in which benefits are truly generalizable to the entire user group: e.g., library staff in general; most books and journals; and perhaps certain basic services, particularly of a short duration, such as standard reference service. Possibly collection components which are specific to an identifiable segment of the user population (as well as significant services to a group) should be financed by that group. (These remarks are pertinent to multidiscipline libraries, rather than to small, specialized ones.)

In addition, the library management may have to redefine the uses which general funding can serve. Perhaps general funding is sufficient only for the "normal" or "average" user of a defined community. Possibly the user with a high volume of demand or from outside the community may have to pay usage fees.

Libraries should assume the responsibility to optimize service through providing expertise on cost-effective selection of sources and on development of interview and query techniques, user aids, and control and evaluation methods. We have not had time or funding to carry out carefully controlled comparisons of cost or content effectiveness among the various services and data bases. However, we plan to study the services and make recommendations to our users for their optimal utilization and application at NBS. As we have said, the successful management of a computerized retrieval facility is both a significant service in itself and an expansion of the library's capability for information service of a sophisticated nature. This role of the library can be recognized by the user group. The library staff members may gradually find themselves becoming more involved as experts in development of user projects. The effect may spill over into all library functions, e.g., collection and other resource development.

Use and Implications of On-Line Information Retrieval for Management†

SAMUEL A. WOLPERT

Predicasts, Inc., Cleveland, Ohio 44106

Received December 15, 1975

The scope, size, and variety of information available through Predicasts Terminal System (PTS) is discussed. Through samples searches in PTS File 16 (*Market Abstracts*) for chemical data, the advantages and disadvantages of search strategies involving natural language and hierarchical codes are compared. The development of on-line information systems is capsuled, and forecasts are made for the future, including comments about the use and implications of PTS and on-line retrieval in general for decision makers.

By the end of 1975, Predicasts Terminal System (PTS) files contained more than one million records and 400 million characters. PTS File 16, *Market Abstracts*, has 90000 records giving extensive abstracts with descriptor statements. File 17, *PTS Weekly*, contains both market abstracts and F&S Indexes representing weekly updates of PTS Files 16 and 18. File 18, *F&S Indexes*, gives brief comments with descriptors, and totals 500000 records on business and economic developments around the world. File 19, CIN, contains the extracts from *Chemical Industry Notes* of the American Chemical Society, approximately 150000 records. File 19 is incremented weekly. File 20, *Domestic Statistics*, has about 150000 records, including forecast abstracts, historical and forecast time series for the United States, and components of the U.S. File 21, *International Statistics*, presently contains forecast abstracts and forecast time series for all the countries of the world (except the U.S.), and by the middle of 1976 will include historical time series as well. *EIS Plants*, File 22, contains descriptive statistics and categorization of the 120000 principal industrial facilities in the United States.

Of all the PTS files, File 16 (*Market Abstracts*) is the one with the best potential use for chemical industry management. File 16 contains digests of articles appearing in hundreds of worldwide journals on new products, acquisitions, capacities, end-uses, market data, technology, production, environment, foreign trade, countries, and regulations for the extractive, manufacturing, transportation, utilities, and other industries. General economics and financial industries are not covered. One of the industry groups reported most intensively is the chemical process area.

† Presented before the Division of Chemical Information, 170th National Meeting of the American Chemical Society, Chicago, Ill., Aug 27, 1975.

While File 16 covers all the countries of the world, about half of the records are from the United States. By type of information, 20% deals with organizations, 30% with technology, 20% with resources, 20% with market data, and 10% with other material. Each File 16 record contains a verbal abstract, a paragraph or more in length, and field descriptors. Included with each of these descriptors are hierarchical codes, which are a powerful tool for low-cost searching (as we will see).

In searching PTS one may use natural language, hierarchical codes, or a combination of both. If we are using natural language, we may first begin with the dictionary to see how many records the file has on a particular word or word combination. The dictionary alphabetically lists all nontrivial words from the abstracts, as well as word combinations and verbal names from the thesauri for the affiliated numeric codes, with the number of times each word or term appears. If we are searching, for example, for records about polystyrene, we discover that the dictionary lists about twice as many entries for polystyrene where codes were not mentioned, as polystyrene records with codes. In the latter case, a human indexer indicated that each of these records was, in fact, about polystyrene. The disparity is accounted for by two facts: first, some of the time when the word polystyrene was used, the human abstractor coded the entry as polystyrene film or foam and not simply polystyrene; secondly, the human indexer considered the polystyrene in some of the abstracts as being trivial, just mentioned in passing and therefore uncoded.

The PTS Thesauri include the numeric product, geographic and event codes with the affiliated verbal names for each code, again listing the number of times each appears. Thus, in the product code (PC) thesaurus we find polystyrene entries under PC 2821310 (polystyrene), 2821311 (regular polystyrene),

2821312 (impact polystyrene), and 2821318 (expanded polystyrene), as well as in other categories. If we use the numeric codes, we could—instead of selecting each of these polystyrenes individually—truncate the last digit and select 282131 to yield the entries for all these polystyrene types. We could further truncate to 28213 for a broader category (styrene resins) for more "hits," or to 2821 (plastic materials) for even more records. We could even select 28 to access all the records in the very broad "Chemicals & Allied Products" category for a number of "hits" greater by almost a factor of 10 than the original polystyrene (2821310) heading.

For another example, to receive information about plastic materials in Africa via natural language, we would have to mention each plastic and its alternative spellings. Thus, we would have to say polyethylene, HD polyethylene, PE, nylon, polyamid, etc. Then all the African countries would have to be listed with the variant spellings for each. At least several hundred terms would have to be cited. To retrieve the same information using hierarchical codes, one would simply select PC 2821 (plastic materials) and the country code for Africa (CC 7) and combine the results. This search could be accomplished in about one minute.

We have estimated costs for eight hypothetical searches, half of which were simple ones using codes and half of which were complex natural language searches. The output formats for each were accession number, journal references, on-line abstract, and overnight printing. The simple searches required six commands and three minutes, as compared to the 24 commands and 12 minutes required for the natural language searching.

Generally speaking, searches using codes yield fewer but more relevant "hits". However, we have found that a mixture of natural language and codes maximizes both efficiency and relevance, and this method is preferred by most of our customers and our own personnel. We have further determined that the average search of the PTS files uses ten minutes of connect time for a cost of about $20.00.

Potential applications for PTS and other on-line systems meet a wide variety of needs. *Information query* can deliver the one or two facts needed by an executive making a policy decision or for a researcher to finish a report immediately. *Orientation* can provide the necessary background material to optimize client contacts or meetings. *Selective Dissemination of Information* (SDI) can periodically deliver custom selected abstracts from current literature for each executive or researcher. It can broaden their coverage, save reading time, and assure that nothing is missed. *Retrospective searches* can provide analysts with literature needed to save many weeks or months of work. *On-line analysis* can provide new insights into relationships and mechanize many tasks.

We have tried to forecast the future for on-line systems in the United States, and have concluded that there is great potential for the users, searchers, and proprietors of on-line data bases. A good searcher must have good knowledge of the subject matter of the data base as well as the retrieval software, and we are convinced that the information retrieval (IR) specialist will have an increasingly important role.

We have determined that most information searches are done for professional and technical employees, an employment type growing more rapidly than the total work force. In 1965 there were 8.9 million such employees. This grew to 12.2 million by 1975, and we expect the total to reach 17 million by 1985. In 1965 there was one librarian/IR person for every 89 professional or technical employees. Either librarian productivity grew or professional use of information lessened, because by 1975 each librarian was servicing 98 professionals. We expect this trend to be reversed in the future as increased use of information by professionals will occur, and the librarian/IR specialist will be serving only 85 professionals by 1985. There were 100 000 librarians/IR personnel in 1965, 124 000 in 1975, and there will probably be 200 000 in 1985. While the average librarian/IR person made five searches in 1965, this will jump to 100 in 1985. Only about 100 searches were made in 1965, but this figure had risen to 10 000 by 1970 and to 600 000 in 1975. We forecast that 20 million on-line searches will be performed in 1985.

In 1965 the only user of on-line systems was NASA. By 1970 they had been joined by other large government agencies. Today most large institutions are beginning on-line literature searching. By 1980 most institutions, large or small, maintaining their own libraries will have terminals. By 1985, institutions which employ professionals, whether they have their own library or not, will have commenced on-line searching.

The information industry is coming of age as the quantity of business and economic information continues to escalate. On-line retrieval affords immediate, relevant, and economical means to secure exactly the data required by librarians and information specialists, researchers, executives, and technical personnel for their decisions.