A tour of information science through the pages of *JASIS*

Following are selected article titles and descriptive material drawn from the pages of the *Journal of the American Society for Information Science (JASIS)* and its precursor title, *American Documentation*, dating from the beginning of the journal in January 1950 until this issue went to press in spring 1999. All citations through the end of 1969 can be found in *American Documentation*; all later citations are from *JASIS*. The “tour” draws on regular and “brief communication” refereed articles. At the end of the tour are provided lists of the “Perspectives” and “Special Topic” issues that have appeared in the journal over the years. (See that section for further explanation of the special issues.)

Selections were made with the intent to illustrate the development of research, technology, and thinking in information science over the years. The pages of the journal have formed the crossroads for most, if not all, major researchers in the field. Famous researchers from outside information science proper, such as Noam Chomsky, Herbert Simon, and George Miller have also appeared in *JASIS*.

Furthermore, every significant area of research in information science is well represented in the pages of the journal. Information retrieval system design and evaluation, description of operational information systems and services, indexing theory and evaluation, search strategy, information seeking research, bibliometric analysis, information policy, and the

economics of information have all been addressed in *JASIS*. Despite the journal's reputation for an emphasis on the mathematical and technical, the first user needs study appears in *American Documentation* as early as 1951.

It is also striking how *early* many crucial ideas in information science appeared in these pages. H.P. Luhn proposes key word in context indexing in 1960, and selective dissemination of information in 1961—the latter now rediscovered in the age of the World Wide Web as “push” technology. William Goffman and associates propose the use of “meta-language” in information retrieval systems in 1964—and mean it in a way close to its current usage. In 1966, Lockheed staff describe a new system that enables a “dialogue” with an “on-line reference retrieval system.” Murray Turoff and Starr Hiltz provide a “progress report” on electronic journals in 1982.

User-centered information system design did not start in the 1980’s, as some would claim. Rather, it dates at least to Mooers’ Law, published in 1960, and Edwin Parker’s 1966 argument that “the system should adapt to the receiver or user, rather than the user to the system” (see below under Information Seeking and Needs). Victor Rosenberg’s 1974 article calling for the introduction into information science research the “intuitive, the subjective, and the experiential” presages by decades the current wave of interest in qualitative methodology.

Just two more examples will be provided here of these early insights. In 1967, Burton Adkinson and Charles Stearns present a remarkably prescient description and forecast of the information scene. Their three stages of library automation (which actually address far more than just libraries themselves) are an excellent summary of what, in fact, has happened in the years since their article.

Finally, Watson Davis, the founder of this Society (under its original name, the American Documentation Institute), describes in a 1951 article the association’s system of “auxiliary publication.” This arrangement, first begun in 1936, enabled researchers to send in additional, supplemental, and otherwise unpublished materials to the American Documentation Institute. Other researchers who wished to use this material could then contact the Institute and have a copy made for their own use. The ability to self-publish such supplemental materials is one of the features most praised about the Internet for researchers today. The technology available today is much more supportive for auxiliary publication than was the case in 1936, but the idea was actually implemented back then.

In the early days, the technology for information retrieval was almost unbelievably primitive by current standards. Coordinate indexing was developed by Mortimer Taube and his associates by using a paper card for each index term. (This so-called “term-entry” system was the converse of
the “item-entry” approach illustrated by Calvin Mooers’ Zatocoding, displayed at the beginning of this anniversary section.) In indexing individual records, the document number of each record would be posted to all the cards bearing index terms appropriate to the document. The cards would then be filed. Upon need to fulfill a search query, the cards for the desired subject terms would be pulled out from their boxes and the document numbers on each card visually compared. A coordinate “AND” between two index terms would be achieved when all document numbers were identified that appeared on both index term cards. In general, in the early days, the ideas galloped well ahead of the available technology.

Whole research specialties have been spawned from the research presented in the journal’s pages. Information retrieval research developed out of the work of Cyril Cleverdon, F.W. Lancaster, Lauren Doyle, and, above all, Gerard Salton, all of whom published in JASIS. (It is not being claimed that JASIS necessarily published the first or only article on the given topics, but articles in the journal have frequently been very influential in the development of information science research areas.) Bibliometric research reached its modern form out of the work of Burton and Kebler, M.M. Kessler, Eugene Garfield, Derek de Solla Price, Henry Small, and Howard White. Similar lists of names could be provided for research in both human and automatic indexing, search strategy, analysis of online database use and users, and information seeking.

We may expect to see some recent JASIS articles prove to be the productive sources of new research specialties and methodological approaches in the years to come—Birger Hjørland and Hanne Albrechtsen’s domain analysis; new approaches to studying information seeking from Elfreda Chatman, Carol Kuhlthau, Barbara Wildemuth, Rob Kling, and Carole Palmer; new kinds of libraries from Richard Lucier, Joseph Janes, and Edward Fox. We may see a revolution in thinking about relevance arising out of Stephen Harter’s 1992 article on the subject. Researchers studying new types of real-world operational information systems appearing today are too numerous to mention; it is anyone’s guess what role multimedia, imagebases, search engines, intelligent agents, and, of course, the Internet, will take in the information retrieval world of the 21st century.

The above are just a few of the names and topics to be found in the pages of American Documentation/Journal of the American Society for Information Science. Indeed, even the lengthier “tour” of the pages of the journal that follows below is just a small part of the full history and contents of the journal. All 50 volumes of the journal were examined to select the items in the tour. I have endeavored to represent a wide range of topics, perspectives, and individuals in the selection of articles, but ignorance,
biases rooted in personal interests, and general human failings will no
doubt be responsible for the omission below of significant articles or the
selection of lesser articles by researchers otherwise represented.

In the following, the descriptions are arranged by subject and chrono-
logically within each subject. Article titles are in boldface to make it easy
to scan down through the listings to see the topic development in each
research area. An unconventional bibliographic format for the article
information is used for ease of reading and follow-up.

To simplify, quotation marks are left out, but all text is taken directly
from the articles; none is supplied by the editor. Text is usually from the
abstracts, but sometimes comes from the text of the articles. (In the early
years the journal did not require abstracts for articles.) Elisions in the
selected texts are indicated by . . . . No text is provided where titles are
self-explanatory.

Information and information science paradigm

1964 Identifying Key Contributions to Information Science. Carlos A.
The present attempt to identify key contributions suggests that we are
far from common agreement on the conceptual, methodological, or
practical contributions to the information science field.

Information science is that discipline that investigates the properties and
behavior of information, the forces governing the flow of information,
and the means of processing information for optimum accessibility
and usability.

25(4), 263–269.
I believe that the notion of man as a mechanical device and the notion
that human behavior can be perfectly replicated by computers are
fundamental wrong. . . . A more holistic approach is needed. . . . We
must begin to look at the interrelationships between various parts
of the information environment. . . . To deal effectively with the
transcendent qualities of human communication we must admit as
scientific evidence the intuitive, the subjective, and the experiential.

1976 Information Science and the Phenomenon of Information. Nicholas
J. Belkin, & Stephen E. Robertson 27(4), 197–204.
. . . the fundamental phenomena of information science are deduced:
the text and its structure, the structure of the recipient and changes
in that structure, and the structure of the sender and the structuring
of the text.
The conventional interpretation of Shannon’s mathematical theory of communication in relation to textual material is unduly restrictive and unhelpful. A reinterpretation which is based on the definition of new symbol sets, comprising approximately equally-frequent strings of characters, is presented.

The concept of entropy, from the second law of thermodynamics, has been used by numerous writers on information theory.

Three meanings of ‘information’ are distinguished: ‘information-as-process’; ‘information-as-knowledge’; and ‘information-as-thing’...

There is a growing recognition that information and communication are interrelated in very fundamental ways.


The theory, which is based on Situation Theory, is expressed with a calculus defined on channels. The calculus was defined so that it satisfies properties that are attributed to information and its flows. This paper demonstrates the connection between this calculus and Information Retrieval, and proposes a model of an Information Retrieval system based on this calculus.

Theoretical information retrieval system design

The lengthening time required for searching larger and larger indexes and for consulting expanding collections of abstracts has become a source of mounting concern. Modern automatic equipment is able to scan and recognize index entries made in the form of holes on punched cards, magnetized spots on magnetic tape. . . .

This article discusses the possibility of exploiting the statistics of word co-occurrence in text for purposes of document retrieval. . . . It is shown that the most strongly co-occurring word pairs, which are therefore ‘associated’ in a statistical sense, can be represented in the form of an ‘association map.’
<table>
<thead>
<tr>
<th>Year</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A generalized method is given for performing information retrieval by computer. . . A block process diagram, a detailed flow chart for application to a computer, and a description of the application are provided.</td>
<td></td>
</tr>
<tr>
<td>1965</td>
<td>The Evaluation of Automatic Retrieval Procedures—Selected Test Results Using the SMART System.</td>
<td>Gerard Salton</td>
</tr>
<tr>
<td></td>
<td>The SMART automatic document retrieval system is used to study association procedures for automatic content analysis. . . There is little overlap between word relationships found through associations and those used in thesaurus construction, and the effects of word associations and a thesaurus in retrieval are independent. . . A properly constructed thesaurus, however, offers better performance than statistical association methods.</td>
<td></td>
</tr>
<tr>
<td>1974</td>
<td>A Clustering Algorithm Based on User Queries.</td>
<td>Clement T. Yu</td>
</tr>
<tr>
<td></td>
<td>The imprecision of some of the concepts which are used in formal models in information science has led to a spate of attempts to apply fuzzy set theory to aspects of information science. An analysis of the various kinds of imprecision that can occur indicates strongly that fuzzy set theory is not an appropriate formalism for these models.</td>
<td></td>
</tr>
<tr>
<td>1981</td>
<td>A Comparison of Two Systems of Weighted Boolean Retrieval.</td>
<td>A. Bookstein</td>
</tr>
<tr>
<td></td>
<td>A major deficiency of traditional Boolean systems is their inability to represent the varying degrees to which a document may be written on a subject. In this article we isolate a number of criteria that should be met by any Boolean system generalized to have a weighting capability.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The proposed similarity measure, based on a fuzzy Hausdorff distance, estimates the mismatch between two possibility distributions.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The system uses a novel architecture to allow more than one system facility to be used at a given stage of a search session. Users influence the system actions by stating goals they wish to achieve, by evaluating system output, and by choosing particular facilities directly.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A new method for automatic indexing and retrieval is described. . . The particular technique used is singular-value decomposition, in</td>
<td></td>
</tr>
</tbody>
</table>
which a large term by document matrix is decomposed into a set of ca. 100 orthogonal factors from which the original matrix can be approximated by linear combination.

This article reports that clusters of co-relevant documents obtain increasingly similar descriptions when a genetic algorithm is used to adapt subject descriptions so that documents become more effective in matching relevant queries and failing to match nonrelevant queries.

These newer techniques, which are grounded on diverse paradigms, have provided great opportunities for researchers to enhance the information processing and retrieval capabilities of current information storage and retrieval systems.

The author explores inter-record linkage relationships of a bibliographic hypertext system through the use of descriptor term co-occurrences. The observed distribution of term co-occurrences using term sizes follows a complex, but regular relationship.

1997  **Clustering and Classification of Large Document Bases in a Parallel Environment.** Anthony S. Ruocco, & Ophir Frieder *48*(10), 932–943.
Development of cluster-based search systems has been hampered by prohibitive times involved in clustering large document sets. We propose the use of parallel computing systems to overcome the computationally intense clustering process.

After a brief review of fuzzy sets, rough sets, and a fuzzy logical model for information retrieval, rough approximations for retrieved data are defined. The approximations are considered for both crisp and fuzzy cases.


**Information retrieval system evaluation**


1968  **Expected Search Length: A Single Measure of Retrieval Effectiveness Based on the Weak Ordering Action of Retrieval Systems.** William S. Cooper *19*(1), 30–41.
The new measure is based on calculations of the expected number of irrelevant documents in the collection which would have to be searched through before the desired number of relevant documents could be found.

1968  **The Browser’s Retrieval Game.** Siegfried Treu *JASIS* 19(4), 404–410.
The user starts with any single term(s) or term pair(s), ‘browses’ through the available index term vocabulary ‘aisles’ by means of a display console, and finally arrives at the game’s objective: Either a query in terms of pair(s) of highly associated terms or the conclusion that the subject of interest is not represented.

1969  **Effectiveness of Information Retrieval Methods.** John A. Swets *JASIS* 20(1), 72–89.
Results of some 50 different retrieval methods applied in three experimental retrieval systems were subjected to the analysis suggested by statistical decision theory. The analysis validates a previously proposed measure of effectiveness and demonstrates its several desirable properties.

The state-of-the-art of evaluation study in information science is analyzed with respect to 1. the scope of evaluation studies; 2. the use of laboratory-type environments; 3. the use of surrogate judges; 4. selection of variables; 5. frequency of study; and 6. comparability of study results.

The Swets model of information retrieval, based on a decision theory approach, is discussed. . . . It is seen that when the variances are unequal, the Swets rule of retrieving a document if its $Z$ value is large enough is not optimal.


1982  **Optimal Values of Recall and Precision.** Tore Olafsen, & Libena Vokac *JASIS* 33(2), 92–96.
When the cost structure is known, the optimal value of recall and the optimal value of precision can be found.

The study finds that there is no operational difference between the relevance-theoretic and the utility-theoretic model of evaluation. It further suggests the need for performance measures based upon a complex set of factors including document and information attributes, the judge, and other environmental factors, particularly the effects of learning during the evaluation process.

The article . . . reports the results of an experiment that used a generalized geographic information system to illustrate how collecting and analyzing data simultaneously from both a system and user views of performance can suggest improvements for information systems.

These problems include the question of the necessity of using real users, as opposed to subject experts, in making relevance judgments, the possibility of evaluating individual components of the retrieval process, rather than the process as a whole, the kinds of aggregation that are appropriate for the measures used . . . the difficulties in evaluating interactive systems, and the kinds of generalization which are possible from information retrieval tests.

The problem of measurement in information retrieval research is traced to its source in the first retrieval tests . . . Finally, it is concluded that the original vision of information retrieval research as a discipline founded on quantification proved restricting for its theoretical and methodological development and that increasing recognition of this is reflected in growing interest in qualitative methods in information retrieval research in relation to the cognitive, behavioral, and affective aspects of the information retrieval interaction.

The TREC (Text Retrieval Conference) experiments were designed to allow large-scale laboratory testing of information retrieval techniques . . . groups within TREC have become increasingly interested in finding ways to allow user interaction without invalidating the experimental design.

Markov chain models are derived from the empirical data on the use of the system.

Relevance
Analysis of the problems of defining the mutual relevancies of queries and document-collections indicates that they essentially involve the problem of conceptual relatedness.
A tour of information science through the pages of JASIS

1964 The Consistency of Human Judgments of Relevance. A. Resnick, & T.R. Savage 15(2), 93–95. The results showed that . . . humans are able to make such judgments consistently.

1975 RELEVANCE: A Review of and a Framework for the Thinking on the Notion in Information Science. Tefko Saracevic 26(6), 321–343. In the most fundamental sense, relevance has to do with effectiveness of communication.

1987 Pictures of Relevance: A Geometric Analysis of Similarity Measures. William P. Jones, & George W. Furnas 38(6), 420–442. A geometric analysis is advanced and its utility demonstrated through its application to six conventional information retrieval similarity measures and a seventh spreading activation measure.

1992 Psychological Relevance and Information Science. Stephen P. Harter 43(9), 602–615. The objective sense of relevance as ‘on the topic’ is curious, in that it is far removed from the way in which we use the words ‘relevant’ and ‘relevance’ in their everyday senses.

1994 User-Defined Relevance Criteria: An Exploratory Study. Carol L. Barry 45(3), 149–159. The results indicate that the criteria employed by users included tangible characteristics of documents . . . subjective qualities (e.g., agreement with the information provided by the document) and situational factors . . .


Design and evaluation of operational information retrieval systems

1950 Multiple Coding and the Rapid Selector. Carl S. Wise, & James W. Perry 1(2), 76–83. Punched cards have permitted us to take the first steps toward developing systems of information analysis capable of permitting searches to be directed to new, unforeseen combinations of entities, concepts and operations . . . The coding method to be considered, as developed by one of us (CSW) for ‘Keysort’ cards, permits coding to be based on a vocabulary of 456,976 concepts, of which any sixteen (or less) may be coded on a single card.

1959 Information Storage and Retrieval Using a Large Scale Random Access Memory. J.J. Nolan 10(1), 27–35. The unit shown . . . is the first really large scale random access memory device. It is the IBM RA-MAC . . . The arm may be moved vertically to any of the 50 disks and laterally to any of 100 concentric recording
tracks. Each track has a capacity of 1000 characters, 500 on the top surface and 500 on the bottom of each disk. Each disk, therefore, has a capacity of 100,000 characters, or a total of five million for the 50 disks.


1961 The Historical Development and Present State-of-the-Art of Mechanized Information Retrieval Systems. Charles P. Bourne 12(2), 108–110. The last ten years of equipment development and the application of mechanization techniques are reviewed for each of several functionally separate approaches, such as punched-card systems, computer systems, and magnetic media systems.

1961 Selective Dissemination of New Scientific Information with the Aid of Electronic Processing Equipment. H.P. Luhn 12(2), 131–138. Improvement of scientific communication is sought through machine assisted dissemination of new information. A service system is described in which a new document is characterized by a vocabulary or pattern of keywords. This pattern is then compared with the vocabularies or profiles characterizing each of the participants of the service. If a given degree of similarity exists between the two, the affected participants are notified by a card carrying an abstract. . . Profiles are kept current by discarding patterns after they have reached a certain age.


1969 MEDLARS: Report on the Evaluation of Its Operating Efficiency. F.W. Lancaster 20(2), 119–142. The Medical Literature Analysis and Retrieval System (MEDLARS) . . . is a multipurpose system, a prime purpose being the production of Index Medicus and other recurring bibliographies. However, the present study concentrated on the evaluation of the demand search function (i.e., the conduct of retrospective literature searches in response to specific demands).

A combination of internally written PL/1 and COBOL programs and the INQUIRE® data management system provide the software required to perform the operations on an IBM 370/155 computer with one megabyte of core storage.


Testing of a Natural Language Retrieval System for a Full Text Knowledge Base. Lionel M. Bernstein, & Robert E. Williamson 35(4), 235–247. . . . a retrieval system which combines use of probabilistic, linguistic, and empirical means to rank individual paragraphs of full text for their similarity to natural language queries proposed by users.


MINDS (Multiple Intelligent Node Document Servers) is a distributed system of knowledge-based query engines for efficiently retrieving multimedia documents in an office environment of distributed workstations.


A new query processing strategy called nonmaterialized relations is proposed which should significantly improve the ability of relational systems to support IR applications without requiring changes to the SQL query language interface.


Imaging: Fine Arts. Howard Besser 42(8), 589–596.

We look at general image quality issues including image capture, resolution, and display, and then turn our attention to the uses of imaging technology for conservation and preservation purposes.


The creation of guides to resources on the Internet specific to particular subjects have raised a number of interesting questions and issues.


There are many approaches, both theoretical and implemented, to using intelligent software agents for information retrieval purposes. These approaches range from desktop agents specialized for a single user to networks of agents used to collect data from distributed information sources, including Web sites.


Indexing systems and techniques


During the past six months, Documentation Incorporated has established an experimental coordinate index under a research program made possible by the Armed Services Technical Information Agency (ASTIA)…. The procedure was not to re-index the reports but to use the subject headings on the TID and DSC cards as the basis for developing appropriate terms for two coordinate indexes, one for each group of cards…. The method of coordinate indexing was first described in two papers prepared about two years ago by Dr. Taube.


…. the employment of a large number of terms, when indexing a collection of documents, must somehow take account of the existence of synonyms. On the other hand, the employment of a comparatively small number, particularly if the notions represented by the terms are not supposed to overlap, makes the indexing process considerably more difficult. These disadvantages can be avoided if a thesaurus is employed…. 
A distinction is made between bibliographical indexes for new and past literature based on the willingness of the user to trade perfection for currency. Indexes giving keywords in their context are proposed as suitable for disseminating new information.

This report describes a method of indexing documents which is based on the assumptions, (1) that a subset of the words in a document can be an effective index to that document and, (2) that this subject can be approximated by selecting those words from the document whose frequencies are statistically unexpectedly high.

Eleven vocabularies were analyzed and integrated.

The PREserved Context Index System (PRECIS) developed for subject indexing for the British National Bibliography is discussed as a basis for various studies relating to classification which could be made from its initial phrases, strings, entries and back-up structure.

The system uses an experimental frame-based knowledge representation language, FrameKit, implemented in Franz Lisp. The initial prototype is designed to interact with trained MEDLINE indexers who will be prompted to enter subject terms as slot values in filling in document-specific frame data structures that are derived from the knowledge-base frames.


Results show that approximately 88% of the access points selected by the Library of Congress or the National Library of Medicine could be automatically derived from title page data.

Visual documents—motion sequences on film, video-tape, and digital recordings—constitute a major source of information for the Space Agency. . . . This article describes a method for automatically selecting key frames from visual documents. These frames may in turn be used to represent the total image sequence. . . . The performance of the abstracting algorithm reduces 51 minutes of video sequences to 134 frames: a reduction of information in the range of 700:1.
Under this algorithm, words and phrases containing geographic place names or characteristics are extracted from a text document and used as input to database functions which use spatial reasoning to approximate statistically the geoposition being referenced in the text.

We discuss whether it is feasible to build intelligent rule- or weight-based algorithms into general-purpose software for interactive thesaurus navigation. . . . The results cause us to question many of the assumptions made by previous researchers in this area.

The salient components of this system include texture feature extraction, image segmentation and grouping, learning similarity measure, and a texture thesaurus model for fast search and indexing.

**Indexing theory and evaluation**

With the obvious deficiencies of traditional subject headings, journal-type indexes, and classification schemes so ever-present, we are being offered three basic solutions: 1. punched cards, both edge-punched and field punched, and either hand or machine sorted, 2. the Batten or Peek-a-boo system, and 3. Coordinated Indexing using Uniterms. There are, in addition, some embryonic systems which store information magnetically, or on film, or by means of cathode ray tubes and which require expensive machines for their operation.

Many of the problems that arise when one tries to construct a device that generates word sequences directly, vanish when one describes sentences in terms of phrase structure. . . .

Library classifications display lists of labels for texts, but rarely tell how to bestow the right label on the right text. . . . In other words, things are assumed created with name tags round their necks, and all we have to do is to catch them and read the tags. . . . This does not quite agree with practical experience. . . .

A study of several thousand entries in a classified bibliography revealed that a large proportion of the title entries contained words identical to or synonymous with words of the corresponding subject heading. It is inferred that a major part of the bibliography studied could have been compiled by a machine procedure operating on titles alone, provided the machine were supplied with a suitable synonym dictionary.

The ... meta-language terms here apply to the written records and their relations, rather than to the subject treated by these records. Examples of such terms are diagram, rehash, report, reprint, preprint, flow chart, analysis, description, discussion, table, definition, abstract. ...

It is well known that any two indexers, indexing one and the same document individually, will select sets of indexing terms which are most unlikely to be identical.

Prompt literatures of organized terse conclusions may increase ability to keep up in a subject, reduce need for translation, and make information available promptly.


An experimental framework for observing an indexing task carried out in a laboratory environment is described. The computer-based interactive Negotiated Search Facility served as the instrument which the experimenter used to monitor the indexing task and the tool which the indexer used to carry out the task.


Conventional information systems, founded on objectivist assumptions, are inappropriate for social scientists.

This article reports on a comparison of the postcoordinate retrieval effectiveness of two indexing languages: LCSH and PRECIS. The effect of augmenting each with title words was also studied.

State-of-the-art information-retrieval systems are found to combine the functional capabilities of the conventional inverted file—Boolean logic—term adjacency approach, with statistical-combinatoric techniques pioneered in experimental information-retrieval research and formal natural-language processing methods and tools borrowed from artificial intelligence.

Design features presented are an access phase, including entry and orientation, a hunting phase, and a selection phase. An end-user thesaurus and a front-end system mind are presented as examples of online catalog system components to improve searcher success during entry and orientation.

1990  All the Right Words: Finding What You Want as a Function of Richness of Indexing Vocabulary. Louis M. Gomez, Carol C. Lochbaum, & Thomas K. Landauer 41(8), 547–559.
The implications of index-word selection strategies for user success in interactive searching were investigated in two experiments. . . . The results demonstrate that searcher success is markedly improved by greatly increasing the number of names per object.

Based on these requirements, information-theoretic considerations, and the Gray code, a classification system is proposed that can classify documents without human intervention. . . . The proposed system can incorporate both classification by subject and by other forms of bibliographic information, allowing for the generalization of browsing to include all features of an information carrying unit.


This article . . . defines the characteristics of indexing that affect retrieval—namely, indexing devices, view-point-based and importance-based indexing exhaustivity, indexing specificity, indexing correctness, and indexing consistency—and examines in detail their effects on retrieval.

The experiment showed that the thesaurus was an excellent ‘memory-jogging’ device and that it supported learning and serendipitous browsing.


Two important issues have been explored: whether thesauri enhance the retrieval effectiveness of Chinese documents, and whether automatic indexing can compete with manual indexing in a Chinese information retrieval system.


Search strategy and evaluation


A principle of sequential optimization in search theory distributes the search effort at each stage so as to maximize the probability of target detection with the effort expended thus far.


Question analysis and search strategy development . . . were characterized as nine decision making steps.


Twenty-nine tactics are named, defined, and discussed. . . .


The way that individuals construct and modify search queries on a large interactive document retrieval system is subject to systematic biases similar to those that have been demonstrated in experiments on judgments under uncertainty. These biases are shared by both naive and sophisticated subjects. . . .


1984 Online Searching Styles: A Case-Study-Based Model of Searching Behavior. Raya Fidel 35(4), 211–221.
The model of operationalist and conceptualist searching styles describes searching behavior of experienced online searchers.

1985  
An Investigation of Online Searcher Traits and Their Relationship to Search Outcome. Trudi Bellardo 36(4), 241–250.
The notion that searching performance can be predicted by or is dependent upon certain cognitive or personality traits has thus become highly suspect.

1986  
Why are Online Catalogs Hard to Use? Lessons Learned from Information-Retrieval Studies. Christine L. Borgman 37(6), 387–400.

1988  
A concluding summary of all results is presented in Part III.

1989  
Analysis of search patterns showed that novices used a heuristic, highly interactive search strategy.

1993  
The results showed that search experience affected searchers’ use of many search tactics, and suggested that subject knowledge became a factor only after searchers have had a certain amount of search experience.

1993  

1994  

1995  
… indicating that a hypertext-based approach to bibliographic retrieval could be appropriate for a variety of searcher experience levels.

1995  
The SLC approach overcomes problems with several searching features that are difficult for children in typical keyword OPAC systems: typing skills, spelling, vocabulary, and Boolean logic.

1996  
Affective questions outnumbered cognitive and sensorimotor questions by two to one. This preponderance of affective micro-information needs during searching might be addressed by new system functions.


**Interface design**

The interface allows users to make requests in a common language. These requests are translated by the interface into the appropriate commands for whatever system is being interrogated.


The following features of online public access catalogs are discussed: the demonstration of relationships between records, the provision of entry vocabularies, the arrangement of multiple entries on the screen, the provision of access points, the display of single records, and the division of the catalog into separate files or indexes.


When an opportunistic searcher encounters an over-determined information system, less than ideal search strategies often ensue. The mismatch can be addressed by reducing the determinacy of the system, thereby making it more amenable to informal problem-solving practices.

**Information seeking and needs**


The creative chemists differ from the ‘noncreative’ in that the former read more technical literature on the job . . . are less influenced in their independence of thought . . . are more inquisitive and have broader cultural interests.

Mooers’ Law: An information retrieval system will tend not to be used whenever it is more painful and troublesome for a customer to have information than for him not to have it.

Four levels of question formation may be isolated and analyzed . . .

We argued that information systems should be designed to maximize the amount of control by the receiver, and that, in general, the system should adapt to the receiver or user, rather than the user to the system.


An aspect of the human use of information that has generally been overlooked in the automation of information services is the human tendency to locate information spatially.

A unit called the whomp is introduced which describes the net effect of receiving a message with both hard information and human stress points. Such messages can be said to ‘produce’ records.

The model focuses on variables which can be manipulated by managers—primarily environmental and situational variables—rather than on variables describing individual attributes.

This article reviews recent progress in modeling human cognitive processes.


The data support the belief that managers of high users of information give more encouragement for information use.
<table>
<thead>
<tr>
<th>Year</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The cognitive and affective aspects of the process of information seeking suggest a gap between the users’ natural process of information use and the information system and intermediaries’ traditional patterns of information provision.</td>
<td></td>
</tr>
<tr>
<td>1993</td>
<td>Environmental Scanning by CEOs in Two Canadian Industries.</td>
<td>Ethel Auster and Chun Wei Choo 44(4), 194–203.</td>
</tr>
<tr>
<td></td>
<td>This article examines the information seeking and evaluative behavior of attorneys as they search the corpus of law for primary authority.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Drawing upon a series of studies that examines the information world of poor people, the author discovers four critical concepts that serve as the basis for defining an impoverished life-world. These concepts are risk-taking, secrecy, deception, and situational relevance.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Social informatics (SI) refers to a multidisciplinary research field that examines the design, uses, and implications of information and communication technologies (ICTs) in ways that account for their interactions with institutional and cultural contexts.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Interview data reveal that scientists undertake individual and cooperative boundary-crossing research. Four research modes are identified.</td>
<td></td>
</tr>
</tbody>
</table>

**Use of information systems**

<table>
<thead>
<tr>
<th>Year</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
</table>
A survey of the bibliographic data bases used in satisfying some of the information requests received by the Bell Laboratories literature searching service was made.

This article examines the usage patterns of the ELHILL retrieval program of the National Library of Medicine’s MEDLARS system.

Volume of searching is directly related to the convenient placement of the terminal in the work place. Slightly fewer than half of all potential searchers actually search for themselves. . . . End-users perform very simple searches, mostly using only the AND operator.

Detailed analysis of 43 user-developed computing applications resulted in a model consisting of five stages: Resource Acquisition, Application Development, Adoption/Renewal, Routinization/Enhancement, and External Adoption.

The results confirm that the major issues related to schools’ use of online and CD-ROM databases involve their role in students’ development of the higher-order thinking skills necessary to plan, design, and conduct competent and credible research in the electronic information age.

Seven managers were interviewed to observe how their electronic documents were organized, stored, and retrieved.

Major changes in research habits included greater reliance on word processing and more work outside of libraries. Problems reported focused on access to computer-based resources, learning to use them, the need for instruction, and inconsistent interfaces.

The article reports on five types of network use, E-mail, electronic discussion groups, accessing remote databases, accessing remote computer facilities, and file transfer.

USENET topics are created, and they evolve, mutate, and become extinct in ways fundamentally different from spoken dialogue.

Scientific and scholarly communication

Scientists, so it is claimed, simply will not any longer be able to handle the flood of information. . . . It is a matter of simple fact, which I know from personal experience and observation and which the reader will be able to check similarly, that scientists did not spend on the average in 1962 more time on reading than they did 12 years ago, though printed scientific output has indeed almost doubled during this period. . . . There must therefore have been a way out between the horns of the dilemma. What is it? Everybody knows it: Specialization.

Merely more documents, more specialties, and more people who use or add to growing knowledge do not constitute an explosive situation. . . . The main idea behind the proposed logical structuring of this field is that the ‘growing literature’ organizes knowledge analogously to the way a learner, by creating models of his relevant environment, is able to take increasingly effective actions.

A large amount of consistency was found between the citing characteristics of the journals in the different scientific fields, with quite clear boundaries between fields and a few well known cross disciplinary journals as cross field information links. The separate disciplines appear to relate to each other in an orderly manner, with a natural sequence: mathematics → physics → chemistry → biochemistry → biology.

We suggest that such evidence supports the traditional belief that the humanist has a general tendency to work alone.

The wide range of influences on the field of communication in its formative years is still reflected by its eclectic ‘core literature’. . . .


If communication in research and development is efficient, then the current cognitive situation in any specialty should fully reflect
all available relevant information. Available evidence suggests that communication in R & D is not in that sense efficient. . . .


1998  Hotel Cosmopolitan: A Bibliometric Study of Collaboration at Some European Universities. Göran Melin, & Olle Persson 49(1), 43–48. There are some country variations, but within each country, the differences among the universities are small, if any.

Bibliometric laws and analysis

1955  Citation Indexes for Scientific Literature? W.C. Adair 6(1), 31–32.

1960  The ‘Half-life’ of Some Scientific and Technical Literatures. R.E. Burton, & R.W. Kebler 11(1), 18–22. It is the purpose of this paper to examine the applicability of a half-life analogy to scientific literature, and to provide at least a small amount of data on which to base certain conclusions.

1962  Compilation of an Experimental Citation Index from Scientific Literature. Ben-Ami Lipetz 13(3), 251–266. Several forms of citation index were compiled with the aid of punched card techniques from 11,000 citations. . . .

1963  Bibliographic Coupling Between Scientific Papers. M.M. Kessler 14(1), 10–25. A single item of reference used by two papers was defined as a unit of coupling between them.

1963  New Factors in the Evaluation of Scientific Literature Through Citation Indexing. E. Garfield, & I.H. Sher 14(3), 195–201. More than one million citations from the scientific literature have been processed by the Citation Index Project at the Institute for Scientific Information.


1973  Co-citation in the Scientific Literature: A New Measure of the Relationship Between Two Documents. Henry Small 24(4), 265–269. A new form of document coupling called co-citation is defined as the frequency with which two documents are cited together.

1976  A General Theory of Bibliometric and Other Cumulative Advantage Processes. Derek de Solla Price 27(5), 292–306. A Cumulative Advantage Distribution is proposed which models statistically the situation in which success breeds success. It differs from the Negative Binomial Distribution in that lack of success, being a non-event, is not punished by increased chance of failure.
A tour of information science through the pages of *JASIS*

1978 Frequency-Rank Distributions. Bertram C. Brookes, & Jose M. Griffiths
29(1), 5–13.

... it was found that about 40% of the citations were for historical reasons, but that in the remaining 60% of the cases, the old paper is still begin [sic] actively used.

The findings show that Bradford’s Law is the reflection of some underlying process not related to the characteristics of the search mechanism or the nature of the literature. The authors conclude that there is instead a basic probabilistic mechanism underlying the law.

An author whose presence holds a class together can be considered as an important element in this network of direct communication, and thus might well be a candidate for special funding consideration.

The analysis assumes that the more two authors are cited together, the closer the relationship between them.

All of these—word frequency, citation frequency, and publication frequency—obey an ubiquitous distribution called Zipf’s law. ... The law has been shown to encompass many natural phenomena, and is equivalent to the distributions of Yule, Lotka, Pareto, Bradford, and Price.


Two types of scholarly ‘migration’ are observed, based on patterns of significant increases and decreases in correlations among authors—‘active migration’ (the individual refocussing of a scholar’s efforts) and ‘passive migration’ (the result of reevaluation of an author’s previous contributions in the context of the rise of new problem areas).


In a series of recent studies, J.T. Dillion [sic] has offered a hypothesis of titular colonicity as a guide for the analysis of scientific and scholarly quality. Formally stated. ... The presence of a colon in the title of a paper is the primary correlate of scholarship. ... The Hypothesis
appears to be supported. . . . Percent titular colonicity is higher in more prestigious and more widely respected journals . . . theoretical research is an order of magnitude more scholarly.


1987  Two Medical Literatures that are Logically but not Bibliographically Connected. Don R. Swanson 38(4), 228–233. This study demonstrates that certain unintended logical connections within the scientific literature, connections potentially revealing of new knowledge, are unmarked by reference citations or other bibliographic clues.

1987  Another Test of the Normative Theory of Citing. M.H. MacRoberts, & B.R. MacRoberts 38(4), 305–306. To put our findings in the terminology of citation analysis, we found that ‘obliteration’ . . . does not uniformly operate in science. We noted three patterns: (1) some work is used but is either never cited or is cited rarely, (2) some work is cited mainly or only through secondary sources, and (3) some work is credited every time it is used.

1990  Applications of the Theory of Bradford’s Law to the Calculation of Leimkuhler’s Law and to the Completion of Bibliographies. L. Egghe 41(7), 469–492. . . . this extension also has an application in determining the size and other properties of the complete unknown bibliography, based on the incomplete one.


1998  Invoked on the Web. Blaise Cronin, Herbert W. Snyder, Howard Rosenbaum, Anna Martinson, & Ewa Callahan 49(14), 1319–1328. Where, how, and why are scholars invoked on the World Wide Web?

**Literature domains**


... it was found that the two disciplines representing the social sciences rarely differ significantly from each other, and the same is true of the two hard science disciplines. However, when hard and social sciences are compared, the differences are often significant.


This new view of knowledge stresses the social, ecological, and content-oriented nature of knowledge. ... The final section outlines important problems to be investigated, such as how different knowledge-domains affect the informational value of different subject access points in data bases.


The research presented here describes a method for automatically identifying sublanguage (SL) domain terms and revealing the patterns in which they occur in text.

New documentary forms

1951  15 Years Experience with Auxiliary Publication. Watson Davis 2(2), 87–89.
Auxiliary Publication, accomplished by deposit of a manuscript with the American Documentation Institute and the providing upon demand of photographic copies of that manuscript either in microfilm or photoprint form, is now an accepted and useful adjunct to scientific and scholarly publishing ... was inaugurated in 1936 and has been in operation ever since. ... Since its inauguration there have been 2,148 documents deposited through April 30, 1951.

... is a brief state-of-the-art report on a specific topic followed by a comprehensive bibliography.

Four forms of electronic journals on the Electronic Information Exchange System (EIES) are described as divergent examples of potential journal forms that could become prominent in the future.
Noting developments in computer networking and microcomputer applications, the author contends that the widespread use of personal computers will accelerate the development of electronic publishing.

The Electronic Medical Record: Promises and Problems. William R. Hersh 46(10), 772–776.


The study identified that during 1997, traditional academic publishers made significant commitments to putting tables of content, abstracts, and the full-text of their print journals on the Web. At the same time, new services and organizations emerged that could ultimately compete with, or eliminate, the need for certain segments of the industry.


Libraries
Meanwhile, because the current interest in information storage and retrieval is necessarily directed to the problems of specialist groups, it has scarcely touched those problems of libraries which arise from their efforts to serve many groups simultaneously. The new indexing methods are restricted in application and cannot yet replace the card catalog. . . .

The application of computers to library operations is discussed. . . . Three phases in automation of libraries are identified: the mechanization of conventional operations such as bibliographical control processes and administrative monitoring systems; the automation of search processes based on subject matter, and the move toward new and different kinds of services that computer technology may make possible. We are in the second phase, and snagged by the difficulty experienced by computers dealing with natural language and subjective ambiguities. To move forward through phase two will require a better dialog capacity between man and machine than presently exists. Before progressing into the third phase a better identification of the purposes that our files of information are to serve will be needed.
ELMS (Experimental Library Management System) is an experimental system for total library management, operating on-line with an IBM 360 through IBM 2260 and 2741 terminals. . . . This paper discusses ELMS features that facilitate user interaction, and may prove useful in similar systems: techniques for tutoring the user . . . ; adaptability for the experienced user. . . .

The performance on an interlibrary loan network is characterized. . . .

A computerized method of assisting monograph collection development by correlating circulation with inventory statistics is presented.

Bibliographic control before and after MARC is reviewed. The capability of keying into online systems has brought an interdependence among libraries, the service centers that mediate between them, and the large utilities that process and distribute data.

Immense cultural changes are taking place in contemporary society. . . . This article discusses these changes and the meaning and depth of their relationship to medical libraries. . . . Only by actively debating and creatively experimenting with these and other ideas will we be able to ensure our continued viability as information specialists and managers in the future.

All these efforts are leading not only to a usable prototype digital library but also to a set of nine principles for digital libraries. . . .

This study suggests how an altered understanding of information can provide the basis for rethinking and potentially redesigning the library’s mission. . . .


We examine what constitutes effective DL use, how faculty members are using DLs, and how useful they find them.

Information policy and standards
1980 The Cultural Appraisal of Efforts to Alleviate Information Inequity. Glynn Cochrane, & Pauline Atherton 31(4), 283–292. This article suggests that action to alleviate information inequity should be guided by the principles of contextualism, incrementalism, motivation of information users, and more knowledge of the absorptive process that is unique to each cultural group.
1985 An Overview of Social Measures of Information. Michel J. Menou 36(3), 169–177. The rise of the information society calls for quantitative and qualitative measures of information activities at the subnational, national, and international levels.
1991 Information Technologies and Social Equity: Confronting the Revolution. Ronald D. Doctor 42(3), 216–228. As a society we are giving inadequate attention to ensuring that as new computer and telecommunications technologies become more pervasive, their benefits are distributed in ways that don’t exacerbate existing disparities between the rich and poor.

An electronic communication system, especially if it is designed to accommodate open access via public terminals can allow information-disadvantaged groups (such as the homeless, and women) to exchange relevant information and engage in political action.

Reading the Bones: Information Content, Value, and Ownership Issues Raised by the Native American Graves Protection and Repatriation Act. Teresa Olwick Grose 47(8), 624–631.

. . . questions have arisen concerning the nature of information as embodied in objects and the moral right to access that information.

SGML and Related Standards: New Directions as the Second Decade Begins. James David Mason 48(7), 593–596.


International information


Information Transfer as Technical Assistance for Development. Marta L. Dosa 36(3), 146–152.


The issues . . . include the assessment of the relevance of the notion of a paperless library to Africa, the expected role of library and information services in narrowing the gap between the information rich and poor, and the possible inappropriateness of information technology and/or advice from developed countries.

Databases on Optical Discs and Their Potential in Developing Countries. S. Nazim Ali 41(4), 238–244.

This in-house system will avoid cost associated with telecommunication tariffs and provide unlimited access.

The transition from a national science model in which the national language is used for publications and other communications, to a transnational model in which a single international language (English) is used and the market is dominated by Anglo-Saxon publishers, has continued in recent decades.

**Economics of information**


Energy, time, and money must be invested to change useless data to productive knowledge, a value-added process.


It is concluded that economic approaches based on ‘information theory’ have not achieved significant practical results in a general sense, but ‘classical’ economic approaches can and should be used. . . .


Information is used by marketplace actors to make decisions or to control processes. Thus, we define information as the ability of a goal-seeking system to decide or control.


Journal costs are thought to consist of ‘first-copy costs’ and ‘additional copy costs’; when costs rise, subscription prices are increased by publishers and some libraries, being faced with fixed budgets, cancel some subscriptions. Publishers then find that the price charged is not correct and increase subscription prices again.

**Curriculum**


The 7 [core master’s] courses are: Introduction to Information Science, Systems Theory and Applications, Mathematical Methods in Information Science, Computer Organization and Programming Systems,
Abstracting/Indexing/Cataloging, Information and Communication Theory, and Research Methods.


History


Perspectives and special topic issues

Beginning in 1980, the Journal began publishing “Perspectives” issues, first under the editorship of Susan Crawford (four issues) and then, in the 33 subsequent issues to date, Lois F. Lunin. In each case the senior editor would work with a co-editor with a specialty in the designated topic of the Perspectives issue. Below, to reduce redundancy, only the co-editors are listed.

In 1994, the Journal began publishing a different kind of series—special topic issues. These would be under the Guest Editorship of a subject expert in the topic of the issue, and papers would be refereed in the same manner as regular articles in the journal.

Lists are provided below of both of these special series, arranged by year and issue number (in parentheses), up through the fifth issue of 1999, when this issue went to press. As the special section occupies most of the issue in which it is published, page numbers are not provided.
Perspectives issues

1980 (3) Online Systems in Science and Technology. Alan M. Rees
1980 (6) Library Networks and Resource Sharing. Barbara Markuson
1981 (2) Computer and Communications Technology. Manfred Kochen
1982 (3) Secondary Information Services: Development and Future. Marianne Cooper
1982 (6) Systems Methodology and Information Research. Una Mansfield
1984 (3) Information Science: Retrospect and Prospect. Robert M. Hayes
1985 (3) International Information Issues. Beth Krevitt Eres
1986 (4) Online Searching. Donald T. Hawkins
1988 (1) CD-ROM for Information Storage and Retrieval. Peter B. Schipma
1988 (2) Integrated Academic Information Management Systems (IAIMS). Marion J. Ball
1989 (3) Hypertext. Roy Rada
1989 (5) Information Science and Health Informatics Education. Marion J. Ball
1990 (6) Author Co-Citation Analysis. Howard D. White
1991 (2) Integrated Information Centers within Academic Environments. George D’Elia
1992 (2) Human–Computer Interface. Donna Harman
1992 (8) Information Technology Standards. Michael B. Spring
1993  (4) Knowledge Utilization. Susan Salasin, & William Paisley
1994  (5) Redesign/Reengineering of an Information Services Division in a Major Health Sciences Institution. Marion J. Ball, & Judith V. Douglas
1994  (8) Indexing. Raya Fidel
1994  (10) Electronic Publishing. Robin P. Peek
1995  (10) Medical Informatics: Information Technology in Health Care. William R. Hersh
1996  (11) Distance Independent Education. Howard Besser, & Stacey Donahue
1998  (11) Internet Issues. Carol Tenopir

Special topic issues
1994  (3) Relevance Research. Thomas Froehlich
1994  (6) Information Resources and Democracy. Leah A. Lievrouw
1994  (9) Spatial Information. Myke Gluck
1996  (4) Full-Text Retrieval. MaryEllen C. Sievert
1996  (9) Electronic Publishing. Robin Peek
1997  (4, 9) History of Documentation and Information Science: Parts I, II. Michael Buckland, & Trudi Bellardo Hahn
1998  (1) Science and Technology Indicators. Anthony F.J. Van Raan
1998  (3) Management of Imprecision and Uncertainty. Gloria Bordogna, & Gabriella Pasi
1998  (5) Knowledge Discovery and Data Mining. Vijay V. Raghavan, Jitender S. Deogun, & Hayri Sever
1998  (7) Artificial Intelligence Techniques for Emerging Information Systems Applications. Hsinchun Chen

1999  (1) Youth Issues in Information Science. Mary K. Chelton, & Nancy P. Thomas

1999  (4) The National Information Infrastructure. Patricia Diamond Fletcher, & John Carlo Bertot