A Simple System for Constructing an End-of-Book Index

RONALD D. FRANCIS and MURRAY GREENWAY

End-of-book indexes are usually constructed either by professional indexers or by authors. This article proposes a relatively simple system that may be used by either. It is possible to construct an index electronically by marking the relevant entries. Those writers who use Microsoft Word will find the use of the system explained on the Help menu. Such electronic systems may appeal to the electronically minded. This article addresses those who prefer a more hands-on system in which they feel a better sense of involvement and control. Further, it gives the author more discretion about including a term, cross-referencing a term, and making sub-categories in the end-of-book index. This article has the major aim of outlining an economical method that involves computer use but does not employ the automatic indexing system.

Keywords: authors, indexers, end-of-book index, indexing systems

INTRODUCTION

Some book publishers ask an author if a professional end-of-book index should be prepared. Clearly, a good index is a valuable addition to any text or scholarly work. Assent to the request often brings the response, however, that the cost of index preparation would come off royalty payments. It is appreciated that skilled indexers bring insight to bear that may escape all but the most experienced author. Further, their experience may enrich a book index as well as providing insights to the author. The opposite is also true: an author will have insight into their special subject that even the most skilled indexer may not have. Most recently, the notion of advanced indexing techniques has been advanced. Notwithstanding, authors may wish to feel, and to keep, control over end-of-book indexing. To this purpose, the presently presented system may be of greater appeal.

From a perusal of the website 'Against Doing the Index Yourself,' it is clear that there is a very wide divergence of opinion about the merit of

authors doing an index themselves. To this end, it is imperative that authors be informed about the systems available and the merits of using an experienced indexer. One of the serious problems of the computer-based index is that it may miss cognate terms. One may need to make more than one pass to find appropriate indexable terms. The number of passes with the proposed system could be one very careful pass or several passes for checking. If and when the ebook becomes more common, then one might expect the search function to replace the index. Given human nature, it is unlikely to happen since print books will still be available for the foreseeable future, and the search constructed by an expert will give cross-references that would otherwise defeat a reader from another discipline.

This article draws attention to the built-in indexing system in *Microsoft Word* and outlines an alternative system that could well appeal to those of a less electronic mindset. To this end, a system has been devised that makes it possible to prepare a book index within several working hours. What follows is the suggestion for a system that may be used by professional indexers as well as by authors. Indexing techniques for advanced data systems — in a paper of that name, has been put forward by Elisa Bertino.[1] This proposed system is offered not as a panacea but, rather, as one that should appeal to those authors/indexers who have a lesser commitment to electronic means. It is also hoped that users will provide their feedback suggestions for improvements to the present writers.

BACKGROUND

Writing in 1989, Gerald Salton expressed the view of a syntactic approach to book indexing, notably using automatic book indexing.[2] This study described the methods by which automatic text and phrase generation might be used. Susan Olason addressed the issue of user index requirements.[3] One of the main traps is to compile an index that will be satisfying to indexers, even if indexers are not the main users. An article by Hazel Bell considers five prospective users of indexes: users, subjects, authors, publishers, and regulators. The notion of being 'user friendly' in the development of an index is paramount.[4] Whether it satisfies the author, publisher, or bookseller is of secondary importance. To this end, it would be useful for any prepared index to be shown to a colleague to check for usability. Indeed, the history of end-of-book

indexing is given in Bell's book[5] and in a review of that work by Judith Douville.[6]

All users have different expectations that are sometimes incompatible. It is this topic that Bell addresses by noting that, because of the existence of regulators, there must be a best way to prepare indexes.[7] Equally, the end user must find the index helpful. In this vein, there are constraints upon what a publisher will require. They have a particular schedule and will advise about the length and comprehensiveness of indexes. Finding such a balance is the basis of Bell's analysis.[8]

Olason's paper starts with the question of how many opinions are there about indexing, of which there are several, but no set of user requirements.[9] One of the main approaches is that of articulating the set of unmet users' requirements. Among her empirical findings is that indexes that do not begin sub-entries with prefix words (prepositions or conjunctions) have higher and higher usefulness rankings. Indexes are ranked as user-friendly when they do not use sub-entry prefix words—and no index is ranked as useless. It is such findings that are an aid to indexers in ensuring that their prepared material is consonant with the wishes of the end user. The author emphasizes the importance of an index, which affects both a publisher's profit and an author's reputation. An exhortation to be sensitive to the needs and wishes of the end users is critical.

The use of a thesaurus may also be a useful addition to the armamentarium of an indexer, a point made well by Vanda Broughton.[10] As Adeline Nazarenko and Touria El Mekki have noted, indexes are expensive, largely because indexing, at its best, is still a manual task.[11] While modern methods of automatic indexing are a help, they do not really address the content problem. The main point of their article concerns the validation of terms. From this topic, it is clear that the exercise of human judgment is indispensable in the preparation of a good index. It deserves emphasis that authors well know which terms are cognate. For example, in psychology, the terms 'Weber fraction' and 'just noticeable difference' will be appreciated as intimately linked, just as the terms '*modus tollens*' and 'denying the consequent' would be to a philosopher.

THE DIFFERENT INDEXING SYSTEMS

There are different systems used to prepare end-of-book indexes. In a 2008 study, Nina Wacholder and Lu Liu compared three different systems:

one was an index prepared by a human indexer and the other two were prepared by computer programs.[12] The findings were that the index prepared by a human indexer and the one using linguistically sophisticated criteria performed significantly better. Plainly, the query language needs insightful input. It also points to the need for human judgment to be an indispensable part of the preparation of an effective index. A second feature of a good index — however, it is collated — is the grouping of like items. Thus, one may have a general heading such as 'Conflict', with subheadings of such items as 'Between professionals,' 'Between professional and client,' 'Conflict of interest,' 'Ethical conflicts,' 'Political conflict,' 'Legal conflicts,' and so on.

Authors will know that there is an automatic indexer function in *Microsoft Word* (albeit with some limitations) and a group search feature in *Adobe Acrobat* and *Adobe Reader*, which are all tools used by professionals to assist them in the indexing of documents or books. To use the Adobe Group search function, one must be logged into the file in which you are creating the index ('the index file') and the file converted to the pdf format ('the PDF'). The screen should show the index file and the PDF on a split screen, side by side. Using Shift Control F, group together all occurrences of a word or phrase that you are interested in, ignoring those outside the range of interest. Scroll down the list and click to a particular page in question so as to determine whether or not the reference should be included in the index or not. At the same time, check whether you require a range or just a reference to a single page. When scrolling the list, skip repeated occurrences of the word or phrase on the same page and those that you decide are within a range.

This system may appeal to some as an invaluable tool — in particular, for checking to ensure that no occurrences of the word or phrase are overlooked in the index — for comparing terminology, and for creating a sense of coverage of a particular word or phrase within the document. The length of time it takes to create an end-of-book index, irrespective of the process adopted, will, of course, depend on the length of the book. Experience by the present writer is for books just under 200 pages in length. The time commonly taken using such a method is about seven person hours — some of the time with the author alone, other times with an assistant. For longer books, it will, of course, take longer, although once embarked upon and armed with appropriate terms, it should not

be proportionately extended. The next section outlines the method used by the present writers using the steps to construct an end-of-book index.

PREPARING AN INDEX

1. Get a copy of the final manuscript as a PDF file and with the pages set and fixed.
2. Author or indexer has a Dictaphone and speed reads the manuscript and slowly dictates the terms. For this purpose, either a voice-recognition system or a Dictaphone may be used, although we need to recognize that learning and voice training are not worth the effort for a one-off job.
3. The case for using voice recognition in the preparation of an end-of-book index is outlined. It was also noted that voice recognition works well for the more simple preparation of an index but less well for updating a multi-line index (such as in a law book or an encyclopedia).

   Alternatively, the author or indexer may have two computers and an assistant. The author speed reads the text and says aloud the term to be entered: the assistant types that term into the index on his/her computer one line per term.
4. When the draft index is complete, the computer sorts the terms into alphabetical order. The author or indexer then scans the terms for duplicates, overlap, and creates cross-references and amends accordingly. Sub-categories may be set up if appropriate. For example, ethics, with sub-categories of personal, professional, research, and so on.

   This process produces a finished index. The next stage is to insert the page numbers using an electronic version of the final page-set text. Where there is a string of page numbers you would, ideally, use the search facility. This counsel of perfection must assume that one has an electronic copy, a point not always satisfied. Ideally, the start of the main discussion of that particular point could be illustrated by making the relevant page number a bold one.
5. Have two computers and an assistant. On the author/indexer computer is the set text; on the assistant's computer is the index.

6. The assistant says the first term, and the author enters it in 'search' or 'find,' and at each appearance the author reads the page number to the assistant, who then enters it in the index. For example, the word Aardvark is the first term. The author types in the term and the computer then finds the first reference. The author reads the page number aloud, which the assistant then enters in the appropriate place.
7. This order of doing things allows the best informed, the author/indexer, to decide whether or not the term is important or if there is a cognate term that needs adding.

With a conscientious assistant, the present writer has found that an index with about 400 terms can, as mentioned, be done in about seven person hours. It is hoped that this method will commend itself to authors and indexers. Any suggestions for improvement would be welcomed.

RONALD D FRANCIS is Professor Emeritus in the College of Law and Justice at Victoria University in Melbourne. He is the author/co-author of twenty-two books and numerous articles and is experienced in thesis supervision and examining. This contribution stems from a need to allow authors to provide indexes when so desired.

MURRAY GREENWAY is the Law Librarian at the College of Law and Justice at Victoria University in Melbourne. He is the author of a book chapter and numerous online guides.

NOTES

1. Elisa Bertino et al., *Indexing Techniques for Advanced Data Systems* (New York: Springer, 2012)
2. Gerard Salton, 'Syntactical Approaches to Automatic Book Indexing' (paper presented at the twenty-sixth Annual Meeting of the Association for Computational Linguistics, Buffalo, New York, 7–10 June 1988), 204–10
3. Susan C. Olason, 'Let's Get Usable!' *The Indexer* 22, 2 (2000): 91–5
4. Hazel Bell, 'Whom Should We Aim to Please?' *The Indexer* 20 (1996): 3–5
5. Hazel Bell, *From Flock Beds to Professionalism: A History of Index Makers* (Hatfield, UK: Oak Knoll Press, 2009)
6. Judith Douville, 'Review of Hazel Bell's Book, From Flock Beds to Professionalism: A History of Index Makers,' *Journal of Scholarly Publishing* 41, 1 (2009): 126–9
7. Bell, 'Whom Should We Aim to Please?'

8. Ibid.
9. Olason, 'Let's Get Usable!'
10. Vanda Broughton, *Essential Thesaurus Construction* (London: Facet Publishing, 2006)
11. Adeline Nazarenko and Touria El Mekki, 'Building Back-of-the-Book Indexes,' *Terminology* 11, 1 (2005): 119–24
12. Nina Wacholder and Lu Liu, 'Assessing Term Effectiveness in the Interactive Information Access Process,' *Information Processing and Management* 44, 3 (2008): 1022–31