ALLAN D. PRATT

Libraries, Economics, and Information: Recent Trends in Information Science Literature

Four trends are noted in the information science literature: (1) much work in this field is being performed at or for academic libraries; (2) the concept of "information" as a process rather than a product is becoming more widely held; (3) both the economic and the conceptual "value" of information are under serious study; and (4) improved measures and models of bibliographic phenomena are being developed which reflect reality reasonably well.

OVER THE PAST SEVERAL YEARS information science and library science have been in some sense "competitive" with each other. Though the boundaries of both disciplines are ill-defined, and though each has aspects which are clearly distinct from the other, certain elements of each are indeed concerned with similar themes. It is the intent of this paper to demonstrate, through a survey of some of the recent literature of information science, the extent to which the two fields are mutually supporting and to suggest that this mutual support may lead to further progress in both.

While there is no generally accepted definition of information science, it is possible to adopt a tentative "definition by enumeration" by referring to the most useful literature survey in the field. A recent (seventh) *Annual Review of Information Science and Tech-*

Allan D. Pratt is a member of the faculty of the Graduate Library School, Indiana University, Bloomington, Indiana.

nology (ARIST VII hereafter) provides comprehensive coverage of all those topics which are, in the judgment of the editors, relevant to information science.[1] It may therefore be assumed to give a reasonably accurate picture of the interests and concerns of the workers in the field; presumably the topics and the relative emphasis among them reflect, at least roughly, their perceived "importance" in some way.

To gain some sense of perspective on the field, it is useful to compare the coverage of different topics in this recent volume with that of the first, published in 1966, *ARIST I.*[2]

Broadly speaking, the chapters of both *ARIST I* and *ARIST VII* may be divided into four sections:

1. Planning Information Systems and Services
2. Basic Techniques and Tools
3. Applications
4. The Profession

In both the first and this recent volume, sections two and three comprise the bulk of the chapters—nine of thir-

teen chapters, with 78 percent of the total references, in *ARIST I;* and ten of thirteen chapters, with 85 percent of the references, in *ARIST VII.* Between the first and seventh volume, however, there have been significant changes in the relative proportions of the sizes of these two sections. The citations in "Techniques and Tools" have decreased from 56 percent of the whole in *ARIST I* to only 25 percent in *ARIST VII.* On the other hand, the citations in "Applications" have increased from 22 percent to 61 percent of the total. Thus in the course of seven years, the applications literature has grown by a factor of almost three, while proportionately there is less than half as much literature relating to "Techniques and Tools."

LIBRARIES

Much more of the recent applications literature is oriented toward "conventional" libraries than was the case earlier. Two of the six applications chapters in *ARIST VII* deal with libraries directly: "Library Automation," which has been a fixture since *ARIST I,* and "Library and Information Networks." Half of the latter chapter's citations explicitly refer to libraries (mostly academic), while the remainder range over such matters as satellite communications, cable TV, the relative merits of different types of equipment, and the like. In addition, a third applications chapter is devoted to "Machine-Readable Bibliographic Data Bases." Though "data bases" (bibliographic or otherwise) did not appear in the index of *ARIST I,* there is nearly half a column of such entries in the index of *ARIST VII.* This interest in bibliographic data bases strongly suggests that libraries are deeply involved in their development and use.

The same increased interest in libraries is evident in two other chapters of *ARIST VII.* A significant proportion of the citations in both "Costs, Budget-

ing and Economics of Information Processing" and "Library and Information Center Management" refers explicitly to libraries (again, mainly academic).

In view of this phenomenon in the literature, it seems safe to conclude that the "conventional" library, far from being displaced by information centers or other new institutions as some predicted, has become the significant "applier" of information science and technology. In the seven years since *ARIST I,* librarians, or at least those who publish for or about libraries, have grasped the nettle firmly.

ECONOMICS

A second theme which is becoming ever more prominent in the information science literature is economics. Threaded through various chapters of *ARIST VII* are the echoes of the remark in the introduction to Part I: "Perhaps the most influential restrictions on information activities are economic constraints."[3] While *ARIST IV* and *V* had chapters on "Library and Information Center Management," as does *ARIST VII,* the latter is the first to have a chapter devoted to "Costs, Budgeting and Economics of Information Processing." Though chapters in previous volumes dealt with the evaluation of information systems, these were primarily concerned with relevance, recall, and similar parameters, not cost/benefit ratios. There is an increasing emphasis on cost accountability even in those organizations nominally nonprofit, e.g., the requirement that Department of Defense information agencies recover half their operating costs from use charges.[4]

This emphasis is becoming evident also within academic libraries. Buckland has sketched the beginnings of a "theory of the library" analogous to the economists' "theory of the firm," which, if developed, may form some basis for a better understanding of the economics

of information.[5] Some may feel that he has exercised excessive caution in attempting to avoid the assertion that libraries are a form of business, but his caution is understandable. If he stated flatly that libraries can and should be managed like businesses, the emotional reaction of his readers might well prevent their understanding the point he is making.

Another obvious indication of this increased interest in economics is the formation of the Information Industries Association (IIA) in 1969. This trade association consists of organizations, mostly profit-making in aim if not in fact, most pragmatically concerned with the value of information. One of the more interesting members of IIA is the FIND service in New York, which provides what is essentially telephone reference service to corporate customers. It claims to be able to "cost-account" reference service, which librarians have traditionally claimed to be either extremely difficult or impossible.[6] The validity of the claim will soon be reflected in FIND's profit and loss statements.

Other evidence of the new emphasis on economic matters was the theme of the 1973 ASIS convention: "Information: Benefits and Costs." This is a good deal more realistic than many previous conference themes, which tended to be somewhat grandiose: "A World of Information" (1972); "Communication for Decision Makers" (1971); or "The Information Conscious Society" (1970).

INFORMATION

In parallel with the increased emphasis on pragmatic economics in "real-world" libraries, there has been a notable increase in theoretical attempts to quantify, or at least to define more accurately, the "worth" of information.

In order to do so, it has first been necessary to overcome the conceptual obstacles presented by the view which Fairthorne termed the "Phlogiston Theory of Information": that "information is the name of some universal essence that can be squeezed out of texts like water from a sponge."[7] There is, encouragingly, less tendency than formerly to speak of "information" as though it were some form of intellectual electricity flowing through mental or social circuitry. The distinction between recorded discourse and the process of becoming "informed" is being more widely recognized.

That this should happen is understandable, because it is difficult to make sense of the notion of the "worth" of information or anything else without considering the question "Worth to whom, and under what circumstances?" While "the user" has always been an important consideration, at least in a statistical sense, in information systems design, it is only recently that the user as an individual has become an important element in the direct measurement of "amount of information."

One of the early papers recognizing the importance of the individual in this context was Ackoff's in 1958.[8] He defined three different kinds of messages: those which inform, those which instruct, and those which motivate a purposeful individual. The "amount" of information, instruction, or motivation related to a particular message received by this individual is a function not only of the message but of the individual's prior state. This same point has been repeatedly made by Fairthorne in various contexts. It was also made, in a tentative and informal way, by Pratt in 1967.[9]

Mitroff, Williams, and Rathswohl have constructed a model of a "Dialectical Inquiring System,"[10] the mathematical formulation of which is based on Ackoff's 1958 paper. They indicate divergencies in their viewpoint from that given in Yovits' 1969 paper in which he proposed that "information is data of use in decision making," and sketched out a model based on that defi-

nition.[11] However, the more recent work by Whittemore and Yovits suggests that the two conceptual structures are rather close, at least in some major respects.[12] There are strong similarities between Mitroff's "purposeful individual" who is moved from state S_1 to state S_2 and the DM (decision maker) of Whittemore and Yovits. The latter authors suggest six different varieties of uncertainty, none of which correspond to the classic "uncertainty" of the Shannon/Weaver model. What Whittemore and Yovits call "executional uncertainty" is broadly the same as Ackoff's "efficiency" variable incorporated into Mitroff's model.

The differences between these two conceptions seem to lie, at least in part, in the differing underlying aims. Mitroff speaks of "inquiry" with its connotation of more or less detached and philosophical ends, while Yovits' concept of a DM has a much more aggressive implication. It is rather as if the former is trying to model Bertrand Russell while the latter is trying to model J. P. Morgan.

Other indications of this more direct involvement of the individual in the "information" process can be found in recent work. Cooper proposes a conceptually simple, though methodologically difficult, measure based on the point that "there can be in principle no better evaluation . . . of a retrieval system's benefits than a subjective evaluation by its users, provided that such an evaluation is made with all due care and is measured judiciously in terms of an appropriate utile."[13] Some means of making his subjective evaluation method operationally viable are described in a more recent paper.[14]

Simon, in a quite different context, brings up the same general theme.[15] His point is that the critical resource is *not* "information" (which Yovits would call "data") but attention—that the decision maker (or inquirer or user) cannot afford to devote enough of his limited time to "attend to" all the data available. Thus, effort needs to be placed not on data gathering, but on organizing it. "Memories [including libraries, etc.] as components of information-processing systems, need to be viewed as stores of *potential* information, which, if indexed effectively, can become available at a reasonable cost whenever it is needed as input to a decision-making process."[16] (Emphasis in original.)

MEASUREMENT

Some recent applied mathematical work is progressing along parallel, or at least related, lines. There is, in addition to the mathematics associated with the cost-accounting problems noted earlier, a growing tendency to concentrate on the mathematical description of phenomena which are in fact quantifiable, a change from efforts in earlier years which purported to measure "relevance" to three decimals. One seems on firmer ground when measuring citations, volumes, journals, circulations, and the like than when measuring relevance.

In particular, there seems to be a growing belief that many bibliometric phenomena and their interrelationships may be described by some forms of exponential or logarithmic functions. The most general statement along these lines is Price's remark that

> . . . it turns out roughly that the true measure of value is not measured by the crude magnitude, but by its logarithm. A collection of related books or journals or papers gains the same degree of utility in going from 100 items to a thousand as in going from 1000 to 10,000 or from a million to ten million.[17]

This characteristic has been implicit for some time in the various studies of the distribution of journal articles in certain fields, under the general name of the Bradford-Zipf distribution. Williams found that the demand for periodical titles was well described by the

lognormal distribution (in which not the value of the variable under consideration, but its logarithm to some convenient base—usually e or 10—is normally distributed).[18] Palmour, Williams' co-worker, has stated that the same function predicted demand for documents at NTIS as well.[19]

Resnikoff and Dolby have found that the five-year circulation pattern of Harvard's Widener Library approximates a lognormal distribution fairly well.[20] (Incidentally, it was found that only 6 percent of the collection had circulated at least once during that time.) Resnikoff and Dolby have also used the Whitworth distribution, the discrete analog of the continuous logarithmic function, to describe the distribution of books over subject classes.[21] Relationships between the Log-Log (Zipf) distribution, the Whitworth, and the log-linear have been pointed out by Griffith and Krevitt.[22] Naranan has generalized some of these distributions in terms of power laws.[23]

These forms of distribution have been found useful, not only in studies of journals and citations but in circulation studies as well. In addition to the Resnikoff and Dolby work noted earlier,[24] Trueswell found similar distributions in his study of libraries at Northwestern.[25] He was able to make reasonably accurate predictions of collection use from his models, and found that something in excess of 50 percent of the collection could be removed with little impact on the user satisfaction level.

It appears that these power-function distributions underlie a large range of phenomena related to bibliographic and information science phenomena. They are of interest not only because they may provide some empirical support for the theories mentioned above but because they may have direct applications in quantifying some aspects of librarianship.

There has been no recent explication of these distributions published for the benefit of those who are more literate than "numerate." Fairthorne's historico-bibliographic essay of 1969 is the only paper which summarizes work up to that time in reasonably nonmathematical terms.[26] There is need of a similar effort covering the last five years because if the applied mathematics is to be in fact applied, it must be understandable to those who would apply it.

SUMMARY

This examination of recent information science literature suggests the emergence of four trends which may be indicators of new directions:

1. Much of the research and development work in information science is being conducted by, for, or under the direction of "conventional" (mainly academic) libraries.

2. There is increasing awareness that "information" is not a property of documents but a complex relationship between (at least) message and recipient.

3. There is serious consideration being given, on both a theoretical and an applied level, to the question of the "worth" of information.

4. There is some progress being made in bibliometrics—the development of means whereby one can measure and predict library-related phenomena with models which reflect reality reasonably well.

These trends, if continued, may raise information science from the low level to which Saracevic felt it had fallen by 1971. In five years of work, he asserted, "The basic literature information problems were not solved, or even dented, especially in comparison to the efforts and money invested."[27] There is room for cautious hope that some small dents may soon appear.

REFERENCES

1. Carlos A. Cuadra, ed., *Annual Review of Information Science and Technology* [Hereafter referred to as *ARIST*] *Volume 7, 1972* (Washington, D.C.: American Society for Information Science, 1972).

2. Carlos A. Cuadra, ed., *ARIST, Volume 1, 1966* (New York: Wiley, 1966).

3. Cuadra, *ARIST, Volume 7, 1972*, p.4.

4. Ibid., p.39.

5. Michael Buckland, "Toward an Economic Theory of the Library," in Robert S. Taylor, ed., *Symposium on the Economics of Information Dissemination* (Syracuse, N.Y.: Syracuse Univ. Pr., in press).

6. Paul Doebler, "Seek and Ye Shall FIND," *Publishers' Weekly* 202:79–80 (Oct. 16, 1972).

7. Robert A. Fairthorne, "'Use' and 'Mention' in the Information Sciences," in Laurence B. Heilprin, Barbara E. Markuson, and Frederick Goodman, eds., *Proceedings of the Symposium on Education for Information Science, Warrenton, Virginia, September 7–10, 1965* (Washington, D.C.: Spartan Books, 1955), p.10.

8. R. L. Ackoff, "Towards a Behavioral Theory of Communication," *Management Science* 4:218–34 (1958).

9. Allan D. Pratt, "An Essay on the Nature of Information," *Proceedings. American Documentation Institute. Annual Meeting, Volume 4.* (Washington, D.C.: Thompson Book Company, 1967), p.3–7.

10. Ian Mitroff, James Williams, and Eugene Rathswohl, "Dialectical Inquiring Systems," *Journal of the American Society for Information Science* 23:374–78 (Nov.-Dec. 1972).

11. M. C. Yovits, "Information Science: Toward the Development of a True Scientific Discipline," *American Documentation* 20:374–76 (Oct. 1969).

12. Bruce J. Whittemore and M. C. Yovits, "A Generalized Conceptual Development for the Analysis and Flow of Information," *Journal of the American Society for Information Science* 24:221–31 (May-June 1973).

13. William S. Cooper, "On Selecting a Measure of Retrieval System Effectiveness," *Journal of the American Society for Information Science* 24:100 (March-April 1973).

14. William S. Cooper, "On Selecting a Measure of Retrieval Effectiveness. Pt II," *Journal of the American Society for Information Science* 24:413–24 (Nov.-Dec. 1973).

15. Herbert Simon, "Applying Information Technology to Organization Design," *Public Administration Review* 33:269–70 (May-June 1973).

16. Ibid., p.272.

17. Derek J. De Solla Price, "Some Remarks on Elitism in Information and the Invisible College Phenomenon in Science," *Journal of the American Society for Information Science* 22:74 (March-April 1972).

18. Gordon Williams et al., *Library Cost Models: Owning Versus Borrowing Serial Publications* (Chicago: Center for Research Libraries, 1968). (PB 183 304).

19. Vernon Palmour, Personal communication, June 21, 1972.

20. H. L. Resnikoff and James L. Dolby, *ACCESS: A Study of Information Storage and Retrieval with Emphasis on Library Information Systems* (Los Altos, Calif.: R & D Consultants Company, 1971). (ED 050 773).

21. H. L. Resnikoff and James L. Dolby, "On Archival Access," in Jeanne B. North, ed., *Proceedings of the American Society for Information Science, Volume 7: The Information Conscious Society* (Washington, D.C.: American Society for Information Science, 1970), p.255–64.

22. Belver C. Griffith and Beth Krevitt, "A Comparison of Several Zipf-Type Distributions in Their Goodness of Fit to Language Data," *Journal of the American Society for Information Science* 23:220–21 (May-June 1972).

23. S. Naranan, "Power Law Relations in Science Bibliography—A Self-Consistent Interpretation," *Journal of Documentation* 27:83–97 (1971).

24. Resnikoff and Dolby, *ACCESS.*

25. Richard W. Trueswell, "Determining the Optimal Number of Volumes for a Library's Core Collection," *Libri* 16:49–60 (1966).

26. Robert A. Fairthorne, "Empirical Hyperbolic Distributions (Bradford-Zipf-Mandelbrot) for Bibliometric Description and Prediction," *Journal of Documentation* 25: 319–43 (1969).

27. Tefko Saracevic, "Five Years, Five Volumes and 2345 Pages of the *Annual Review of Information Science and Technology*," *Information Storage and Retrieval* 7:138 (Sept. 1971).