

Serials indexing: from journals to databases

Caroline Barlow

The user pool for these publications is vast, with access varying from professional-mediated searching to search-engine keyword searching by individuals at their home PC – just one consideration to be borne in mind during the indexing process, a process which differs in many aspects from back-of-the-book indexing. For example, terminology that develops over time needs to be cross-referenced in detail. Locators have to include year or volume number. Indexing of journal articles for databases requires the use of controlled vocabularies and thesauri. Newspaper indexes may include reference to columns, individual sections and different editions. And the cumulation of serials indexes needs careful management if consistency is to be maintained, something which may militate against interim indexing through the year, although this does have advantages. Negotiation of staged payments is acceptable. An experienced serials indexer helps you through the maze.

Introduction

Indexing of continuing publications has a long history. For example, the Inspec database (<http://www.theiet.org/publishing/inspec>) is going to celebrate its 40th anniversary in 2009, and the Institution of Electrical Engineers (now IET) first published *Science Abstracts* in 1898 and *Physics Abstracts* from 1903. In July 2008, Inspec held a presentation for the 10-millionth item to be added to the database (10 gigabytes of data in total). This item was in the field of nanotechnology. To give an idea of the scale of such an undertaking, Inspec expected to abstract and index 632,000 items in 2008 from 4,400 journals (1,600 of which are indexed cover-to-cover), including *The Indexer*, and 2,200 conference proceedings, as well as numerous books, reports and dissertations, approximately 60 per cent of the material being supplied in electronic format.

This database is only one of a number of international secondary abstracting and indexing services in many subject areas; other examples include MEDLINE (<http://medline.cos.com>) and Emerald Abstracts (<http://info.emeraldinsight.com/abstracts/index.htm>). This latter company and many smaller services are based in the UK.

From this we can deduce that there is an army of indexers who work both in-house and as self-employed contractors to keep these databases current and consistent. CAS (Chemical Abstracts Service, <http://www.cas.org>) regularly advertises for people to analyse, abstract and index the world's chemical and life-science literature, including the evaluation of chemical structures. Their specifications require candidates to have BS, MS or PhD qualifications in chemistry or one of a number of chemistry-related subjects. 'The ability to read and translate Japanese, Korean, Spanish, French, Italian, Polish, Portuguese, German, Russian, or other Slavic languages is highly desirable.' My own hesitant French gained me employment in the 1960s while I was living in Africa.

Where are all these indexers, you might ask? As they are generally tied to a specific company, they are trained internally to very exacting standards and probably do not consider indexing outside these jobs. Consequently, members of our indexing societies rarely meet them. So the indexing fraternity

is larger and more disparate than expected, and also includes library cataloguers and indexers. It is from this field of serials indexing that many developments in indexing and indexing technology emanate.

Journal, rather than book indexing

When deciding to extend your indexing portfolio to include serials indexing you need to consider the ways in which this will differ from back-of-the-book indexing. The main differences stem from the continuous nature of the publication and also of the index. This type of indexing requires tighter control of terminology, more cross-referencing and careful consideration of the best form of locator to use. Looking at previous indexes or experimenting with the indexing of journals and magazines that you receive yourself can enable you to gain an insight into this. The differences will soon become apparent. What follows is a quick run through these differences, pointing out any pitfalls that may arise.

Users

When indexing anything, a consideration of the likely users is essential. For serials, the user pool is extensive and can include researchers, students, writers, journalists, secondary publishers, historians, genealogists, librarians, information professionals and other indexers, as well as the subscribers and editorial staff themselves. As there are many types of user, the indexing can cover items not always perceived as index-worthy, such as advertisements, small news items and obituaries, which are particularly relevant in newspaper indexes.

Nature of serials

The main advantage of this type of indexing work is the fact that serials go on and on, and can be a source of regular work. This may not equate to a consistent monthly cheque, but could mean that some work comes every week, which can form the basis of an indexer's portfolio. However, with the advent of electronic publishing, many journals editors

are dispensing with indexes and indexers in favour of keyword searching on their websites. As we know, this is not a reliable or recommended way forward and many indexers are working to reverse this trend. The problems associated with keyword searches are outlined in the next section.

Subject headings and keywords

Often in journals the authors of the articles will prepare abstracts and append keywords to aid the indexer. Inevitably a large number of authors will produce many variants of keywords relating to the same subject, and there may also be spelling variations. (The editor may also recommend particular keywords to be used.) These keywords are often used to prepare an issue index, but are rarely used by secondary publishers, although they do use and edit the abstracts. Here is an example of a keyword index from one issue of a wastewater treatment journal.

activated sludge 1, 11, 29, 39, 49, 67
 activated sludge process 95
 aerobic 145, 245
 aerobic degradation 195
 bacteria 39
 effluent 229
 effluent minimization 183
 effluent treatment 49
 filamentous bacteria 29
 filaments 39
 heavy metals 239
 heavy metals in soil 239
 kraft mill effluent 103, 123
 kraft mills 49
 kraft wastewater 145

It is immediately obvious that this list of keywords has not been edited in any way and would certainly benefit from an indexer's guidance. Sadly, it is not atypical.

The main way to overcome the shortcomings of keyword lists compiled by more than one person is to use a thesaurus, controlled vocabulary, authority file or alphabetical list of preferred terms and their relationships. These can be compiled for use by specific journals and databases, or one of the many thesauri already available in a wide variety of subjects can be used. A good source of reference of these is the Taxonomy Warehouse, where there are thesauri for subjects from aerospace to zoology (<http://www.taxonomywarehouse.com>). Such lists are also a great help in developing indexing skills in any specialism. Terms are added to thesauri only after considerable debate and review of surrounding terms.

As an example of a thesaurus entry, the term 'indexing' might appear as follows (NT is narrower term; BT broader term; TT top term; RT related term (or cross-reference)).

indexing

NT book indexing
 journal indexing
 BT library and information science
 TT computer applications
 RT thesauri

These terms can also be shown in a hierarchical display:

computer applications – cont.

- library and information science
- • indexing
- • • book indexing
- • • journal indexing

Cross-references

Consistency becomes even more important in the indexing process when considering serial publications because terminology evolves over time and new concepts are introduced. For example, aged people became elderly people and are now older people, and there are many other terms connected with disability, race and gender that need careful consideration over periods of decades and centuries. Watergate and 9/11 are also terms that have evolved to acceptance, though they might have appeared under different terms initially. All these need to be cross-referenced in a continuing index.

Groundbreaking research is often recorded in the letters pages of relevant journals and the terminology might not be the same as that used in preliminary research or even the term settled on years afterwards. An example is the discovery of the structure of a third allotrope of carbon by Harry Kroto and his teams at the University of Sussex, Brighton, UK, and Rice University, Houston, Texas, USA. The allotrope was named buckminsterfullerene after the architect of the Expo '67 geodesic dome structure, a closed hemispherical shape of hexagons and pentagons, which could be a possible structure for C₆₀, with a carbon atom at each vertex (Kroto et al, 1985). The alternative name of 'soccerene' was suggested, but later rejected, named after the first 'official' FIFA football (of similar construction) used in the 1970 World Cup in Mexico, some 15 years before the discovery of the C₆₀ structure.

Locators

There are more facets to include in locators for serials than in stand-alone documents like books. Indications of year, volume, issue, as well as page can be represented using bold, italic, underlining and a variety of punctuation marks. Examples are **08**:1.3–5 (*Sidelights*) and **25**.154–60 (*The Indexer*).

It is perhaps more important in serials indexes to give exact page ranges so that the whole of an article can be reproduced, without adverts and extraneous material, when ordering from a library service. For newspaper indexes it can be useful to divide the page into columns, or quadrants, and indicate this with letters or numerals, and also indicate the length of coverage by S for small, less than 6 inches, M medium 6–18 inches, L large over 18 inches. Thus, a locator 7/7/06 M3:2 indicates the date, length, column and page of the item.

Regular sections of a publication may need to be indicated by abbreviations, for example, book reviews (BR), clinical papers (CP), letters (L), practice (PR), resources (R), or these items could be included under these headings as subheadings, for example:

book reviews

Evidence-based therapeutic support
Immunisation against infectious disease
Safeguarding children and young people

A similar system is used in the index to *Sidelights*.

Various problems occur when the pagination starts at page 1 for each issue of the journal for a particular year, which means that the issue number has to be included in the locator, as in *Sidelights*. Also, there may be unnumbered pages, special issues that do not follow the volume and issue numbering of the rest of the year, detachable sections which can be pulled out and kept, and inserts that are not attached. Abbreviations could also be included for the different sections of a newspaper which are not part of the main newspaper pagination, or indeed for the different editions of the same newspaper. Pictures, graphs, tables, maps and the like can be indicated in the usual way by using abbreviated suffixes, such as fig., illus., t and m. All these need to be tested in your indexing software to check that they are easy to use and are consistent.

What to index

Correspondence can be very important, as are corrections made to previous papers in earlier issues. These all need to be related to the items to which they refer and which have been indexed in previous issues.

Inevitably there are parts of the journal that you are asked not to index, and these can include advertisements, book reviews, committee minutes, contact addresses, diary dates, editorials, flyers, gossip columns, letters, photographs, publication announcements and puzzles. Yet some of these would be of great interest to future researchers.

This point is particularly evident in the indexing of newspapers. Interests for researchers could range from the age, occupation, place of work and residence of people involved in news stories to buildings' dates; crime reports; editorials covering current stories; election results; entertainment; first and last stories, such as the last trolley bus; letters to the editor; local sports events; natural disasters; obituaries; railways and roads; social trends; visits of celebrities and even the weather. Photographs and advertisements can also be of interest.

NEWSPLAN is the cooperative programme for the microfilming and preservation of local newspapers and for making them accessible to users. It involves public, academic and national libraries, archives and the newspaper industry. It has received a Heritage Lottery Fund award for its work. The British Library Newspaper Library plays a leading role in the support of NEWSPLAN, which is based on regions corresponding to the Regional Library Systems of the UK and Ireland (<http://www.bl.uk/aboutus/acrossuk/worknat/news/index.html>).

Introductory notes

Introductory notes are essential in the indexing of serials. It is particularly important to note the locator system, any

department or regular column abbreviations, unusual punctuation and the exact meanings of any italic, bold or underlined text. Some good examples of introductory notes are given in Pat Booth's *Indexing* (2001: 177–9).

Separate indexes

It is not uncommon to provide indexes for authors, subjects, cases, statutes and statutory instruments, but for journals the editors seem particularly prone to ask for separate indexes for such things as abbreviations, books reviewed, chemical names, companies, trade names, workpiece materials, processes, and even miscellaneous, which catches everything else. However, for the users this entails a lot of extra work and unless they are dedicated subscribers to the particular journal it is likely that they will give up before finding the information they are seeking.

It is always wise to check your layouts with your indexing software before you start indexing so that you are able to make best use of any short cuts provided to aid what can be a convoluted and repetitive process.

Author indexes

The usual problems occur when indexing many authors. Although house styles may ask for initials only, it is better to give full names to differentiate as much as possible, though names, their order and the number of initials can vary greatly for individual authors. Particular attention has to be paid to spelling and accents as you would expect, but you are also dependent upon the editing of the journal itself for accuracy, unlike subject indexing when you can more easily spot inconsistencies. It is very important to the authors that they are credited with the work and there is much debate about whether to include the first author only, the first five authors, or indeed all of the authors. The order in which they are listed is not necessarily either alphabetical or according to seniority. In-house reporters and journalists may or may not require indexing, and in all this the indexer has to be guided by the editor.

The way in which authors are indexed need not be just a long list of names, but can be associated with the titles of their work. However, this can lead to a massive index as the following example shows.

COOREUITS, P., LADEVÈZE, P. AND PELLE, J.-P., Mesh optimization for problems with steep gradients, No.2, 1994, pp. 129–44.

PELLE, J.-P, see COOREUITS, P.

PELLE, J.-P, see LADEVÈZE, P.

LADEVÈZE, P., see COOREUITS, P.

LADEVÈZE, P., PELLE, J.-P and ROUGEOT, P., Error estimation and mesh optimization for classical finite elements, No.1, 1991, pp. 69–80.

ROUGEOT, P., see LADEVÈZE, P.

From *Engineering Computations: International Journal for Computer-Aided Engineering and Software*. Cumulative index Volumes 8–12, 1991–5 (MCB University Press)

Company indexes

In some business areas the names of companies, their divisions and location seem to change quite frequently and the tracking of mergers and acquisitions is difficult, but essential, in serials indexing.

Rhodia
 Albright & Wilson bid
 independence from Rhône-Poulenc
 Rhodia Food
 Rhodia Organique Fine
 Rhône-Poulenc
 Rhodia independence
 Rhône-Poulenc Agro
 Rhône-Poulenc Yuka Agro
 Sanofi
 Synthélabo proposed merger
 Sanofi Santé Nutrition Animale
 Sanofi – Synthélabo
 Synthélabo
 Sanofi proposed merger
 From *Chemistry & Industry* (1999)

Subject indexes

The aim of the serials indexer is always to produce a unique record for a particular article so that the searcher can identify a specific piece of information. As many journals are very subject-specific this can be a real challenge. The use of qualifiers, leading to subheadings, sub-subheadings and sub-sub-subheadings can be useful, but would need editing when the extent of the index becomes apparent. The alternative is to be left with long strings of locators which would entail going back to check the relevant articles, which is both time-consuming and uneconomic. This is an important point to consider in the case of cumulating indexes, which is treated in the following section.

Some organizations provide a template or form to fill in and this is especially useful in database indexing. A style sheet can be developed for individual journals and there are checklists available for guidance in *Last but not least: a guide for editors commissioning journal indexes* (SI, 2004: 4–5).

In addition to the more conventional subject index, there can be many variations.

Example 1: The terms are taken from a keyword list and used in rotation for all terms chosen.

Mechanical properties
 Aggregate; Hybridisation; Quantitative petrography;
 Rapakivi granite; Subvolcanic 197
 Quantitative petrography
 Aggregate; Hybridisation; Mechanical properties;
 Rapakivi granite; Subvolcanic 197
 Rapakivi granite
 Aggregate; Hybridisation; Mechanical properties;
 Quantitative petrography; Subvolcanic 197
 etc.

Example 2: The terms are freely chosen and the article title and reference form the subheading.

INDUSTRIAL RELATIONS

Arbitration outcomes in the service sector; an empirical assessment (P.K. Mills and D.R. Dalton) 5(2) 1994: 57–71

INDUSTRY see MANUFACTURING INDUSTRY

INFORMATION EXCHANGE (see also COMMUNICATIONS)

Intrinsic service quality determinants for pharmacy customers (M.-B. Hedvall and M. Paltschik) 2(2) 1991: 38–48

From *International Journal of Service Industry Management*

Example 3: There are several subject indexes with predetermined headings using article titles as subheadings in chronological order. This method is nearer to classification than to indexing.

SUBJECT INDEX – WORKPIECE MATERIALS

Aluminium

see Non-Ferrous Metals

Ceramics, Glass Ceramics

Honing of oxide ceramics 10

Ultrasonic machining of silicon nitride 120

Non-Ferrous Metals

Characterisation of micromachined surfaces by atomic force microscopy 59

Machining commutators 64

Silicon Nitride

see Ceramics, Glass Ceramics

SUBJECT INDEX – PROCESSES

Boring

see Turning, Boring & Milling

Milling

see Turning, Boring & Milling

Turning, Boring & Milling

Characterisation of ABN800 6

Super-abrasive reaming of bores in cast iron 52

Characterisation of micromachined surfaces by atomic force microscopy 59

Machining commutators 64

A micro machine tool for precision machining 105

SUBJECT INDEX – MISCELLANEOUS

Conferences, Seminars, Exhibitions, Visits, Training, Anniversaries

A showcase for natural stone 39

International Surface Colloquium 63

Euro PM96 111

Industry

Transatlantic expansion by Unicorn 112

From *Industrial Diamond Review*

Cumulation

There are two main ways in which you may be asked to cumulate a serials index. The easier of the two is to be given the total range of issues to index over a short period of time. Alternatively you may be asked to cumulate an index that has been indexed on an annual basis by you or someone else. This latter method can lead to long strings of locators which have to be researched and can be a very expensive exercise. It is often more economic to index again. Even if the previous index is your own, unless you have retained the pre-edited version of your index with qualifiers and multiple subheadings that you prepared to preserve the uniqueness of each entry, this can be a long job.

As mentioned in the cross-references section above, the problem of terminology changes requires careful consideration. Noting the dates when you changed the terminology becomes useful in this instance.

Retrospective indexes are sometimes prepared for special occasions. *The Builder Illustrations Index 1843–1884* (Wheatley Medal winner) only covers illustrations over 40 years in *The Builder* and it took a team of indexers ten years to prepare 840 pages including six separate indexes.

Indexing manuals

As there are so many variants involved in serials indexing it is often worthwhile to construct a basic manual or checklist for each serial to include house style, rules, exceptions, queries and their answers, changes in terminology and their dates, cross-references and double entries.

If you have used or adapted a thesaurus or similar authority file, which may be based on previous work done by other indexers, keep it in a safe place and make a note of changes you have made to these and the reasons for making the changes.

Business aspects

The work of the serials indexer tends often to be concentrated at the end of the year. A method of spreading the work could be advantageous. However, it is often better to index all the issues for the year within a short space of time, or in batches, for better continuity and consistency.

It is possible that the index will appear in the December issue, and then you will need to work from proofs before publication. Indexes can also appear in the January issue, as a detachable insert or as a separate publication.

The year could start with a subscription to the journal so that you are able to follow its development weekly, monthly or quarterly. Approval for the index, however, might not be forthcoming at the start of the year. Negotiations can be made for staged payments during the year if you think the project warrants it. For example, you might be asked to prepare monthly or six-monthly provisional indexes to aid the editorial staff, but then you need to warn them that the index might change during the final editing stage.

If you are asked to tender for serials indexing it is best to obtain, say, three issues of the publication and produce test files. This also enables the editor to see your work and test software compatibility. This is also a good time to prepare

an indexing manual for the title and perhaps suggest changes that would help your indexing process, although the editor may not always welcome these. As you are essentially tendering for a job that could continue for several years, it is worth the extra effort involved in the initial stages.

Conclusions

Serials indexing can be a challenging addition to the indexer's portfolio of work and involves many different considerations. Database indexing is particularly useful as it can provide a regular source of work. Although I have not covered specifically the transfer of serials indexes to the relevant publication website, this is often done. It may be simply a PDF of the index to a given year as prepared for the printed publication, or (as with *The Indexer*) might be a continuous cumulative version.

References

- Booth, Pat F. (2001) *Indexing: the manual of good practice*. Munich: KG Saur.
- Kroto, H. W., Heath, J. R., O'Brien, S. C., Curl, R. F. and Smalley, R. E. (1985) C₆₀:Buckminsterfullerene. *Nature* (London), **318**, 162–3, 14 November, 1985.
- Society of Indexers (SI) (2004) *Last but not least: a guide for editors commissioning journal indexes*, Sheffield: Society of Indexers.

Resources

- American Society of Indexers (ASI) <http://www.asindexing.org/site/asires.shtml>
- Beare, Geraldine (1999) *Indexing newspapers, magazines and other periodicals*. SI Occasional Paper OP4.
- Diakoff, Harry (2004) Database indexing; yesterday and today. *The Indexer* **24**(2), 85–96.
- Greenhouse, Shelley. The future of database indexing. *Key Words [ASI]* **8**(4), 124–6, 132 (<http://mysite.verizon.net/vze2bpts/ASIkwardarticle.htm>).
- The Indexer* <http://www.theindexer.org>
- NEWSPLAN Programme. <http://www.bl.uk/collections/nplan.html>
- Sidelights* Index 1999 to 2005. <http://www.indexers.org.uk/files/index.pdf>
- Society of Indexers (SI) <http://www.indexers.org.uk/>.
- Resources for indexers (SI Members' page) <http://www.indexers.org.uk/index.php?id=270>
- Taxonomy Warehouse <http://www.taxonomywarehouse.com>
- Weaver, Carolyn G. (2002) The gist of journal indexing. *Keywords* **10**(1) January/February, 16–22 (<http://www.asindexing.org/site/keypast.shtml>).

Caroline Barlow is a freelance information scientist specializing in indexing in the fields of health and safety, physical and materials sciences. She is a chemistry graduate and started her indexing career in 1973, based mainly on database indexing. Email: caroline@barlow-indexing.freereserve.co.uk