

Catalogue and Index

Periodical of the Chartered Institute of Library and
Information Professionals (CILIP)
Cataloguing and Indexing Group

Editorial

December 2011, Issue 165

By the time you read this, Christmas will be just a matter of a few days away, so we bring you an issue full of tasty treats to consume while toasting your tootsies in front of a log fire. (Yes, we're in Fantasyland!) This issue is all about indexing, so if you have ever considered a career change, herein you will find some ideas and advice to contemplate while you sip your mulled wine. And it's not all back-of-the-book indexing – we also have an article about indexing conference material.

Just to add a bit of variety and spice to your reading, we have reviews of what seem to be three extremely good books. There might be just enough time to add them to your wish list. Ideal holiday reading if you're not interested in the latest airport blockbusters.

May we wish you all the best for a safe and happy holiday season – whatever your religious bias (or lack thereof) and here's hoping for a more prosperous 2012 than many of us have had this year. We will be back with you around March.

Cathy Broad

library@ethicalsoc.org.uk

Heather Jardine

heatherjardine402@hotmail.com

Please note that views and opinions expressed in the articles are those of the authors and not necessarily those of CIG.



Contents

2-3 Time for a change

Ann Kingdom

4-6 Troubles shared?

Bill Johncocks

7-8 A triumph for team indexing

Ann Kingdom

9-12 Indexing conferences at the British Library

Sarah Hammond

13-15 A librarian considers indexing

Cathy Broad

16-18 Book reviews

19 Noticeboard

Time for a change? Use your transferable skills

Ann Kingdom Chair, Society of Indexers

Freelance indexers come from a very wide variety of backgrounds, but for obvious reasons indexing has always traditionally attracted library and information professionals. Indeed, there have always been strong links between the Society of Indexers and CILIP. Many prominent members have played key roles in both organisations – the late Ken Bakewell being just one example – and a representative of the CIG has usually sat on the Wheatley Medal panel. For readers of *Catalogue and Index* facing uncertainty in their current post, fed up with the nine-to-five grind or the distraction of office politics, or keen to use their skills in retirement, what better career move than retraining as an indexer?

The Society of Indexers runs a well-established distance-learning course (Mary Pigott and Ken Bakewell were both involved in earlier editions), originally delivered via printed booklets and relying on postal services to transmit assignments between students and markers. For the fourth edition (launched in 2010), a web-based platform offers seamless access to study materials, a wide range of useful resources, and interactive exercises and quizzes that enable students to review their progress. No longer forced to work in isolation, they have plenty of personal contact and support through an email discussion list and online tutorials, which give much more indexing practice to reinforce the theory.

High standards are expected in order to gain Accreditation and the status of professional member of the Society of Indexers (MSocInd); after all, a carelessly and inconsistently constructed index is no help to the subsequent user. However, anyone with a background in library and information work should have no difficulty in achieving high marks in all four modules, the content of which is summarised below.

Training in indexing: module content

Module A. Introducing Indexing: indexers, users and documents

- basic indexing terminology
- the functions and characteristics of indexes
- what users want from indexes
- what kind of people make good indexers
- the indexing process
- the role of authors and other document originators
- document production and categorization
- creating bibliographic references

Module B. Choosing the Words: term selection and the formation and arrangement of headings and subheadings

- selecting concepts for indexing
- choosing appropriate index terms
- forming headings and subheadings
- devising cross-references

Module C. Names and Numbers: filing order, proper names and locators

- rules for arranging index entries
- compiling multiple sequences of indexes
- indexing proper names
- presenting locators (page numbers)

Module D. Finishing Touches and Beyond: index presentation, specialized forms and applications

- indexing books and journals (periodicals)
- index layout, presentation and delivery
- embedded indexing
- website indexing
- thesaurus construction
- indexing for computer-based search systems
- team and cumulative indexing

The fact that librarians are eminently suited to take up indexing was ably demonstrated by Marie-Pierre Evans, a qualified librarian, who was presented with the Society's Betty Moys Prize at its annual conference on 3 September as the most outstanding student to complete the course in 2010. While stuck in traffic jams going to and from her job, the prospect of never having to commute again was very tempting. Indeed, the long-contemplated move to North Wales would prove no impediment to her career if she could work from home. Indexing had been at the back of her mind for some time; she knew that with her library training and experience she would have the right skill set. 'I spent quite a few years cataloguing and classifying materials in English, French and Russian and I thought the course would be a doddle,' she joked 'but I was impressed by its rigour. It gave me a really good foundation in both theory and practice and I'm really enjoying my new-found freedom working from home. Those tedious traffic jams are now a distant memory.'

To find out more about the Society of Indexers and its training course, visit www.indexers.org.uk

More than half the members of the Society of Indexers today have been librarians at some stage of their careers, so there must be some similarities between our two professions. At the very least, it seems that both appeal to the same type of person, although there is an ongoing debate about the transferability of skills between them. CIG members will of course be only too familiar with the problems in UK librarianship; they might perhaps be interested in a review of the challenges currently being faced by the junior profession.

Happily, there is no dispute over seniority. If librarianship required only the prior invention of writing, indexing is normally considered to be a child of printing. That's historically true, but it wasn't inevitable. Indexing doesn't actually depend on uniformly paginated editions; all it really needs is a stable text, incorporating some structure that will allow readers to identify locations unambiguously, a distinction that has assumed more importance recently. Instead of fixed page lengths, that structure could be line, section or verse numbers, or dates in chronologically-ordered works like Pepys' *Diary* and Boswell's *Journal of a Tour to the Hebrides*.

It's not about page numbers

Perhaps the archetype of such stable, structured texts is the King James Bible, whose fourth centenary has been suitably celebrated this year. In that definitively stabilised text, positions can of course be identified independently of pagination, with a precision of a few sentences at worst, by citing the familiar sequence of book, chapter and verse. It has become so universal that probably nobody has ever said 'open your Bibles at page 847'. Unfortunately though, this familiar referencing system wasn't completely standardised until almost a century after Gutenberg began to produce uniform editions and, in any case, attempts actually to *index* the Bible in the fifteenth century might have resulted in immolation as predictably as did the first translations. The early tools for accessing the text were therefore the less controversial concordances; simple lists of actual term occurrences with their locations.

That distinction from concordances defines what human indexers do to this day. Rather than just listing where terms appear, they analyse each passage for its likely significance to the intended readership; add alternative terms which the reader might be expected to think of (regardless of what appears on the relevant page); provide links to other entries related either semantically (synonyms or hierarchies) or contextually, and split unwieldy numbers of postings into manageable chunks using sub-entries. Unlike us, computers are hopeless at divining meaning, so they can't index; they're brilliant, though, at spotting the spaces between words. So, for almost fifty years IT professionals have been trying to foist on readers an alternative way of accessing text, and this alternative tool is founded on the concordance principle: not displaying meaning but just words and where they are used.

Dumbing down retrieval

Around forty years ago, in the early days of machine sorting, when permuted titles were the best that could be achieved, full-text searching was seen as the Holy Grail of information retrieval. Now we have it, and quite frankly, it sucks. Keyword-based retrieval is so limited because it's based on a wholly false model of human communication. We all know that writers don't restrict themselves to repeating unvarying noun phrases, but instead delight in a rich, allusive, convoluted and capriciously varied vocabulary, from which readers discern meaning and gain understanding at a completely different level from matching keywords. We all know, but the IT industry has been very reluctant to accept it.

In fact, keyword retrieval pulls off the difficult trick of combining low precision with low recall; it finds too much but doesn't find everything. Search for 'tigers' and you will miss the key anecdote where one is called a 'maneater', but you still get all the worthless occurrences like 'except tigers' or 'as we saw when discussing tigers' as well as the false drops: 'fought like tigers'. With more and more non-fiction books being multi-author compilations, only a human indexer can link, say, 'farm waste pollution' in Chapter 3 with 'agricultural runoff problems' in Chapter 14. Serious books always deserve serious indexes and books intended to be understood by people should be indexed by people.

Sadly, keyword retrieval is now all too familiar from online searching. Of course, Google is wonderful for finding a cheap flight to Alicante or discovering who said that a foolish consistency was the hobgoblin of little minds; Wikipedia is invaluable for checking the differences between DNA and RNA or identifying the capital of Mozambique. But books don't just impart facts; they offer challenging insights and if we replace the habit of attentive, sceptical reading with establishing truths by consensus, and that of engaging with arguments by pasting extracts, it's not alarmist to suggest that civilisation might be in peril. A literate society is a free society and the converse surely applies.

Equally sadly, modern publishing is controlled by people for whom the distinction between a concordance and an index has become blurred. The ability of computers to extract word occurrences from machine-readable text and alphabetise the result in a second or two is impressive enough to have persuaded some authors and publishers that this might serve in place of an index. After all, unless you're actually trying to find the treatment of a subject, one list of terms accompanied by page numbers looks rather like another. Even highly gifted authors can forget that the index is a tool to benefit the reader and one about which their opinions don't automatically deserve to be especially privileged. Also, concordance-generation packages abound, their devisers' siren voices decrying meticulous indexing as outdated and needlessly time-consuming. The superficially plausible threatens to displace the lastingly valuable.

Limits on re-use

There are still two more key challenges facing indexers. The first is that twenty-first century indexes are expected to be capable of being re-purposed; they won't be restricted to the backs of bound volumes and supplements to periodical runs, but required to give access to the same text on websites and in eBooks. Nevertheless, human indexing is fundamentally unchanged by the requirements for more flexible delivery because it comprises two distinct stages. The first, the selection of entries, is intellectual and can't be automated while the second, generation of the sorted list, is both mechanical and error-prone and where possible always should be automated. Indexers are conservatives but not Luddites. We already use computers for what computers are good at; transferring files, finding occurrences, sorting, suppressing duplication, validation and formatting the results.

An unpaginated text can be given a device-independent index either by embedding the terms electronically in hidden form behind the text discussion; by using the markers we discussed before (whether pre-existing or automatically-generated sequence numbers for sentences or paragraphs) or by inserting hyperlink anchors so that readers can click on index terms or locators and be taken instantly to the start of the corresponding treatment. Emerging technologies based on XML promise to be able to combine all three, with different stylesheets for different display formats (including printed books).

The techniques exist, but indexers are not always being given the chance to use them and our readers aren't often being given the choice. The effort to cut time and expense seems to have made some publishers forget that they're in the business of selling enlightenment, not fizzy drinks. Books must be fit for purpose and unindexed non-fiction books simply aren't. The fact that dominance in the eBook reader market is being determined by sales of non-indexworthy, read-once-and-throw-away 'airport novels' has tempted many publishers to offer full-text search because it's easy, but it could also be persuading a rising generation of potential readers that keyword occurrences are a substitute for analytical indexes, while the Internet suggests that written material is for cutting and pasting into assignments, not for attentive reading.

There's currently some discussion about whether indexes can cope with print-on-demand, allowing the assembly of ad hoc compilations by selecting chapters from several previously published and previously-indexed books. At present, there are difficulties, not just because a good index fits the book for which it was designed as a key fits a lock, but because (although embedded indexing will preserve accurate locators in this form) writers don't normally write with this kind of fragmentation and recombination in mind. Controlled vocabularies are excellent for collections and perhaps for such reassembled fragments; their granularity has always been too coarse to suit individual books. But I have to admit that even indexing may have its limits!

Finally, of course, our profession is disadvantaged by the iniquitous tradition of financing index production from authors' royalties, so that most indexes are made by authors and most are made badly (though not only by authors). The pressure to perform a task for which most authors have no more aptitude than indexers have for cover illustration, say, is inevitably greater at times of squeezed profit margins.

It's not all gloom; more English language books are indexed than is the case even in most of continental Europe, and the spread of multi-authorship has convinced many authors that indexing needs a professional to reconcile chaotic terminology. Formerly, of course, that professional might have been a really attentive editor.

Saving the world?

Indexers and librarians alike labour with the express purpose of saving readers' time. Both may sometimes have been tempted to make their users jump through unnecessary hoops in order to get their cleverness noticed and our profession certainly has often been short-sighted and timorous (for example over assisting author-indexers, embracing new technology and by emphasizing training to the exclusion of external relationships). Indexers are few, scattered, isolated and hopelessly disadvantaged in negotiating with global publishing corporations and their prestigious IT consultants. Mutual respect and understanding is a first essential of all cooperation and, though relations with CILIP are generally excellent, our Society was offended recently when the editor of *CILIP Update* incorrectly implied that unpaginated eBooks weren't indexable. Nevertheless, librarians, indexers, copy-editors and proofreaders are natural allies; we all depend not just on writing but on attentive, even critical, reading and must make common cause in its defence.

Bill Johncocks is a science indexer, indexing tutor and editor of the SI newsletter, *Sidelights*; also a former MIIInfSc. Since going freelance, he has lived on the Isle of Skye.

A triumph for team indexing: 2011 Wheatley Medal

Ann Kingdom Chair, Society of Indexers

At the Society of Indexers annual conference at Keele University, the Society's President, Professor John Sutherland, presented the Wheatley Medal for 2011 to a team of indexers from the Netherlands, Caroline Diepeveen, Pierke Bosschieter and Jacqueline Pitchford-Belder, for their work on the index to *The Encyclopedia of Jews in the Islamic World* (edited by N. Stillman and published by Brill in 2010). According to the chair of the judging panel, Professor Cathy Shrank of Sheffield University, this was an exceptional index and a most deserving winner. 'The size of the task (indexing a five-volume work) was immense in itself, and indexing an encyclopaedia (which is itself a kind of index) requires something different from indexing other works. In this case, this difficult task was accomplished with scholarly authority.' The judges also noted the excellent use of reader-friendly devices, the consistent practice of translating foreign language terms to make the text accessible to a less specialist audience, the attractive lay-out, which was also easy on the eye, and the adroit use of double-postings. They singled out the scrupulous editing for particular praise. 'None of us found any clues to the fact that this index was compiled by a team. The overall editing [by Caroline Diepeveen] to combine three people's work must have been very thoroughly done. This is no mean feat in an index of this size. In short, the panel was deeply impressed.'

Joan Dearnley was highly commended for her index to *The Edinburgh Companion to Virginia Woolf and the Arts* (edited by M. Heller and published by Edinburgh University Press in 2010). The judges felt that she had risen ably to the various challenges it posed: a work focused on one person, requiring careful thought on how to break down the subject into multi-faceted categories; a multi-authored volume whose contributors had different styles and often used different terms; and, most significantly, a text dealing predominantly with abstract concepts. 'The index for this *Companion* more than met these challenges,' said Cathy. 'The indexer showed a remarkable ability for analysing the text, and produced an index which was responsive to, and usefully reflected, the main subject of the book. This was a considerable accomplishment.'

Finally, Barbara Hird was commended her indexes to the *New Cambridge History of Islam*, vol. 1 (edited by C. F. Robinson) and vol. 4 (edited by R. Irwin, both published by Cambridge University Press in 2010). These indexes were submitted as separate entries; the panel – unaware that they were by the same person – found it impossible to judge between them. 'It was with some relief that we discovered that we could roll them together,' said Cathy. The indexing task for both these volumes was complex, owing to their size and detail. 'The indexes showed scholarship, and were comprehensive, clear, and helpful, with the glossing of foreign terms well-placed to aid the less-specialist reader. The indexer also dealt admirably with the challenge of transliteration and diacritics. This was exemplary, meticulous work.'

Overall the judges felt there was a very strong field this year, and extend their thanks to all the publishers and indexers who had submitted work. However, sadly, there was not a single scientific or medical work amongst the eighteen entries.

Biographical information on the winners

Caroline Diepeveen studied political science and international relations at the University of Amsterdam and the London School of Economics, gaining a bachelor degree in political science and a masters degree in international relations, specializing in the modern history of the Middle East (both University of Amsterdam). After brief careers in research and management consultancy, she took up indexing after relocating to Scotland, following her husband who works in the oil industry. She received her Society of Indexers accreditation in December 1997 and has been a freelance indexer ever since, specializing in social sciences, modern history, international law, Islam and Judaism. She is now an advanced professional member of the Society of Indexers and since 2000 has lived and worked in Middelburg (Netherlands).

Pierke Bosschieter started her working life in the public health service. In need of change, she studied library and information science and worked for several years as an archivist. Food and cooking have always been her great passion, so when an opportunity arose to start a vegetarian catering business with a friend, she jumped at the offer. When after a few years of catering the work proved to be too strenuous, she had to search for a more sedate means of support. A lifelong avid reader, she at last found her *métier* in indexing. She started indexing in 2005, specializing in the humanities and, of course, in food-related subjects. She publishes articles about indexing on a regular basis, as one of the editors of *The Indexer*.

Jacqueline Pitchford has a background in library and information science and received her Society of Indexers accreditation in 2009. Prior to starting her own business, Index 'n' Things, Jacqueline gained over 12 years' experience in desk research and knowledge management working for leading international management consultancy firms such as McKinsey & Company and KPMG Consulting. Apart from offering indexing services, Jacqueline carries out desk research assignments on a regular basis. She has experience in the following subject areas: agriculture, consumer goods, culture, economics, environment, history, international relations, religion, social sciences, and transportation and logistics.

The index for the *Encyclopedia of Jews in the Islamic World* was their first index as a team and all three are founding members of the Netherlands Indexing Network (NIN).

Joan Dearnley holds degrees in history and fine art and a postgraduate diploma in librarianship. Several years working in academic libraries was followed by a period looking after her two small children, during which she gradually developed a portfolio career of freelance and part-time work (piano teacher, art gallery assistant and school librarian), while also studying for a masters degree in fine art. It was whilst 'indexing' books for a local first-school library that she remembered hearing a radio programme about indexing and the Society of Indexers. She qualified as an indexer in the late 1990s, attending her first conference in 1999, where she found herself volunteering to help run the local Yorkshire Group. So began what has become a continuing involvement with the work of the Society, as well as a successful indexing career. She currently serves on the training and CPD committees, has contributed to the development of the Society's training course and is a workshop tutor. Joan became a Registered Indexer in 2002 and is now a Fellow of the Society of Indexers.

Barbara Hird read Classics at Cambridge University and was a Rotary Graduate Fellow at the University of Athens. After several years as a civil servant, and then at home with small children, she trained as an indexer, and has been a Registered Indexer, and subsequently a Fellow of the Society of Indexers, since 1987, specializing in the ancient world, late antiquity and the middle ages. She is a past Vice-President of the Society and has chaired its Training and Qualifications Board. She was previously awarded the Wheatley Medal in 1999 and 2008 and was highly commended in 2009 and 2010.

Background information about the Wheatley Medal and the Society of Indexers

The Wheatley Medal was established in 1961 by the then Library Association and the Society of Indexers to recognize and encourage excellence in indexing. The panel of judges includes representatives from the Society, from the Chartered Institute of Library and Information Professionals (CILIP) and from the academic community. Further information and a list of past winners is available on the Society of Indexers website (www.indexers.org.uk).

The Society of Indexers was established in 1957 and is the only professional association of indexers in the UK. It aims to promote indexing, the quality of indexes and the profession of indexing. In addition to its well-respected online distance-learning course, it runs a programme of workshops at various venues throughout the UK, and publishes a quarterly journal (*The Indexer*) and a series of occasional papers on specialized aspects of indexing. Publishers and others searching for an indexer need look no further than the Society's directory of Indexers Available, online at www.indexers.org.uk, where they will also find considerable advice on commissioning indexes.

For further information about the Wheatley Medal and about the Society of Indexers, contact the office (admin@indexers.org.uk) or visit the website: www.indexers.org.uk

Indexing conferences at the British Library

Sarah Hammond Conference, English Language and Thesis Cataloguing Team

Before I came to the British Library in 2008 I had worked for three years at the National Railway Museum, where my duties included cataloguing new acquisitions. This was my first introduction to cataloguing and classification and I found that I enjoyed it greatly, so when I had completed my MA in Librarianship at the University of Sheffield and was looking for a job, I was pleased to see an indexing post come up at the BL. My previous knowledge of cataloguing in Aleph at the museum proved to be extremely useful as that's the same software used by the BL, so I was considered safe to be let loose on the catalogue fairly early on. I work exclusively on conference output in a team of paraprofessionals creating records from a huge variety of sources. My official title is "Conference Team Support Member" and that "support" relates to the professional cataloguers whose cataloguing duties include conference material.

I work at the British Library in Boston Spa, part of which is the Document Supply Centre (DSC). The DSC has produced the Index of Conference Proceedings for many years, as it has long been acknowledged that often the first place new research is presented is at a conference, long before it is written up as a peer-reviewed journal article. This Index is produced in print on an annual basis and remains a popular product.

Conference material is drawn from serials, monographs and monographs-in-series and may consist of full proceedings, selected papers, abstracts, poster sessions and extended reports. Currently only print items are selected from but I am working to develop a robust method of selecting from ejournals that I hope will extend the coverage of the index in the future.

There are several other organisations that produce an index of conference proceedings but the British Library's publication remains a favoured product as it is arguably more comprehensive than any other index and it is tied to actual DSC holdings, so you don't have far to look to get your hands on them (rather than being simply a bibliography where you are left to your own devices to find the material). Having flicked through a few of the others I would also argue that it is more user-friendly too. The selection process we use is in part a manual one and I'll illustrate that later by describing a typical workflow.

Hundreds of journals, trade papers and magazines are received every day by the library and once they are checked in on the LMS they are routed to appropriate departments. The one we work most closely with is our Current Awareness Service who have responsibility for selecting high use titles whose tables of contents will be scanned and keyed in to be ETOC – electronic table of contents. The material that is selected for ETOC is therefore searchable at article level in online databases, including our own British Library Direct and also ZETOC, which is provided by Mimas at the University of Manchester from British Library data.

Some titles are automatically routed to my team based on the fact that they have been a rich source of conference material in the past but much is still sorted by hand. A group of us do a daily "conference sort" which to the untrained eye may look like a group of browsers in the WH Smith magazine shelves who have no intention of buying. What we're actually doing is flicking through tables of contents to identify conference material for inclusion in the Index. When I first started doing this I was incredibly slow at it, especially when it came to foreign language publications, but you soon find that you "get your eye in" and can speed through scores of items in a quarter of an hour. Words will start to pop out at you: conference, symposium, congress,

workshop, colloquium, proceedings, meeting, tagung, journée etc. Also a giveaway is a short date span and a geographic location, e.g. 23-25 August, San Diego, California. A perk of this job is that I do get to see a huge wealth and variety of material and have discovered a few gems that I like to read on my breaks. My favourite is The Horn Book, a US publication that discusses and reviews books for children and young adults. I have found many wonderful titles that I've then bought for my daughter thanks to this serendipitous find!

Once we've selected items for indexing we take them back to our desks and create analytic records for them following British Library in-house rules, which look something like this when completed:

CNF a Bioengineering and biotherapies (Meeting)
n (4th :
d 2007 Sep :
c Nancy, France)
245 a 4th Meeting on bioengineering and biotherapies and 3rd meeting of Lorraine
Center of Cartilage Engineering /
c editors: S. Muller ... [et al.].
260 b IOS Press,
c 2008.
300 a 1 v.
591 a Papers.
653 a Bioengineering
653 a Biotherapies
653 a Cartilage engineering
700 a Muller, S.
710 a Lorraine Center of Cartilage Engineering.
945 Biomedical materials and engineering, ISSN 0959-2989 ; vol 18 no 4/5 2008
FIN a Y
d 20090211

Now before anyone is offended / shocked / upset / bemused at what may look like us playing fast and loose with AACR2 I'll just take this opportunity to remind you that these records are added-value to the catalogue rather than being the backbone of BL cataloguing rigour. Whilst these records sit in the catalogue, they are primarily for publication through the Index and the online databases and serve to give researchers access to more material – material that would be difficult to find by any other means. Users of Zetoc will be familiar with the search functions, one of which is specifically for conferences. Another bonus of selecting items for which to create records is that these publications go back to Current Awareness Service and are added to those for which ETOCs will be created, so the end result is even more searchable content for the researcher.

So what you see are several local fields, the first of which is the CNF which is the DSC Conference Holdings field. It is much like a 111 field (Main entry – meeting name) and is used to record the name of the conference, followed by a conference term in parentheses (i.e. meeting, conference, symposium etc) and subfields record conference enumeration (if any), dates and location. There is some degree of freedom in choosing where to draw the 245 from and this is necessary because often the conference material is not a whole journal issue; the issue may well be selected papers from the conference with regular papers, book reviews, correspondence etc making up the rest of the issue (in this case the 300 \$a records a page span covering the conference material only). So the 245 can be drawn from a title page, a section heading, the table of contents and sometimes from an article title. If there is more than one source it's prudent to draw from the fullest and most descriptive one as this is one of the places from which to elicit index key terms for the 653 fields. The 653 terms are used in the Index to alphabetically order the entries, so that's how you search for conferences so it's vital to get the best key terms you can find.

These index terms can also be drawn from the CNF field and the 710 which is where conference organisers are recorded. Failing any meaningful terms in any of these fields, at a push they can be drawn from the 945 field; this is the DSC shelving information which records information that identifies the specific journal and issue for shelving and retrieval purposes. Where an editor is given we can check these against the authority files but where none exists we put them in the 700. This all means that when a record is saved, Aleph always gives us a gentle rebuke that there is no BNB number and perhaps we'd like to consider authority control, and these are both over-ridden.

In addition to these fields we also add a 773 and LKR (linker) field which are holdings and items information needed to fulfil Aleph and MARC21 system requirements.

We use MARC Report to check that we've dotted each "i" and crossed each "t" but these records never receive their full approbation due the fact there are no 650s and the 260 field never has a \$a Place of publication.

Monographs and monographs in series with a UK imprint are catalogued to a higher level by the professional cataloguers and so have LCSH and DDC. The support team index monographs in series with no UK imprint; this can follow the same structure as outlined above but where higher level records can be found from Library of Congress or OCLC records we can derive records and therefore produce catalogue records at a higher level (our analytics being level 7), expanded to include the DSC conference fields. We can then run these through MARC Report and it's the only time I get a full thumbs-up because the subject work has been done.

One of the enjoyable aspects of this work is the huge variety of material I get to see, from journals on dance therapy, environmental engineering, ceramics and biomedicine to hospitality management, naval safety, aquaculture and toxicology. For someone like me with an acquisitive, magpie-like approach to information consumption it's a great job; I get to glance through the cutting edge of every subject under the sun. I've indexed conferences on piracy, midwifery and differential calculus in the space of a morning!

Material that isn't in scope for inclusion are expos and exhibitions but also conference output lacking in content. This covers the glossy spreads of photographs of people at awards ceremonies at a conference or sketchy reports of how everyone had a great time. Having seen so many examples of these I have come to the conclusion that it is impossible to take an interesting photograph at a conference, especially from a session: backs of heads and somebody at the front giving a PowerPoint presentation. Having said that, my favourite group shot is of conference attendees at a food hygiene event: a school-photograph arrangement of a hundred people wearing identical white coats, Wellingtons and blue plastic disposable caps – even their own mothers wouldn't have recognised them!

And finally, my take-home advice for anyone wondering what career path to follow is inferred from seeing which organisations hold conferences in which locations: if it's money you're after, become a dentist in the USA. They only ever hold conferences in Hawaii or Florida. Clearly business is good.

For more information:

British Library: The Conference Collections

<http://www.bl.uk/reshelp/atyourdesk/docsupply/collection/confs/index.html>

Zetoc <http://zetoc.mimas.ac.uk/about.html#inside>

A Librarian Considers Indexing

Cathy Broad

Librarian, Humanist Reference Library and Archives

Last September I attended the annual conference of the Society for Editors and Proofreaders, during which I took part in a workshop on indexing. I am a qualified proofreader, and I am hoping that this talent will open up a new career that will be useful in supplementing my paltry state pension, whenever I get it (they keep changing the dates) and my even paltrier work pension. I am considering taking up indexing to add another string to my bow, so I thought that this would be a good chance to find out what is involved and whether it would suit me.

The main advantage of both proofreading and indexing is that the work can be done from home. No more tube travel, crammed in like sardines, with your face in someone's smelly armpit and being forced to listen to someone else's questionable taste in music. Or, if you commute by train, no more putting up with morons who insist on shouting into their mobile phones even though they are in a so-called Quiet Coach. Working from home has the advantages that you can avoid these annoyances, you can set your own timetable, work as much or as little as you want or need to, take breaks whenever you want to, and you get to make a living from reading books. What's not to like?

Moreover, I thought that with my library qualifications and experience, and being the sort of person I am, indexing would be the sort of career in which I could flourish. I am very careful and precise, and I have an eye for detail. Other people may call me pedantic if they wish; I prefer to think of it as an unwillingness to tolerate stupid mistakes.

The workshop was led by Ann Kingdom, Chair of the Society of Indexers, with assistance from experienced indexer Liza Furnival, who demonstrated one of the three most commonly-used indexing software programmes. Unfortunately, time constraints made it impossible to look at all three or to go into much detail about the one that we did see.

So, what exactly is an index? And what is not an index?

Ann began the workshop by explaining what an index is, and perhaps more importantly, what it is not. A good index is not a mere concordance of words that appear in the text. A good index lists concepts, as well as names. A full-text search ability is increasingly often being used to replace the index in e-books, but this is not the same as an index, however much publishers might like to try to persuade us that it is an acceptable substitute. An index picks out the important occurrences of a word, not every occurrence, which is what the full-text search does.

Ann also explained why indexes are important, why professional indexers should be used, and how indexing is done. She explained the terminology and provided examples of how things are carried out. We worked through a number of exercises, which gave us the opportunity to evaluate both good and bad indexes – and there were some really bad ones. We learned about matters such as sub-headings (which are used to divide up subjects, rather than having long strings of page numbers under just one heading); cross-references (which are used to provide alternative terms, just like See and See also references in libraries) and introductory notes (which are given at the beginning of the index to explain matters such as the use of italics or bold type for illustrations).

Indexing software can take on the routine work of alphabetizing and sorting but an indexer is still needed for the fine detail. A human brain is required in order to make judgements about the important occurrences of a concept, and to differentiate between the senses of a word, something that a computer is not yet capable of doing.

Automated indexing has its uses but it also has limitations. It cannot differentiate between homographs, nor between homonyms. It cannot tell which of several meanings apply to words that are spelled the same but have different meanings and sometimes different pronunciations. It cannot tell the reader whether 'flake' refers to a happy accident, a worm, the ends of an anchor or the tail fin of a whale. It cannot tell the difference between 'entrance' (meaning a doorway) or 'entrance' (meaning to fascinate). It cannot tell you whether 'book' is used as a noun or a verb.

In addition, automated indexing is not capable of making a judgement about the importance of a concept – whether a word mentioned in a paragraph is just another occurrence of the word (in which case it need not be indexed), or whether the whole paragraph is important enough to justify an entry. It cannot tell you whether the word is used in the correct context. It might be another occurrence of the right word, but if the word is used in the wrong context, then it should not be indexed. Automated indexing also cannot realise that an important concept might be the subject of a paragraph without that concept being mentioned by name. Nor can it make the connection between the concept and a synonym. It takes a human intellect to make that sort of connection.

Authors often index their own books, but sometimes they are too close to the subject to be able to see it from a reader's point of view. Unfortunately, they often have to do it for financial reasons – the cost of an index is frequently paid for from the author's fee, and it is common to find a clause in the author's contract making them responsible for the index. Understandably, many authors do not want to go to the extra expense of hiring a professional indexer and so they do it themselves, seldom with as good a result as would be the outcome if they had hired a professional indexer to do the job.

How do they do it?

Next, Ann spoke about the practicalities of indexing. Obviously, the indexer needs to know the subject. You should not index outside your area of expertise. This is a very important point to bear in mind if you are considering indexing as a career. Know your limitations. Don't agree to index a book on nuclear physics if your last contact with science was in secondary school. If you are familiar with the subject, you will already know the most likely terms that a reader will use to access the information. Knowing the subject will make it easier and quicker for you to do the job.

On a practical level, the indexer needs to know how long the book is, how much space is available for the index, the audience the book is aimed at, and the house style of the publisher. This information is provided by the publisher, and the indexer needs to bear it in mind at all times. Not doing so can lead to a lot of wasted time and effort.

The indexer has to read the book carefully, either on paper or on screen. You have to put yourself in the reader's shoes, and think of what the reader will want to search for. In this regard, indexing is very similar to the practice in cataloguing of allocating subject headings and keywords, as in doing those tasks, a cataloguer is always mindful of what terms the reader will use as access points. You need to analyze each paragraph, decide what it is about, what is important in it, and what will be of interest to the reader. The intention is to make it as easy as possible for the reader to quickly locate the information they want by using as access terms the words that the reader will use, and arranging them in appropriate headings and sub-headings. The intended readership and the space available will dictate how detailed the indexing terms will be. Indexers usually put in a lot of terms and then have to edit the index to make it fit into the allowed number of pages.

Once the index is compiled, it needs to be checked. You have to make sure that there are no spelling mistakes, that capitalization, punctuation and alphabetical arrangement are all correct, and that all the sub-headings are sufficiently specific to enable the reader to quickly locate relevant information. You also need to make sure that all your cross-references match and don't just lead the reader in a circle.

Indexing software

Then we had Liza's demonstration of an indexing software system. The big three of indexing software are Cindex, Macrex and Sky Index. Like library management systems, they all do pretty much the same thing, and most indexers choose one system and stick with it. They have more in common than they have differences, and all of them seem to be in a state of constant improvement, all vying to be top dog. It seems to be the case that you pay your money and you make your choice.

The indexer works through the text page by page, inputting entries and page numbers. The system remembers terms, so there is minimum typing after the first time a term is input, and terms are automatically slotted into alphabetical order. You can set parameters for style and layout. For example, you can set up the system so that it alphabetizes word-by-word or letter-by-letter, whichever way is dictated by the house style of the publisher. You need to be very accurate, as a typographical error will result in a new term. A good indexer should be able to deal with around ten pages of text an hour.

There is an increasing use of computer files with references to paragraph and line rather than page numbers, with a code for the location, which can then be translated into a page reference when the final proof is done. This system is increasingly used for e-books.

Summing up

The workshop also included information on how to become an indexer, and Ann gave us lots of follow-up information, which provided details of training in indexing, the conventions used in indexing and evaluating the index.

All in all, I would say that if you need or want to find a second career, if you are happy to work from home and do not need the social stimulation of the workplace, and if you are prepared to spend a lot of your time marketing your services, then indexing might be something you could consider. On the other hand, if working on your own leaves you feeling lonely and out of touch with the world, or if you are even slightly diffident about blowing your own trumpet and chasing work, then perhaps you should try something else. As with any type of freelancing work, it takes a lot of effort to keep your profile before the gaze of those whom you are hoping will provide you with work. If you're not prepared to spend about 30% of your time chasing down those jobs, then maybe you should think of another alternative career. And don't forget that, as with any kind of freelancing work, you are responsible for all the financial nasties that, as an employee, are the responsibility of your employer. So you must be prepared to deal with tax matters, VAT, National Insurance, and all the other facets of self-employment. If you can't commit to that, then best stick with libraries.

Book reviews

eBooks in Academic Libraries

Ksenija Mincic-Obridovic

Oxford : Chandos Publishing Ltd, 2010 ISBN 9781843345862

Reviewed by Helen Garner, Sheffield Hallam University

This book surveys both the development and future of eBooks in academic libraries. The author uses her experiences as Cataloguing Department manager at the University of Auckland to provide insights into both benefits and issues associated with providing access to eBook collections and monographs. Throughout the book, she provides extensive research and evidence of eBook publishing practice and user surveys. It is very readable, has an international perspective and covers a wide range of examples. It has some interesting sections covering cataloguing practice in this area. It is worthwhile reading for anyone working with eBooks.

In the introductory chapter, the author surveys trends in eBook publishing and usage since 1998 and describes the problems of maintaining metadata and providing access to eBook content. She describes how scholarly eBook publishing has developed to meet the needs of academic libraries. In chapter two she explores the reasons why momentum in eBook publishing has been slower than in journal publishing. She examines the range of eBook monograph types that exist: from digitized versions of print texts to database type structures with full-text search functionality. She outlines the importance of understanding the many different types of eBooks, in terms of functionality, pricing models and resource discovery options.

Chapter three discusses purchasing models and the availability of eBook titles. It examines the impact and reaction to Google Books by the publishing industry. Developments in free eBook content are examined. Government eformat, Project Gutenberg, Internet Archive, Hathi Trust and content from open access repositories are all discussed. The impact of eBook readers in 2009, thus making eBook format mainstream, is also covered.

Chapter four provides a really good overview of managing and developing eBook collections. It considers how well ILS type systems facilitate the administration of eBooks. It looks at the range of subscription options available to libraries: purchase, subscription, patron driven, pay per use and lease to own. Issues associated with platforms and interface, viewing options, printing limits and authentication processes are discussed. This chapter offers a good discussion of the problems faced by cataloguers when dealing with eBook metadata and discusses the approaches and standards.

Chapter five offers an extensive review of research into usage of eBooks and surveys of eBook users. It compares data from five key studies: the University of Denver study in 2005, the SuperBook project funded by Wiley, Emerald and CIBER in 2006, a Springer survey of eBook usage in 2008, an ebrary survey in 2008 and the JISC national eBook observatory project in 2009. Analysis from these surveys is really interesting and useful.

Chapter six looks at future developments and the potential of eBooks to become interactive learning tools. It describes a number of multimedia non-linear texts which have been designed to facilitate online learning. The author concludes that libraries have an important role to play in maximising the potential of eBooks.

Book reviews

Metadata for digital collections: a how-to-do-it manual

Steven J. Miller

London: Facet Publishing, 2011 ISBN: 9781856047715

Reviewed by Alice Laird, Senior Information Executive, ICAEW

This book gives the reader exactly what it promises - a practical guide to metadata for digital collections - what it is; and how it can be used to provide intellectual access to the ever-increasing number of digital collections being created by libraries, museums and other organisations. I suggest that the book is of value to anyone with an interest, at any level, in the practical application of metadata, but it is particularly aimed at metadata practitioners; at users of out-of-the-box digital collection software packages; and at Library and Information students and instructors as an introduction to the world of metadata practice.

The book is organised into chapters that are designed to build up the reader's knowledge and understanding, starting with an introduction to the basic concepts and functions of metadata. The author then introduces the reader to Dublin Core (DC) and resource description, followed by chapters looking at DC elements and examples in greater detail. He then examines controlled vocabularies, XML basics, MODS (the Library of Congress's Metadata Object Description Schema) and the VRA (Visual Resources Association) core categories. Once he has established a solid foundation of understanding, the author goes on to examine metadata interoperability issues and provides very practical advice on designing and documenting a metadata scheme and introduces the emerging environment of linked data and the semantic web. The book is illustrated throughout with clear examples. The companion website <http://www.neal-schuman.com/metadata-digital-collections> provides access to a range of resources to support the reader.

I read this book from cover to cover. It took me several reading sessions as there is much to absorb but each chapter held my interest and taught me something relevant. I am currently involved in researching a metadata project for my library and the practical nature of the book provided me with an approach and framework for my own project. Even the chapters that were not directly relevant for my work were helpful and the final chapter, on linked data and the semantic web, included the clearest explanation I have found to date on this subject. I really liked the end of chapter summaries and bibliographies - the summaries helped consolidate learning from each chapter, and the suggestions for further reading were practical and inspiring. The book does have a slight US-bias but I don't think this detracts too much from its usefulness. I thoroughly recommend it.

Book reviews

Facilitating access to the web of data: a guide for librarians

David Stuart

London: Facet Publishing, 2011 ISBN: 9781856047456

Reviewed by Alison Felstead, Head of Resource Description, Bodleian Li-

David Stuart defines the web of data as “structured data in a machine-readable format that is being made publicly available online by individuals and organizations from every sector of society.” His book provides a very useful introduction to this subject and its associated concepts and terminology, including open data, the semantic web, data silos and embedded semantics. The level is well-judged, and Stuart manages to make the more technical aspects of his material understandable to his readers by the use of examples to illustrate the concepts. There is an extensive bibliography reflecting the wide-ranging examples included in the book.

Unfortunately, the book is let down by careless editing, resulting in too many typographical errors, and poor grammar (“Every data its user,” an updating of Ranganathan’s law “Every book its reader,” being the worst offender). The text is sprinkled with acronyms and initialisms, and whilst many of these are explained when they first appear, some are not. One example – URIs – is defined when it first occurs as “Uniform Resource Identifiers” but later on as “Universal Resource Indicators.” The URL for the Many Eyes website varies, and the inclusion of these and many other URLs in the body of the text tends to break up its flow.

I would like to see subsequent editions issued with the errors corrected, and the addition of appendices listing initialisms and their definitions (for ease of reference) and web addresses. Many of the screen shots used to illustrate specific examples are small and poorly-reproduced, and the index hides some terms under others (for example, “triples” and “triplestores” are found under the index entry “RDF”, which assumes that the reader already knows that these are concepts related to the Resource Description Framework).

Stuart is passionate in advocating that librarians should engage with the web of data, in order to ensure that they do not become obsolete, and also to help their users find and access the information they require wherever it may be found. He concludes the book by presenting nine practical steps that the reader can take to become a “data librarian,” and it was reassuring to see that only the last of these involves becoming a computer programmer. I finished the book feeling less like a fly and more like a spider, inspired to explore the web of data and possibly to try a little spinning of my own.

NOTICEBOARD



The CIG conference will be held on 10th-11th September 2012 at the University of Sheffield. This will be followed by a half-day introductory workshop on FRBR on 12th September. A call for papers and more details will follow in the new year. If there is anything that you would like to see included in the programme please do not hesitate to contact any member of the CIG committee (<http://www.cilip.org.uk/get-involved/special-interest-groups/cataloguing-indexing/about-us/pages/committee.aspx>).



**Visit to Foyle Special Collections Library and Maughan Library,
King's College London, Tuesday 14th February 2012, 2pm - 4pm**

The Foyle Special Collections Library houses some 170,000 printed works, as well as maps, slides, sound recordings and manuscript material. Its collections, built up over centuries by purchase, gift and bequest, contain incunabula, many unique items and cover all subject areas. You will be able to view a selection of books from these collections, some items from the FCO Historical Collection, and from the historical medical collections. These examples will be used to illustrate some practices of rare book cataloguing. The visit will also include a tour of the Maughan Library, a listed building with many interesting architectural features and used to house the Public Record Office. Places are limited to a maximum of 20 people. To register interest or to book a place, please email Katrina Clifford on k.clifford@kingston.ac.uk.



Catalogue & Index is electronically published by the Cataloguing and Indexing Group of the Chartered Institute of Library and Information Professionals (CILIP) (Charity No. 313014)

Subscription rates: free to members of CIG; GBP 15.00 for non-members & institutions

Advertising rates: GBP 70.00 full-page; GBP 40.00 half-page. Prices quoted without VAT.

Submissions: In the first instance, please contact the Editors: Cathy Broad (library@ethicalsoc.org.uk) and Heather Jardine (heatherjardine402@hotmail.com)

For book reviews, please contact the Book Reviews Editor: Neil T. Nicholson, Cataloguing & Metadata Services Team Leader, National Library of Scotland, e: n.nicholson@nls.uk

ISSN 0008-7629

CIG website: <http://www.cilip.org.uk/specialinterestgroups/bysubject/cataloguingindexing>

CIG blog: <http://communities.cilip.org.uk/blogs/catalogueandindex/default.aspx>

Tags from the CIG blog on specific areas of interest:

**[authority control](#) [book reviews](#) [Catalogue and Index](#) [cataloguing](#) [CIG activities](#)
[CIGS](#) [classification](#) [committees and working groups](#) [conferences](#)
[Dewey](#) [digitised material](#) [Dublin Core](#) [events](#) [folksonomies](#) [linkblog](#) [MARC](#)
[metadata](#) [news](#) [RDA](#) [Semantic Web](#) [social software](#)
[standards](#) [taxonomies](#) [UDC](#)**