Learned Publishing (2001)**14**, 123–130

From herbals to Hotbot: the development of journal indexing

Hazel K. Bell

ABSTRACT: The development of journal indexing since the early alphabetical tables to scientific works, culminating in current sophisticated cumulations, is traced. Standards for preparing indexes which have gradually evolved to cope with the increase in volume and complexity of the literature are described.

Hazel Bell

Introduction

The earliest scientific works written partially or entirely in the vernacular were the herbals of the late 15th century. These were medico-botanical texts describing medicinal plants, their properties and healing powers, giving recipes for the preparation of remedies and drugs, and sometimes containing also descriptions of animals and minerals from which medications could be prepared. They were also most properly and thoroughly indexed (by Conrad Gessner), containing extensive name and subject indexes, often in more than one language and sometimes even in more than one script: Roman, Greek, German blackletter, and Hebrew all were used.[1]

Periodicals first appeared in Britain in the 17th century, with indexes. The *Philosophical Transactions* [of the Royal Society of London for the Improvement of Natural Knowledge], the first English-language scientific journal, began as a natural outgrowth of Henry Oldenburg's work as secretary and rapporteur of Royal Society affairs.[2]

> He had so many inquiries to answer from correspondents abroad and at home who wanted to know about the Society's work and about scientific news generally that it seemed simpler to him to print a monthly report, and by its sale to get some return for all the work he was doing for the Society. . . . The outcome was the first monthly journal to appear in Britain that was profitable enough to last more than three issues. . . . Papers read before the Society, letters, summaries and abstracts of books of special interest to the virtuosi, – the scientific news of the day was summed up in this journal.

In 1678 the *Transactions* included the earliest known British periodical index: 'A general index or alphabetical table to all . . .

from the beginning [1665] to July 1677'. A modern critic reports of it:[2]

> This index is in absolute alphabetical order. The entries are set out in 'run-on' paragraphs with much variation in the number of references given according to the specificity of the headings. 'Aches' has only one reference whereas the 'Air' references fill a column and a half. All multiple entries are arranged in chronological order by date of publication of the issues referred to; or, if several references relate to one article, in sequential order by page number. Though there are occasional internal connections and local groupings under what amount to subheadings within these blocks, no systematic subordinate rearrangement of the material, alphabetical or otherwise, appears to have been attempted at this time. Some degree of cross-referencing is evident, both of the see also type from one used heading to another, and of the see type from an unused heading to the preferred term.
>
> What strikes one, upon examining the entries grouped under 'Air', is the attempt by the writer to connect his material. The index has been written by a person familiar and at ease with the subject matter. The items being referred to in order of their publication, each sequence of entries under a particular heading is moulded into a brief, but detailed, history of the progress in that area of the Society's concerns. Guidance is offered to the reader on what should be consulted; he is advised of mistakes in the original; and occasionally reference is made to material not in the running sequence. The outcome is more essay-like than 'indexical' in the sense that would be understood today, having a sequential interdependence which suggests that this index has been written to be read. . . . The Royal Society's index was compiled at a time before the atomization of reference sequences had become mentally sustainable in other than an occasional or random fashion. The Royal Society was engaged in promoting those processes of observation, experimentation, analysis and inference that would produce such a sustained capacity, but empiricism had not yet overturned the older mental processes exemplified in Isaac Newton's theological writings. The index itself must have been an instrument in that revolution as well as one of its symptoms.

The index has been written by a person familiar and at ease with the subject matter

Over the next two centuries, we are told:[3]

> The 'back-of-the-book' index, very much as we know it today, was well established by the end of the seventeenth century when the first scholarly journals were published. Cumulative and collective indexes to periodicals came a bit later – around the late eighteenth century. By the nineteenth century, the scholarship and information explosion had so grown that retrospective indexes were being initiated. William Poole, working in the library of Yale University, noticed that many of the library back-issue periodicals were never used, and concluded that indexing them would make them more useful. He developed and published such an index, 154 pages, in 1848. It was a great success, and was succeeded by the *Index to periodical literature*, published every five years; then by a monthly *Co-operative Index to Current Numbers of Leading Periodicals*, and by the *Annual Literary Index*, the *American Library Index* and the *American Library Annual*, all produced in cooperation with volunteer members of the American Library Association. In 1901 the indexing of general periodicals became professional with the beginning of Wilson's *Readers' Guide* to periodical literature. Other early and prominent indexing services had their origins in this period; *Index Medicus* was first published in 1879 by a surgeon in the Civil War.

Card indexer at work

From 1946 to 1974, the novelist Barbara Pym helped to edit the publications of the International African Institute in London, including four annual issues of the journal *Africa*, for which she compiled the indexes. Her assistant gives us a picture of the journal indexer of the period at work:[4]

> She would sit crouched over a battered, wooden four-drawer card index of antique

design, like Miss Clothier [a character in a Pym novel], 'moving the cards here and there with her fingers, as if she were coaxing music from some delicate instrument'. Occasionally I would catch a murmur, like the recital of some litany: 'Abortion, administration, age-sets, agriculture, amulets, ancestors, animal husbandry . . .'.

She loved the mystique of certain aspects of her job; best of all, she enjoyed the art of indexing. The great ongoing indexes were the cumulative, annual indexes for Africa. The real challenge was the Index of Tribes and Languages. Variations of tribal names gave delightful opportunities for cross-referencing, and the special characters had to find their formal place in the scheme of things – the longtailed S (= sh) or the Bushman 'click' (!Kung would come after Z).

I think what she enjoyed most about indexing, apart from the pleasure of putting words into a certain order, was the peaceful, enclosed space an indexer inhabits. It requires a certain sort of concentration: you need to withdraw, as it were, into the world of that which is to be indexed.

Author-indexing

It became common practice in science and technology for authors of journal articles themselves to provide the abstracts and keywords for indexing appearing on the title pages. The Engineers Joint Council ran workshops in 1952–5 to train engineering society members in the techniques of source indexing. In 1961 the American Institute of Chemical Engineers announced a plan to accompany articles in *Chemical Engineering Progress* with 'catalog cards' containing the abstracts and bibliographic citations of articles as well as keywords. The American Institute of Physics Documentation project had stimulated participation by physics journals in aid-to-indexing projects which requested authors to fill out indexing forms upon acceptance of their manuscripts for publication.[5] By 1965 80% of authors in the six AIP journals were participating in the programme.

With the approach of the era of electronic transfer of information from authors to editors, Virgil Diodato examined 'Author and source indexing and abstracting of journal articles'. Authors' capabilities for this were debated: letters appeared in *Science* on the assignment of index terms by the authors of articles. One reader advocated, 'Since the greatest authority on any item of literature is the author, is he not the one best able to classify the item properly?'[6] Another counterclaimed, 'Authors are seldom qualified to do indexing, because of their non-objectivity and inability to see how their work fits in with others in the field.'[7]

letters appeared in Science *on the assignment of index terms by the authors of articles*

Standards and criteria

In 1960 the Society of Indexers issued a set of 12 Standards for indexes to learned and scientific periodicals, distributed in leaflet form to editors and publishers.[8] They specified:

- The index to a journal 'should be issued in or with the final part of the periodical, failing which it should be published as soon as possible after the publication of the final issue';
- 'a dictionary index combining subject, author and title entries may be preferable to separate subject and author indexes';
- 'articles should be indexed under authors, titles, significant words in titles, and under the subjects dealt with in the text';
- 'the names of all co-authors should be printed in the index, cross-references only being provided if space is limited';
- and included ways of dealing with abbreviations, treatment of special features, bibliographies, and cumulation.

Ferriday, surveying periodical indexes in 1963, found the index to *Nature* 'distressing', showing a 'manic thoroughness': a single book review received seven separate entries in the index.[9] 'The inversions and convolutions of the title are marvellous and masterpieces of typography . . . the system of subject indexing by title-words is, if carried out by the *Nature* pattern, absurdly wasteful and can be quite ineffective.' Even worse was deemed the index to the *Journal* of the Royal Society of Arts, for which 'the whole

basis of the index is to accept the first key word only as the point of entry . . . the user is expected to remember the title of the article he is looking for'. Ferriday reserves praise for the *British Technology Index*, 'where the subject headings are as specific as possible, and its techniques therefore more relevant to the problems of indexing single periodicals'.

In 1964 the British Standards Institution (BSI) published its first *Recommendations for the Preparation of Indexes for Books, Periodicals and Other Publications*. It preferred a single index to separate name and subject indexes, and placing the index at the beginning of the volume of the periodical indexed. A survey of the subject indexes to 30 periodicals in the field of science and technology conducted in 1965 noted widely varying patterns of practices and principles in these.[10]

It preferred a single index to separate name and subject indexes

Information masses

In 1965 J. Edwin Holmstrom described the information retrieval problems caused by what we came to call the information explosion, and the concomitant necessity for proper indexing:[11]

> A vast apparatus of abstracting and bibliographical publications has grown up for the purpose of supplying scientists and technologists with at any rate an indistinct awareness of what has been published in the fields that concern them, and of making it possible for the original sources of the items that have the relevance they want to be 'retrieved' . . . To assist this process it is important that the indexes to particular books should cover not only items in those books themselves but also bibliographical references cited incidentally in them.

By 1963 'the problem of communication in science, pure and applied' was recognized as having arisen:[12]

> partly because of the volume of scientific literature . . . and because of the steady melting away of the hard-and-fast boundary walls that formerly separated one scientific or technical speciality from another . . . we are approaching the situation in which a really up-to-the-moment awareness in one's special field calls for a simultaneous knowledge of what is going on in several other fields. . . . There is therefore a great need for indexes and other information-processing compilations that will enable pure and applied scientists to survey not only what is going on in their own field but in marginal fields as well.

Journal indexes began to be cumulated not just for a single periodical, but collated over particular disciplines. The *British Technology Index* was founded by the Library Association in 1961 as a partial solution to such problems, covering 400 British technical journals to produce a monthly subject-index to 2,000–3,000 articles per issue, published within the preceding seven weeks, with annual cumulations.[13] It aimed to provide access to technical papers in over 300 British journals.

Textbook guidance

Specific guidance on the indexing of periodicals as distinct from book indexing was offered in a volume published in 1972:[14]

> Many periodicals publish their own indexes, but too often these are mere contents lists which do not attempt to analyse the contents of the articles in detail, or even to enter the titles under suitable subject-headings. It is rare that these so-called indexes are compiled by experienced indexers, and very few of them cover more than a single year at a time, while most are not compiled on scientific principles. In some subjects, learned societies and research bodies have attempted to overcome the lack of good indexes by compiling and publishing detailed indexes in their own fields. The most successful ventures are those in such subjects as Chemistry, Medicine, Engineering, and other technical and scientific subjects. Even so, the specialist libraries in these subjects continue to compile their own indexes as well, because they need still more detailed information and, what is more important, they require indexes compiled from their own point of view.

> . . . A research institution concerned with the motor industry will have a very different interest in an article on paints and varnishes from that of the British Colour Council. The points of interest to the one are quite different from those which appeal to the other. This forces organisations to index for themselves the more important articles, and means that any central system of indexing must be supplemented by individual indexing for outstanding items.
>
> The indexing of periodicals is based on the same principles as those for the indexing of books, but it involves a stricter discipline, a wider knowledge and unswerving consistency. . . . [it] is an operation carried out over a long period – possibly performed by several people – and covers a greater amount and variety of material, generally speaking. . . . It is essential to build up a standard list of subject-headings to ensure uniformity in every instalment of the index.

The increasing size, too, of cumulative periodical indexes was recognized as a difficulty. 'The larger the index the more need there is for detailed subdivision, and for cross-references, while the problems of sub-arrangement increase rapidly with the size.'[14]

In 1976 the BSI's *The Preparation of Indexes to Books, Periodicals and Other Publications* specifications for periodicals concerned identification and bibliographical details, single/multiple indexes, and terminological changes.[15] In that year, reviewing library journals, Graham Jones deplored the lack of consistency in their indexing practice.[16]

> What are we to say of those journals which over a period of ten years exhaust almost every possible option as to what to include, what use to make of form headings such as 'obituary', and what abbreviations to introduce into the index? Perhaps only this: that a prefatory note explaining indexing practice would not only help the user but would tend to stabilize practice over a period. . . . But he who has to search through a long periodical run in which indexes are sometimes bound at the front of the volume, sometimes at the back, and sometimes omitted altogether, might well wish for instructions directed as much at library staff as at binders.

The advice on the subject of indexes to scientific journals tendered by Maeve O'Connor in 1978 was:[17]

> If issues contain 20 or more articles, an alphabetical list of every author of every paper can be extremely useful for users of the journal who wish to retrieve an article which they remember appeared in a certain issue and had a particular scientist as one of the authors (not necessarily the first). . . . An index of advertisers is sometimes useful in journals that carry advertisements.
>
> Most journals provide a *volume author index* and *subject index*, or a combined author–subject index. Sometimes there are separate indexes for abstracts, biological taxonomy, book reviews, editorials, corrections, correspondence, obituaries, news columns that report research data, sponsors or advertisers, and the date on which each issue was published. . . . All these, however, can usually be included under appropriate headings in the subject index. . . . Corrections should be listed in at least two places in the subject index of the volume.

Geoffrey Hamilton offered practical advice to an ALPSP seminar in 1986.[18]

> He was a persuasive pleader for the importance of volume indexes, to cover not only the main articles but policy statements, the setting up of new committees and information about members, letters to the Editor, book reviews, even advertisements 'should you ever wish to refer to these again. If there is no easy way to find out what's in a journal it won't be used'. Before starting on an index, he said, it is necessary to decide on its scope, how often it should appear and in what depth. The manpower/time involved are likely to be underestimated. Problems that may arise include: page reference difficulties if there are supplements with different sequences or the index is to cover more

any central system of indexing must be supplemented by individual indexing

than one volume; changes in terminology over time; specialist terminology for which the indexer may need guidance.

The BSI's next revision, *Preparing Indexes to Books, Periodicals and Other Documents* (1988), listed under its index entry for periodicals, as subheadings, 'location references' and 'cumulative indexing', the latter with sub-subheadings, 'assigned keywords' and 'cross-references'.[19]

intellectual indexing, vocabulary control, and structured search techniques are even more important in electronic data files than in printed files

Current precepts

Reviewing the indexes to the Journals of the American Jewish Historical Society in 1999, Bella Hass Weinberg considered their significance, coverage and format.[20] She detected inconsistency of coverage, poor vocabulary control, a lack of continuative headings, and imprecise filing. She affirmed current principles for good serial indexing:

- 'The term "analytical index", when applied to a retrieval tool for serials, implies more than *title-derivative indexing*, i.e., extracting words from the titles of articles';
- 'An alphabetical subject index by its very nature scatters related terms, and therefore good vocabulary control and a careful syndetic (cross-reference) structure are essential';
- The maintenance of a thesaurus is essential for a large serial-indexing project.
- 'Current indexing standards and manuals recommend that *see also* references be placed at the head of an entry so that users can immediately see suggestions for other headings that may be better than the ones they started with.'
- A single A–Z sequence for an index is commendable: 'separate sequences of authors, titles, subjects, and/or book reviews create problems for users';
- A correspondence table of issue number and year prefacing the index can compensate for omission of dates in the locators.

E-publishing and e-journals

Alarm bells began to ring for indexers with the advent of electronic publishing and the internet in the early 1990s. Addressing a conference of the American Society of Indexers in 1993, Tom McFadden advised:[21]

> It is tempting to suppose that, because the Internet is already in machine-readable form, indexing the Internet need involve nothing more than asking a machine to read it. . . . Experience has shown that intellectual indexing, vocabulary control, and structured search techniques are even more important in electronic data files than in printed files, precisely because of the great size of the databases and the genuinely remarkable power of the searching algorithms. [Intelligent human] indexers still have a job to do in the brave new world of electronic publishing.

Electronic publishing has indeed brought an increasing trend towards fast-access, no-paper journals, with the concomitant rise in an industry providing single-article document delivery and indexing services for electronic searching. Max McMaster, addressing the Australian Society of Indexers in August 2000, provided a useful distinction:[22]

> Journal indexing and database indexing of journals are not the same thing. Database indexing involves the allocation of a few subject terms from a specialized subject thesaurus, and these terms, together with the bibliographic details, are included in an electronic database. Journal indexing, whether for an annual or cumulative index, can involve the creation of author, title and subject indexes: the subject terms being selected from the terminology of the individual articles, rather than from a thesaurus or controlled vocabulary.

He went on to consider, 'What types of items should be excluded from journal indexes? How many indexes should be included? Multi-authored articles and the problem of cross-references; Confusion over authors' names; Should initial articles be ignored in filing? and General or specific subject headings.'

On the internet, search engines such as AltaVista, Excite, Hotbot, and Yahoo! 'offer a directory structure which sorts into subject categories to narrow down the search possibilities'. However, 'estimates suggest that no single search engine indexes more

than one sixth of available web pages': some index only the first 17 words of the page, only on 'the keywords in the current primitive meta-tagging, or only sites that have been registered with them'.[23]

It seems that no one can cope in a timely manner with retrieval of the amount of information coming out, so indexes tend to be restricted to the fixed past of publications rather than the ongoing present.

no one can cope in a timely manner with retrieval of the amount of information coming out

Do the right thing

Printed indexes to printed journals must still be provided. As Richard Jones, for nine years the indexer of the weekly *BMJ*, writes:[24]

> The *eBMJ* has been enormously successful, but subscriptions to the printed journal have not declined. As long as the printed journal is published it needs an index to be bound in with it.
>
> I try to keep to a weekly schedule. I index almost all items, including the fillers and snippets; I include the smaller items as they are not indexed by Medline. I used to index using Mesh headings, but I find that the content and range of coverage in the *BMJ* means that a purely scientific controlled language is inadequate.
>
> I make a cumulation with each volume and use this as an authority file to make sure I am using the same headings. I make entries generously for authors, subjects, series, and author cross-references.

Learned Publishing (quarterly) is likewise on the side of the angels: the final issue of each volume includes the index. The first three issues can be indexed any time after their publication; the contents of the final issue are added from the issue proofs; and the index is typeset along with the corrections to the rest of the issue.

Many journal indexers find October and November extremely busy, speaking of 'the journal season' when the publishers want their indexes cumulated to include the last issue of the year. Journal indexes should be fully detailed, appearing in the final issue of the volume, not subsequently. The fullness and promptitude of your journal index indicate the value you place on its contents, and their worthiness to be resought. A journal volume should not have a separate, subsequently published index to its contents any more than a book does. David Lee rightly urged an ALPSP seminar on indexing: 'An indexless journal has to be one which is bought, read and virtually discarded. . . . Build indexes into your journals; don't leave them as optional extras.'[25]

Help is at hand. The Society of Indexers publishes an annual directory, *Indexers Available*, offering practical advice on commissioning an index, and a classified list of professional indexers with their subject and genre specialisms.[26] It can be consulted on the Society's website at: www.socind.demon.co.uk

The 2000 edition, No. 18, shows 110 indexers specializing in the indexing of journals and periodicals (learned and professional). The Society also publishes a series of guides including the informative *Indexing Newspapers, Magazines and Other Periodicals*,[27] as well as *Indexing the Medical and Biological Sciences*,[28] which includes a chapter on journals covering types of journals (specialized, current opinion-type, abstracting, and those relying on advertising), journal indexing methods (choice of matter to be included, of headings), picture indexes, author indexes, and cumulative indexing.

Acknowledgements

Thanks are due to the following for permission to quote from their work: E. J. Coates, Martha Cornog, Virgil Diodato, Jane Dorner, Hazel Holt, Graham Jones, Richard Jones, David Lee, Max McMaster, Maeve O'Connor, Bella Hass Weinberg, Hans Wellisch, and the editors of *The Indexer*.

References

1. Wellisch, H.H. Early multilingual and multiscript indexes in herbals. *The Indexer* 1978:11(2) Oct., 81–102.
2. Johnson, P. Printed indexes to early British periodicals. *The Indexer* 1989:16(3) Apr., 147–55.
3. Cornog, M. A history of indexing technology. *The Indexer* 1983:13(3) Apr., 15–27.
4. Holt, H. Barbara Pym as indexer. *The Indexer* 1987:15(4) Oct., 236–7.
5. Diodato, V. Author and source indexing and abstracting of journal articles. *The Indexer* 1984:14(2) Oct., 91–4.
6. Clark, J.R. The author as indexer. *Science* 1961:133(3458) Apr., 7, 1040.

7. Black, H. Authors as indexers. *Science* 1961: 133(3464) May 19, 1634–5.
8. Standards for indexes to learned and scientific periodicals. *The Indexer* 1960:2(2) Autumn, 634.
9. Ferriday, P. Indexing of periodicals. *The Indexer* 1964:4(2) Autumn, 34–8.
10. Narayana, G.J. and Ramaswami, K. Index to a periodical volume. *The Indexer* 1965:4(3) Spring, 59–66.
11. Holmstrom, J.E. The indexing of scientific books. *The Indexer* 1965:4(4) Autumn, 123–31.
12. Coates, E.J. Aims and methods of the *British Technology Index*. *The Indexer* 1963:3(4) Autumn, 146–52.
13. Singleton, A. Technical indexing at *B.T.I. The Indexer* 1974:9(2) Oct., 37–49.
14. Collison, R.L. *Indexes and Indexing*, 4th rev. edn. London: Benn, 1972, 124–8.
15. British Standards Institution. *The Preparation of Indexes to Books, Periodicals and Other Publications*. 1976.
16. Jones, G. 'This incredible stream of garbage': the library journals, 1876–1975. *The Indexer* 1976:10(1) April, 9–14.
17. O'Connor, M. *Editing Scientific Books and Journals*. London: Pitman Medical, 1978, 81–2, 142.
18. Melland, A. Producing a society's journal. *ALPSP Bulletin* 1986 July, 4–7.
19. 1988 British Standards Institution, *Preparing Indexes to Books, Periodicals and Other Documents*.
20. Weinberg, B.H. Indexes to the Journals of the American Jewish Historical Society: significance, coverage and format. *Judaica Librarianship* Spring 1999–Winter 2000:10(1–2), 14- 17.
21. McFadden, T.G. Indexing the Internet. *Learned Publishing* 1994:7(1) Jan., 17–27.
22. McMaster, Max. Journal indexing: some issues and maybe some answers. In *The August Indexer: Proceedings from the 2nd International Conference of the Australian Society of Indexers*. AusSI, 2000.
23. Dorner, J. *The Internet: A Writer's Guide*. London: A. & C. Black, 2000.
24. Jones, R. Personal communication.
25. Lee, D. Indexing: the necessary art. *Benefit from Effective Indexing*. ALPSP, 1981.
26. *Indexers Available*. The Society of Indexers, 2000.
27. Beare, G. *Indexing Newspapers, Magazines and other Periodicals*. Society of Indexers, 1999.
28. Blake, Doreen *et al. Indexing the medical and biological sciences*. Society of Indexers, 1995.

Hazel Bell
139 The Ryde
Hatfield
Herts Al9 5DP
UK
Email:hkbell@dial.pipex.com