The scientific information environment in the next millennium

Rafael Ball

The author

Rafael Ball is Head of the Central Library, Forschungszentrum Jülich GmbH, Germany.

Keywords

Information services, Science forecasting, Paradigms

Abstract

The evolution of the information environment in research and development at the beginning of the next millennium is dominated by the change of medium from printed to digital information and by the diversification of the information environment. This change of paradigm (the so-called end user revolution) is accompanied by extended efforts for literature and information searches. The lack of transparency in the information environment, the bewildering variety of media and the enormously expanded methodological efforts for literature searches by scientists have to be balanced by information specialists in the libraries. In the next millennium, scientists will only search for information in one single electronic search user-interface for all available sources and media at one time. Knowledge and information have always been the most important resource in science. However, the change in the methods of science has also been accompanied by a change in the method of scientific information retrieval. This article shows how a scientist's efforts, tools and methods have changed in the past few decades for the verification and retrieval of literature and information required against the background of changing media. An outlook shows possible constellations in the next millennium.

Electronic access

The research register for this journal is available at
http://www2.mcb.co.uk/mcbrr.lm/.asp

The current issue and full text archive of this journal is available at
http://www.emerald-library.com

Before the computer made its entry into libraries, the information resources (whether monographs, journals or others) as well as their holdings lists (catalogues) were available in hand-written or printed form. Shelving of the literature and above all structuring of the meta-information (that is of the catalogues) followed a (not always uniform) classification system and varied from library to library. The scientific user of a solid traditional library (pre-digital library) had to be familiar with the classification scheme for shelving the collections, with catalogue structures (author catalogue, alphabetical subject headings catalogue, classified catalogue, geographical catalogue etc.), and had to know the ordering modalities for all the collections not on open access.

Users acquired such knowledge during an introduction (for example, a guided tour of the library with an explanation of the catalogues). Special knowledge required for scientific literature searches was mostly imparted to the user by a tutorial at the beginning of his university studies. It was possible in a lecture course of two hours per week for just one term to learn how to use the entire holdings of a large university library and to be able to use the relevant reference literature for the respective specialised information. Equipped with these tools, the scientist was in a position to cope with all issues occurring in connection with literature and information searching throughout his scientific life. The medium was uniform (printed or hand-written), aids were not required and it was sufficient to be able to read.

With the further development of reproduction techniques and of technology in general, the print-adequate media first made their entry into the libraries (microfilm, microcard, and microfiche). This medium, too, was easy to master and did not constitute a qualitative difference from the familiar printed medium. The only complication was the technical handling of microfilms, microfiches or microcards since technical tools were required for reading (for example, a microfiche reader).

The next step – and here we experienced a qualitative jump – consisted of making digital information available in libraries. Similar to the introduction of microfiches, it was initially meta-information which was made available in this new form. Digital information has been

available online since the late 1960s as databases containing bibliographic information, specialised databases, fact databases and, only at a later stage, full text databases. This information was first distributed on magnetic tapes as fixed data carriers, then on floppy disks and from 1985 as CD-ROMs. This resulted in two different learning steps for the scientific user: the online use of databases was first restricted to search specialists. The scientific user had to make great efforts to familiarise themselves with the retrieval languages and the usually complicated login. End use was an alien concept and the scientists were neither willing nor able to use this information themselves. Consequently, there was no basic change in the effort expended in literature and information searching. The study of catalogues and reference literature only had to be supplemented by discussions with the searcher. But it was always the searcher's task to carry out the online search. A completely new situation has been created, however, by the introduction of digital information on floppies and CD-ROMs. The information was no longer centrally stored on a host server and had to be queried online, but was directly available in the libraries in a decentralised manner. It was distributed over a wide area and thus the libraries became responsible for operating and updating. And what is more: after these databases had been initially operated and supervised by librarians alone, something happened which one could call an end-user revolution. The scientists were allowed to and even forced to use the CD-ROM databases on their own (Schatz, 1997). At the beginning, there were no networks in universities and research institutions, and the databases on CD-ROMs were mostly available as stand-alone solutions. What did this end-user revolution mean in concrete terms of additional searching effort for the scientists? How did the method of literature and information retrieval change for the user? What are the resultant advantages and disadvantages? Let us summarise again: for their literature search, the library users previously had to know the classification system of the holdings and the structure of the catalogues, they had to be familiar with microfiche/microfilm reader techniques and had to call upon the online searcher for the online databases. The new freedom of being allowed to personally

use CD-ROM databases was purchased at the expense of having to cope with an immense number of retrieval languages, search interfaces and login procedures. Thus, the diversification of information supply is undoubtedly, at the same time, a gain and a loss in terms of content, techniques and media. The effort required for a literature search thus increased continually. The progressive introduction of digital full texts further increased the effort since it was now necessary to master relevant reading programs. Since the early 1980s, therefore, there has been a complexity of coexisting traditional and electronic information offered to scientists. Only a command of all these tools made it possible to ensure a more or less complete search. Up to then, the libraries had done little to help the scientist cope with this complexity. It was only the hesitant breakthrough of new management ideas and emerging customer orientation which led to considerations of structuring the information supply for users or special user groups.

The establishment of the Internet as a serious information source since the beginning of the 1990s has initiated a further development step in the information environment. Since then the offers have followed in quick succession in a confusing manner. We find today a hopeless muddle of traditional literature collections, digitised holdings (online or on CD-ROM), metadata (catalogue data, for instance) in the traditional form (hand-written or printed), on microfilm, microfiche or microcard, on CD-ROM as an online database (local, regional, national, international library system) and an additional multitude of the most diversified intricate Internet offers available to the scientist searching for literature and information. This includes electronic journals, digital monographs, pre-print servers and abstract databases, fact databases and bibliographic databases. Since it has meanwhile been recognised that structuring is necessary, especially the commercial providers have bundled their information into complete offers. Various publishers and agencies as well as academic societies and scientific institutions, for example, offer their respective "private packages" composed of a mix of e-journals, facts, news, full-text offers, fact databases and other value-added products. Thus, the information is only apparently bundled; in fact, however, it is

completely fragmented for the end user. The information environment has become even less transparent and more complicated.

We started out from the question of how great the effort is for a scientist searching for information, and we come to the conclusion that it is immense today. It is therefore imperative to provide relief so that scientists can concentrate again on their principal task. (To my knowledge, there have as yet been no empirical investigations concerning the time required for regular literature searching by scientists.) And there is hardly any doubt about who should be the infrastructure facility providing relief for the scientists: namely the libraries. It may therefore be doubted whether libraries are really obsolete (Butler, 1999) or whether they will not be needed more urgently than ever as the only possible clearing house in a multidisciplinary science environment.

During the past 20 years, the digitisation of the information sources has clearly improved information supply and the diversity of information as well as access to literature and information. Modern technology and the activities of information producers have brought about this increase in performance. On the other hand, it is necessary to substantially reduce the concomitant lack of transparency and confusing variety as well as the resultant clearly increased effort for literature and information retrieval by the scientists themselves. Information diversity in terms of content, media and techniques must not be obtained at the expense of scientific efficiency and user performance. It is thus a central task of libraries to perform this work for the scientists. In the ideal case, a scientist will only search for information in future in a single electronic search user-interface (Alper, 1998). He will then find a transparent, well-structured and nearly complete information offer throughout all the media, all the contents and techniques, prepared by his library. Information specialists have screened and structured the information supply, while in the background technicians and computer scientists operate search mechanisms that need not interest the scientific user any more. The effort that a scientist expends on literature search and information searches to support, prepare and postprocess his scientific, experimental or theoretical work must be perceptibly reduced. It must at least not

exceed the effort formerly involved in traditional literature search using a card catalogue. Libraries must eliminate the drawbacks of the end-user revolution. The information environments of various publishers, academic societies, institutions and other producers must be integrated and unified. This also applies to offers by different commercial, non-profit and private providers, as well as offers of different types such as monographs, series, journals, databases, online offers, mixed or blended information and offers with different media structure (print, microfiche, CD-ROM, digital, electronic). Moreover, due to the interlinkage of the humanities, social and natural sciences, only a central library can offer the orientation know-how urgently required for interdisciplinary science. However, it is precisely this interdisciplinarity which is forward-looking, and future societal issues can only be solved with its help.

The role of central information facilities and libraries can be summarised as follows. Central libraries:

- screen the confusing diversity of information supply;
- clarify the lack of transparency;
- disentangle the offers;
- reduce the search effort;
- integrate the information supply in terms of content, techniques and media;
- simplify the technical handling of literature searches;
- support scientists in the acquisition of new media;
- create an interdisciplinary frame for literature supply to all those involved in science, teaching and research;
- regard themselves as providers of services for science, research and teaching;
- relieve the scientist wherever this is possible;
- need to be accepted by science and must be established as equal partners of science in the research scene.

References

Alper, J. (1998), "Assembling the world's biggest library on your desktop", *Science*, Vol. 281, pp. 1784-6.

Butler, D. (1999), "The writing is on the Web for science journals in print", *Nature*, Vol. 397, pp. 195-200.

Schatz, B.R. (1997), "Information retrieval in digital libraries: bringing search to the Net", *Science*, Vol. 275, pp. 327-34.