TOWARD CUSTOMIZED INFORMATION SERVICE

S. Raman Venkataraman
University of British Columbia
Science Division/The Library, 1956 Main Mall, Vancouver, B.C., Canada

Information needs of scientists and technologists vary across a broad spectrum, from general to specific topics depending upon whether the information is needed for innovation or problem solving. Information is gathered from both formal and informal systems. The growing formal information cannot be used unless it is easily and precisely retrievable. Secondary services have been and are still attempting to provide bibliographic control to major scientific and technical publications. Secondary services have gone through major changes such as from discipline based services to mission based services to problem based services. With the recent technological changes information services are entering a new era of providing information tailored to meet specific needs of individuals.

This paper reviews the present state-of-the-art of formalized information services and attempts to predict the future trend.

TOWARD CUSTOMIZED INFORMATION SERVICE

S. Raman Venkataraman
University of British Columbia
Science Division/The Library, 1956 Main Mall, Vancouver, B.C., Canada

INTRODUCTION

Information needs of scientists and technologists vary across a broad spectrum, from general to specific topics depending upon whether the information is needed for innovation or problem solving. Both scientists and technologists gather information from two main sources, formal and informal.

The purpose of this paper is to review the present role of the formalized information system in satisfying the selective or specific information needs of the end user, and to discuss future trends.

INFORMATION NEEDS

Progress in science and technology rests on research. Research is cumulative and is built upon previous work. Therefore scientists and technologists seek information when they wish to: (1) select or devise a new field of research, (2) check the state of the art in a specific field, (3) know the names of those involved in their own research area and, (4) find a solution to a specific problem.

GATHERING INFORMATION

Information is gathered both from informal (information exchange among colleagues and friends) and formal (document distribution) systems.

Formal sources may be comprehensive but not necessarily timely in providing current research information. On the other hand, informal sources can be more timely in providing information about current research but will not be comprehensive, so that the systems complement each other. Colleagues may be a good source to the published literature in a specific field but formal sources can provide names to contact for referral for those entering a new subject field (Brickson et al, 1984).

GROWTH OF FORMAL INFORMATION SYSTEMS

Originally information was gathered through private correspondence and from occasional publications of books and pamphlets. With the invention of printing in the 15th century, an enormous advance in the dissemination of information through the formal system took place. With the replacement of manual printing by more technically advanced methods, the formal information system has grown year by year in size, range and international character.

ORGANIZATION OF PUBLISHED INFORMATION

The growing body of published science information cannot be used unless its components are easily and precisely retrievable through adequate identification and tagging. Of the different forms of publication in the sciences, journal articles,

conference proceedings, dissertations and reports are the major sources of current information. Libraries do not and cannot provide comprehensive bibliographic control to publications other than books. Secondary publication services (indexes and abstracts) provide bibliographic control of major scientific and technical publications. Indexes make the user aware of the available literature, while abstracts provide a means of appraising the value of the available literature, and of determining whether the original publication is worth acquiring.

ORIGIN AND GROWTH OF SECONDARY PUBLICATIONS

The early history of periodical indexes was characterized by the following pattern of development. In the 17th century, individual indexes to individual periodicals were published. Between the 17th and early 19th centuries, individual indexes were replaced by cumulative indexes to individual periodicals and finally by collective indexes to more than one periodical title. In the middle and late 19th century, current ongoing indexes were published to more than one periodical title. Also early in the 19th century the prototype of the modern abstract journal was published (Kornick 1962).

Many sources have documented the growth in the number of secondary publications, but no comparative figures are available. Nine active abstract journals were reported at the end of 1799 and the number of abstract journals increased to 209 by 1920 (Manzier 1977). The International Federation for Documentation listed 1,500 abstracting journals in 1969 (IFD 1967). In 1963, the National Federation of Science Abstracting and Indexing Services (NFAIS 1963) reported 1,855 titles which were meant to be guides to searching the literature of science and technology. Approximately 2,500 secondary publications were reported in 1978 (Bearman 1978) based on data from a world inventory.

With the enormous growth in the primary literature secondary publications have been growing in size. In 1971 308,976 abstracts were listed in Chemical Abstracts; in 1979 436,887 abstracts were listed, an increase of 41% in eight years. The total number of abstracts listed in CAS up to the end of 1979 was 8,068,661 (Chemical Abstracts Service 1980). This proliferation is not unique to Chemical Abstracts. Engineering Index listed 60,154 references in 1973 and 106,795 references in 1983, an increase of 77% in 10 years. During the same period, there was an increase of 50% in the number of abstracts listed in Physics Abstracts (1973: 81,352 abstracts, 1983: 121,808 abstracts).

SUBDIVISION OF SECONDARY PUBLICATIONS

With growth in their size, secondary publications became bulky, cumbersome and time consuming to scan. Some secondary publishers, like CAS and ISI, introduced cumulative indexes, but over the years these cumulative indexes also became bulky and the extent of coverage was wide. In order to broaden the access and bring together all publications on a given topic, many secondary publications introduced subject headings under which the citations were grouped. Chemical Abstracts now groups the abstracts under 80 sections, and Physics Abstracts now has 61 subject sections. In 1947 Excerpta Medica started with 8 sections, in 1983 it was published in 46 sections.

The splitting of secondary publications and grouping of citations under specific subject headings has not provided any solution to the limitations inherent in manual searching. Manual searching is based on preassigned index terms or controlled vocabulary and hence cannot be manipulated at the time of searching. Moreover, with the rapid growth in the amount of information a manual search has become more time consuming and incomplete. Therefore a need for a more advanced method of literature searching has become important.

COMPUTER PROCESSING OF A&I PUBLICATIONS AND BIRTH OF SECONDARY SERVICES

The use of computer technology in the 1960's to publish hard copies of indexes and abstracts led to the development of machine readable forms of secondary publications (databases). The computer based retrieval system permitted storage of multiply indexed records. This provided a multiple access approach involving a large number of index terms in various logical relationships which was not possible in the manual system. Secondary services such as current awareness and retrospective retrieval also became possible.

EVOLUTION OF SCIENTIFIC RESEARCH AND SECONDARY PUBLICATIONS/SERVICES

Besides technology, changes in the emphasis of scientific research present a great challenge to secondary information systems and services.

Discipline Oriented

Secondary publications, since their birth have been traditionally discipline oriented or subject based. They covered publications on a specific subject exhaustively. Some of the significant discipline oriented secondary publications are Geologisches Zentralblatt (1911-), Chemical Abstracts (1907-), and Engineering Index (1896-). Until World War II most of the A&I publications remained discipline oriented and unidisciplinary.

Mission Oriented

During World War II United States government agencies mobilized science and technology to support the national defence missions. Since then scientific research has been directed considerably more towards solutions of problems like space exploration, energy, nuclear science, pollution, etc. Secondary publications such as Nuclear Science Abstracts (1947-), International Aerospace Abstracts (1961-) Science and Technical Aerospace Abstracts (1963-), Pollution Abstracts (1970-), Environmental Abstracts (1970-), and Energy Abstracts (1976-) were introduced to report promptly on the literature relating to new scientific missions. The users of the mission oriented information systems, in contrast to the discipline oriented system users, are specialized and need information from many subject areas to reach their goals. Consequently the subject covered by mission oriented secondary publications is interdisciplinary in character, transcends the conventional and arbitrary division of science, and covers new fields of science of unproven stability and indefinite duration.

Problem or Crisis Oriented

Vincent Giuliano (1978) identified three major eras in information. The first era, oriented towards discipline based service, began in the 19th century and continues to the present. The second era of mission oriented service began during World War II and continues today. The change in the world energy balance, growing competition among industrial nations, increasing visible effects on growing population of natural calamities, advances in sciences and technology, growing pollution etc. have all been creating new problems or crises. Solutions to these problems need an information service which can quickly make available a package of specialized information in a very short period from a broad range of information sources. For example, a mining company investigating a new mining product must be aware of, in addition to the traditional mining research and patent literature, the effect of their operations on the environment as well as government regulations. Another example is specialized information on diseases needed by medical professionals (King 1979).

It became apparent in the 1960's that neither a single large discipline based service nor a single mission based interdisciplinary service could have the capability of providing information in a crisis or problem oriented situation. Many of the discipline and mission oriented services had grown large and most of them were incomplete in their coverage of information relating to a specific problem of interdisciplinary nature. Moreover the various services covering the information could not be integrated due to lack of standardization. So the third era of information of problem or crisis based service began in the late 1960's and has been receiving growing attention since then.

PROMIS, an automated problem-oriented medical information system, divides medical knowledge into two bodies: (1) knowledge concerning the individual patient - knowledge that used to remain scattered in multiple and largely unedited medical records; and (2) knowledge concerning populations of people or kidneys or bacteria or diseases or drugs or surgical procedures - knowledge that is usually found in medical libraries, textbooks and journals. Its overall objectives are: (1) to facilitate good patient care by making available in a few minutes to the individual physician a complete updated list of problems of a patient and by providing in sequence all the narrative, laboratory and other data pertinent to these problems, (2) to make immediately available all relevant epidemiological and other research studies, (3) to make possible a medical audit, (4) to make possible a business audit to assess the physical, financial and time resources that go into the solution or management of a given problem and (5) to provide problem oriented records that can serve as a basis for clinical medical education (PROMIS 1978, Dubois and Davis 1978).

ENDEX/OASIS, an Environmental Data Index and the Oceanic and Atmospheric Scientific Information system, has been developed as a one stop computerized service providing references to technical literature, data and research dealing with the environmental sciences and marine coastal resources (ENDEX/OASIS 1976). Petroleum Resource Data System, GEOTHERM (Geothermal Resources File), CRIB (Computerized Resources Information Bank), NCRDS (National Coal Resource Data System) are some more examples of problem-based services (Meyer 1977).

CUSTOMIZED INDIVIDUALIZED INFORMATION SERVICE

Bearman (1983) believes that we are now in the fourth era of customized, individualized information services tailored to meet the specific need of individuals.

Selective Dissemination of Information (SDI)

The goal of Selective Dissemination of Information (SDI) is to supply current items of relevant information. SDI, in both concept and practice is not new (Landau 1969). This goal remained unattainable until Luhn (1958) first proposed SDI as a new application of automated information retrieval systems, to keep a group of professionals continuously informed of the latest publications in their field of interest. The subscriber to a SDI service submits a profile narrowly defining the specific areas of his/her interest. The system responds by reporting periodically all documents matching the profile.

SDI services have now become a major vehicle for providing periodic updates or current awareness. There are many SDI services serving the scientific and professional communities. A Canadian nation-wide SDI service called CAN/SDI, operated by CISTI (Canada Institute for Scientific and Technical Information), provides current awareness service from input tape services. Services are now available for 17 databases. The United States National Oceanic and Atmospheric Administration also provides a similar individualized current awareness service to both NOAA and non-NOAA professionals (ENDEX/OASIS 1976). Many secondary publishers also have started publishing current awareness bulletins based on standard profiles. For example CA Selects is a specialized series of topical abstract literature on one well-defined area of chemical interest and is published by Chemical Abstracts Service. The topics represent various areas of active interest in biochemistry, chemical engineering and chemistry. The publication began in October 1976 with six topics (Blake 1978). In the current year CAS Selects is available for 149 topics (CAS 1985). Apart from this CAS also provides many custom-defined information services based on individual profiles.

Microcomputers or personal computers with the capability of accessing remote databases have made possible the downloading of retrieved results. Local indexing is possible with an intelligent terminal which can combine edit control, a prompt and check for data entry, and efficient storage for later processing. Using microcomputers, it should be possible to download online search results based on a broad search strategy and then to create files tailored to individual needs using those references already extracted from databases (Rowell 1984, Thompson 1978). Information Research Centre (IRC), the Markets and Research Service Department of Bell Communication Research Inc., was assigned to provide centralized information research support to the marketing and strategic planning organizations within the seven regional companies who are Bellcore's owners. IRC using an IBM-PC, downloads search results into a file, then edits the file and incorporates it into reports or documents to be delivered to client demanding a new set of skills and a new perspective on the role of information professionals (Pritcher 1985).

FUTURE TRENDS

Researchers who have access to microcomputers are becoming increasingly interested in downloading, restructuring, storing and reusing A&I online search results.

The main concern of secondary publishers is not that a single user wants to store some references on his/her own microcomputer after reformatting and editing. The use by individuals will be acceptable as long as it is paid for. The American Chemical Society permits a customer to download search results from CAS ONLINE for the following purposes: (1) editing or changing the format of CAS search results to create input to other search files or systems (2) combining CAS search results with output from other search files or systems to create a consolidated file of search results and (3) editing the search results to remove irrelevant answers. The results must be for personal use only and not for sale (CAS ONLINE). Many publishers do not permit downloading of search results from their database even for individual use. Secondary publishers are mainly concerned with users who want to capture a large portion of a database, repack and market a new product as it is difficult to monitor such use.

Due to downloading, secondary services are being confronted with two issues, property right and copyright. As a result of the application of electronic technology to the publication of primary literature the future of the secondary service itself is being questioned. The secondary services are facing all the economic pressures of the business world.

Several predictions relating to the future of secondary services have already appeared in the literature (Butler 1979; King 1979; Lancaster and Neway 1982; Neufeld 1982, 1983; Williams and Brandhorst 1979). One of the predictions is an integrated cooperative system or information supermarket (Williams and Brandhorst 1979). Engineering index is planning to evolve into a new type of information service - a one stop 'information supermarket' (Landau 1983).

Significant changes have already taken place in the function of secondary services. Some A&I publishers like ACS and ISI have become vendors of online services. Some vendors like BRS have become publishers of primary literature. The American Chemical Society, in order to provide timely and comprehensive information services, have joined with West Germany's Fachinformationszentrum Energie, Physik, Mathematik GmbH to form the first link in STN international, an international on-line network for scientific and technical information. They are also planning to establish an STN centre in Tokyo.

For survival, A&I services must adapt to market needs and user preference. One of the main functions will be reducing unmanageable amounts of information down to a smaller number of items that are of interest to the user. Customers prefer information tailored to their specific requirements at a price they can afford. New technology and software are available to filter out the unwanted information. In the past 25 years the secondary services have been attempting to provide the end users with more and more customized and problem oriented products. Microcomputers have provided a tool for the end users to create files tailored to their individual needs. Efforts are being made to use Artificial Intelligence: (1) to enable the information retrieval system to answer questions which go beyond simple Boolean combination of words in the records and (2) to develop a natural language interface to data retrieval systems in order to make the retrieval system more user friendly (Smith 1980).

The secondary services have to continue to provide, enhance the quality and increase the number of customized and problem oriented products. In what form the products will be delivered, who will be delivering them and how they will be delivered will depend upon the following factors:

(1) Price the clients are willing to pay
(2) Price at which the services can afford to sell
(3) Advances in the applications of microcomputers and artificial intelligence to information retrieval
(4) Whether the primary or secondary publishers are going to allow the end user to edit and reuse the downloaded online results

and (5) whether the primary or secondary publishers are going to allow intermediaries to create and market the customized or individualized information products.

REFERENCES

ACS COMMENT. "Changing Times at Chemical Abstracts Service", Chemistry & Engineering News. Vol. 62, No. 49 (1984), pp.36-37.

BEARMAN, T.C. "Secondary Information Systems and Services", in Williams, M.E. ed. Annual Review of Information Science and Technology, White Plains, NY. Knowledge Industry Publications Inc., Vol. 18 (1975), pp.179-208.

BEARMAN, T.C. "Toward the Fourth Era of Information Service", in Neufeld, M.L. et al. ed. Abstracting and Indexing Services in Perspectives: Miles Conard Memorial lectures 1969-1983, Arlington, Va. Information Resources Press, 1983, pp.53-55.

BLAKE, J.E. et al. "CA Selects - A Specialized Current Awareness Service", Journal of Chemical Information and Computer Sciences, Vol. 18, No. 4 (1978), pp.187-190.

BRICKSON, T.K. et al. "Scientific and Technical Information Transfer: Issues and Options", Rand Note 2131, 1984, pp.111.

BUTLER, B. "Beyond the Library - U.S. Online Trends", in (Proceedings of the) 3rd International Online Information Meeting, 1979, pp.385-92.

CAS Chemical Abstracts Service International Catalog. Columbus, OH, Chemical Abstracts Service, 1985, 29p.

CAS ONLINE Terms and Conditions.

DUBOIS, R.M. and DAVIS, L. Automated Problem-Oriented Medical Information System (PROMIS) Data Tape. Springfield, VA., National Technical Information Service, PB 81-198186, 1978.

ENDEX/OASIS User's Guide to ENDEX/OASIS: Environmental Data Index and the Oceanic and Atmospheric Scientific Information System. Washington, National Oceanic and Atmospheric Administration, 1979, 66p.

GUILIANO, V.C. Into the Information Age. A Report by Arthur D. Little Co. for the National Science Foundation. Chicago, American Library Association, 1978, 134p.

IFD: International Federation for Documentation. Abstracting Services. New York, International Publication Service, 2nd edition, 1969, 2vols.

KING, D.W. "The Information Community: Its Dilemma, Opportunities and Challenges", NFAIS Newsletter, Vol. 21, No. 2 (1979) pp.18-28.

KORNICK, D.A. A History of Scientific and Technical Periodicals: The Origins and Development of the Scientific and Technological Press, 1665-1790. New York, Scarecrow Press Inc., 1962, 274p.

LANCASTER, F.W. and NEWAY, J.M. "The Future of Indexing and Abstracting Services", Journal of the American Society for Information Science, Vol. 33, No. 3, (1982), pp.183-189.

LANDAU, H.B. "Document Dissemination", in CUADRA, C.A. and LUKE, A.W. ed. Annual Review of Information Science and Technology, Chicago, Encylopaedia Britannica Inc., Vol. 4, (1969), pp.229-270.

LANDAU, H.B. "Evolution of a secondary service to an Information Supermarket", On-Line Review, Vol. 7, No. 5, (1983), pp.407-411.

LUHN, H.P. "A Business Intelligence System", IBM Journal of Research and Development, Vol. 2, No. 4, (1958), pp.314-319.

MANZIER, B.M. The Abstract Journal, 1790-1920: Origin, Development and Diffusion. Metuchen, NJ., The Scarecrow Press, Inc., 1977, 132p.

MEYER, R.F. "Petroleum Resources Data Systems", Journal of the International Association for Mathematical Geology, Vol. 9, No. 3, (1977), pp.281-99.

NEUFELD, L. "What's Ahead for Secondary Services", American Society for Information Science, Vol. 9, No. 2, (1982), pp.40-43.

NEUFELD,M.L. "Future of Secondary Services", On-Line Review, Vol. 7, No. 5, (1983), pp.421-26.

NFAIS National Federation of Abstracting and Indexing Services, History and Issues 1958-73. Philadelphia, PA., National Federation of Abstracting and Indexing, NFAIS Report No. 5, 1973, 23p.

PRITCHER, P.N. "Integrating Microcomputers and the Information Professionals: Strategies for Training Information Managers to the Use of IPC", Online. Vol. 9, No. 2, (1985), pp.15-22.

PROMIS LAB Functional Specification of a PROMIS Instance System. Volume 2. Springfield, VA., National Technical Information Service, PB 81-198194, 1978, 340p.

ROWELL, P.P. "Scientific Literature Currency and Organization using a Microcomputer", Online. Vol. 8, No. 1, (1984), pp.18-21.

SMITH, L.C. "Artificial Intelligence Applications in Information Systems" in WILLIAMS, M.A. ed. Annual Review of Science and Technology, White Plains, Knowledge Industry Publications Inc., Vol. 15, (1980), pp.67-105.

THOMPSON, G.K. "The use of Microcomputers by Secondary Services and Individual Researchers in the Social Sciences", in Information Age and Perspective. Proceedings of the ASIS Annual Meeting, Vol. 15, (1978), pp.341-43.

WILLIAMS, M. and BRANDHORST, B. "Future Trends in A&I Data-base Publication", Bulletin of the American Society for Information Science, Vol. 5, No. 3, (1979), pp.27-28.