Indexes and indexing


The much-anticipated 15th edition of The Chicago manual of style is here. In this review I focus on Chapter 18, ‘Indexes’. In the USA it is the de facto standard for indexing. First, however, I would like to provide a brief overview of this new edition with an emphasis on material outside of Chapter 18 that is of interest to indexers.

One of the first things I noticed when I first looked through the book was its visual appeal. While the overall format of the chapters is the same as previous editions, with numbered paragraphs, the pages are not visually as dense. Subtle changes in typefaces have made a dramatic difference. This is particularly noticeable in Chapter 14, ‘Mathematics in Type’. I should add that indexers will find Table 14.1, ‘Mathematics signs and symbols’, very helpful.

Chapter 10, ‘Foreign Languages’, provides a bounty of information for indexers working in the humanities. Special characters and capitalization rules are addressed on a language-by-language basis. Transliteration guidelines are presented for Arabic, Chinese and Japanese, Hebrew, Russian and South Asian languages. This is followed by discussions of classical Greek, Old and Middle English and American Sign Language.

The Manual of style continues to be primarily addressed to writers and editors of scholarly books and journals. However, the growing prevalence of electronic publications is discussed throughout the Manual. In Appendix B, ‘The Publishing Process of Books and Journals’, flowcharts outline the steps involved in the editorial and production stages for print and electronic publication. Beginning indexers not familiar with the publication process will find these diagrams, and associated reference chapters, enlightening. A far more detailed discussion is presented in Appendix A, ‘Design and Production’.

One index entry that caught my eye is ‘abbreviations, specific’. This is a lengthy entry (over 8 column-inches) that ranges from abbreviations for biblical references to time. Indexers working in many fields will benefit from the guidance offered for abbreviations of specific types of terminology.

A glossary of key terms appears in Appendix A. These include traditional publishing words (‘folio’, ‘recto’, ‘verso’) as well as electronic publishing terms (‘DTD’, ‘metadata’, ‘XML’). The definitions are clear and succinct. Many include see also references to other terms in the glossary.

Chapter 18, Indexes

One of the most welcome changes in Chapter 18 can be summed up with one word: clarity. I have used the indexes chapter from the 13th and 14th editions as a text for my indexing classes. In these publications intervention is always required (18.4).’ If you are beginning to sense that there is a change in tone, you are correct. The writing in this edition is more succinct than in previous editions. For example: ‘Cross-references should be used with discretion; an overabundance can be irritating’ (18.14). So true.

The next several sections describe the components of an index, including main headings, subentries, locators, cross-references and index style (run-in/indented). Section 18.13 addresses the format of inclusive reference locators. In the 14th edition this information was not included in the indexes chapter; readers were referred to another chapter where two types of abbreviation systems for inclusive numbers were presented. Although the 15th edition acknowledges that there are various ways to present inclusive page numbers, the Chicago preference is stated clearly: there is no doubt that the Chicago style for ‘234–235’ is ‘234–35’. This is typical of the clarity found throughout this chapter.

Naturally, many indexers are wondering if Chicago has changed its preference in regard to alphabetizing. The answer is ‘not really’.

Chicago, most university presses, and many other publishers have traditionally preferred the letter-by-letter but will normally not impose it on a well-prepared index that has been arranged word by word (18.56).

Chicagostill does not adhere to a strict letter-by-letter system. However, it clearly explains both letter-by-letter and word-by-word systems. It provides an order of precedence for each system and describes the handling of a parenthesis, comma and other punctuation marks. Indexing software programmers will easily be able to incorporate the 15th edition’s alphabetizing style into their programs. The most important change in the alphabetizing rules is that commas, regardless of their usage, are treated in the same way.

Chapter 18 includes guidance for the handling of personal names, foreign personal names, organization and business names, and names of places in indexes. A very nice addition is a summary of punctuation use in an index; students will surely welcome this presentation. The following marks are discussed: comma, colon, semicolon, period, parentheses, em dash and en dash.

The ‘Mechanics of Indexing’ sections have been moved toward the back of the chapter. The section begins with a discussion of the schedule (18.100): ‘Anyone making an index for the first time should know that the task is intensive and time-consuming.’ Moving along, the sections address various index specifications such as the format for footnote/endnote locators, author indexes (‘Preparing an author index, though somewhat mechanical, takes more time than often supposed’), and citation format for illustrative material.

In regard to author indexes, Chicago tackles the indexing of multiple authors’ names (18.116). Should Jones, Smith, and Black 1999 share one index entry, or should three entries appear? And what about Jones et al.?
Chicago recommends the following procedure: Make separate entries for each author whose name appears in the text. Do not index those unfortunates whose names are concealed under et al. in the text.

Sound advice is offered about editing an index. A distinction is made between the editing tasks of an index author and an index copy-editor. Section 18.136 includes a copy-editing task list. Typographical considerations for the layout of an index are also outlined and the chapter concludes with examples of various display formats for indexes.

The Index to the 15th Edition
The index to this edition is much longer and more comprehensive than the index to the 14th edition. There are 75 pages of dense index, resulting in a 9 per cent index, compared with 50 pages of index (6 per cent index) in the previous edition. Over the years the index to The Chicago Manual of Style has provided the most comprehensive example of Chicago’s preferences for the indexing of technical material. In previous editions the indexes have been in the indented style, with main headings and subheadings. The avoidance of sub-subheadings led to many index entries with long strings of undifferentiated page references.

I found no examples of long strings of page references in this new edition; on the contrary, index entries reflect analysis and precision. Topics are sensibly broken down into useful subheadings and sub-subheadings. Cross-referencing, particularly see also references, is appropriate and helpful. It is obvious that the indexer was given a free hand to write an index that reflects the complexities of this text. It is perhaps the most comprehensive and eloquently designed index to a technical work that I have ever seen. Readers should have no problem locating specific information in this index.

Also of interest is the format of the index. In previous editions of the Manual of Style, the recommendation for three-level indexes was to present them in indented format. Now, the recommendation is: ‘In an indented index, sub-subentries are best run in’ (18.28). While this format is not new to Chicago (it was described in the 14th edition (17.141)), that the format is now recommended is new. How does it work? Quite well. Had the sub-subheadings been formatted in indented style, the index would be considerably longer. The pages would be visually complex with multiple indentation levels and inevitable turnover lines. The run-on format for sub-subheadings pulls the index entries together and creates a visually satisfying page that is easily scanned. Unlike in previous editions, all bad breaks between columns of index entries are accompanied by ‘continued’ lines. Therefore, the top line of every column clearly indicates the reader’s location in the index. Put quite simply, this index is a showcase for the presentation of dense technical indexes.

Overall this revision of the indexes chapter is most welcome. The writing style is crisp and clear. Indexers and editors now have precise guidelines for the presentation of indexes from The University of Chicago Press.

Nancy Mulvany, freelance book indexer

Indexing the Medical Sciences. Doreen Blake, Michèle Clarke, Anne McCarthy and June Morrison, Sheffield: Society of Indexers, 2002. vii, 84 pp. 21 cm. Index. ISBN 1-857157-24-1 (pbk): £17.50 (£15 for members of indexing societies). (Overscans £20/£17.50.) (Society of Indexers Occasional Papers on Indexing no. 3.)

Medical indexing, while immensely fascinating and rewarding, can at times seem somewhat daunting. And of course, as in all subjects, there are plenty of challenges. The introduction to this book states that it ‘aims to provide guidance concerning the pitfalls and possibilities in medical indexing’. A book of this size (84 pages) cannot be comprehensive, nor does it set out to be. Instead, it advises and guides the indexer through many of the issues that crop up regularly. Drawing on the combined experience of four authors and a number of contributors, it succeeds admirably in achieving its aim.

Do you need to check whether the prefixes in onomeric names should be in italic or roman (and whether they should be ignored in the filing order); or how to handle Greek characters; or how to use adjectival headings appropriately? You’ll find the answers to these and many more questions, along with further nuggets of information, in this publication. Issues relating to terminology are encountered frequently in medical indexing, and this area is covered to some depth. Discussions range from the problems of inconsistency in multi-author books and the use of synonyms, to details of biochemical, chemical and scientific names and pharmaceutical terminology. Other specific problems that indexers will be familiar with are also covered – including the use of abbreviations and acronyms, gene nomenclature, hyphens, prepositions, singular/plural headings . . . and so on.

Although general indexing principles are common to those of other disciplines, the choice of headings is not always as straightforward as it may seem, given the relatively structured nature of many medical texts. Some of the ‘ traps’ for the unwary are pointed out and solutions offered. The importance of careful use of cross-references is a recurring theme, as is the need to evaluate each text individually. The specific indexing requirements of a wide variety of books and journals are considered, and in a new chapter since the last edition there is a useful introduction to electronic indexing. Indexing of names and medical eponyms is discussed and there is a handy reminder on editing matters. A selection of book and Internet resources is also included.

Readers are advised and informed in a friendly, reassuring style. For example, a section on density of entries notes that medical textbooks often require very detailed indexing, but that there are also situations where a topic is discussed in great detail and one entry may cover several pages. The indexer who ‘worries that this is not good enough and who spends as much time on this one entry as on 20 or 30 ‘obvious’ ones’ is told ‘don’t despair – we all do it’.

Although this book can be used to answer specific questions (and a comprehensive index is included), I think it deserves a more thorough reading. It is a very practical book, with a wealth of friendly commonsense advice, and offers much encouragement to those with less experience.

Louise Bleazard, freelance indexer


This book is a collection of articles about software used in indexing. It is divided into five sections: Part 1 covers dedicated indexing software. Part 2 embedded software, Part 3 online and web indexing software, Part 4 database and image management software and Part 5 voice-activated, automatic and machine-aided software.

The press release for this book states that ‘objective review of the software products used in indexing’ will be provided. Approximately 70 pages of Part 1 are devoted to dedicated indexing software products: CINDEX, MACREX, Sky Index, Authex Plus and mINDEX. These are the programs most frequently used by freelance book indexers. With the exception of one article written by Michael Wyatt, a review of CINDEX and Sky Index, there are no objective reviews. It is both disappointing and shocking to see so much vendor-written material. Readers hoping for a reissue of Linda Fetter’s Guide to Indexing Software, known for thorough and unbiased assessment, will be disappointed. Even the first chapter, ‘Choosing a Program for Indexing’, is vendor-written.

No fewer than 13 pages of Part 1 are devoted to software comparison charts (Tables 2.1–2.17) filled out by the vendors. These charts cover various features that we expect to find in dedicated indexing software. Unfortunately, there is nothing to indicate how these programs comply with accepted standards. We have no independent verification of how well they deal with the specifications outlined in the charts. While the charts themselves were designed by the editor, Sandi Schroeder, the content of the charts is vendor-written.

In order to round out the presentation, professional indexers using the various software products were asked to write brief product recommendations for the various products. These recommendations are their own and are not endorsed by ASI, nor were they solicited or endorsed by the software vendors.
Yet in the case of my own two endorsements (MACREX and SkyIndex), they certainly were marketing pieces solicited and endorsed by the vendors. The byline for every article begins with ‘Review by’. Even the marketing pieces I wrote include a ‘Review by Nancy Mulvany’ byline. These are not software reviews. As a former member of the Computer Press Association I am well aware of the difference between a formal review and informal description. I did enjoy reading how my colleagues use their indexing programs. However, I am not sure that the new indexer trying to decide which program to buy will find this type of narrative discourse helpful.

The last chapter in this section discusses utility programs that can be used with files produced by the dedicated indexing programs. Again, the lead author is the vendor of these programs. Most interesting to me is a rather esoteric application that involved the need to change 10,000 names in an index file. Thankfully we don’t run into problems like that too often.

The next section of the book, Part 2, deals with embedded indexing software. It begins with an article by Peg Mauer, ‘Embedded Indexing’, that covers embedded indexing in general and Microsoft Word and FrameMaker in particular. The remaining chapters are devoted to ‘hands-on’ discussions of various indexing programs that make embedded indexing such a chore. These chapters will be of interest to indexers working with these particular programs that make embedded indexing such a chore.

Part 3, ‘Online and Web Indexing Software’, is composed of four chapters, two of them vendor written. The first chapter, by Kevin Broccoli, is an excellent introduction to the tools that are available for web indexing. Unfortunately it was reprinted from a 1999 publication by the Indexing SIG (Special Interest Group) of the Society of Technical Communicators. Some of the URLs referred to now no longer work properly. Despite that, Broccoli’s article provides a good introduction to these software tools. The next two chapters are vendor-written presentations about HTML/Prep and HTMLIndexer. Jan Wright wraps up the section with an article about indexing in RoboHELP.

Part 4 is composed of three chapters about database and image software. The first article, ‘Software for Database Indexing’, written by Shana Milkie, is an interesting presentation of results derived from a survey of database indexers. Milkie’s table, ‘Database Indexing Software Survey’, lists a wide variety of programs and vendors, none of which are discussed in this section of the book. Readers seeking information about specific database indexing programs will need to look elsewhere.

Next Kim Schroeder discusses software for image indexing. Her article includes screenshots from image database management software as well as a discussion of what indexers and archivists need. The last article in this section of the book is rather baffling since it has nothing to do with either image or database indexing. ‘Writing an Index with Microsoft Excel’, by Seth Maislin, is another reprint from STC’s Indexing SIG publication. The author even states that using Excel to write an index ‘is not pretty, and it is not recommended’. So why is this article even present in this book?

The last section of Software for Indexing includes four chapters that deal with voice recognition software, automatic indexing, custom indexing software for indexing newspapers, and machine-aided indexing. Charles Anderson’s article ‘Indexing with Voice Recognition Software’ is excellent. He discusses using Dragon Naturally Speaking with dedicated indexing software. Next, Jan Wright discusses the trials and tribulations of generating an automatic index with Sonar Bookends. Her conclusion is: ‘Unfortunately, Sonar’s claims of producing professional indexes in minutes do not bear out under real-world conditions, and I would be embarrassed to print the final result in any publication.’ Alice Redmond-Nicol’s brief article about NewsIndexer from Access Innovations is an introduction to the specialized field of newspaper indexing and the development of the NewsIndexer program. ‘Machine-Aided Indexing and Automatic Filtering’, by Marjorie Hlava, is the last chapter in the book. These last two chapters are also vendor-written.

Who is the audience for this book? New indexers looking for unbiased, thorough and independent evaluations of dedicated indexing software, upon which their businesses will be built, will not find that information here. For those curious about the use of embedded indexing software, Part 2 does provide nitty-gritty discussions of real-world use of those products. People interested in indexing HTML material will find the discussions in Part 3 a very good starting point.

With a few exceptions I find most of this book disappointing. Many of the chapters are reprints from other publications with little or no updating. However, the amount of vendor-provided material is inexusable. Of the 25 chapters, 5.5 are vendor written; that accounts for 22 per cent of the book. The absence of thorough and objective reviews of dedicated indexing products especially troubles me. These are small-market products that will not be reviewed in the mainstream computer press. I remember when I was shopping for my first indexing program. This was long before the appearance of CINDEX and Sky Index. There were several other programs on the market at that time. Victoria Agee’s article, ‘Why I Use MACREX’, reminded me of many of them. Had I relied on vendors’ claims, I would have purchased a product that was destined to disappear from the market. But I was lucky; I discovered Linda Fetters’s column, ‘Electronic Shoebox’, in the ASI Newsletter, and I found helpful guidance from an objective review of the American Society for Information Science and their publisher, Information Today, would so thoroughly blur the line between objective review and product marketing.

Nancy Mulvany, freelance book indexer

The 20th century cumulative index, The Indexer, Volumes 1 to 21 1958–1999. Howe: Indexing Specialists (UK) Ltd, 2003. PDF format (154 pages). Capacity required 997 Kb. Available as: A single copy direct download (€12.95) 3.5-in floppy disk (50 Hz.) A cumulative index to the volumes of any respected periodical is always to be welcomed. It offers easy location of an item that one remembers having already read ‘some time, but when?’. It can, as stated in the preface to this index, perform ‘a unique role in making . . . knowledge available that is not necessarily obvious from consulting the individual volume indexes’. It can provide a historical perspective, showing the introduction, development and flowing of a new topic. It can gather under a person’s name all the articles that he or she has written during the period, as well as items concerning him or her, giving a helpful survey of the person’s activities in the subject field. It can even rekindle an interest in past (supposedly ‘dead’) topics and so lead to them being taken up once more and extended. The potential for saving time and ensuring reliable search results is considerable – no more checking through separate volume indexes, which may have been compiled by several different indexers over the years (15 in this case) who may have been subject to varying editorial policies, who may have employed a range of styles and forms of heading, and who may (quite validly) have used different words to represent the same topic.

The layperson (as well as the novice indexer) may think that all that is needed to produce a cumulative index is to merge the contents of the individual volume indexes, subject to the permission of the original indexers or other copyright-holders. But high-quality cumulation, as Wellisch (1995: 134) points out, ‘cannot be done by mechanical amalgamation of existing entries’ and requires a depth of indexing that will almost always have to be at a higher level than for an index to the same material covering only ‘a single volume’ (p. 136). Booth (2001: 220) lists the defects that ‘mere amalgamation’ is likely to produce in a cumulation, and Mulvany (1994: 145), referring to a cumulative index to a series of books having their own indexes, provides confirmation of differences found between indexes of the same material compiled by persons with different editorial policies, who may have used different words to represent the same topic.

The 20th century cumulative index of The Indexer is a good starting point.
produce a coherent, well-structured and consistent cumulative index, a considerable amount of editing is necessary.

When a large number of volumes is to be covered, there is rarely time to read the whole text (as an indexer normally does), but it is almost always essential to check some, even many, of the original articles. Knight (1979: 156) maintains that each individual entry should be checked with the text before it is included in the cumulation, and that ‘recurrent material’ (items on the same topic) should be traced in all the volumes so that comprehensive and consistent entries can be prepared (p. 157). This requires the cumulative indexer to identify headings (including personal names) that use different words or forms of words but which could relate to the same topic or person, to check the original texts and to harmonize the multiple headings into a single one.

This cumulation covers 21 volumes (40 years) of The Indexer (from the time of its first publication). It brings together a wealth of material encompassing a wide range of subjects (because indexing relates to the whole universe of knowledge) and so should provide access to much of the fascinating (as well as the mundane) information that has been generated by indexers and other contributors over the years. It offers the possibility of discovering the ‘whole story’ of important topics like the use of computers and the arrival of dedicated software, the development of formal training, and the establishment of indexing societies in various countries. It provides, too, the opportunity to look up well-known names such as ‘Bakewell, K G B’ and ‘Wallis, Elizabeth J’ and so to scan the long lists of subheadings that indicate the contributions that they have made to indexing. Searching is made easier by the A–Z navigation guide (enabling quick transit to, say, the ‘F’ sequence of headings) and the ‘Find’ facility that locates a particular word that appears anywhere in the entries.

The preface explains the use of bold type to indicate volume numbers, and the abbreviations that feature in the entries (e.g. f = illustration/photo) are helpfully repeated at the bottom of each page. Some useful conventions employed in the separate volume indexes have not been followed in this index. It would be useful, for example, to have an explanation of the styling of the various kinds of entry, e.g. italic for book and journal titles, and capital letters for regular features such as book reviews. Also, whereas the volume indexes have indicated the titles of articles by placing them in single quotes, this is not done in the cumulation; there is a resulting loss of clarity, particularly when a title has no author’s name in parenthesis, leaving only the initial capital letter as a possible way of distinguishing an article title from a subject heading (except that proper name subject headings also start with capitals). Nor do the preface indicate which types of content have not been indexed; volume indexes normally exclude, for example, bibliographic references listed at the end of articles, and the contents of the ‘Indexes reviewed’ and ‘Publications received and noted’ sections.

According to the preface, the team ‘experienced difficulty merging together the separate volume indexes, because [of the] variation in style . . . and differing forms of presentation over the years’. The resulting entries give the impression of being only lightly edited. In fact, it is evident that many of the original entries have only been interfiled, not merged under harmonized headings. For example, the entries for two well-known writers are scattered in each case under several headings:

Herbert, A P, author/indexer 12 144
Herbert, Alan (APH) 6 108–15

then, after an entry for ‘Herbert, Patricia’:

Herbert, Sir Alan (A P H) 5 133 [alphabetized by ‘Sir’]
Humorous indexes of 6 108–15 [already indexed 3 lines above]

Langridge, D W . . .
Langridge, D W L . . .
Langridge, Derek . . .
Langridge, Derek W . . .

and those relating to the use of computers in indexing are dispersed under synonymous or overlapping headings such as:

computer uses in indexing . . .
computer-sided indexing . . .
computer-assisted indexing . . .
computer-supported indexing . . .
computerized indexing . . .

Similarly, some items relating to computer programs are indexed under ‘computer programs’, some under ‘programs (computers)’ and others under ‘software’. There is a see also cross-reference from ‘computer programs’ to ‘software’, but not from ‘software’ to ‘computer programs’ (there is one to ‘programs, computers’); ‘programs, computers’ has no cross-reference to either ‘computer programs’ or ‘software’.

A search for items relating to rules for filing (sorting) headings reveals that there is a heading ‘filing’, with subheadings for ‘ALA rules’, ‘BLAISE rules’, and ‘British Library rules’, but there is also the more specific heading ‘filing rules’, which contains no reference to any of these.

Duplication in subheadings and sub-subheadings often occurs, e.g.:

Pym, Barbara
No fond return of love . . .
indexer in 12 26, 109
indexers in 13 26

and:

Wace, Michael . . .
SI representative Wheatley Medal Sub-Committee 8 75, 114, 219
Wheatley Medal
sub-committee . . . 8 75, 114, 219

The interfiling of subheadings under keywords extracted from titles makes for some awkward subheading sequences containing mixed types of heading, e.g.:

abbreviations . . .
British Standard . . .
dictionaries . . .
Everyman’s Dictionary of . . .
filings . . .
Gale reference works . . .
Initials and . . .
medical . . .
organizations, World guide to abbreviations of (Bustress) . . .
Oxford dictionary of . . .
periodical titles, British Standard . . .
reviews of works on . . .
scientific . . .
serial and periodical titles standards . . .

The heading that immediately follows this sequence is an inverted book title:

Abbreviations of organizations, World guide to (Bustress)

but this is already listed above as the inverted subheading ‘organizations, World guide . . .’, so is unnecessary. Also in the sequence are subheadings that refer to the same or overlapping topics: ‘British Standard’, ‘periodical titles, British Standard’ and ‘serial and periodical titles standards’; ‘dictionaries’ and ‘Gale reference works’ (which include dictionaries), as well as ‘Everyman’s Dictionary . . .’ and ‘Oxford dictionary . . .’ (which could be listed as sub-subheadings of ‘dictionaries’). It is not stated whether the original periodicals were consulted at all. Perhaps not, as there are some uncorrected errors that have been carried forward from the volume indexes. For example, under ‘Book reviews’ one of the authors of Information filing and finding is given as ‘Smith’ whereas it should be ‘South’. Going back to the original text would also have enabled the discovery that, for example, under the heading ‘abstracts’, the subheadings ‘archaeological’ and ‘archaeology, indexing’, with their separate locators, both refer to the same topic (in different volumes) and so in the cumulation the locators could have been brought together under a single subheading. Similarly, under ‘copyright’ the three subhead-
ings ‘of indexes’, ‘in indexes’, and ‘law of’ refer to related items, so might well be reduced to one subheading.

There are introduced errors, too. For instance, the third heading in the ‘A’ sequence of the index is ‘From Abantiades to Zygophyllum (Forey)’. This is incorrect in two respects – its placing and its typographic. The correct placing (because it is the title of an article) is under ‘F’, and the typographic styling should be ‘From Abantiades to Zygophyllum (Forey)’. Both were correctly presented in the original index.

Presentational inconsistencies include:

- different orders in joint authors’ names, e.g.:
  - Adams, Douglas and John Lloyd . . . [forename first]
  - Aitchison, Jean and Gilchrist, Alan . . . [surname first]
- ‘s’ and ‘z’ variants (maybe copied from earlier entries, but odd together), e.g.:
  - British National Bibliography . . . computerized . . . cross-referencing criticised . . . mechanized databases . . .
- alternative spellings, e.g.:
  - computer programs [heading]
  - computer equipment see also computer programmes [cross-reference]
- cross-references that lead nowhere, e.g.:
  - aristocracy, titles see titled persons
  - noblemen, titles see titled persons
  - peerages, indexing see also titled persons
- capitalization: most ‘subject’ headings are all in lower case, but not all, e.g.:
  - Telephone Directories television
- Web indexing websites see also web indexing

One of the expectations of any index is that it will contain (either as headings with locators, or as see cross-references) the words that users look up when requiring information. Some significant terms do not appear in this index. For instance, while ‘copyright’ is there, ‘moral rights’ is not (even though the topic has been written about in the periodical). Other expected, but ‘no-show’, terms include ‘authority files’, ‘search engines’, ‘virtual libraries’ and ‘whole-document indexing’.

The preface refers to the work of ‘unravel[i]ng the intricacies of the British Standards’, but does not say to what extent the standards were intended to be observed. Certainly, the recommendation of BS ISO 999 cited above, concerning the gathering together under one heading of all references to a concept and of variant spellings, has not been followed; as a result, users will themselves have to think of possible alternative terms and variations.

This cumulation will to some extent meet the expectations identified at the beginning of this review. It certainly provides, at a very reasonable price, rough and ready access to much of the valuable content of 40 years of The Indexer. It supplemements, rather than replaces, the separate volume indexes. Its limitations – the lack of harmonised headings and subheadings, the inconsistencies, gaps, and errors – mean that it will not be regarded as a model index.

References


Indexing aids


There is no one who, from time to time, when writing, does not search their mind for a better word or phrase to express what they want to say or who does not look for a substitute for words used more than once. The writer could turn to the Thesaurus, described by Peter Roget, the first compiler, as ‘words and phrases classified and arranged so as to facilitate the expression of ideas and assist in literary composition’. In other words, a book wherein lists of words are arranged according to related concepts. Or the writer could pick up this Dictionary of Synonyms and Antonyms, a near cousin of the Thesaurus, which sets out, in a different approach, some 170,000 synonyms and antonyms in alphabetical order. One can, in this Dictionary, immediately look up the word in question and find another word having the same meaning, or nearly so. Exact synonyms are rare, and since words listed can have slightly different meanings from each other, the advice is given, if uncertainty exists, to make use of a dictionary to ascertain the exact meaning of a word.

A single entry is divided into two or more parts where the word is capable of different meanings, and two separate entries occur when words owe their different meaning to history (e.g. lock, a bolt; lock, of hair). At the end of many entries, an antonym is given where appropriate.

Lists of related words are contained in an Appendix. These are most interesting, but why confined to air/space vehicles, cattle breeds, cheeses, collective nouns, dog-breeds, herbs/spices, jewels/gems, wine-bottle sizes, and zodiac signs (some having almost one hundred entries)?

An indication of the usefulness of the Dictionary is the fact that it has been reprinted each year (with one exception) from 1990 to the present reprint.

British Goudie, freelance indexer

Indexing technology


This book will be a very welcome addition to the bookshelf of anyone working with any form of digital publishing. It is by far the most thorough, authoritative digital publishing reference source available. Editor William Kasdorff, Past President of the Society for Scholarly Publishing, has pulled together an impressive array of experts to produce this definitive guide.
The Columbia guide to digital publishing is composed of 15 chapters, a glossary, bibliography and index. Kasdorf wrote the first chapter, ‘Introduction: Publishing in Today’s Digital Era’. It is an excellent overview of the role of digital publishing in various facets of the publishing industry. This chapter sets the groundwork for the rest of the book.

Other chapters address topics such as: the technical infrastructure, mark-up, content management, digital rights management, e-books, archiving issues, legal issues, accessibility, and international issues. Of particular interest to me is John Strange’s (Blackwell Publishing) chapter, ‘Organizing, Editing, & Linking Content’. He begins with a discussion about the role of XML and DTDs and moves on to the importance of metadata and types of linking. This chapter, like the others, is full of URL references for further study. Also, words and phrases in bold type indicate that the term is defined in the glossary.

If you are wondering about the meaning of XML, DTD, and URL, this book’s glossary will provide answers. The glossary is not only extensive, but within each definition section numbers are provided that indicate the location within the book where the term is discussed, as in the following example:

**glyph.** A unique character shape or a graphic representation of a character. See sections 3.3(ii), 2.5, 6.7(i).4

Using the running footer (bottom, recto pages) it is very easy to locate the section references in glossary entries. Indexers working with documents that discuss digital publishing will find the glossary in this book extremely helpful.

Thad McIlroy (Arcadia House) wrote Chapter 6, ‘Composition, Design, & Graphics’. This is a must-read for those interested in typography and design in digital production. Topics covered include: typography, graphic types and file formats, color processing, page production, image capture and processing, workflow and printing. Like the other authors, McIlroy discusses real-world, hands-on issues, such as: characters look okay on the computer screen, but they don’t print correctly. I enjoyed his discussion of dashes (hyphens, en-dashes, and em-dashes). It is the best summary I have read, including this caution:

It is a sign of typographic inexperience to use, instead of an em dash, a pair of hyphens (--) or a hyphen with spaces around it (-). Improper use of dashes is common on the Web, where the spaces often disappear and the dashes can’t be distinguished from hyphens.

Are indexes discussed? Yes, here and there. There is only one reference in the index to ‘indexes’. A logical place to discuss indexes is in regard to mark-up and granularity. Granularity refers to the level of detail to be tagged in a document. For example, a low level of granularity would be to tag the beginning and ending of a bibliographic reference. A high level of granularity would include tags for the authors, date, journal title and other elements within the reference. In the chapter about electronic books (Chapter 11) there is an interesting discussion about supporting advanced functionality through the identification of specific components of the content: ‘possibly including very fine-grained classification of domain-specific items: for instance, names of persons, places, and countries in a history book; or diseases, organisms, drugs, organs, anatomy, and treatments in a medical book’. Sounds a bit like an index to me!

The most detailed discussion of indexing is in the editor’s Preface. William Kasdorf’s appreciation for a book index is heartening.

The third major avenue into the Guide is the Index. This is a particularly notable feature of the Guide. Indexes are taken for granted in print and, unfortunately rarely provided online. Whereas the Glossary provides a shallow, topical view of the content, and the Table of Contents provides a logical structured view, a good index provides an intellectual view of the content unavailable by any other means. It is the result of an intelligent reading by an indexer trained in recognizing and documenting the interrelationships of the intellectual content; the indexer not only notes topics and subtopics, but also makes judgments about them, selecting the most important and relevant sections to direct readers to. Simple searching for terms does not substitute for this; it provides too many inappropriate or unhelpful hits, and can miss the most relevant section if the term does not occur in it. Much effort continues to be made to compensate for this in search engines, but these only approximate the function of a well-made index like the one print or online users of this Guide have available.

The next section of the preface, ‘How This Book Was Done’, is fascinating and includes a discussion of the indexing process. Findability has become the battle cry of indexers, information architects, and taxonomy designers. The Columbia guide to digital publishing is a showcase of findability, starting with a very detailed Table of Contents, cross-referencing within chapters, and an extensive glossary and index. There is also an online version of The Guide, which has not been assessed for this review. It is available by subscription at www.digitalpublishingguide.com

Nancy Mulvany, freelance book indexer

**Other subjects**


This is an angry book written by angry and frustrated people, with a touch of sadness and arrogance thrown in for good measure. It is based on a conference that took place at the University of London on 12–13 March 2001 and was in response to an article that had appeared in July the previous year, published by the New Yorker and written by the American novelist and essayist Nicholson Baker. In this article, Baker lamented the fact that newspapers were being discarded from libraries in favour of copies in the form of microfilm, and the serious shortcomings this type of preservation engenders.

Newspapers have always been and continue to be the Cinderellas of the print industry. The book remains prime, magazines and journals come a poor second, and newspapers are a thumping third. The reasons are obvious: books are compact, bound, written by well-known people and take up space in libraries in an orderly fashion. Newspapers and journals, produced day in and day out, week in and week out, take up space, disintegrate, grow in an alarming fashion and are difficult to handle and store. That they are the only print medium that give us an insight into how people thought and acted on any particular day over the last 300 years in a continuous thread seems to have been forgotten. They are ephemeral and few of us have the resources to keep them for any length of time. It is because of this that it should be the duty of national libraries worldwide to keep copies, in their original form, for users to consult.

Of course, many would argue that keeping original newspapers these days is an obsolete way of preservation, particularly as technology enables us to copy print to a high degree of accuracy, but this simply is not true. Provided the original is in a good state, there are no illustrations, it is in black and white and it is in a Roman script, then there may be a case for microfilming the contents. But most newspapers and journals are illustrated, if only with advertise ments, and many are not in pristine condition and are creased, crumbling at the edges and have suspect bindings. Conservation is expensive, and the cry of ‘acidic paper’ seems to justify the fact that these newspapers and journals will eventually eat themselves to death. The only way to preserve is to copy electronically. But this too is a trap. Yes, acidity is something to watch, but then microfilm too is subject to deterioration and its shelf-life is expected to be much less than that of the objects it is copying. The machines on which they are used are also subject to problems of breakdown and difficulty of handling. Anyone who has tried to scroll down slowly,
let alone across a page of newsprint, will know just how difficult and frustrating the process can be.

Digitization is one way forward, but OCR (optical character recognition) is still struggling to accurately distinguish between type scripts. At the Fourth National NEWSPLAN conference held in April last year at the British Library, it was pointed out that Gothic lettering, particularly when it appeared on a page that also had Roman style, never mind three different languages, as was the case in Finland during the 18th and 19th centuries, presented huge problems for OCR translation. Olive software did demonstrate their remarkable system, which looks very promising in that it can indeed show whole pages of newsprint which were perfectly readable and searchable on any topic, advertisement or photograph. A very exciting development.

All of this, however, fails to take into account the necessity to see, feel and use original material. Thanks to modern technology and the wonders of the internet, one can now have virtual tours of museums and see objects in the round, but no one in their right mind would consider disposing of the original object in favour of digitization. So it is with newspapers. It is often imperative to look at an article in the context on the page; to see how the news is laid out, what graphs and illustrations, including photographs, are there and the placement of advertisements. Image/text is a growing area of academic interest and serendipitous, often fortuitous, findings of information can only be achieved by looking at the whole page of a newspaper, not just a section or scrunch-up image of the whole.

All the articles in Do we want to keep our newspapers? cover the problem of disposal and preservation in one way or another. The keynote paper by Nicholson Baker is a rallying cry to libraries and academics alike, that by librarian Karin Wittenborg is contentious in her comment that she doesn’t think ‘that all newspapers are worth saving in original form in their entirety’. Others put forward the case for keeping the original papers and allowing limited access to them, but at the same time making better digital copies available; and finally Mike Crump, Director of Reader Services at the British Library, puts forward the British Library’s policy with regard to newspaper disposal. He makes the point that all libraries dispose of material even when they act as archives and that the BL ‘aims to be the nation’s guardian of the national printed archive’. A pity then that as far back as 1996/7 the newspaper library’s News sheet wrote that some ‘60,000 volumes of mainly broadsheet newspapers’ could be disposed of, even if these included American, European and Commonwealth newspapers. Never mind that these are foreign papers: quite often academics, foreign or indigenous, need to do cross-readings of papers to come up with a balanced view of a particular event or paper. Baker is particularly critical of the BL’s policies, particularly in the USA, where large sums of money were devoted to the process, with the obvious corollary that much space would be saved in libraries. Baker argues that the storage costs could have been much less than that of the huge microfilming projects. Recently he has been treating with the British Library over its policy on foreign journals, and has formed a company to buy runs from libraries and store them. Quite what he hopes to gain from this is not even clear to the author himself, but it is an effort to save some lambs from the slaughter, and he must be hoping there are some other agnophiles in the world.

He is also particularly horrified at the destruction of books that has taken place, the raison d’être here being the likelihood of the paper turning to dust; by microfilming, the intellectual content of the book is saved. The title of the book refers to the test made on paper by folding and folding again the corner of a page, which in the end breaks off, a practice the author sees as an unusual misuse of paper corners! A further horror story is the removal of acid stains by an immersion in DEZ (diethyl zinc) gas, a process that led to a number of false starts, and was never very successful, leading at worst to strange smells in the book, and at worst to explosions in the chamber.

The other process that may be applied to books is digitization, where again much space may be saved by transferring the intellectual content to disk, provided the book is discarded. Most of us have suffered from the horrors thrown up by OCR at its worst. One chapter is headed ‘subterranean convulsion’, which sounds like a satirist’s Bush-ism, but is how ‘subterranean convulsion’ came forth on the computer. In reality, there may be value in the digitizing of books, but the best of technology is not applied to them, and the treatment of illustrations is a problem.

Baker is not precious about books. He has a number of friends who are conservationists, and there are also many scholars to whom any manhandling of the traditional artefact is anathema. Although much of the tone of this book is in enthusiasm ‘over the top’, speak, much cooler prose could legitimately be devoted on the behalf of historical bibliographers.

The author has been assiduous in tracking down those who have brought about the current situation. Quite a few of the texts written by heroes whom older librarians will have studied are quoted – Vernor W. Clapp and Fremont Rider for instance. The connection with the CIA, NASA and other American organizations is inter-
est. Where possible the author has interviewed those involved; there must have been some squirming on those hot seats. There are also heroes – the UK's Peter Waters for instance. But no one has quite got into gear as far as this author. His book, despite some tedious repetitiveness (looking for further mountains after piling Pelion on Ossa?) is well worth reading. He has a sharp humour (much effort has gone into his chapter headings) and a lot of guts. The book rarely mentions indexes, though indexers will think of their own value in the preparation of such in both microfilming and computerization projects. The book's own index, if letter by letter, is very good, by the way.

A reviewer has labelled Nicholson Baker 'the Erin Brockovich of the library world', which is not an unreasonable title.

David Lee, freelance indexer and bibliographer


Gagingwell Oxon. Godelingwelle e. 1173. 'Spring or stream of the kinname'. OE gadeling 'giving + wella'. Adolf (m.) German: composed of the Germanic elements adal noble + wolf. . . . The association with Adolf Hitler (1889-1945) has meant that the name has hardly been used since the Second World War.

If you find this sort of information fascinating, you will enjoy this book. Claimed as 'the definitive guide to surnames, first names and place names of the British Isles', it brings together three separate Oxford dictionaries, of surnames (first published 1988), of first names (first published 1990), and of English place-names (first published 1991) with the incorporation of new sections on Irish, Scottish and Welsh place-names. Its appearance reflects a recent growth of interest in onomastics (names studies). The long-established English Place-Name Society is more active than ever and the growth of interest in names research has expanded some of the derivations in its dictionary of place-names (1936 and later editions); for example, Ekwall derives Halliday from 'Cissi's Ceaster' (Roman town of a chieftain called Cissa); the new dictionary gives this derivation plus a possible alternative, an Old English word cisse meaning a gravelly feature. Over 12,000 place-names in England alone are included; the criterion for selection is inclusion in the popular touring atlases published by the Ordnance Survey and other bodies. Some places in Ekwall are omitted, but others not in Ekwall are now there, including some quite small hamlets. However, as the criterion for inclusion is etymological, the dictionary cannot replace a gazetteer such as Collins Britain atlas and gazetteer (HarperCollins, 1999) as a reference book for indexers of topographical material.

Ann Hudson, freelance indexer and tutor in indexing


A book of limited interest to indexers as such, this is valuable to the expert in new spheres of watermark study, printing papers from the 1996 Conference. Paul Needham writes on 'Concepts of paper study', which is a delightful paper, in the other sense. Several writers take watermarks of particular authors (John Donne and Benjamin Franklin for instance) and solve puzzles in the dating of editions, or manuscripts. Music is a particularly interesting if esoteric study to several authors here. New aspects of technology are explored in relation to watermarks – phosphorescence watermark imaging, DYLUX instant access imaging paper, for instance. Later essays cover theoretical aspects of watermark study. Two of the editors describe the Thomas L. Gravell archive, which is now available on the internet. Gravell is one of the heroes of watermark study. There is now quite a decent set of watermark databases on the World Wide Web. The last essay, perhaps of most interest to the classifcationist in any indexer, is Ted-Larry Pebworth's 'Towards a taxonomy of watermarks'.

This is not an easy book to read, but it is an important one in the field. There is an index, which includes some broad headings like 'paper mills', 'repositories' and various group entries under 'watermark'.

David Lee, freelance indexer and bibliographer

The author, a well-known authority on slang, has arranged this book by topic, and a glance at the contents page confirms his statement in the introduction that the ‘central strands of the slang vocabulary’, which were once defined by the Edinburgh Review as ‘sex, money and intoxicating liquor’, form its mainstays, along with a ‘proliferation of body parts and functions’. Starting with ‘men’, ‘women’, ‘the body’ and its parts, we proceed through insults, swearing and oaths, food and drink, and sexuality to ‘race relations’ – a chapter each being allotted to the Jews, Irish, Dutch, French and Chinese – and the story continues with crime and the police, culminating in ‘imprisonment’, ‘execution’ and ‘death’. With the exception of drug terminology, the seamy side of life is certainly fully represented here.

Lucien Green receives a mention in the acknowledgements ‘for his invaluable help on the index’. This is printed, in this paperback edition, in very small type to fit into 48 pages (out of a total of 393). A typical page contains some 30 indexable slang terms, which are printed in bold, making them easy to spot once the index has led you to the right page. The index entries correspond exactly to the terms in the text, so the index cannot have been difficult to compile. Most of the entries are brief, consisting of one or two words, only a few, such as ‘one of these days you’ll wake up dead’, requiring more than one line in the index. But would a user really look for the above expression under O, rather than W for ‘wake up dead’? Again, ‘take no prisoners’ is not easily found among a large number of entries beginning with ‘take’, and there is no entry under P for ‘prisoners, take no’. This is the disadvantage of always using the term strictly as found in the text, rather than indexing by the headword most likely to be sought by the user.

Christine Shuttleworth, freelance indexer

Publications received and publications noted


