

XIndex: a cautionary tale

Denise Sutherland

The author describes the many pitfalls of working with CUP's Xindex application.

In mid-2012 I indexed the book *The archaeology of Australia's deserts* by Dr Mike Smith, published by Cambridge University Press (CUP). The publisher requested a linked index,¹ where identification codes (IDs) are used instead of page numbers to mark positions in the text. This sort of indexing can be done at an early stage in the process, simultaneously with proofreading and copy-editing. It is by definition page-independent. I had not done an index like this before, but my teacher, Glenda Browne, felt I could handle it, so I ploughed on.

I was offered a few options by the CUP editor: use CUP's own indexing application XIndex (created under subcontract by the software house Aptara), use my own indexing software plus a plug-in (Cindex + WordEmbed in my case), or prepare a 'traditional' index without adding codes or tags to the text.

After looking over XIndex, I opted to use it. On the surface it looked like a reasonable indexing package, and the chapters of the book came embedded into the software, which made me feel more secure about keeping unbreakable links between the text and the indexing tags.

In August 2012, I received the archaeology manuscript loaded into XIndex. You can't index anything else with XIndex, or open any files with it, you can only access the chapters of that one particular book, which appear in a menu on the left of the screen. The program autosaves your progress, and it works on any computing platform.

As the manuscript does not have any page breaks or page numbers, identification codes are placed around every 50 words or so through the whole manuscript. These codes consist of a chapter identifier, and then a consecutive number. For example, 1_176 is the 176th tag in Chapter 1, and 8_5 is the 5th tag in Chapter 8. Titles and headings also get their own tags, which means they can occur more closely together than elsewhere in the text (see Figure 1).

To create an index entry, you select the term or section of

text, and use either a button or keyboard shortcut to create a main, sub-, or sub-subheading entry. XIndex automatically selects the 'best' identifier tag to be assigned to your new entry (more on this later). It is also possible to type in a new term for an entry, using a word that doesn't appear in the text, but this has to be associated with a term from the book. (In other words you can't just add a new term by itself, you can only add it as an option under a new entry from the book – see 'Alternate Entry at Same Locator' in Figure 2.)

Within a few days of starting it was clear that XIndex was a problem, but I decided to press on, rather than redo all my work (not a wise decision!). Here are the main problems I found with the software:

- There is no undo.
- There is no copy and paste, either from the manuscript to the index entries, or between index entries.
- You cannot print anything.
- Any change to main entry text automatically propagates to all other locators for that entry, and you can't switch this off.
- To add a subheading entry, there has to be a main heading selected for that section. You can't add a subheading entry with a new main entry simultaneously.
- There is no keyboard shortcut for delete. You have to move through three menus, and click 'Yes' on the dialogue warning box that pops up. The 'Edit' menu does allow you to delete a single locator/range, but not the entire entry.
- It is difficult to turn a subheading into part of the main entry, involving a complex set of steps. It is also difficult to turn a main entry into a sub of another heading.
- When loading a new chapter, the ID markers revert to being 'hidden', so you have to keep toggling them 'on' if you want to see them.
- There are no keyboard shortcuts for adding bold or italics to entries: you have to type in the HTML codes for each one.
 - You can't add indicators for figures, tables, photos and so on.
 - There is no sense of the length of each chapter, just one long scrolling screen of text. As you can't print them out, it's hard to assess the length of each part of the book, which makes estimating the time required for work difficult.
 - You can't easily export the index midway to see how the layout and length are going. To get around this, I had to make a copy of the entire XIndex application, go in and export

archaeologists 1_148 in the Australian arid zone is to more fully exploit the intrinsic potential of late Pleistocene sites as archives of the cultural and economic landscapes in which they are embedded.

The 1_149 Politics of Practice

Other 1_150 issues concern the contemporary political and economic situation in Australia's deserts. Biologists, ecologists, archaeologists, and anthropologists all benefitted from the 4WD revolution that swept remote Australia from the 1960's onwards. But the same forces that opened up the desert for archaeological research also placed more 1_151 constraints on where, when and how this research could be done. Since the passage of the

Figure 1 An example of XIndex tags placed into the text²

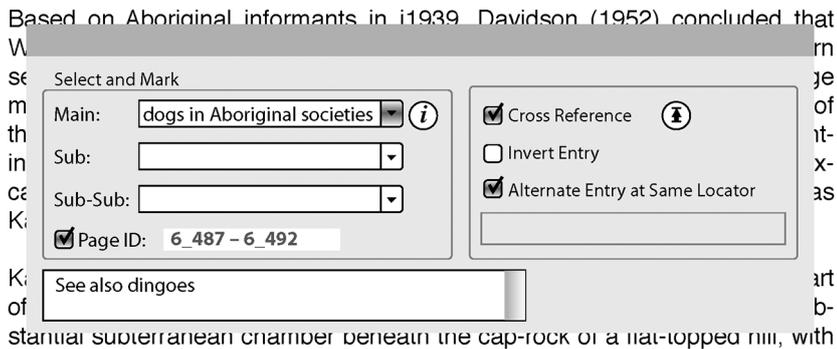


Figure 2 An example of the XIndex index entry fields

the ‘final index’ from this duplicated file, which could then be opened in Word.

- The ‘search the book’ function is very flawed. You can’t copy and paste terms. If, as a result of the search, you want to enter an index entry for that position in the text, you have to close the search box, then add the term to the index, then start a new search, type in the term again (no copy or pasting!), and search again, for the next instance. If it’s a book-wide search, repeat ad nauseam through each chapter of the book.
- When selecting a range of text, the markers weren’t always transcribed correctly into the ‘new entry’ box.

There are also problems with the final export of the finished index to Word. I found I had to do all the major editing of the index in Word, as editing in XIndex was so slow, problematic and unsatisfactory. These are just a few of the issues with exporting the finished index:

- XIndex won’t allow you to export the final index with any ‘non linked see refs’, such as ‘see also individual site names’, as there wasn’t a main entry for ‘individual site names’, so these had to be edited out before exporting the final index, and added manually in the final Word document.
- Letter dividers (A, B, C, D etc.) for each alphabetical section are not added.
- The program generated format errors, such as capital letters in some *see also* references, and lack of spaces between some terms.

While indexing this book, I provided copious amounts of feedback to Aptara, and they said they would put in place all my suggestions in the next upgrade. We shall see.

Apart from all the interface problems that merely made life painful, my biggest concern with the XIndex set-up was that the markers in the text are set so far apart – roughly every 50 words or so – and that when I selected a word or section from the text to index, the program would bring up an ID number that was often placed quite a distance from the term. I sent yet another query email to Aptara. They gave me the following response regarding this issue (in August 2012):

XIndex application picks up the Page ID depending on

the text selected by indexer while marking the term. There are three scenarios:

- 1 **For a single word selection:** XIndex picks the nearest ID from left or right depending on which is nearer to selected text.
- 2 **For multiple word selection:** XIndex picks the nearest ID to the left or right depending on which is nearer to selected text.
- 3 **For multiple word selection with selected word count more than 70 words:** For Range entries, left-most entry is ‘start ID’ and right-most is ‘end ID’.

Note: For all the cases, IDs are picked from text outside the selection text.

Read that last sentence again. The ID tags are selected from *outside* the selected text. As the IDs can be up to 50 or so words away from the indexed text, it means that a page break can easily occur in that gap between tag and indexed term – which could generate a lot of ‘off by one page’ indexing errors in the final book. I raised my concerns with the author, but there was nothing we could do about it. Aptara assured me all would be well.

In November 2012, a few months after submitting the final index, I heard from the author. It was not good news. He had the final book page proofs, and had done a spot check on the index entries. There were around 30 per cent errors, with ‘off’ page numbers. What a surprise.

CUP was not amused. The author was not amused. I was not amused. Between us CUP and I had to check and correct every single locator by hand, from the page proofs. And it was necessary, because there were a *lot* of errors introduced by XIndex during the conversion from ID tag to page number.

XIndex might seem to work reasonably for authors who are indexing their own books (or for Aptara’s own indexing services), but it is woefully inadequate as a tool for professional indexers, not least since there are several other more effective ways to do the job, using indexing software in combination with programs such as WordEmbed, DexEmbed, or simple mark-up of the PDF file. And its failure even to handle the conversion of ID tags to page numbers presents serious problems for anyone using it, professional indexer or not.

Notes

- 1 CUP call it an ‘embedded index’, although this is not strictly correct.
- 2 The illustrations are by Denise Sutherland, and are accurate mock-ups of actual XIndex screens.

Denise Sutherland has only been indexing for two and a half years, but is completely hooked! She is also a graphic designer, editor, author and puzzle writer. She is a member of the ANZSI ACT Region Branch committee. Denise is the author of many puzzle books, including Solving cryptic crosswords for dummies. You can find her online at www.sutherland-studios.com.au, email: denise@sutherland-studios.com.au