects, interactions & who should not take aspirin, including children younger than 16 years old

Aspirin for pain and fever (Medicine)
17 May 2013 – Learn all about aspirin, its use to relieve pain & fever, side effects, interactions & who should not take aspirin, including children younger than 16 years old

Medicines and treatments for pain (Condition)
14 Nov 2012 – Find out about pain medicines and treatments including risks, and how to use them safely.

Common medicines and their brand names (Consumer publication)
01 Jun 2007 – Learn about active ingredients in drugs such as painkillers sold under different brand names.

Figure 7 Search results for aspirin on the NPS website

Aspirin for pain and fever

Active ingredient: aspirin

Brand names include: Dispirin, Aspro, Solpirin

What is aspirin?

Aspirin is used to relieve mild to moderate pain, fever (a temperature higher than 38.5°C) and reduce inflammation. It is one of a group of medicines called non-steroidal anti-inflammatory drugs (NSAIDs).

The main effect of NSAIDs is to reduce inflammation, which the body produces in reaction to an injury or infection, as a way of healing the body.

Low-dose aspirin (up to 150 mg a day) can also be taken for a long period of time to act as a 'blood thinner', helping to prevent a heart attack or stroke – see Aspirin (anti-clotting-medicine).

Note about medicines names [\[Show \]](#)

Find out more about who can take aspirin, the side effects of aspirin and the interactions with aspirin.

References

- Rossi S, ed. eAMH online. Adelaide: Australian Medicines Handbook, January 2013.
- Sweetman S, ed. Martindale: The complete drug reference [online]. London: Pharmaceutical Press. (Accessed 27 March 2012).

Latest information - aspirin (simple pain reliever and fever medicines)

Audience:

All (36) Health professionals (1) Consumers (35)

Content type: All content types (36)

Aspro Clear Effervescent Tablets (Medicine)

23 Jan 2015 – Find out about Aspro Clear Effervescent Tablets. Plus information and tips on how to use medicines wisely and safely.

A- A+  

 Tweet  G+1  Recommend  0

Related searches

Select any of the terms below to do a new search using that term.

Broader terms

- ▶ simple pain reliever and fever medicines

Narrower terms

- ▶ Aspro (Tablets)
- ▶ Aspro Clear Effervescent Tablets
- ▶ Aspro Clear Extra Strength Effervescent Tablets
- Show 4 more ...

A+ A-

Recommend this page



CPD activities

- ▶ Analgesics in persistent pain (National Prescribing Curriculum module)

DRUG MISUSE: Implications for pharmacists

Free CPD activity

Start now ▶

Related medicines

- ▶ Aspro Clear Effervescent Tablets
- ▶ Aspro Clear Extra Strength Effervescent Tablets
- ▶ Aspro Tablets
- Show 4 more ...

Organization and navigation support

A controlled vocabulary can be used to support information architecture, which is the 'design of the structure of information systems, particularly websites and intranets, including labelling and navigation schemes' (Browne and Jermeý, 2004: 130). Files in websites are often arranged into directories and subdirectories that are based on navigational structure and that group related content based on a taxonomy (Browne and Jermeý, 2004: 59–61).

For example, content on the NPS MedicineWise website (www.nps.org.au/) is grouped into the following categories: Medicines, Conditions, Health Topics, and Medical tests. These categories can be easily seen in the site structure and may be used for navigation within the site (see Figure 6).

On the NPS MedicineWise website, the thesaurus is used to provide other navigational functionality. A user looking for information about aspirin on the site will find suggestions for related search terms, including broader and narrower terms, that are derived from the NPS thesaurus (Figure 7).

This functionality is not limited to the search results page. It can be found on any page of the website. For example, a user landing on the aspirin page will see also see related search suggestions and a 'stream' of search results on the same topic (Figure 8). All of these functions are driven by the application of keywords from the thesaurus as metadata.

Metadata can also be used for faceted classification within websites. Faceted metadata classification is defined as 'breaking subjects into standard component parts (facets) and presenting these to users as search options' (Browne and Jermeý, 2004: 129). An example of faceted classification can be found on the Choosing Wisely website (www.choosingwisely.org.au/) (Figure 9).

Figure 8 Search suggestions and results for aspirin on the main page for the item

Display by: Specialist medical organisation | Medicine branch | Medical test | Medicine or treatment | Condition or symptom

Expand all | Collapse all

Allergy tests	+
Blood glucose test	+
Carotid artery ultrasound	+
Cholesterol and lipid tests	+
Computed tomography (CT) scan	+

Figure 9 Faceted classification on the Choosing Wisely website

Web indexers' skills and roles

As mentioned earlier, the advantage of the online environment is the ability to provide the user with multiple search tools. Including several types of indexing on a website can be a good thing. Some users may prefer to browse a BOB-style index or site map, others may want to navigate using the site hierarchy, and there are some who prefer to use a search box (hoping for results refined by application of metadata!).

Browne (2007) commented that BOB-style website indexes satisfy general design principles for usability in that they offer recognition rather than recall (they are browsable and do not require the user to consider spelling or word formatting), are a familiar tool for book users, give the user control and freedom (they can move through the index as they wish), and help users to recognize or recover from errors (by providing cross-references and parenthetical qualifiers). These usability heuristics could also be applied to metadata indexing using controlled vocabularies. Providing different access points to information with use of both preferred and non-preferred terms (synonyms or cross-references) from a thesaurus in a search tool and the ability to browse using a hierarchical navigational structure are some of the usability features offered by this method. The key point for both methods is the grouping of similar content by the indexer. This is where the traditional skill of indexing transcends the boundary between print publications and the online environment of websites. Browne and Jermeý state that:

web indexers need traditional indexing skills such as the ability to analyse the subject of documents, to describe that subject in appropriate language for an index, to think of alternative access points (that is, other ways of describing a topic), and to create references to or from headings. (Browne and Jermeý, 2004: 18)

The indexer's skill in providing access points for multiple audiences is also important. For example, the NPS MedicineWise website contains information for both consumers and health professionals. Constructing the NPS thesaurus so that it will provide keywords that can be used by both audiences can be challenging. The fact that the website is constantly changing, with new content added regularly, means that the thesaurus must also constantly change. For example, addition of new content to the NPS MedicineWise website on drugs that treat indigestion and stomach ulcers required major adjustments to the thesaurus to provide both

new information architecture and new keywords for metadata. Before the new content was added, there was a thesaurus entry that looked like this:

indigestion and stomach ulcer medicines
UF: antacids
BT: digestive system

Adding new concepts for specific classes of drugs in this category changed both the hierarchy and terms used in this section of the thesaurus to this:

indigestion, reflux and stomach ulcer medicines
BT: digestive system
NT: antacids
NT: heartburn and reflux medicines
NT: histamine H2 antagonists
NT: other medicines for indigestion, reflux and stomach ulcers

After adjustments were made to the thesaurus, metadata on existing pages on the website had to be adjusted and metadata applied to the new pages to reflect the changes. Changes like this happen constantly. Indexing for websites like this is never 'finished', it is a 'work in progress'.

Although professional indexers maintain the NPS thesaurus, metadata on the NPS MedicineWise website is applied by multiple users. This could be described as 'distributed indexing' (Browne and Jermeý, 2004: 129). Indexing done by multiple indexers, some of them novices, can create problems with consistency. In cases like this, professional indexers must provide quality control and/or mentor others in the art of indexing. Helping novice indexers to get the granularity or 'level of detail at which information is viewed or described' (Browne and Jermeý, 2004: 129) right can be difficult. In most cases, web indexes are linked to webpages rather than to specific paragraphs, lines or words, so it is a case of determining the level of detail for indexing from a full webpage. If indexers have input into the creation of the webpage, specifically the length and content of the page, they can often influence the 'findability' of the page. For example, an indexer may suggest to a content creator that breaking a large webpage with multiple concepts on it into smaller pages covering each concept would allow more accurate and specific application of metadata. And, again, the indexer's skill in considering multiple audiences can help with this task. Involvement of indexers from an early stage of website production, both in determining information architecture and in considering possible search terms, can be of great benefit.

Conclusion

Website indexing may be done using multiple methods and tools, but the basic art of indexing still applies. As websites are often organic and constantly changing, the work of the web indexer is never done. In fact, the need for website indexing may only increase in the near future as new

initiatives for collating and linking data appear. For example, the development of the semantic web, schema for use of microdata, and the Dublin Core Metadata Initiative (DCMI) may all require the skills of human indexers. Indexers' skills are needed in the digital realm.

Acknowledgments

This article is based on a presentation by Mary Coe and Alexandra Bell at the Write/Edit/Index conference, jointly hosted by ANZSI and IPEd, in Canberra, Australia on 6–9 May 2015.

Note

The websites cited were viewed in January 2016. Please note that the NPS MedicineWise website, which is referred to extensively in this article, will be redeveloped in 2016. Screenshots may not reflect the current appearance or functionality of the website and URLs may not provide active links.

References

- Browne, G. (2007) 'Changes in website indexing.' *Information Wissenschaft und Praxis* 58, 437–40, www.webindexing.biz/PDFs/CWI.pdf
- Browne, G. and Jermy, J. (2004) 'Website indexing' [online text], Auslib Press, <http://webindexing.biz/website-indexing-2nd-edition/>
- Hedden, H. (2010) *The accidental taxonomist*. Medford, N.J.: Information Today.
- Hedden, H. (2004) 'A-Z website indexes explained', SitePoint, 6 December.

Mary Coe has worked as a book, database, and website indexer for 25 years. In addition to operating a freelance indexing business, Mary is employed as a website indexer and metadata analyst at NPS MedicineWise. She holds a Master of Information Studies (Applied Research) degree from Charles Sturt University. Mary is an active member of ANZSI. Email: coe.mary@gmail.com

Enhancing a subject vocabulary for Australian education

Philip Hider, Barbara Spiller, Pru Mitchell, Robert Parkes and Raylee Macaulay

This article presents a case study focusing on the process of improving subject access to a collection of resources for teachers in the higher education sector. It describes how existing controlled indexing vocabularies in the education field were evaluated and how the Australian Thesaurus of Education Descriptors (ATED) was selected as the closest to the requirements. However as a general education vocabulary it lacked some of the specificity needed to describe higher education resources, so a process of incorporating further terms followed. The study highlights how established vocabularies can not only be used in different environments, but can themselves be strengthened through the process.

Background

The Australian Government Office for Learning and Teaching (OLT) has provided considerable amounts of funding (\$57.1 million in 2012–15; OLT, 2014) to projects investigating higher education learning and teaching practices in Australia. Its online Resource Library (www.olt.gov.au/resource-library) is a key means of disseminating the outcomes of these projects.

Presently, in 2015, the Resource Library houses resources from over 600 projects, funded not only by the OLT, but also by its predecessor organizations, including the Australian Learning and Teaching Council (ALTC, 2009–11) and the Carrick Institute for Learning and Teaching in Higher Education (2004–09).

In order to find these resources and apply the recom-

mendations emanating from these projects, it is vital that educators and other stakeholders are aware of, and effectively able to use, the Resource Library. However, anecdotal evidence indicated a lack of awareness of the Resource Library among Australian academics, and problems with consistently being able to find and retrieve relevant resources from the database. This situation prompted the OLT to commission a project in 2014 to redesign the Resource Library, in order to improve access and usability.

The project, conducted by a team of academics and librarians from Charles Sturt University, the University of Wollongong and the Australian Council for Educational Research (ACER), was completed in 2015. It aimed to deliver a new, re-indexed database for the OLT, which provides improved access to the content of the Resource Library for end users.