Book Reviews

Theory and Application of Information Research. Edited by Ole Harbo and Leif Kajberg. London: Mansell Publishing, 1980. 235p. £16.00. ISBN: 0-7201-1513-2.

This book reproduces twenty-one papers presented at the Second International Research Forum on Information Science, which was held at the Royal School of Librarianship in Copenhagen during August of 1977. The title of this work may be misleading since the majority of the papers could better be described as the foundations of information science. The papers that advanced the theory of information science were the exception, and the contributions dealing with practical applications were even rarer. The contributors included many familiar names: Kathleen T. Bivins, Anthony Debons, William Goffman, Manfred Kochen, Allan D. Pratt, and Hans H. Wellisch from the United States; Nicholas J. Belkin, J. M. Brittain, B. C. Brookes, Robert A. Fairthorne, J.-M. Griffiths, M. H. Heine, S. E. Robertson, B. C. Vickery, and T. D. Wilson from the United Kingdom; and many names from Europe that may be less familiar on this side of the Atlantic. The forum was organized into five sessions: general models of information science, information science in relation to other scientific disciplines, measurement, the information retrieval process, and the future tasks of information scientists in Europe. Within the book, the distinction between these sessions generally is not obvious. Appendixes give the forum program, summarize the discussions of the papers, and report on group discussions.

In the introduction, it was stated that it was hoped that the forum would *bridge the gap between theory and research on one side and practice on the other*. The book does not fulfill this hope, but it does present a good collection of papers dealing with a variety of aspects in information science. The view that the main problems of information science are cognitive rather than technical is evident in many of the papers. However, Bradford's law, Shannon's theory, and the epidemic model are addressed in several of the papers. With a few exceptions, the papers are quite readable and do not require a mathematical background to be understood and appreciated. The summaries and group discussions are disappointing, possibly because several of the authors were unable to attend the forum. Kathleen Bivius was the only American contributor present. There is no index, although one would have been helpful.

The book is valuable and should be part of any library collection covering information science. Anyone interested in information science should be able to find several highly relevant papers. However, only a limited number of scholars will find it necessary to read the entire work.—*Edward T. O'Neill, Matthew A. Baxter School of Information and Library Science, Case Western Reserve University, Cleveland, Ohio.*

Personal Documentation for Professionals—Means and Methods, by V. Stibic. Amsterdam: North-Holland Publ. Co., 1980. 214p. $29.25 (Dfl 60.00). ISBN: 0-444-85480-0.

While there have been many a number of books written on the design, development, and use of large-scale database systems, there have been few that focus on the control of one's own personal collection of reprints, memoranda, reports, drafts, slides, and related miscellanea, which accumulate so rapidly in any professional "information-handler's" office.

Stibic's book addresses this problem in a thoroughly professional and competent manner. His first two chapters introduce the general nature of the problem, and discuss professionals' information needs and sources. The third, "Document De-

scription," covers the record structure, abstracting, subject descriptions, keywords and classification methods, and their various combinations. The fourth chapter details the various technical means for storage of original documents, microfilm, and such control mechanisms as card indexes, peek-a-boo cards, and computer-supported indexes.

All of these chapters draw on the experience and practices familiar to users of large-scale systems. Stibic recommends the use of ISO and other standardized practices, and endeavors to emphasize the need for constructing one's own system in accord with generally accepted design principles.

Stibic is careful to point out, however, that if one is in fact designing a *personal* documentation system, then personal idiosyncrasies and preferences can be built into it. It is not necessary to use an established and standardized vocabulary or classification system without modification. One may alter it to suit one's own purposes. However, the *structure* of the system (whether descriptors, classification numbers, or other means) must be controlled; otherwise the system will become useless.

The next four chapters are case studies of different systems. The first is a card index technique used by an individual. The second describes a computerized index to support the documentation needs of a project team. (Essentially an augmented KWIC index, published quarterly.)

The third case study is one of particular interest to many professionals at the moment—the use of a personal computer as an indexing control system. The system, though not explicitly identified, is roughly comparable to many of those available in the U.S.; a microcomputer with 64K RAM, a display of 80x24 lines, two floppy disks with 512K bytes/disk, and an 80-character-per-line printer. The indexing is done via a faceted classification system of about 250 terms, which are hierarchically linked, providing automatic up-posting from specific to generic terms. A hash-coding technique is used to minimize the storage space required on the disk, and searching is performed by simple serial searching of the index records.

The fourth case study is an examination of the upgrading of the manual card index described in the first study to a system supported by a large main-frame computer, using a terminal in the professional's office. A combination of automatic keyword extraction and manual classification is used for indexing. Complex Boolean searches are possible with this system.

Stibic concludes with a chapter on future prospects, touching briefly on such things as internal and public viewdata/teletext systems. He also provides a checklist of desirable features of "a multi-purpose personal work station." Such a station is not merely a special-purpose device used to aid in some parts of one's work, such as retrieval, but is an integral part of all of one's work; computer, calculator, text-processor, mail-dispatch system, calendar, in/out box, and so forth.

The author, a scientist of long standing with Philips in Holland, has provided a valuable guide to this area. There are two relatively minor points of criticism, however. Whether it was the author's or the publisher's choice is not clear, but there is an excessive use of italics throughout the text. This lavish use seems more appropriate to teenagers' romantic novels than to a serious work. In this case, it is more distracting than helpful. Secondly, but more understandably, the extensive references Stibic gives are frequently to documents not easily available in the U.S. Some are OECD papers, some refer to the German DIN standards, and some to internal Philips technical reports. These are minor points, however, regarding an excellent book. It is recommended not only for the information professional, but for anyone who is seriously concerned with the problem of keeping track of what one needs to know.—*Allan D. Pratt, University of Arizona Graduate Library School, Tucson.*

Viewdata Revolution, by Sam Fedida and Rex Malik. A Halsted Press Book. New York: Wiley, 1979. 186p. $34.95. LC: 79-23869. ISBN: 0-470-26879-4.

Sam Fedida is the inventor of Prestel,

the British Post Office's viewdata system. With this as his license, he and Rex Malik have written a 186-page volume explaining the Prestel system. Prestel is a series of databases, which are accessed by a keypad similar to a calculator. The common television takes on the characteristic of a CRT for viewing alphabetical and numerical information. The connection to the computer is by telephone, and, in Britain, the post office is in charge of the telephones. Overall, in spite of several printing errors, this book does provide information about the system.

The authors explain the types of information that will be available on the Prestel system, such as "Buying a Car," "Houses for Sale," "Entertainment," "Education," "An Evening Out," and "News." They have also devoted individual chapters to electronic mail, electronic funds transfer, and education, explaining how each works in the system.

The authors stress the benefits and attributes of their system almost to the point of redundancy. In each of the chapters, the manner in which the information is going to be accessed is repeated. Despite the repetition, the primary focus is what Prestel will do for the betterment of mankind.

The uniqueness of Prestel is the simplicity of its access process. According to the authors, being able to access the information in one's own home will make Prestel a major tool for dissemination of information for many agencies and businesses. At times, the "hard sell" is very obvious throughout the volume. However, the diagrams are good and help to explain the authors' points.

The problems Fedida and Malik anticipate in the electronic mail and protocols are realistic. In the chapters "Future I" and "Future II," the authors go off on a tangent, using a time line, on what they see in the future. Again, it is basically a repetition of what was said in the previous chapters, only from a futuristic point of view. Here, the reader gets a distinct feeling of what is really bothering them now in the system; that is, government bureaucracy. They cite the different groups trying to control the information by

means of legislation. They delve into the problem of uniformity of standards. Television is an example. What will be standard for convertors and adapters for the computer hookup? This is a real problem that was well explored throughout the work.

This volume is good for librarians who are interested in cable, telecommunications, and computers. However, be aware of its poor organization. There are numerous printing errors that affect its readability. Nevertheless, if a person can wade through these errors and the repetition of ideas, he/she can obtain some useful information from this text. There is a distinct feeling throughout this work that it was put together hastily. Nonetheless, there is a dearth of information on this subject, and this book will serve some useful purpose for libraries.—*Robert Miller, Memphis/Shelby County Public Library and Information Center, Memphis, Tennessee.*

ALA Filing Rules. Filing Committee, Resources and Technical Services Division, American Library Association. Chicago: American Library Assn., 1980. 50p. $3.50. LC: 80-22186. ISBN: 0-8389-3255-X.
Library of Congress Filing Rules. Prepared by John C. Rather and Susan C. Biebel. Washington, D.C.: Library of Congress, 1980. 111p. $5. LC: 80-607944 ISBN: 0-8444-0347-4. Available from Customer Services Section, Cataloging Distribution Service, Library of Congress, Washington, DC 20541.

These two works represent the culmination of over a decade of effort within the library profession to overhaul the techniques by which entries are arranged to form catalogs. The impetus for this work came from recognition that computer technology would soon be enlisted to perform the arrangement of entries for the production of catalogs, and that filing rules current at the time would be impossible to implement in their entirety on the computer.

Although the original intention was to develop rules appropriate for the arrangement of entries by computer, those at the Library of Congress and the ALA commit-

tee working on the problem soon realized that, from the point of view of catalog users, it would be very undesirable to have different sets of filing rules in operation depending on the physical medium of the catalog. Therefore, the scope of the effort was broadened to rules that could be applied both manually and by machine using headings that were formulated according to more than one set of cataloging rules.

Now that we have these new rules, the question arises whether they are better than what preceded them. The criteria for "better" ought to be whether the rules make entries easier to find both for known-item searches and browsing within the complex device called a library catalog. Or to state the same criteria negatively: it should be more difficult to lose an entry in the catalog if it has been filed according to the rules. The evaluation of these rules against other possible approaches to catalog arrangement ought to be centered on observation of the needs of a variety of both experienced and unsophisticated catalog users and on measurement of the effectiveness of the alternative approaches to meet these needs.

The complex problems of filing clearly exemplify the need for research as recently expressed by Herb White in his columns in *American Libraries*. Lacking any empirical data on which to base an evaluation, we must rely on our professional judgment and personal biases to argue the case for the new rules.

To this reviewer, it seems that common sense supports a set of rules that are simple, consistent, and easy to explain to library users. The need for simplicity and consistency directly implies the "file-as-is" principle (i.e., file exactly as the heading is visually constructed, not by some interpretation of it), which should be applied even at the cost of having to search in more than one place in the arrangement; e.g., numeric digits and numeric words, Mac and Mc, Müller and Mueller.

The file-as-is principle has been more consistently applied in the ALA rules than the LC rules, the latter undoubtedly a result of the anticipated complexity and size of LC's catalogs, although there is no justification argued for these departures from the basic principle.

Of specific interest to readers of the journal is whether these rules can be implemented for computer sorting of catalog entries. Do the rules succeed in meeting their original objective?

The ALA rules certainly appear to be amenable to very straightforward systems analysis and programming. For this the committee and its chairperson, Joe Rosenthal, need to be commended. From some sources there are already claims of systems that fully implement the new ALA rules, which certainly could be the case. However, it would be interesting to know how these systems deal with the following, which seem to be potentially troublesome:

• The lack of consistent support in the MARC format for handling initial articles when the rules call for ignoring initial articles in corporate names other than personal or place names, title subheadings ($t subfield), and subject headings. The English articles obviously present no problem, but the table of articles in appendix 2 shows more than thirty words that can be both an article and the cardinal numeral 1. In addition, the footnote, "In Hawaiian, the 'O emphatic' must be carefully distinguished from the preposition O, but O also serves the Hawaiian language as a noun and a verb (each with several meanings), an adverb, and a conjunction," must surely give pause to the diligent systems designer.

The recent Library of Congress practice of dropping nonfiling initial articles from heading fields still does not solve the problem of initial articles in the several million MARC records that already exist in library catalogs.

• The requirement that Roman numerals be filed numerically presents an opportunity to construct an interesting but not overly complex algorithm. However, although the MARC format makes the identification of Roman numerals in heading fields fairly straightforward (the $b subfield), the identification of Roman numerals embedded in a long title is much more ambiguous. For example, does *IV* mean "4" or "intravenous"?

• The rules require that punctuation in an Arabic numeral that is included to increase its readability is to be ignored in filing, but decimal points are significant in determining the numeric value of the number (i.e., .003 files before 1). How does one specify an algorithm to deal with the title, "5.000 kilomètres dans le sud"? Using European practice, this number is obviously 5,000, but why not 5 according to the computer algorithm?

• The special rule for nonroman alphabets (rule 7) is interesting: "If, in the arrangement of bibliographic records, it is necessary to distinguish access points containing characters in different nonroman alphabets, scripts and syllabaries (cf. Rule 1, Order of Characters) the following order of precedence is used. . . ." There follows a table beginning with Amharic and ending in Tibetan. That is the entire rule. Systems designers who have implemented this rule clearly have transcendent skills! Reliance on the MARC language code in the 008 field has both theoretical and practical problems.

• The introductory text advises libraries to include in the file information notes and references that explain filing practices to catalog users. However, the rules do not specify where these references are to file in relation to other headings. Admonishment to provide these at "appropriate points" is not much help.

• The ampersand is ignored in filing (for which we should be grateful). But, by including the optional rule 1.3, which allows filing the ampersand "as its spelled-out language equivalent," the ALA committee has put systems designers in the position of having to explain why this rule cannot be implemented on the computer—at least not until the MARC format includes a code for language of the field (not a likely development, and even then not all ambiguity would be eliminated). Interestingly, the Library of Congress treats all ampersands as a character filing between blank and the letter A.

• The optional rule 9.1, which allows the inclusion of "the role of a person or a corporate body in a legal action in arranging access points," presents a problem when the rule requires suppression of all

other relators. How is the computer programmed to recognize a legal action? Is there a finite list of such relator words? Differences between AACR2 and previous cataloging practices further complicate the use of this option.

Admittedly, many of these problems are marginal in terms of the number of entries in a catalog affected, but to a systems designer, even though there is only one instance, it must be accounted for in the computer programs if the system can claim a "full" implementation of the rules. Clearly, full implementation will require some changes in the MARC format before all rules can be applied absolutely consistently and unambiguously.

The Library of Congress rules, although applying similar principles, depart significantly from the ALA rules in detail and complexity. A full analysis of the implementation problems would require much more space than this review will allow. Suffice it to say that although the Library's LIBSKED program has been under development for twelve years, and its strengths and limitations have undoubtedly influenced the development of these filing rules, there are elements in these rules that have not yet been implemented in LIBSKED, and several where no one has yet figured out how to do it.

Although the work on these rules is complete, there are two more projects the profession should undertake that would be most useful for those concerned with catalog development. In both sets of rules, there is mention in the introduction of the need for a brief version of the essential rules, which could be handed out to catalog users. Why did the committee not develop such a brief guide and include it as an appendix to the rules? Those of us who work on computers are familiar with the reference cards for programming languages put out by computer manufacturers. A similar format for the filing rules would be very useful.

Another more difficult but equally useful project would be the publication of a standard design implementation of the ALA filing rules expressed in terms of the MARC format. Such a design would include the MARC fields and subfields

necessary for each possible entry from a bibliographic record and a description of any special processing required for particular data elements. The design would be expressed at a level that is independent of programming languages and computer hardware. We need a standard reference that translates the filing rules into the language of the MARC format. The ALA rules, in some tantalizingly brief instances, begin this process.

Both sets of filing rules are significant improvements over those previously available to systems analysts. Reference librarians should find these rules easy to explain to beleaguered catalog users. For their simplicity and relatively slight departure from the "file-as-is" principle, the ALA rules are to be recommended. The Library of Congress rules, in their attempt to retain the classificatory structures that support the browsing user, further complicate the task of the user performing a known-item search. Library research has indicated that the preponderance of catalog searches in research libraries are known-item searches.—*John F. Knapp, Ringgold Management Systems, Beaverton, Oregon.*