

## AUTOMATED SUBJECT RETRIEVAL IN THEOLOGICAL LIBRARIES

Hans Arns

The period in our civilization which started with Gutenberg's invention some 450 years ago is now rapidly coming to an end. The printing press burst asunder the bounds of knowledge. It altered our ways of access to the past, to our environment, to our own workings, and its effects have reassigned our very place in the universe. Printed information is growing at a rate which planet earth no longer can sustain and the sequence in which any aspect of knowledge has progressed can only be grasped after long study. The inability to discern these myriads of voices preserved in the printed word is fragmenting our society even more than the unfortunate divisions that took place in Christianity in the 16th century. This growth of knowledge has also made us aware of its elusiveness and have put a question mark behind many of our former certainties. Between the Scylla and Charybdis of agnosticism and fundamentalism the modern theologian has to row his boat to find an interpretation of the eternal message that offends neither intelligence nor faith. This entails hard intellectual work which cannot be accomplished without access to a multiplicity of written sources. Traditionally the theologian has depended on the library (preferably his own) to supply all that he could not afford to buy himself. That time is now receding into the past. Scholars can no longer afford to purchase much beyond their basic needs. Our libraries on the other hand are in the race to keep up-to-date in so many exploding areas of knowledge and can therefore not provide for all their needs either. This exponential growth in publication is happening at a time when our ability to acquire it in the traditional manner is decreasing constantly. So what are we to do?

As always necessity is the mother of invention and for libraries all faced with the same problems two developments are now coming to the rescue which are changing the face of traditional information provision. These two developments are cooperation and an increased reliance on the spin-offs of modern technology.

Cooperation and mutual dependency have become lifesavers in our era. Collection development policies for theological consortia as well as the ANZILA Newsletter 15

establishment of local and national union catalogues have contributed greatly to the range of theological research materials now available in Australia. The result in our decade has been to consider all Australian libraries as parts of an interlinking network and resource. One expression of this has been the *Conspectus* project, an initiative through which each major Australian library analyses its collections according to a commonly accepted methodology. Its purpose is to identify subject strengths which can then be listed in a national finding tool pinpointing where these subject strengths can be found.

The second development which is breaking down the walls between our libraries is due to incredible changes in communication technology. A basic computer and modem now provides access to major databases all over the world. ABN dial-up membership for example now allows any library access to the riches of the National Library's national database which contains records of over 7 million titles and nearly 13 million holdings from contributing libraries throughout Australia. It also provides holdings and locations for theological and related journals supplementing the nearly 35000 titles listed in the 1991 edition of AULOTS.

The time when our libraries were small and librarians near-infallible guides to their contents is also over now and the modern user has to depend on well constructed catalogues and other bibliographical tools to gain intellectual access.

The wide acceptance of uniform cataloguing rules of subject headings like those of the *Library of Congress* or the *Religion Indexes Thesaurus*, standard classifications such as *Dewey Decimal Classification* now in its 20th edition have standardised library practice and thus prepared us for the uniformity which is part of any close collaboration. This development has not altogether been without loss of denominational perspective (or terminology) in the description of our collections but it is difficult to resist the benefit which this uniformity bestows.

With the advent of computer technology we are leaving behind many of the cumbersome processes with which our catalogues were compiled. We now can get most of our cataloguing data ready-made on line from the Australian national database or from CD-ROM products like OCLC

or BiblioFile in machine readable format. We transfer these data into our own database and a library application programme provides all the traditional access points of the paper catalogue. In addition we can search our automated catalogues by keywords and many other ways depending on the sophistication of our software. As well as providing improved intellectual access the computer allows enquirers to complete their searches at a remarkable speed.

While revolutionising access to monograph collections cooperation and modern technology have also provided better access to journals. Librarians have long been battling to acquaint users with the riches stored away in the annual volumes of theological indexes of all kinds. However many users still prefer the internal and rather limiting chain of references provided within their own reading. Perusal of these indexes is time-consuming and researchers are not helped by the fact that most of the resulting citations cannot be provided by their own library anyway. But things have been changing fast. For some years now a number of databases including *Religion Indexes* have been available for searching by modem through *Dialog Information Services*. Unfortunately their high usage and telecommunication costs prevented anyone from becoming proficient in searching this mass of data satisfactorily. More recently a solution has been found in the development of CD-ROM technology.

CD-ROM is a revolutionary information storage medium. It is the computer version of the audio compact disk that we play on our stereos. The acronym CD-ROM means "Compact disc, read-only memory". Its technology allows 250,000 pages of printed pages to be packed on to a single disk. One disk can hold the equivalent of 1500 standard 5.25 inch floppy disks of text or images. Their storage capacity is not their only advantage. Computer programmes are applied to the disks that allow the text to be searched in highly sophisticated ways in a matter of seconds. This medium is not as vulnerable as paper and microfilm and it offers the possibility of extending the life of text encoded on them. In the coming decade CD-ROM disk technology will probably launch as dramatic a change from paper as was papyrus when it replaced the clay tablets used in antiquity.

A growing number of CD-ROM databases are of value to the scholar in the area of theology and related disciplines. Foremost among these is the *Religion Indexes database* mentioned above which contains now 25 years (and soon 40 years) of indexing on one disk of *Religion Index One: Periodicals*, *Religion Index Two: Multiauthor works*, *Index to Book Reviews in Religion and Research in Ministry*. Distributed by H.W. Wilson with that company's powerful software it allows for a variety of simple or sophisticated search methodologies. It provides the searcher with a direct link (if needed) to Wilsonline through Dialog just in case references added since the last annual update are needed. A great advantage of the CD-ROM format is the lifting of time-constraints forced on the researcher by the high cost of on-line searching. Bible references can be searched by book, chapter and verse, although somewhat awkwardly and only for the last few years. The American Theological Library Association is however engaged in much retrospective indexing which will gradually increase this facility also. It is unfortunate that the *Catholic Periodical and Literature Index* which started in the 1930's and which provides a Catholic perspective in any area of theology still remains unavailable either on-line or on CD-ROM. Negotiations are now under way to incorporate this huge and complimentary finding tool into the *Religion Indexes* database in the near future. It will be interesting to see how the very different subject terminologies of these two will be integrated or linked.

A second important database on CD-ROM is named *Rex* and contains thirty years of abstracting by *Religious and Theological Abstracts*. It provides citations with full abstracts and can be searched by author, title, journal title, keywords (in abstracts) or any combination of these. Some 226 journals are indexed (not all over the 30 year period) but 164 overlap with *Religion Indexes*. It is a pity that the accuracy of its citations and its transfer to machine readable format have left much to be desired. However its price tag is modest in comparison with *Religion Indexes*. Many other databases are available either on CD-ROM or on line in our Academic and State libraries. Useful ones include *Humanities Index*, *Social Sciences Index*, *Social Sciences Citation CD edition*, *Philosophers Index 1940-1990* (available in 1991) and others. The Australian *AUSTROM* database on CD-ROM contains a long backrun of *Australian Public Affairs Information Service (APAIS)* which indexes a large number of Australian journals and newspapers.

Three other major types of databases available on CD-ROM should be noted at this stage. The first one should be involved at the beginning of any research project and I refer here to *Dissertation Abstracts*. It used to take up long runs of annual volumes which were cumbersome to search. Reduced to CD-ROM format its current edition has bibliographic citations and 350 word abstracts to over 120,000 doctoral dissertations from 1984-1987 while the backfile edition has a staggering 700,000 citations from 1861-1980 and 140,000 abstracts from 1980-1984. These doctoral dissertations hail from academic institutions all over the world. To obtain the maximum benefit from these databases it will be wise to get some help from a librarian in constructing a search methodology. A little lateral thinking in penetrating this mass of information could well mean the difference between disappointment and an exciting beginning to one's research. In addition to providing you with a useful bibliography which may be printed, or, if the facility is available transferred to your floppy disk, University Microfilms in Ann Arbor, Michigan can provide the researcher with any of these theses in microform and even in paper or hardback format.

The second database is of a type which will be appearing in greater numbers in years to come i.e. of huge volumes of printed text transferred to CD-ROM disks. An example of this type is the *Thesaurus Linguae Graecae*. This project undertaken with mighty expense and great care was born in 1972 when a small group of classicists in California convened to develop guidelines for computer-based procedures to collect, store and manipulate massive amounts of Greek text. It now contains 61 million words of text deriving from more than 3,000 classical and early Christian Greek authors between Homer and 600AD taken from the most up-to-date scholarly editions. Texts from Migne's *Patrologia series Graeca* are replaced by more accurate editions such as the *Sources Chretiennes* editions whenever they become available. This database includes the Septuagint, New Testament writings, Josephus, Philo and all the Christian Greek writers to the year 600. The database is constantly being updated and added to. Texts used are listed in *Thesaurus Linguae Graecae Canon of Greek authors*, published by Oxford University Press and already in its third edition. The database may be accessed by a variety of software packages and is available at an institutional licence (which includes updates) of US\$500 - for 5 years (US \$300 - for individuals). It is a major research tool at Macquarie

University's Ancient History Documentary Research Centre. A project by the Packard Humanities Foundation in Los Altos, California is under way to collect major amounts of classical and early Christian Latin authors for storage and use on a CD-ROM. These two databases are only two in a medium that is going to change the face of our libraries. In 1991 the *Index Thomisticus* is being published containing all the works of Thomas Aquinas and the list of new works and collections is growing constantly especially since improved text-scanning techniques are available now.

The third type is the Biblical research tool. It ranges from CD-ROM disks containing a number of modern Bible translations or critical editions in the original languages to text-analytical tools, Bible dictionaries, language aids and wordprocessing packages for text in Greek, Hebrew or other oriental languages. An example of this type is being produced by the Centre for Computer Analysis of Texts at the University of Pennsylvania. It will include various biblical and related materials presently being used or produced by the Computer Assisted Tools for Septuagint Studies Project at that University. It will retail at a cost of less than US \$50.

Others already on the market include CD-WORD available at a cost of US\$595 - containing Nestle-Aland's Greek New Testament, Rahlfs Septuagint Greek edition, a variety of English translations, two important Bible dictionaries, Bauer's as well as Liddell and Scott's Greek lexicons, the one volume English Kittel, three Bible commentaries including Harper's and the Jerome Biblical commentary, maps, illustrations and lexical aids.

As a postscript I must mention versions of encyclopaedias such as Grolier's are now available on CD-ROM, some with illustrations in colour, maps and sound material which are considerably cheaper and more versatile than their paper predecessors. To keep track of databases on CD-ROM one may consult the annual volumes of *CD-ROM in Print*.

A word also needs to be said about electronic networks proliferating in the United States and Europe. An example is the *BITNET Electronic Communication Network*. This network links personal and departmental computers at some 1300 sites in 38 countries for electronic exchange of non-commercial information. Gateways also allow exchange of electronic mail between BITNET and many other networks.

The *Users directory of computer networks* also lists all the institutional members of BITNET. This network also provides excellent help to the novice user. In Australia AARNET already allows direct access to the catalogues of many academic libraries.

The addition of a modem to one's computer should be considered a top priority for researchers and institutions. A scholar's workstation thus created is a tool of such multiple capabilities that most scholars have yet to begin to understand the range of services the microcomputer can render. With such a workstation the scholar can browse the shelves of academic libraries around the world and have access to colleagues and experts without the benefit of a sabbatical, travel grant or visa. It is a valuable consideration in the current Australian economic climate.

To conclude I would like to emphasise the need for cooperation in future development. The options will be many and we need to consult on complementarity of all kinds. We need mutual help to increase awareness and help in using these resources. We need to combine to ensure survival of older materials in the new formats in commercially little viable areas like theology and to deal with questions of copyright and access. An organization like ANZTLA will therefore be of vital importance to represent our future interests. We can be sure that our libraries will be shrinking in size and that the numbers game will become irrelevant. In fact we need to prepare for a near future in which scholars themselves can search in and request all their information needs from major databases while sitting at their own computer terminals, download results into a portable reader and take this information wherever they go.

Whether all these developments really constitute progress or will lead us to even further cultural distintegration is debatable. Somehow I feel however that our Christian heritage will have a vital part to play in the development of new ways of dealing with perceived realities. My advice to librarians is therefore to be aware and be ready to read the signs of the times!

*Hans Arns*

*Librarian, Catholic Institute of Sydney.*

*This paper was delivered in Sydney at the ANZTLA Conference, July 1991.*