GEOSCIENCE INFORMATION SOCIETY

Proceedings • Volume 6 • 1975

QE
48.85
G43
GEOL

PROCEEDINGS OF THE TENTH ANNUAL MEETING

OF THE

GEOSCIENCE INFORMATION SOCIETY

OCTOBER 21ST, 1975 SALT LAKE CITY, UTAH

RETRIEVAL

OF

GEOSCIENCE INFORMATION

PROCEEDINGS

Volume 6

Geoscience Information Society

1976

GEOSCIENCE INFORMATION SOCIETY

PUBLICATIONS

PROCEEDINGS, V. 1	(covers 1966, 1967 meetings), published 1969 (ed. Graves) HANDLING GEOSCIENCE DATA AND INFORMATION. (Reprinted)	$ 3.20
PROCEEDINGS, V. 2	(1971 meeting), published 1972 (ed. Graves) TOWARD THE DEVELOPMENT OF A GEOSCIENCE INFORMATION SYSTEM.	$ 2.50
PROCEEDINGS, V. 3	(1972 meeting, Minn., MN), published 1973 (ed. Phinney) (no title)	$ 5.00
PROCEEDINGS, V. 4	(1973 meeting, Dallas, TX), published 1974 (ed. Wheeler) GEOSCIENCE INFORMATION.	$10.00
PROCEEDINGS, V. 5	(1974 meeting, Miami Beach, FL), published 1975, (ed. Morrison) GEOSCIENCE INFORMATION.	$10.00
PROCEEDINGS, V. 6	(1975 meeting, Salt Lake City, UT), published 1976, (ed. Hall) RETRIEVAL OF GEOSCIENCE INFORMATION.	$12.00

DIRECTORY OF GEOSCIENCE LIBRARIES: U.S. & CANADA. (2nd. ed.) 1974 (ed. Parker and Walker). $ 3.50

GIS NEWSLETTER. Subscription for the current year (1976); 6 issues. $10.00

ENVIRONMENTAL GEOLOGY IN THE PITTSBURGH AREA, ed. R. D. Thompson. Geological Society of America, Field trip guidebook no. 6, 1971. ~~(Available from: Marjorie Hooker, 2018 Luzerne Avenue, Maryland 20910, U.S.A.)~~ $ 3.00

GEOLOGIC FIELD TRIP GUIDEBOOKS OF NORTH AMERICA, A UNION LIST INCORPORATING MONOGRAPHIC TITLES, 2nd. ed., 1971. (Available directly from the publisher; make checks payable to: Phil Wilson Publishing Co., Box 13197, Houston, TX, 77019). $20.00

ORDERING INFORMATION

Geoscience Information Society
c/o American Geological Institute
5205 Leesburg Pike
Falls Church, Va. 22041

Orders must be prepaid. Make check to Geoscience Information Society--mail order to address at left.

The papers in this volume of the Society's Proceedings were presented orally at the Tenth Annual Meeting in Salt Lake City, Utah on October 21, 1975.

Symposium I, The Location and Retrieval of Energy Resource Information, was co-chaired by Louis Cima and Carol Alexander.

Symposium II, Data Retrieval for the Student, Researcher and the Environmentalist, was co-chaired by Jack Morrison and Vivian Hall.

Papers are arranged in order of their presentation at the meeting. One paper scheduled was not presented and one paper given was not submitted for publication, therefore they are not included.

Papers appear in this publication as they were submitted, editing and corrections have not been made due to limited time and staff available.

Vivian S. Hall
Program Chairman 1975
Geoscience Information Society

For Sale by the GIS Secretary, c/o American Geological Institute, 5205 Leesburg Pike, Falls Church, Va. 22041, for $12.00. Orders must be prepaid. Please make checks payable to Geoscience Information Society.

CONTENTS

SYMPOSIUM I:

The Location and Retrieval

of Energy Resource Information

Data Tagging in Information-Accessing Services
Containing Energy-Related Data

Fred A. Tate and David F. Zaye*
Chemical Abstracts Service
The Ohio State University
Columbus, Ohio 43210

INTRODUCTION

At present, specific types of data and specific data values are difficult, if not impossible, to locate in the primary published literature and the dimensions of this problem continue to grow as the primary data base increases.[†] Location through direct searching of the primary literature is seldom possible, and secondary services have not developed economical and effective means of providing this needed access. However, Chemical Abstracts Service (CAS) is experimenting with a new processing methodology which would be capable of providing significant improvement in data accessibility. CAS is developing a technique for locating data in the primary literature based on the inclusion of "pointers" in the CAS Data Base. Initial development is to be based on experimental use of a special version of the CAS computer-readable file entitled ENERGY, which includes CA sections specifically devoted to energy. In this pilot effort, the pointers will consist of an energy-oriented set of unique, computer-verifiable codes --

---

* To whom correspondence should be directed.

† The Chemical Abstracts Service (CAS) defines chemistry as the study of the composition and structure of matter, of the transformations it undergoes, and of the theories and laws thereof. On the basis of this definition, CAS considers data to be measured or calculated values of the properties of substances, their interactions, and their interrelations with energy.

called Data Tags -- to indicate the presence in the corresponding primary document of selected types of recorded data. These tags will be assigned by the CAS primary-document analysts, and they will be placed in this special ENERGY file for use in automated searching. This so-called data tagging would thus make the existence of these data visible to those who use this CAS service.

Projections for 1976 show that approximately nine percent of all CA coverage will be contained in the ENERGY service which includes the seven CA sections identified in Figure 1. Illustrated in this figure are only a few of the hundreds of types of data which may be tagged within these sections. Since CA is discipline-oriented, energy-related studies are also found in many other CA sections. Reaction kinetics and mechanisms in organic chemistry, macromolecular chemical transformations and processes, and biochemical nitrogen fixation are a few of the types of subject matter in the other sections which contain calculated and experimental data and which have importance in energy-related general research.

Concept Tags. In combination with the evaluation of data tagging, CAS is also investigating the tagging of certain other types of information contained in the source documents. Such tagging would be done by "Concept Tags." Whereas Data Tags indicate the presence of types of numerical values, Concept Tags will denote the existence of concepts, or information, which may be narrowly limited in appearance or may be reported in many widely separated parts of CA coverage. Data Tags and Concept Tags are expected to provide complementary avenues of access in automated search by supplying additional useful detail which can simplify the identification of pertinent information. The mechanism for concept tagging is similar to that indicated above for data tagging, i.e., a unique code acts as a pointer to this information. Although this paper will discuss data tagging primarily, data tagging as used in the text of this paper is intended to represent varying combinations of Data Tags and Concept Tags.

PROBLEMS OF INCORPORATING DATA IN CA ABSTRACTS

Data Content in Abstracts. CA abstracts, like abstracts of most secondary services, seldom include much data. This practice results from the fact that most data are associated with a specific substance and/or type of energy and that a large portion of the indexed concepts or substances are not specifically included in the corresponding abstracts. The huge cost of listing all of the individual items of data recorded in primary documents covered by CA prevents such practice. Such cost results not only from the expense of recording and printing the data in CA, but from the tremendous amount of effort necessary to detect and correct transcription errors. Numerical data, unlike written words, is almost entirely lacking in the elements of context and redundancy which facilitate rapid proofreading and which, in the case of a published error, often enable the reader to "guess" the correct meaning. Even when a primary document does provide context and redundancy (as in the reference to specific numerical values corresponding to points in a graphical presentation), such context and redundancy are generally lost early in the document-analysis process during the course of extracting the individual values. Consequently, those who must deal with the publication of numerical data are faced with extensive and costly processes of editing, proofing, and correction.

Data Content in Volume Indexes. CA abstracts are intended to identify the primary bibliographic citation and to record the nature of the objectives, the methodology, and the results included in the corresponding primary document -- all of which are couched in the jargon of the original, reflecting the author's emphasis. Many substances are identified in CA subject indexes as a result of direct analysis of the original document, but the corresponding CA abstracts are not expected to record all of the points of access given in the indexes. Thus, to include in the abstracts all specific data values found in the primary literature, a great deal of the indexible content which presently does not appear in the abstracts would have to be added. This would require a substantial increase in the average cost of preparing and delivering abstracts. This difference between CA abstract content and CA subject index content focuses on the difference in intent between abstracts and indexes. The abstract is intended to help the user determine whether the nature of the corresponding document is likely to be useful in the context of an existing need or interest. It reflects the emphasis of the original document and provides for inflections not possible in the comparatively terse entries of a broad-range subject index. Thus, CA users must consult the original source to obtain experimental detail. This of course implies that data should be judged in the framework of the report in which it is recorded. It is only within this context that the usefulness of the data can be appraised. Also, because methodology, reporting units, and terminology for various types of data are sometimes peculiar to the subdiscipline or technology toward which the original document is directed, data values are not always addressed in the same way. The CAS practice of following the jargon of the source document in the CA abstract does not permit standardization of data terminology or units of measure. Moreover, any variation from author terminology would reduce the direct relationship between the abstract and the original publication. Such a relationship is often especially useful in comparing an English language abstract with a non-English original.

CAS emphasizes detailed volume subject indexing. This is based on careful analysis of source documents by a corresponding subject expert. And as most CAS input is prepared by full-time CAS staff and as this in-house percentage is steadily increasing, the consistency of assignment of Data Tags by CAS document analysts can be readily controlled within the CAS data base. Also, because of the nature of the CAS computer-based production system, Data Tags can be economically and reliably recorded and incorporated in a variety of output packages for evaluation of the usefulness of such tags.

Direct Indexing of Data. In contrast to the vernacular of the primary literature and the CA abstract, which usually consists of the highly specific jargon of a narrow subject field, the CA volume subject indexes utilize standard, controlled terminology which is intended to provide access keys to individual details of the original document. This is illustrated by the CAS practice of using specially designed nomenclature as a part of the document-analysis process to specifically identify each indexed chemical substance. It is not economically feasible to include all such indexed points in the abstract as well as in the index. Thus, many subject entries are specifically identified only in the subject indexes.

If data types (with or without data values) were to be directly accessible in CA volume subject indexes (i.e., if the data types were to appear as primary points of entry in the printed index listings), the addition of a relatively long entry in the index would be required for each type of data for each appearance in the primary literature. It is obvious that providing such direct access to individual data would be prohibitively expensive due to the following points: (a) the number of entries for substances in indexes each year (index entries dealing with substances currently run at about one million per year, corresponding to about 350,000 different substances); (b) the frequent appearance of more than one item of data per substance; and (c) the impact of printed index size upon cost.

DATA TAGGING

Data tagging could greatly improve the usefulness of information-accessing services as tools for locating published data, both for feeding established data collections and for identifying limited sets of data for one-time use. For reasons of economy and efficiency, as discussed above, information-accessing services are not able to record or index individual items of data which are contained in their corresponding primary literatures. As an alternative, CAS is undertaking a stepwise program to develop and evaluate the use of Data Tags in a CAS service that would indicate the presence of specific types of recorded data in primary documents. This is based on the postulate that Data Tags, if suitably recorded in information-accessing services such as those provided by CAS, would significantly improve access to corresponding data values by pointing to those source documents which include specific types of data. If useful in CAS services, the technique should also be useful in other information-accessing services.

CAS will investigate tagging at various levels of detail to find suitable balances among cost, timeliness, reliability, and usefulness of Data Tags. CAS will also investigate the potential of Data Tags by creating special evaluation packages which will be used in a "pilot plant" framework. The objective is to permit "tuning" of the tagging technique to optimum performance by obtaining user reaction to tagging in actual searches being conducted on the evaluation packages. During this experiment no Data Tags will be included in regular CAS services.

In addition to assessing the utility of the alternative uses of Data Tags, CAS will evaluate the economics of including Data Tags in various CAS accessing tools. Because the feasibility of including Data Tags is inextricably bound into the usefulness of the tags, the various needs for data and the patterns of acquisition must be considered.

The Experiment. CAS is evaluating the utility of data tagging in CAS information-accessing services. The objective of the experiment is to seek a sound basis for including Data Tags in a selected set of complementary CAS services. Tradeoffs between the usefulness of tagging in a computer-readable CAS service and the expense of generating these tags will be explored. The experiment will consist of two phases: (1) an operational evaluation of data tagging from the user's and from the processing standpoints; and (2) an assessment of the effects of broad range tagging on energy data tagging. This approach to the development of data tagging has been selected because it will permit experimentation with different levels of data tagging content and format without including within the permanent record of CAS services any of the variations which are employed in practice during the development process. Since all CA content will go through the computer, the experimental Data Tags can be added to the file record without separate follow-on processing. Later the inclusion of tags in the printed services can then be based on sound experience.

Operational Evaluation in Energy-Specific Data Tagging. The first phase of the experiment will assess the usefulness of Data Tags for a defined data user group within the community and will also assess the effects of data tagging upon the cost and efficiency of CAS operations. These investigations will be based upon the evaluation of a special version of the CAS computer-readable file ENERGY. This file, as mentioned earlier, corresponds to the CA sections that include publications on fossil fuels (i.e., coal, peat, petroleum, oil, shale, and tar sands) and their derivatives; electrochemical, radiational, and thermal energy technology; nuclear phenomena and technology; thermodynamics; electrochemistry; and propellants and explosives. The primary literature covered by ENERGY contains many varied data types and the current need for improved access to this data makes this a priority area.

Source documents covered by ENERGY include journals, patents, reports, and proceedings of conferences and symposiums. Searchable information includes abstracts; names and affiliations of authors, patentees, and patent assignees; source document bibliographic citations; document titles in English; molecular formulas; CA Issue (i.e., Keyword) and Volume Subject Index entries; CA Section Numbers; and CAS Chemical Substance Registry Numbers. Special evaluation packages of ENERGY will be prepared which contain Data Tags in addition to these items. These Data Tags will be associated with either a specific chemical substance or type of energy. The tags will provide the capability of automatically identifying all documents which include specific types of data as well as the substance or energy types directly corresponding to these data. The data tagging will be carried out as an integral part of the CAS document-analysis, information-recording, and editorial operations in routine production. This will permit reliable identification of the differences in efficiencies and in costs introduced by the addition of data tagging to the processing efforts. A major portion of the document analysis for this energy-related segment is already done in-house and much of the editorial staff is already practicing one-time document analysis. These factors will simplify the management of the effort and the evaluation of processing impact.

At first, as was mentioned earlier, tagging of the ENERGY service content will focus on the types of data which constitute the primary concern of a selected audience chosen to participate in the evaluation of energy-specific Data Tags. Such tagging is here called "selective." Other types of data which are "non-selective" for the audience but which may appear in the corresponding primary documents being analyzed will not be tagged in the initial evaluation packages. At this stage of development this selective tagging will allow the main emphasis to be placed on the usability of selected energy-specific Data Tags in searches and on reliable, economical processing methodology.

The participating audience will be supplied with issues of the evaluation package which include the Data Tags. Feedback from the audience will then be used to help in tuning the level of tagging detail to optimize Data Tag utility. CAS staff will experiment with the first issues of this package to develop "tips" to help the cooperating audience best utilize Data Tags in its own search environment.

<u>Full Range Data Tagging</u>. In the second phase of the experiment "full range" data tagging will be considered. Full range implies tagging not only the selective or energy-specific types of data of primary concern to the participating group, but also other types of data recorded in the corresponding primary documents. These additional data types will then be identified in the routine document-analysis process as well and will be assigned Data Tags so that the usefulness of full-range

tagging in user searches and its impact on production operations can be examined. Tagging additional types of data will be introduced in stepwise increments so that possible thresholds can be identified at which the level of data tagging detail has impact on the efficiency of document analysis and consequently serious economic impact. This phase will have another objective and that will be to identify the value of the added diveristy of the non-energy-specific Data Tags in the search process.

Energy-Specific Data Tags in Other CA Sections. A further extension of this experiment will involve identifying other CA Sections in which energy-specific Data Tags may be useful. The amounts of energy-specific and other types of data in such sections will be quantified. Sections dealing with catalysts and reaction kinetics, surface chemistry and colloids, electric and magnetic phenomena, as well as unit operations and processes are candidates for this investigation.

The quantification task will identify the types and frequencies of data items that are recorded in the source documents covered in selected CA sections and subsections. The selections will be aimed at sampling what are judged to be typical segments of CA coverage in terms of the range of data types and the concentrations of data values given in the corresponding primary literature. Both data-rich and data-poor segments will be sampled, and consideration will be given to data types which are widely distributed and to those which are clustered in narrowly limited sources. The types of source document -- journal paper, patent, report -- will be considered in sample design.

The usefulness of Concept Tags will be considered as an alternative and as a useful addition to Data Tags. Since in some cases the concept being tagged may also have numerical data associated with it, the interrelationships between data tagging and concept tagging will be examined. These interrelationships will also be examined to determine the tradeoffs in using Concept Tags, Data Tags, or both from the standpoints of the user and/or the processor.

The usefulness of tags and the economic feasibility of such tagging will be evaluated from the results of the experiment. This evaluation, as well as evaluation of the potential best mixture of Data and Concept Tags, must be conducted in advance of any operational introduction of energy-specific tagging so as to assess its economic consequences in overall CA coverage. This evaluation is necessary to plan the implementation of tagging in a stepwise fashion for a broad range of data types across the full subject content of CAS services. The impact on timeliness of coverage, cost of staff training, reliability of recorded tags, and range of audiences served, as well as the cost of the routine inclusion of tags in CAS services, must be assessed.

CONCLUSION

Increasingly more and more data is being reported in the primary literature of science and technology. The volume of this literature is so large and extensive that it becomes very difficult, if not impossible, to locate data through direct searching. While data is included in papers or patents, often it is not recorded in either the corresponding abstract or the index entries because it does not reflect the primary intent of the study being reported.

Chemical Abstracts Service is conducting an experiment designed to alleviate this problem. This experiment will attempt to evaluate the usefulness and economic feasibility of including Data Tags in secondary information files. Data Tags are computer-verifiable codes which denote the presence of specific types of numerical or factual data found in a primary document. The purpose of using these tags is to identify potentially useful data so that it would be visible in the secondary information record for the document.

The experiment will involve the tagging of energy-related data found in a special version of the CAS computer-readable file entitled ENERGY. A selected group of data users will be chosen to participate in the experiment for evaluating the usefulness of the tagging. The experiment will also attempt to identify other types of energy-related data that may be present in documents covered by CA in sections not included in the ENERGY package. The economic feasibility of incorporating Data Tags in these sections will be investigated.

ACKNOWLEDGMENT

The development of many techniques and applications described in this paper are partially supported by the National Science Foundation under Grant SIS-75-03491. Chemical Abstracts Service, a division of the American Chemical Society, gratefully acknowledges this support.

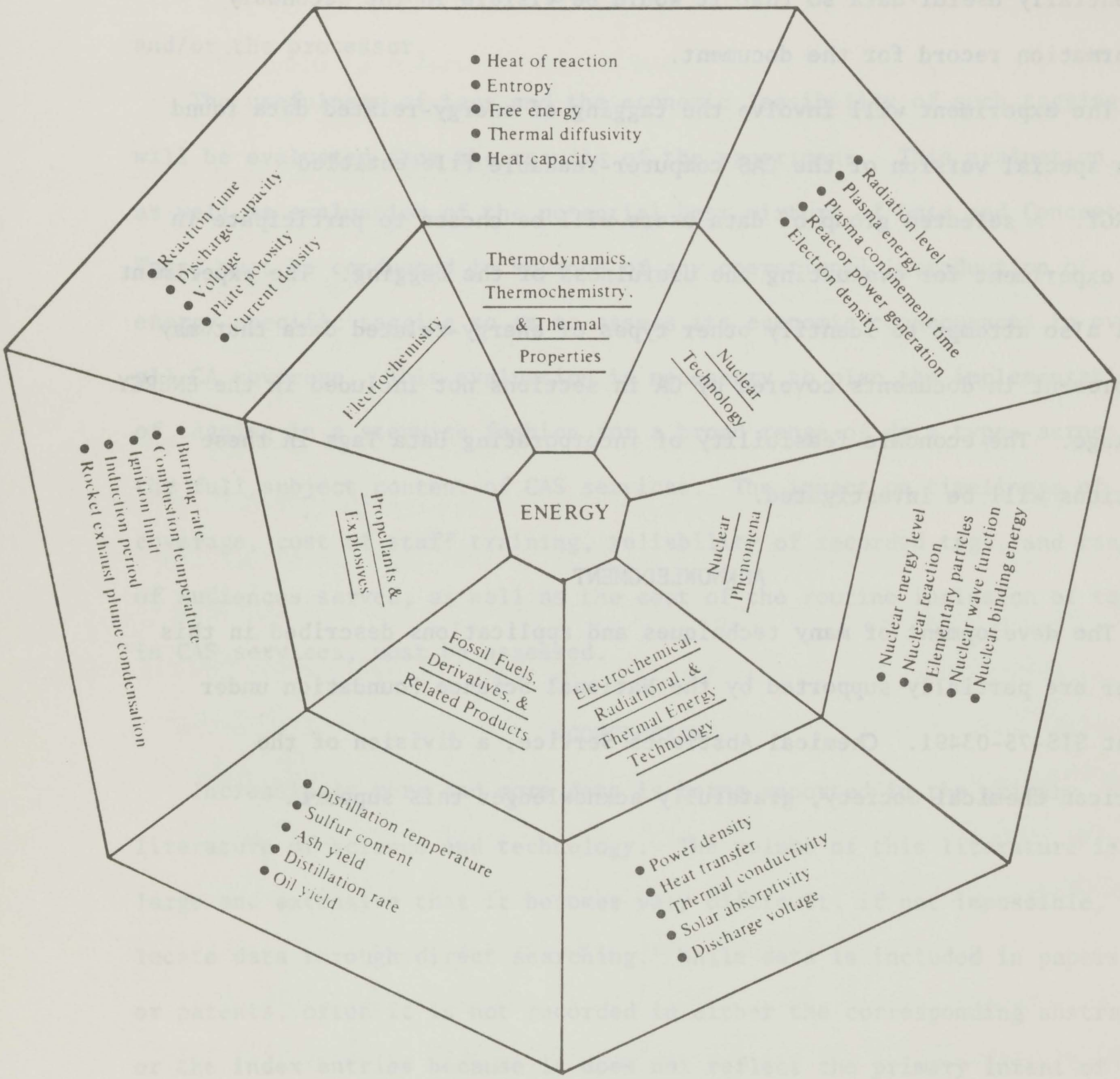


Figure 1. CA Sections Comprising the ENERGY Service. CA Sections are underlined and examples of data types from each section are shown.

CONSIDERATIONS FOR THE DESIGN OF A NATIONAL INFORMATION CENTER FOR RENEWABLE ENERGY SOURCES

by

R. W. Thorpe and J. W. Kuipers
QEI, Incorporated, Bedford, Massachusetts

A. The Role of Information Support Services in the Renewable Energy Field

It is of greatest importance for the present and future national welfare that we develop new sources for low-cost energy. For the research, development, and application processes which must be carried out to achieve this goal it has been proposed that a national information center be set up to provide essential information services for the energy field. Such a center would provide necessary support for the research, development, and application processes and would help to speed up these processes so that energy program goals could be achieved in a substantially shortened time period.

The study which we are conducting at this time will closely examine the various factors that should be taken into account in the design of a national information center for energy sources. The portion of the energy field which is of particular interest to us is that concerned with renewable or non-depletable energy sources. By our definition, we are excluding sources such as fossil fuels, coal, oil, as well as nuclear fission processes, from our consideration. The energy sources included in the renewable energy field encompass the following:

nuclear fusion processes

solar processes

geothermal processes

tidal and wind processes

It is recognized that no sharp line can be drawn to delimit the "renewable energy" area but the intent is to concentrate on the information needs of that part of the energy field that will continue to be of importance to us for an unlimited period in the future.

Information is essential for all aspects of human productive endeavor. Information is needed to gain understanding, to reach decisions, to solve problems, or to generate new information. It is the goal of an information system, any system for communication of scientific and technical information, to improve and enhance the productivity of the users of the system. What is needed in the study which we are conducting is to examine the principal facets of the information system which is now in use in the renewable energy field and to determine what steps can be taken to improve the principal functions of this system in order to assist the users of the system in improving their productivity.

B. Some Basic Characteristics of an Information System for the Renewable Energy Field

1. Possibilities for a Centralized Machine-Supported System for Renewable Energy

Why do we not take steps at once to build up a comprehensive machine data base to include all areas of the renewable energy field and make this instantaneously available to all the thousands of users in the field by means of communication networks and low-cost terminals? Would there not be advantages in consolidating input/output and processing costs in one system? Should we not attempt to eliminate paper production and handling? Isn't this approach do-able and within the state-of-the-art?

We believe there is no question that machine supported centralized files can be implemented and maintained to serve a large user population. It is probably true, too, that a centralized machine-supported data base will be an important element in a national information center at some stage. However, to rush into an immediate implementation of a centralized data base system will not particularly assist in alleviating the principal problems. Unfortunately, energy information systems are very complex and are required to perform many vastly different tasks on a wide range of different types of information. Thus, it is necessary to examine closely the basic aspects of any information system on energy sources to determine precisely what the problems are and what steps can and should be taken to solve these problems. Therefore, in the following paragraphs we shall describe certain important facets of any energy information system, in order to elucidate the basic problems and to lay the foundation for proposing steps to solve these problems.

2. Communication Channels in a Scientific Field

In any given field of science and technology, there are two major communication channels: the formal and the informal. (2). Formal communication involves the use of published literature both primary and secondary. Informal communication channels involve direct interpersonal communication. (4) (7) (17). Both of these channels are of importance and should be recognized, used, and exploited where possible to achieve desired objectives in the design and usage of an information system.

3. Functions in the Information Transfer Process

In the formal communication channel, different functions in the information transfer process can be identified. (1). The following functions characterize various steps in the information transfer cycle:

Cycle Step 1. Research and Development

Cycle Step 2. Composition and Recording

Cycle Step 3. Printing and Distribution

Cycle Step 4. Library Acquisition and Storage

Cycle Step 5. Organization and Control

Cycle Step 6. Presentation to User

Cycle Step 7. Assimilation by User

Recycle to Step 1. etc.

Communication by these functions from author to user can take place with the use of a variety of media: journal articles, research reports, conference papers, etc. Communication of a particular research finding to a number of users may involve a variety of communication functions and different media and modes.

4. "Current Awareness" and "Retrospective Search" Functions in the Information System

In the information transfer process, a particular purpose mode may be exercised by system users. The two most important purpose modes are concerned with 1) "current awareness" and 2) "retrospective search" of archival information items in the system. In the first mode a system user desires to be made aware of the new information items which may have entered the system. In the second mode he desires to find and use specific information items contained in the system. The acceptability of a particular system design will relate to the ability of the system to satisfy the user's wishes with respect to these modes.

5. Multi-disciplinary Aspects of the Field

The renewable energy field encompasses a wide variety of disciplines and specialized technical areas. The field covers a spectrum of subject areas from controlled thermonuclear fusion to solar photovoltaic conversion and includes the associated applications technologies as well. A plan for a national information center must provide for the needs of system users in each of the various specialized subject areas.

6. Different Requirements for Different User Groups in the Field

A national information center must serve a variety of users in particular disciplines and specialized areas. It must also provide for user's interests at differing levels of generality. We suggest that the needs of the following types of users must be served:

1) R & D and technical users in various disciplines.

2) High-level administrators in operations in the renewable energy field.

3) Mid-level administrators and supervisors.

4) Legislators, government and public officials.

5) The general public.

Each of the user types indicated has a different kind of information requirment.

7. Screening, Review, Evaluation, and Analysis in the Information Transfer Process

A certain amount of screening, review, evaluation, and analysis is now being carried out in various stages of the information transfer process. The screening, selection, and editing of articles for the journal literature, the generation of state-of-the-art reports, the compilation of tables of physical data, are examples of important current review and analysis processes. For the needs of a national information center, this activity becomes extremely important. It may well be the key activity in the system which can make a successful operation possible.

8. Relation of Information System Characteristics and the Design of a National Information Center

In the foregoing discussion we have attempted to highlight certain characteristics of an information system as it might apply to the requirements of the renewable energy field. We have mentioned channels of communication, informal and formal channels, transfer functions, review and analysis capabilities, and modes for current awareness and retrospective search. These are the important aspects of an information system which will want to be examined in much detail. We will want to find ways to improve rates of communication and transfer. We will want to satisfy the user convenience requirements and purpose modes. We will have to be aware of system complications because of differing user interests and areas of specialization. We will want to examine these factors and propose steps which will make it possible for an improved information system to impact on the system user to improve his productivity.

C. Discussion of Principal Problem Areas That Need to Be Taken into Account

The preceding discussion has provided background about certain characteristics of an information system which must be kept in view in attempting to achieve an improved system design. Before we can begin to think about improved designs, however, we must look more closely at the specific problem areas.

1. Problems Concerned with the Nature of the Information and the Data in the Field

We can summarize the problems associated with the information and data in the renewable energy field under the following points:

- High information volume and high per unit costs for information acquisition and retrieval
- Multidisciplinary nature of the information implying dissimilarities among information pieces.
- High level of complexity in both organization and content.

- Wide-ranging nature of information coverage.
- Information sources are widely scattered among hundreds of journals and data bases.

The principal implications here are that sufficient resources must be allocated to deal with the information flowing in the system. The complexity and multidisciplinary nature of the information require that only people with a sufficient knowledge and understanding in a specialized area can properly analyze this information. The use of general indexers and abstractors for complex and specialized information will not be possible since the results will not be satisfying to system users.

2. <u>Deficiencies in Information System Configuration and Operation</u>

If we examine the state of affairs in information systems which are now in use, the following problems are most apparent:

a. There are serious delays in the reporting and publication of results of R & D and of the results of the application of technology.

b. Little provision is made in any existing system for the reporting of R & D in progress.

c. There are high levels of difficulty and inconvenience for system users when they attempt to access and obtain information products now being generated.

It is clear that in both the formal and informal channels of communication there are serious delays and deficiencies. These represent prime targets for attention and resolution for any proposed improvements in system design.

3. Problems Concerned with the Management of Information Support Systems

Information systems now in operation (and this is the case with many systems supported and sponsored by the government) must deal with the following difficulties:

a. There is lack of management planning, cognizance, and control of information support services.

b. There is lack of integration of various information support activities and services.

c. There is lack of adequate funding for information support services.

Too often information support services are viewed by management as optional and readily dispensable peripheral activities which are capable of being set up or terminated at will, as a particular problem need arises or recedes. There is not sufficient recognition that information services are a continuing requirement which must be fully integrated with an R & D program or a technology application program and that costs for these services should be covered as a definite part of the funding supplied to carry out the original R & D task or program.

D. Possible Steps to be Taken in Addressing the Principal Problem Areas

Our study is still in an early phase and we are still compiling and organizing background and particulars about information handling processes in the renewable energy field. We are at a point, however, where certain generalizations can be made. We list below some of the major guidelines which we believe are supportable and which should be given serious consideration in a plan to set up a national information center for the renewable energy field. The following steps should be considered:

1. Institute a positive information support program with direction from the highest management level.
2. Provide an adequate organizational and managerial framework for information support activities.
3. Provide continuing funding in direct proportion to the total operational activity.
4. Institute an intensive review, evaluation, and analysis capability at various organizational levels.
5. Conduct all information analysis only in specific centers where subject expertise can readily be made available.
6. Centralize processing and operations which will tend to speed up alerting, announcement and accessing processes.
7. Use information analysis centers as interaction points with "invisible college" networks.
8. Use information analysis centers to generate news and reviews of research in progress for particular disciplinary areas.
9. Fully exploit all existing and available information support resources.

E. <u>New Contributions for a National Information Center Design</u>

In the broad picture of problems, difficulties and possible steps to be taken to alleviate the problems, what specific new contributions can we hope to achieve in this project? Although the items listed below are tentative we believe they do merit careful evaluation. They do represent a new emphasis which we believe would result in an overall information system which would be more satisfying and productive for system users.

Proposed New Contributions in Our System Design

a. The establishment of a management framework which would give information support services a continuing and integral role in any R&D or technology program.

b. The institution of information analysis centers at various locations for the specialized areas and disciplines in the field, in order to provide new system products and outputs. The various analysis centers would be supported by a centralized operation for production and processing.

c. The use of information analysis and evaluation centers to provide alerting and announcement of newsworthy items in the field, to assist in generating specialized bibliographies, to generate state-of-the-art reviews, etc.

d. The use of information analysis centers as a primary interaction point with members of the "invisible college" for a specialized area. The use of workshops and seminars to accomplish the evaluation of information in a given area. The rapid release and publication of the results

of review and evaluation.

Acknowledgements: In writing this paper the concepts expressed to us by F. W. Lancaster have been most useful. Prof. Lancaster has served as a consultant for our study project.

References

1. Market Facts Inc. *Development of Statistical Indicators for Scientific and Technical Communication*. Interim status report, June, 1975.

2. Lancaster, F. W. and A. M. Brown. *Conceptual Alternatives to the Scientific Journal*. Westat Research Inc., 1969.

3. Babo, Kertz, Morris, and Yeager. *A Survey of Fuel and Energy Information Sources*. Vol. I, II. MITRE Corporation, 1970. NTIS PB 197-386.

4. Crane, Diana. *Invisible Colleges. Diffusion of Knowledge in Scientific Communities*. University of Chicago Press, Chicago, 1972.

5. De Solla Price, Derek. *Little Science, Big Science*. Columbia University Press, New York, 1963.

6. King, Donald W. and Edward C. Bryant. *The Evaluation of Information Services and Products*. Information Resources Press, Washington, D.C., 1971.

7. Korfhage, Robert R. "Informal Communication of Scientific Information." *Journal of the American Society for Information Science* (January-February 1974): 113-117.

8. Kuipers, J. W., F. W. Lancaster, and R. W. Thorpe. *Effectiveness and Cost-Effectiveness Considerations for NASIC Information Services Operation*. QEI, Incorporated, Bedford, Massachusetts, 1973.

9. Lancaster, F. W. "The Cost-Effectiveness Analysis of Information Retrieval and Dissemination Systems." *Journal of the American Society for Information Science* 22 (January-February 1971): 12-27.

10. Lancaster, F. W. and J. Jenkins. "Quality Control Applied to the Operations of a Large Information System." *Journal of the American Society for Information Science* 21 (September-October 1970): 370-371.

11. Landau, H.B. *Research Study into the Effective Utilization of Machine-Readable Bibliographic Data Bases*. Final Report. Auerbach Corporation, Philadelphia, Pennsylvania, 1969.

12. MacLachlan, James. *A Plan for a Publication Network for Rapid Dissemination of Technical Information*. University of California, Berkeley, June, 1973. Distributed by ERIC, ED 086 210.

13. Marschak, Jacob. Economics of Information Systems. University of California, Los Angeles, November 1969, NTIS AD-699890.

14. Murdock, John. "A National Program for Energy Information." Bulletin of the American Society for Information Science (June-July 1974): 12-13, 32-33.

15. United States National Laboratory, Oak Ridge. Inventory of Current Energy Research and Development, for the Subcommittee on Energy of the Committee on Science and Astronautics, U.S. House of Representatives, U.S. Government Printing Office, Washington, D.C., January 1974.

16. Cooper, M. "Current Information Dissemination: Ideas and Practices", Journal of Chemical Documentation, vol. 8, no. 4, November 1968, pp. 207-218.

17. Heenan, W.F. and D.C. Weeks. Informal Communication Among Scientists. George Washington University, Biological Sciences Communication Project, 1971.

18. Allen, T.J. Managing the Flow of Scientific and Technological Information, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1966.

Experimental Design of Information Systems for Crises Management

John Sherrod
Informatics Inc.
Rockville, Maryland

W. David Penniman
Battelle Columbus Laboratories
Columbus, Ohio

INTRODUCTION

The United States has proceeded through a succession of crises during the past decade which is unparalleled in the Nation's history. Crises have occurred in such disparate areas as civil unrest, energy and environmental pollution. There is reason to believe that crises will continue to erupt in our complex society, perhaps with even greater frequency and with far-reaching consequences for our citizens. (2)

If this frightening trend continues, an unprecedented range of interdisciplinary responses will be necessary to cope with these critical situations. Certainly this era of crises has significant implications for science and, ultimately, scientific and technical information.

Ready availability of relevant information and subject expertise has been important in the past and will be essential in the future to assist crisis-related organizations in their decision- and policy-making activities. Just as crises of national proportions have placed increasing demands on managers and decision makers in government and industry, parallel pressure has been placed upon information system designers to create functional systems within a compressed time frame to meet these urgent requirements. It is the purpose of this study to formulate strategies and develop innovative techniques that will facilitate access to relevant scientific and technical information in a crisis situation.

DEFINITION OF CRISIS

There is no universal standard as to what constitutes a crisis. Therefore, for the purposes of this study, a crisis is defined as events of sufficient criticality to evoke a programmatic response at the national level. More specifically, the events must have sufficient social impact to result in corrective action by the Congress calling for new or modified Federal programs.

Particular interest in this study is given to those situations in which the legislation includes requirements for significant information programs. A few examples of such legislation from the recent 93rd Congress would include: the Juvenile Justice and Delinquency Prevention Act of 1974; the Solar Heating and Cooling Act of 1974; the Fire Prevention and Control Act of 1974; and, the Geothermal Energy Research, Development and Demonstration Act of 1974.

OBJECTIVES OF THE STUDY

A serious information deficiency has existed for the succession of crises that have occurred. Crisis managers often have had to make key decisions and establish basic policies with a minimum of information support. The major objective of the study is to identify means for substantially correcting this situation.

It is impossible to predict accurately where and when the next crisis will occur. Once a crisis is recognized, however, it is essential to reduce the length of time between that recognition and crisis action. Without knowing the exact nature of the crisis, it is not clear what the information service should be or how it should operate. The principal attribute of any such system, however, must be the capacity to respond rapidly to any crisis situation.

The particular system under study and design here is based on the referral principle. However, newer technology and developing management concepts should add improved features to the older concepts of referral.

Referral services are simple in concept and have been rendered on a more or less informal basis by libraries for many years. Unlike the normal library reference service, which generally provides either substantive or bibliographic information in response to an inquiry, a referral service acts as a clearinghouse to locate for the inquirer information sources that are likely to have reliable and relevant information, whether published or not.

The National Science Foundation was the first to recognize the value of this form of interface between the user and information as applied on a broad scale. They, in cooperation with the Library of Congress, established the National Referral Center for Science and Technology in 1962.[3] It met with immediate success. The Center's first directory was published in 1965 and sold more than 18,000 copies.

As successful as the National Referral Center has been, there are a few basic problems associated with its operations and some questions left unanswered with respect to ways for possibly improving upon the general referral concept.

The ultimate success of any referral activity depends in large measure on the accuracy and completeness of the information resources inventory contained in the data base. Since there are no standards or widely-accepted definitions of what constitutes an information resource, collecting input by questionnaire, as is the current practice, often leads to incomplete or inaccurate responses. This problem was particularly acute in the early stages of the Referral Center.

The Center also has had trouble helping users to understand that it is a referral service and not a reference service. The advantages of a referral activity over traditional bibliographic services are not easily recognized by the average information seeker. The publication of book-form directories, while almost the antithesis of a referral service, has served as an excellent marketing device. However, to obtain maximum utilization of the Center, much more aggressive marketing will be required.

The objective here is not to assess the performance of the National Referral Center for Science and Technology. However, the lessons learned from a little more than a decade of operation should prove useful in pinpointing steps that can be taken to improve the usefulness of scientific and technical information when short lead times are critical.

The function of the system is to achieve appropriate couplings between organizations involved in crisis management and the various information resources potentially available to them. This function is illustrated in Figure 1. The coupling function would be carried out using techniques closely related to traditional referral techniques but improved in a number of ways. Specifically the system will be designed to utilize:

1. Improved data gathering and indexing techniques to maximize the quality of the information and data files;
2. Computerization of the system to facilitate experiments with automatic referral of requests;

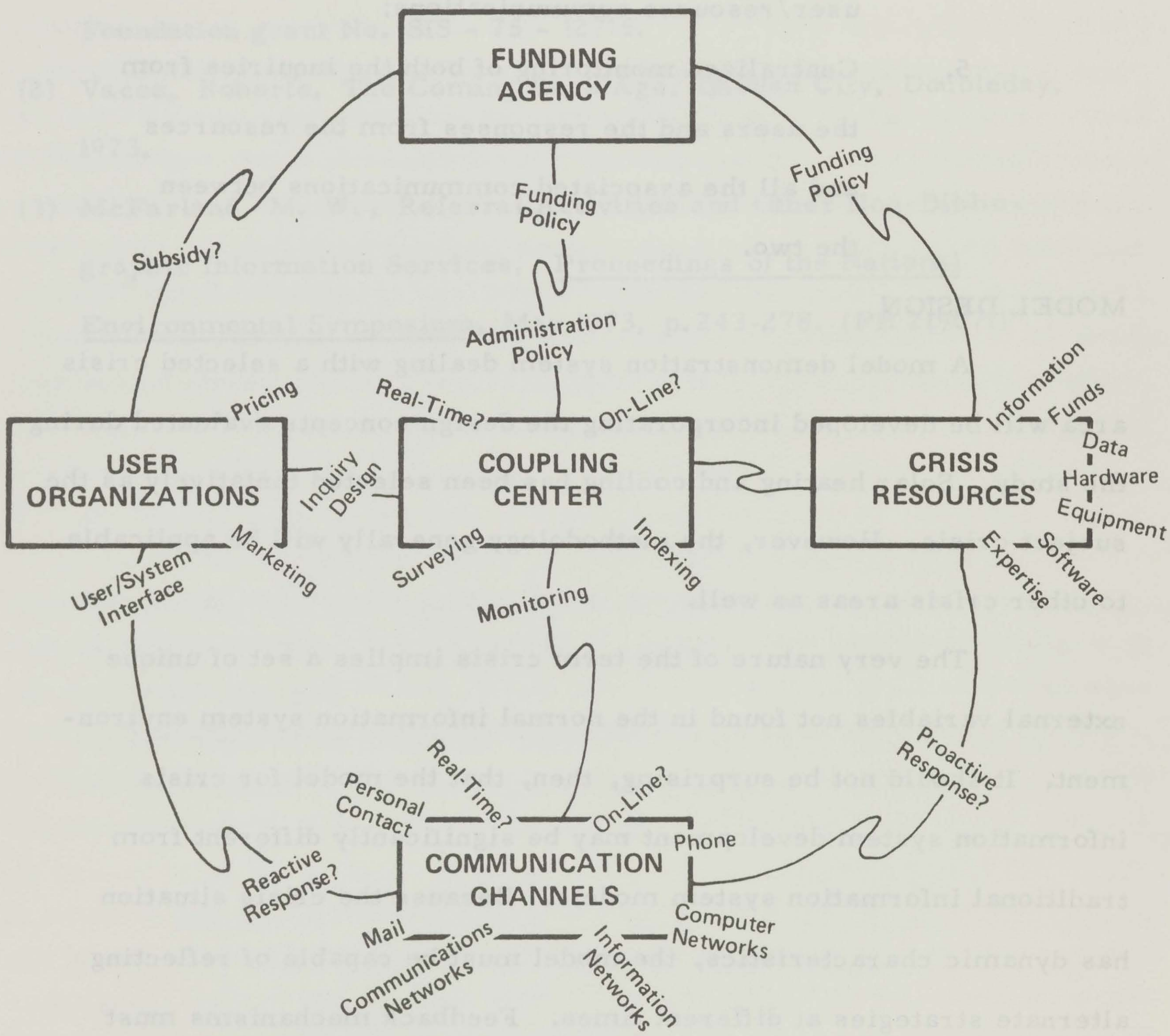


FIGURE 1 THE MULTIPLICITY OF CONSIDERATIONS

3. Decentralized responsibility for referral response to enhance rapid and direct communication between the user and the information resource;
4. Message-switching techniques to further enhance user/resource communications;
5. Centralized monitoring of both the inquiries from the users and the responses from the resources and all the associated communications between the two.

MODEL DESIGN

A model demonstration system dealing with a selected crisis area will be developed incorporating the design concepts evaluated during the study. Solar heating and cooling has been selected tentatively as the subject crisis. However, the methodology generally will be applicable to other crisis areas as well.

The very nature of the term crisis implies a set of unique external variables not found in the normal information system environment. It should not be surprising, then, that the model for crisis information system development may be significantly different from traditional information system models. Because the crisis situation has dynamic characteristics, the model must be capable of reflecting alternate strategies at different times. Feedback mechanisms must be reflected in this model to depict the complex relationships that exist between users, resources, facilitating organizations and communication channels as illustrated in Figure 1.

For the type of crisis defined here, the variables have a number of alternative values or options. When the modeling effort is complete, the selection of the optimum approach for development of a specific crisis information system should be made much easier.

REFERENCES

(1) This report was prepared with the support of the National Science Foundation grant No. SIS - 75 - 12715.

(2) Vacca, Roberto, The Coming Dark Age. Garden City, Doubleday, 1973.

(3) McFarland, M. W., Referral Activities and Other Non-Bibliographic Information Services, Proceedings of the National Environmental Symposium, May 1973, p. 243-278. (PB 219071)

Development of a Geothermal Thesaurus*

Jessie J. Herr
Information Research Group
Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720

Introduction

A thesaurus of terminology associated with geothermal energy is being developed by the Lawrence Berkeley Laboratory's (LBL) Information Research Group. This project is being carried out in collaboration with two groups: LBL's Geothermal Information Group (known as GRID for historical reasons), which is preparing the National Geothermal Information Resource, a data compilation (Phillips 1975); and the Energy Research and Development Administration's (ERDA) Technical Information Center (TIC), which is doing the indexing for ERDA's Energy Data Base (Alexander 1975). The Geothermal Thesaurus is a thesaurus for use in indexing and retrieval; but, in addition, it contains information that permits its use as an interface between indexing systems.

One way of viewing an information-retrieval thesaurus is as an interface between its users and an information storage and retrieval system. In retrieval, a thesaurus serves to translate the query, originally in natural language, into the vocabulary actually used for describing documents in the system and to suggest alternative ways of describing the subject of the search within the context of the particular system. During indexing, a thesaurus is used to translate the vocabulary of each document into the regularized vocabulary of the system and to indicate the document's relationship to others in the collection.

A thesaurus for use in an information storage and retrieval system basically consists of a collection of terms along with information about those terms. The most important type of information contained is the relationships between the terms, which are used to thread through the thesaurus to find the most appropriate term to represent a concept. Other types of information included are text, such as scope notes, and codes, for instance, those representing a broad categorization of the terms.

The Geothermal Thesaurus is an information-retrieval thesaurus and contains the types of information required for its use for vocabulary control and assistance during subject indexing and retrieval. However, in the Geothermal Thesaurus the concept of a thesaurus has been broadened

*Work performed under the auspices of the U.S. Energy Research and Development Administration.

beyond that of an interface between an indexer or searcher and the system governed by the thesaurus to include the role of interface between information systems. Its traditional role is carried out in large part through the network of links between terms within the geothermal vocabulary. Its role as interface between systems is based on inclusion in the thesaurus of links between the geothermal vocabulary and the vocabularies of other systems.

The linking of the geothermal vocabulary to other indexing vocabularies is a natural consequence of two inter-related aspects of the current status of data-base development. The first aspect is that of the increasing availability of bibliographic data bases in machine-readable form. The second is the recognition that there is overlap in the coverage of the various data bases and that, because they are in machine-readable form, there exists the possibility that material prepared for one data base with one set of users may be manipulated in order to satisfy the needs of another user population and be included in another data base.

The Geothermal Thesaurus contains links to a number of vocabularies; the most important and the most thoroughly developed links are those to the ERDA indexing vocabulary. These links are provided in anticipation of an active interchange of bibliographic data between the highly specialized National Geothermal Information Resource and the broadly based ERDA Energy Data Base. Because both data bases use the same indexing style, namely coordinate indexing based on a controlled vocabulary of indexing terms called descriptors, the major variance in indexing is associated with differences in the specificity of the vocabularies. These differences arise primarily from the facts that the National Geothermal Information Resource concentrates on geothermal energy and data, while ERDA's Energy Data Base is concerned with all aspects of energy sources and utilization and is bibliographic.

The development of the Geothermal Thesaurus will be discussed, beginning with an outline of its subject scope, sources and methods used in compiling the list of terms, and those parts of the thesaurus structure that are related to its use for vocabulary control and assistance during indexing and retrieval. Then, the techniques being used to link the Geothermal Thesaurus to other vocabularies will be described and some examples of their use given. Finally, thesaurus processing software developed at LBL that permits extension of the concept of a thesaurus will be briefly mentioned.

Subject Scope of the Geothermal Thesaurus

The subject scope of the Geothermal Thesaurus includes the characterization of geothermal resources, including geographic distribution and geological and hydrological properties; detection and evaluation of geothermal resources; extraction of geothermal fluids, including reservoir engineering, drilling techniques, and topics associated with the transport of the fluids, such as corrosion; utilization of geothermal resources for both electric-power generation and non-electrical uses, such as district heating; legal, economic, and environmental aspects of geothermal utilization; and basic laboratory studies related to the characterization and utilization of geothermal resources.

Compiling the List of Terms

A variety of sources have been used in compiling the list of candidate terms. One group of sources has been used to ensure that the vocabulary reflects that of the literature. In this group are indexing from TIC and GRID, indexing resulting from the LBL translation of the indexing for a 14,000-item bibliography (Summers 1971), and titles, abstracts, and indexing of geothermal papers retrieved in searches of a number of machine-readable data bases. Previous vocabulary efforts (thesauri, glossaries, indexing vocabularies) have been used to place the geothermal vocabulary in a more general context; the most important of these is the INIS Thesaurus (INIS 1974) upon which the ERDA indexing vocabulary is based. Data formats developed for the International Geothermal Information Exchange Program (Clark 1975), GRID specifications of frameworks for the analysis and description of data, and reports on research needs are important guides for terms to be used in data description. Finally, numerous reference works, review articles, and discussions with scientists and engineers engaged in geothermal research have provided an overview of the subject and the general outline of the thesaurus.

The tendency has been to include, rather than exclude, terms of unknown usefulness, and to provide paths through the thesaurus to make these terms accessible. The inclusion of questionable terms is feasible because of the limited size of the thesaurus at this time, and is desirable because the terminology requirements for data description and storage are under investigation.

Although the inclusion of terms is fairly unrestricted, those that are entered into the thesaurus are regularized in form in accordance with standard thesaurus-development practice. The list of terms will be reviewed periodically. Those that are used infrequently will be eliminated from the thesaurus so they won't impede locating those terms that are used; heavily used terms will be subdivided, if possible, to ensure effective retrieval.

At present, the Geothermal Thesaurus contains nearly 1300 terms, with about 2300 relationships between terms, more than 600 lines of text, and 2300 links between terms and codes. Nearly all subjects within the scope are represented by terms of sufficient specificity for general subject indexing. More detailed terminology is available for those subjects upon which GRID has concentrated (geological setting of geothermal resources, geochemical and geophysical exploration, and thermodynamic properties of aqueous solutions of interest in geothermal systems).

The Structure of the Geothermal Thesaurus

The relationships that have been used in the Geothermal Thesaurus are the following (*indicates links used to relate the Geothermal Thesaures to other vocabularies):

Links between terms

USE	USE instead of (USE...AND...)
UF, UF+	Used For, Used For in combination
SEE	See (SEE...OR...)
SF	Seen From
BT	Broader (more general) Term
NT	Narrower (more specific) Term
RT	Related Term
*GT	Goes To (translates to)
*CF	Comes From (translates from)

Links between terms and text

SN	Scope Note
DF	Definition

Links between terms and codes

SC	Subject Category
*SO	Source of the term

Information about Terms within the Geothermal Vocabulary

USE references lead the thesaurus user from a term that is not acceptable in the system to one that is. Often the terms are synonyms; to avoid dispersing information in the system, one of the pair is chosen as the term to be used. For instance, ALUMSTONE and ALUNITE are synonyms for a mineral of interest in geothermal systems and the Geothermal Thesaurus contains the entry,

ALUMSTONE
USE ALUNITE

which is read as, ALUNITE is to be used instead of ALUMSTONE. All term–term relationships are reciprocated, and the entry for ALUNITE includes the reciprocal of the ALUMSTONE entry,

ALUNITE
UF ALUMSTONE

The USE reference is also used in connection with abbreviations, for instance,

UNITED STATES OF AMERICA
USE USA

USA
UF UNITED STATES OF AMERICA

Occasionally, a term is designated as unacceptable for use and the concept is to be represented by a combination of terms. In such a case, the reciprocal UF+ is used to indicate that the term is used in combination with other terms as a replacement for an unacceptable term.

The SEE reference refers from an unacceptable term to more than one alternative terms, any one of which would be acceptable, for instance,

CONDUCTIVITY
SEE ELECTRIC CONDUCTIVITY
OR THERMAL CONDUCTIVITY

with the reciprocals

ELECTRIC CONDUCTIVITY
SF CONDUCTIVITY

THERMAL CONDUCTIVITY
SF CONDUCTIVITY

The Broader Term and Narrower Term relationships show class membership or geographic inclusion. A BT relationship leads the user to a more general term and an NT leads to a more specific term, for instance,

OFFICE BUILDINGS	COMMERCIAL BUILDINGS
BT COMMERCIAL BUILDINGS	NT OFFICE BUILDINGS

In addition, the Thesaurus displays all of the broader or narrower terms for a particular term, this is called display of the term hierarchy. Continuing the previous example, COMMERCIAL BUILDINGS has the broader term BUILDINGS (the number following the relationship label indicates the level in the hierarchy),

COMMERCIAL BUILDINGS
BT1 BUILDINGS
NT1 OFFICE BUILDINGS

which leads to the fully developed BT hierarchy for OFFICE BUILDINGS and the NT hierarchy for BUILDINGS,

OFFICE BUILDINGS	BUILDINGS
BT1 COMMERCIAL BUILDINGS	NT1 COMMERCIAL BUILDINGS
BT2 BUILDINGS	NT2 OFFICE BUILDINGS

The BT relationships assist the user in broadening a search and the NT's help in narrowing a search, or making it more specific.

The Related Term relationship is used to link terms that are not related by class membership, but indicating their existence may be helpful to the user in some way, for instance,

SALINITY	BRINES
RT BRINES	RT SALINITY

Terms that are sometimes, but not always, related by class membership are entered as Related Terms, instead of as a BT/NT pair.

Although the meaning and intended use of a term usually are defined by the terms linked to it, occasionally, further information will assist in the effective use of the thesaurus. This information is contained in text called scope notes, which are used to make the scope of use of a term absolutely clear.

As a normal part of thesaurus development, definitions have been obtained for many of the terms included in the thesaurus. Whenever possible, these definitions have been stored as a part of the information associated with the terms. Not all terms have definitions; however, an attempt has been made to include definitions for terms that are particularly important in geothermal energy or whose definitions are not readily available. Definitions are included as a special type of text (rather than as scope notes), so that they can be selectively printed.

Terms are assigned to one or more subject categories, which are represented by four-character codes. These codes are used to list the terms by subject ctaegory to provide an alternative to the alphabetical thesaurus as a means for finding the most appropriate terms. In the Geothermal Thesaurus, three types of subject categories have been used: (1) mission-oriented categories, such as geothermal exploration; (2) discipline-oriented categories, such as chemistry; and (3) categories representing long lists of similar terms, for instance, minerals.

Links between the Geothermal Thesaurus and Other Vocabularies

All of the relationships and other types of information described above are associated with the use of the Geothermal Thesaurus as an interface between an indexer or searcher and the data base. Next, we turn to those parts of the Thesaurus related to its use as an interface between its vocabulary and that of other systems. The term–code link Source and the term–term links Goes To and Comes From are used to tie the Geothermal Thesaurus to other vocabularies. The Source code identifies the vocabulary to which a term belongs. At present, the most significant source code is that which indicates that the term is a part of the ERDA indexing vocabulary. Other source codes are being included in anticipation of the use of the thesaurus to assist in screening machine-readable data bases for geothermal-related material.

The Goes To and Comes From links are used in conjunction with the source code indicating the ERDA indexing vocabulary. They will be used to translate the highly specialized vocabulary being used by GRID into the more general vocabulary used for ERDA indexing. Similarly, they will be used to translate ERDA indexing into the GRID vocabulary; however, only a partial translation will be possible because it will be from the general to the specific.

There are three general situations in which these inter-vocabulary links are being used. In the first, the vocabulary for GRID is more specific than that for ERDA. As an example, the ERDA terminology includes GEOTHERMOMETERS, while that for GRID also includes two specific types of geothermometers, GEOCHEMICAL THERMOMETERS and ISOTOPE GEOTHERMOMETERS. In this case, the Thesaurus contains the following entries:

GEOCHEMICAL THERMOMETERS	ISOTOPE GEOTHERMOMETERS
SO GRID	SO GRID
GT GEOTHERMOMETERS	GT GEOTHERMOMETERS

with the reciprocals,

GEOTHERMOMETERS
SO ERDA GRID
CF GEOCHEMICAL THERMOMETERS
CF ISOTOPE GEOTHERMOMETERS

With these entries, if an item indexed with GEOCHEMICAL THERMOMETERS by GRID is to be included in ERDA's Energy Data Base, this term will first be replaced by GEOTHERMOMETERS to make the indexing compatible with the rest of the Energy Data Base.

Another case in which the Goes To relationship is utilized is that of terms that are unambiguous within the limited context of geothermal energy, but which would be ambiguous in the broader context of the Energy Data Base. Such terms are qualified in the ERDA vocabulary to make their meanings quite explicit; however, this qualification is unnecessary and artificial within the geothermal-only scope. An example is the term FAULTS, which appears in the ERDA vocabulary as GEOLOGIC FAULTS. To permit each to use the most appropriate form, while maintaining compatibility, the Geothermal Thesaurus contains

FAULTS	GEOLOGIC FAULTS
SO GRID	SO ERDA
GT GEOLOGIC FAULTS	CF FAULTS

Finally, the Goes To relationship is applied to cases in which there are differences in the meaning of terms, some of which are common to both vocabularies. For instance, both systems define a series of pressure ranges. The ERDA vocabulary defines HIGH PRESSURE to cover the range from 100 to 1000 atmospheres, while GRID subdivides the range into ELEVATED PRESSURE, for 100 to 500 atmospheres, and HIGH PRESSURE, for pressure greater than 500 atmospheres. To provide compatibility, the Thesaurus includes the entries,

ELEVATED PRESSURE	HIGH PRESSURE
SO GRID	SO ERDA GRID
GT HIGH PRESSURE	CF ELEVATED PRESSURE

An item indexed by GRID with the term HIGH PRESSURE would not be modified; however, the term ELEVATED PRESSURE would be transformed to HIGH PRESSURE before input to the ERDA system.

Thesaurus Processing Software

The Geothermal Thesaurus is being processed using software developed at LBL. The major purpose of this software, as with most thesaurus processing systems, is to reduce the clerical work associated with thesaurus building. The software performs tasks such as: checking the spelling of terms being entered into the thesaurus; reciprocating all term–term and term–code links; checking the consistency of term–term links; and generating the hierarchies for BT and NT relationships.

In most thesaurus processing systems, the types of relationships are embedded in the program, that is, the program specifies the types of relationships that can be used. In LBL's system, the thesaurus builder specifies the structure, or types of relationships, to be used and this may be different for every thesaurus. This feature was included to permit the addition of nonstandard relationships, such as Goes To and Comes From, and to permit the processing of thesauri developed using other software packages, each of which defines a different set of relationships.

A thesaurus publishing program is under development to permit the flexible formatting of the printed thesaurus and the selection of the information to be printed. With it, if desired, definitions could be printed in an auxiliary publication rather than in the thesaurus proper, and information, such as term sources and the vocabulary-translating relationships, which is not of interest to most thesaurus users, could be printed only in the thesaurus editor's master copy of the thesaurus.

Summary

At present, the Geothermal Thesaurus contains 1300 terms. Nearly all subjects within the scope are represented at a specificity level sufficient for general subject indexing; some topics are covered in the detail required for the description of data and for capturing text-type data, such as the geological setting of a geothermal field. The Geothermal Thesaurus contains information, usually associated with information-retrieval thesauri, which is designed to assist its users in finding the most appropriate term(s) to represent a concept. In addition, it contains information that relates its vocabulary to other vocabularies. In particular, it has links to the ERDA indexing vocabulary, which will be used to provide compatibility between a highly specialized, data-oriented geothermal-energy information system and a broadly based, bibliographic system concerned with all aspects of energy. Thesaurus software developed at LBL performs the normal processing functions and permits flexible definition of the structure of a thesaurus.

References

C. G. Alexander, "ERDA Information Programs," to be presented at the Tenth Annual Meeting of the Geoscience Information Society, Salt Lake City, 21 October 1975.

A. L. Clark, J. A. Calkins, E. Tongiorgi, and E. Stefanelli, "A Report on the International Geothermal Information Exchange Program" presented at the Second United Nations Symposium on the Development and Utilization of Geothermal Resources, San Francisco, May 1975.

INIS Section, International Atomic Energy Agency, INIS: Thesaurus, IAEA-INIS-13 (Rev. 7), July 1974.

S. L. Phillips, "The National Geothermal Information Resource," to be presented at the Tenth Annual Meeting of the Geoscience Information Society, Salt Lake City, 21 October 1975.

W. K. Summers (comp.), Annotated and Indexed Bibliography of Geothermal Phenomena, New Mexico State Bureau of Mines and Mineral Resources, Socorro, 1971.

TECHNICAL INFORMATION PROGRAMS OF THE FEDERAL ENERGY ADMINISTRATION

T. M. Albert, Director
National Energy Information Center
Federal Energy Administration
Washington, D.C. 20461

The Federal Energy Administration was established in May 1974 under Public Law 93-275. It was set up to take action to conserve energy supplies; to insure fair and reasonable consumer prices, and fair and efficient distribution methods for energy supplies; to expand usable energy sources; and to develop policies to meet the energy needs of the nation.

To accomplish the tasks set before the FEA and to get a handle on the energy problem today requires the collection and evaluation of great quantities of numerical data-engineering and statistics, as well as the entire gamut of related R&D information. The extent of the energy-related R&D information is well known. The amount of documents runs into the millions.

Engineering data exchange is necessary if we are to properly turn research into working systems. Numerical data on materials of all kinds is obviously extremely important. And, finally, in order to understand our energy problems - to forecast, to plan - we must collect tremendous quantities of statistical data of all kinds: reserves, production, transmission, transportation, demand and consumption, conservation, and financial.

At FEA, many data collection forms have been designed and utilized and, of course, large scale computer systems are used extensively. Data is collected on a wide range of topics; for example: oil imports, crude oil allocation, sites needed for energy facilities, gasoline market shares, cost equalization, fuel availability, petroleum production and pricing, electricity generation, coal reserves, asphalt production, production of solar collectors, and much more.

To give some idea of the scope of the problem, I will provide some numbers. In the Federal Government alone, more than 43 agencies carrying on more than 350 separate programs, collect some form of energy-oriented statistics. The quantity of individual items of numerical data collected is staggering. No one has taken an inventory, but FEA has made a beginning and is proposing to do more. The amount of duplication is probably quite extensive.

FEA alone has something in the neighborhood of 70 statistical data-collecting systems. All involve data-collection forms, surveys, and data processing. Details are available on 28 of these.

The 28 systems involve both monthly and quarterly data collection. Approximately 3,680,000 records are collected and processed monthly, and approximately 361,000 quarterly. We can estimate that each of the records is approximately 100 characters long. For the

28 systems alone, this extends to about 45.6 million records annually and about 4.6 billion characters annually.

The data-collection effort of the U.S. Bureau of Mines (BOM) is more extensive in items of data collected than is FEA's. One data-collection form alone at BOM involves the collection of approximately 1.2 million fields of data. The data-collection effort of the Federal Power Commission is very extensive; consider also the U.S. Geological Survey and many universities and private organizations. We could go on, but I think the point is made, and we have not even considered the quantities of data on materials and engineering.

Considering the extent of the energy crisis today, some attempt at controlling the available information must be made. To provide effort in this direction, the National Energy Information Center (NEIC) in FEA was established and organized to serve as an information center and a national clearinghouse for technical energy information for federal and state agencies, the Congress, industry and the public. We are attempting to establish a point of control or a conduit for the massive quantities of information available. The NEIC will collect, index, process, retrieve, disseminate, and publish information about energy, and will establish national programs for exchange of such information.

As a clearinghouse, we are concerned with two aspects of information handling: first, to provide answers to individual, unique requests; and, second, to provide for the problem of exchange, storage and use of large quantities of data or data bases.

The NEIC is organized in three divisions: Systems Services, Reference and Referral, and Technical Information Services. The Systems Services Division functions include micrographic use, development and maintenance of automated on-line bibliographic/text data bases, FEA data base administration, and FEA forms management.

The Reference and Referral Division handles response to mail and telephone queries for technical information, referring to a proper source if necessary, and doing limited research as necessary. The Technical Information Services Division functions include: production and distribution of publications, collecting, indexing and cataloging pertinent reference material; surveys and the updating of the Federal Energy Information Locator System (described below); and standards development.

There is also a staff function concerned with liaison with the states and other federal agencies.

More detail about some of NEIC's work may be of interest. For example, our people have available for use more than 70 on-line bibliographic data bases; NEIC is also a station on the ERDA-RECON network.

The NEIC is in the process of setting up a Federal Inter-Agency Council for energy information. This group would provide for better cooperation and coordination in the federal government concerning energy projects and data/information collection and exchange. We are also exploring the possibility of a State-Federal Committee along the same lines.

Realizing that no one group can know everything we are attempting to establish a national referral system which would cover energy information sources and contacts. This is to encompass the private sector as well as the federal and state agencies.

A task force has made a lengthy survey in the federal government which has resulted in the Federal Energy Information Locator System (FEILS). Basically, we have information describing the energy-related efforts of 43 agencies and encompassing over 350 programs within these agencies. The file includes details such as a description of systems, data bases, data collection forms used, a list of output reports, and contacts. The information will be avaiable in published form in December. The file will be available as an automated on-line searchable system in about 90 days.

FEILS will continue to be updated and expanded. We hope to increase the amount of detail we have on the data in the data bases and to add abstracts of the output reports and publications.

Within 60 days there will be available in published form a Directory of State Government Energy Agencies.

At FEA, we will have completed by November, an FEA data element dictionary. In an on-line system, which is already operating, we have three linked files which describe FEA's data collection forms, the data elements on the forms, and the systems which utilize the data. As a future possibility, we are looking into the possibility of including under the same system other agencies' data element dictionaries.

At NEIC, our approach is to first establish files of data about data. What exists, what is it, and where. The data element dictionary and FEILS are examples of this. Along the same line, we are exploring the feasibility of collecting and establishing a file of citations for every energy-related document in the agencies covered by FEILS.

The programs mentioned above describe some of NEIC's efforts. Our goal is to provide a highly knowledgeable source for facts on energy economics, resources, and operations; and a source and referral point for R&D information. It is a massive job. It is also a job we can do better if the members of the energy community, including yourselves, know about us, work with us, and utilize our services. By working with us, I mean that we hope you will notify us of pertinent papers, activities, and data bases; work with us to establish data exchange programs as appropriate; and help us to build and update our referral system.

In addition to the NEIC programs, of course, other parts of the FEA such as Conservation and Environment, and Energy Resource Development (ERD) have important on-going programs. For example, ERD is drafting legislation which would create a national system for the acquisition and storage of rock and core samples. Facilities would be provided in a number of locations for storage of the material, and the indexing system for control and retrieval of the material as well as information about the samples would be developed by the U.S. Geological Survey.

Finally, I wish to urge the members of the Geological Society of America to utilize the services of the NEIC. If you have a question or need a fact get in touch with us. The address is:

The National Energy Information Center
Federal Energy Administration
Federal Building, Room 1404
Washington, DC 20461
202-961-8685

The energy community covers a wide range of disciplines and knowledge. The community can't afford to be compartmentalized if this country is to solve its energy problems. Perhaps the National Energy Information Center can make it easier for all of us to work together.

National Geothermal Information Resource

S. L. Phillips, J. A. Fair, F. B. Henderson, and T. G. Trippe
Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720

That development of the United States energy resources is essential to our nation's progress and economic well-being cannot be questioned. Based on the availability of domestic energy sources, industrial and electrical utility plants are modified, private and public transportation is affected, and rules and regulations concerning energy are changed. A very thin line separates an uneconomical energy resource from a viable reserve upon which so much of our current and future lifestyle depends. The need for sound data is obvious.

In order that geothermal energy may more readily contribute its share to the energy needs of the United States, there must be an effectve transfer of information covering all aspects of geothermal exploration, evaluation, and utilization. The gap between the time when data are obtained and when it becomes public knowledge must be minimized. Of equal importance to compilation and dissemination are both the evaluation and the critical review of the status of numerical data.

In this work, numerical data are defined as the magnitudes of specific physical or chemical properties of well-characterized geothermal gases, solutions, and rocks and minerals (Ref. 1). These data are extracted from the literature, evaluated, and recommended values together with estimated limits of reliability are determined by the National Geothermal Information Resource (GRID).

Besides evaluation, timely reviews on the current status of information are important in a rapidly growing field such as geothermal science and technology. The reviews provide the national geothermal energy program with annotated and indexed references to unevaluated data for use in preliminary estimates, calculations, and modeling.

In this report, the principal GRID categories for compilation of geothermal information are described, and examples given of current activities. Included also is a brief description of our computer-aided document file.

Introduction

The Lawrence Berkeley Laboratory is sponsored by the U. S. Energy Research and Development Administration to establish a National Geothermal Information Resource. The objective of the GRID program is mainly to collect, organize and disseminate evaluated data on the following 6 major categories of geothermal science and technology: (1) physical chemistry, (2) exploration and evaluation, (3) utilization, (4) environmental aspects, (5) institutional considerations, and (6) reservoir characteristics. For example, current GRID activities include the following: (1) development of a thesaurus of geothermal data descriptors. This is part of an effort centered around dissemination of data in computer-processable form on a multilateral and global basis. (2) A review and evaluation of numerical data covering the basic physicochemical properties of geothermal vapors, solutions, and rocks at elevated temperatures and pressures. For example, the density of sodium chloride solutions. (3) Current bibliographic information on exploration, with a review of the status of data on geochemical methods. (4) Electrical and non-electrical uses of geothermal energy. An example is the compilation and evaluation of fundamental data on hot water (brine) transport and space heating applications. (5) Computer-aided updating of environmentallyoriented geodetic maps.

We have found that the geothermal literature generally falls into one of the following classes: (1) Unevaluated numerical data covering basic physical chemistry. The change in density of sodium chloride solutions at elevated temperatures and pressures is an example of laboratory-generated data. (2) Evaluated and analyzed information which is scattered through the literature; this is compiled and indexed by GRID so that it will be available for quick retrieval by the user. Geochemical exploration is an example of this class of information. (3) Data on important aspects of geothermal development that are neither widely disseminated nor evaluated. An example of this class is geothermal subsidence.

The discussion which follows centers around a description of the information contained in these six major categories and examples of results of the work.

Physical Chemistry

Geothermal physical chemistry covers laboratory measurements and theoretical calculations of basic scientifc data on gases, solutions, and rocks covering initially temperatures to 500°C and pressures to 500 bars. (See Table 1). Such data are important to understanding and predicting the behavior of a reservoir which is used for either electrical or other applications. For example, data on the solubility of silica and calcium carbonate are needed to predict scaling in pipes caused by precipitation due to a change in operating temperature or pressure as compared with that at depth. Density data are important in calculating partial molal volumes and in calculating and evaluating hydrostatic pressure relationships. For example, density data are needed to convert pressures into the

Table 1. GEOTHERMAL PHYSICAL CHEMISTRY
(See Ref. 2, 3)

Basic scientific numerical data on gases, solutions, rocks at temperatures to 500°C, pressures to 500 bars; covering the evaluation and status of:

A. SOLUTIONS

1. Sodium, potassium and calcium chlorides

a. Density	e. Enthalpy change
b. Vapor pressure	f. Heat capacity
c. Heat of solution	g. Electrical conductivity
d. Partial molal heat content	

2. Ferric chloride, ferrous chloride, sodium sulfate

a. Heat of solution	c. Heat capacity
b. Partial molal heat content	d. Enthalpy change

B. ROCKS AND MINERALS

1. Heulandite - clinoptiloite; chlorites; illite and montmorillonite clays; pargasite - ferropargasite amphibole:

a. Enthalpy	d. Heat capacity
b. Entropy	e. Porosity
c. Free energy	

2. Others (e.g., zeolites, micas, glasses)

C. ROCK-SOLUTION INTERACTIONS

1. SiO_2 solubility vs. pH
2. SiO_2 - silicates - H_2O equilibria

D. ENERGY CONVERSION

1. Isobutane properties

E. CORROSION AND SCALING

equivalent of hydrostatic depth for $NaCl-KCl-CaCl_2$ solutions; this permits estimating the maximum temperature likely at a given depth in a convecting hydrothermal system. While there are data available for the density of chloride solutions, it is contained in a large number of publications. The data are both difficult to access and largely unevaluated. See Table 2 for typical uses of density data.

Table 2. Typical applications of density data to geothermal systems.

Direct Use	Calculated
Depth-boiling point curves	Partial molal volume
Fluid inclusion thermometry	PVTX
Mineral solubility	
Hydrodynamics	Free energy
Reservoir modeling	

Additional information on the needs for basic scientific data are found in A Recommended Research Program in Geothermal Chemistry (Ref. 2), and the Conference on Thermodynamics and National Energy Problems (Ref. 3).

Exploration and Evaluation

Geothermal resource exploration and evaluation activities center around reviews of the status of data, and indexed abstracts on the methodology used to locate and evaluate geothermal resources covering the following principal categories:

Geology	Hydrology
Geochemistry	Land-Use Factors
Geophysics	Resource Assessment
Drilling	

These techniques are used to: (1) locate subsurface regions which contain a high temperature anomaly, and (2) estimate the extent of the reservoir, its temperature and its potential for use in power production or non-electrical applications. See Table 3. As shown in the table, geochemical data are an important aspect of exploration methodology.

Geochemical data are used mainly (1) to infer the temperature of the reservoir at depth from chemical analysis of surface waters. The SiO_2 and Na-K-Ca content of brines are two examples of commonly used chemical geothermometers. See Table 4. Geochemical sampling and gas and water analyses are based on standardized methods, for example, see Reference 4. (2) To provide information on the sources of reservoir original and recharge water by measuring isotope concentrations, for example, the ratio $^{16}O/^{18}O$. (3) To predict the likelihood of problems which may be encountered in utilization of the fluid. Silica precipitation which can cause scaling, low fluid pH which enhances corrosion, and H_2S content which may pose an environmental problem are three examples.

Additional information on geochemical exploration methodology is obtained from a number of publications. For example, D. E. White's section on "Geochemistry Applied to the Discovery, Evaluation, and Exploitation of Geothermal Energy Resources" (Ref. 8).

Environmental

The GRID compilation of environmental aspects of geothermal energy currently covers a review of the status of data on man-related effects to Air, Water, and Land compartments of the environment due to the exploration and utilization of geothermal energy. See Table 5. Generally, a review of the status of geothermal environmental data will center on the following information: (1) statement of the problem, (2) effects of the environmental quality parameter to Air, Water, or Land, (3) methods

TABLE 3.

Typical Methods for Geothermal Exploration

(See References 5-8.)

Typical Data Sought	Commonly Used Method
SURFACE	
Site Location	Literature search; observation of hot springs, fumaroles; infrared aerial photography.
Site Evaluation	Rock hydrothermal alterations.
Measurements:	
Geochemical	Dissolved constituent concentrations; isotope ratios.
Geophysical	Electrical resistivity; temperature gradient; gravity; seismic; magnetic; drilling (< 1 km depth).
Inferred (Predicted):	
Reservoir Temperature	Dissolved constituents.
Recharge Water Sources	Isotope ratio.
Reservoir Depth	Electrical resistivity.
Category of Reservoir:	
Vapor	Electrical resistivity; Cl concentration.
Hot Water	Electrical resistivity; Cl concentration.
Power Potential of Field (Megawatt-years)	Electrical resistivity; reservoir temperature; surface temperature gradient.
Model of Reservoir	Data from measurements.
SUBSURFACE	
Category of Reservoir	Drilling (≥ 1 km depth).
Vapor	Fluid composition.
Hot Water	Fluid composition.
Temperature of Reservoir	Drilling.
Rock Properties:	Drilling; core samples.
Porosity	Laboratory testing.
Permeability	Laboratory testing.
Density	Laboratory testing.
Hydrothermal Alterations	Petrological examinations.
Enthalpy	Ratio of heat flow over fluid mass flow.
Model of Reservoir	Data from drilling samples.

Table 4. Commonly Used Aqueous Chemical Geothermometers (Ref. 5, 6, 8)

Quantitative Temperature Estimates

SiO_2

Na - K - Ca

Qualitative Temperature Estimates

$CaHCO_3$

Cl - dilution

Temperature Indicators

Mg

Mg/Ca

Na/Ca

Na/K

Na/Rb

Na/Li

$Cl/(HCO_3 + CO_3)$

Cl/F

Cl/SO_4

Dissolved volatiles (B, NH_3, Hg)

Trace Elements (Hg, W, Sb)

used for control and abatement, (4) mechanism and pathway of the parameter, (5) instrumentation for monitoring and measuring the environmental effect, and (6) recommendations in areas where data are needed because it is either currently insufficient or unavailable.

Table 5. Typical environmental parameters associated with geothermal energy resources. (Ref. 9, 11, 12)

AIR:	H_2S, NH_3, Toxic Metals (e.g., Hg, As, Se)
WATER:	H_2S, Boron, Silica, Toxic Metals
LAND:	Soil Degradation, Noise, Land Use
SUBSURFACE:	Subsidence, Seismicity

Additional information on environmental aspects of geothermal energy may be obtained from An Environmental Study of the Wairakei Power Plant by Axtmann (Ref. 9) and the Bibliography of Geothermal Resources (Ref. 10).

Utilization

This category contains information on the current and expected electrical and non-electrical uses of geothermal energy including the following:

- Hot Water (Brine) Transport
- Space, Process, and Agricultural Heating
- Power Generation
- Binary Cycle Power Generation
- Corrosion, Erosion, and Scaling
- Resource Evaluation

Institutional

This category covers Federal, state, and local organizational, legal, and regulatory considerations in the development of geothermal energy:

- Land Use
- Exploration and Production
- Operating Regulations

Industry Financial Incentives
Sale of Geothermal Power
Fluid Transport

Reservoir Characterization

Reviews and evaluation of data relevant to the development and production of wells covering:

Porosity
Artificial Stimulation
Natural Recharge
Artificial Recharge
Modeling
Well Tests and Measurements

Additional information on utilization, institutional, and reservoir characterization is contained in "Geothermal Energy: National Proposal for Geothermal Resources Research" (Ref. 13), and "National Plan for the Development of Geothermal Energy" (Ref. 14).

Documentation File (GEODOC)

GEODOC is a computer based file which contains the descriptive cataloging and indexing information for all documents processed by the National Geothermal Information Resource Group. This file (along with other GRID files) is managed by DBMS, the Berkeley Data Base Management System. Input for the system is prepared using either the IRATE Text Editing System with its extended (12 bit) character set, or punched cards. (Ref. 15).

Each record in the GEODOC file contains the descriptive cataloging, abstracting, and indexing information corresponding to a single document. The information within a given record is subdivided into "data elements". Table 6 lists the definitions of all the data elements which may appear in a GEODOC record. Some data elements (e.g., authors name) can occur repeatedly within one record. An "m" in the third column of Table 6 indicates that such multiple occurrences are allowed. The "LBL tag" for each data

Table 6. GEODOC Data Elements
(Ref. 15)

LBL Tag	INIS Tag	m*	n*	Data Element Definition
SC	008			document short code: unique identifier for document
TY				type of document/bibliographic levels/literary indicator
DES-CAT		m	n	delineates information for one bibliographic level
BL	009			bibliographic level indicator
PT	200			primary title (translated into English if necessary)
PS	201			primary subtitle (translated into English if necessary)
TA	620			title augmentation
L	600			language (for non-English document)
OT	230			original title (non-English) or journal/series title
OS	231			original subtitle (non-English) or journal/series subtitle
ED	250			edition
CODEN				journal CODEN
AUTHORS		m	n	delineates author-affiliation group
AU	100	m	n	author's name
AN	100			author note (ed., comp., eds., comps.)
AA	100	m		author's affiliation
AC 700				affiliation code
CE	110	m		corporate entry
CC	710			corporate code
DG	111			academic degree
SPO		m		sponsor
SPC				sponsor code
SCN		m		sponsor contract number
RN	300			report or patent number
SN	310			secondary numbers
INT	320			International Standard Book Number or Patent Code
PUB	402			publisher
PUP	401			place of publication
PUD	403			publication date
COL	500			collation (volume, issue, page)
N	610			note
COT	210			conference title
COP	211			conference place
COD	213			conference date
AV				availability and price

Table 6. GEODOC Data Elements (Con'd.)

LBL Tag	INIS Tag	m*	n	Data Element Definition
REL-REF		m	n	delineates information for one related reference.
RL				relator
RLR				relationship and reference
RSC				related short code
ABSTRACT		m	n	delineates one abstract
ABS				abstract
ABSO		m		abstract source
INDEX		m	n	INDEX.1 general indexing. INDEX.2, 3,...N splits
CQ		m		category/qualifier
TICC		m		TIC category
DE	800	m		descriptor from thesaurus
DD	800	m		data descriptor from thesaurus
ID				identifier
PD	810	m		proposed descriptor
CONTROL			n	internal LBL data elements
LA		m		local availability
BR				borrow/return
DCSO		m		descriptive catalogers initials, date and comment
AISO		m		abstractor-indexers initials, date and comment
DATA-FILE		m		data file name
POT		m		data descriptor for potential data
IN		m		data descriptor for included data

*m-This data element may have multiple entries

*n-This data element contains no value and need not be entered on input. It serves to delineate a group of data elements.

element is shown in the left hand column of Table 6. These tags are used to label the data elements within a record.

The data elements bear certain hierarchical relationships to each other; that is, some data elements are subordinate to, or, conversely, parents of, others. Figure 1 is a diagram of this hierarchical structure. The structure is also indicated in Table 6 by indenting the tag names of subordinate data elements and placing them after their parents. The data elements may be input to the system in any order except that subordinate data elements must follow the occurrence of the parent to which they are associated.

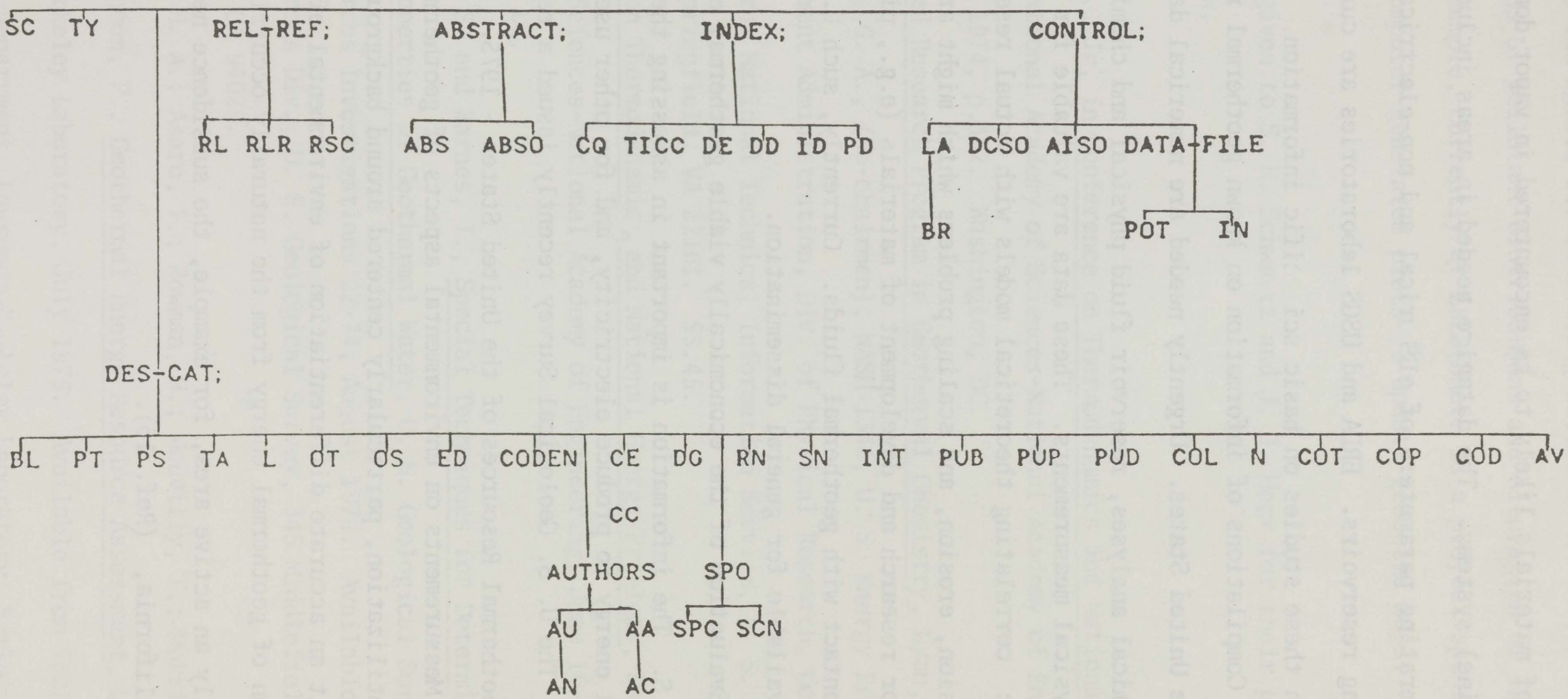
The GEODOC descriptive cataloging techniques are modeled after those of the International Nuclear Information System (INIS) of the International Atomic Energy Agency (Ref. 15). The data elements have been chosen to correspond as closely as possible with those used by INIS to facilitate information exchange on a multilateral and global basis. An INIS tag is shown in Table 6 for those data elements whose internal format is identical or almost identical to the GEODOC format. This is primarily for convenience in referencing INIS documentation.

Other GRID files similar to GEODOC but with appropriately defined data elements will be used to store and retrieve numerical data (e.g., physical chemistry data).

Conclusions

After reviewing the current geothermal resources information in light of the data requirements, some general observations are appropriate. While data currently available are satisfactory in some respects, there are still some inadequacies. The optimum future data activities should include:

Fig. 1 GEODOC Record Structure (Ref. 15)



(1) Laboratory measurements of the thermodynamic and transport properities of materials likely to be encountered in vapor-dominated and hot-water (brine) systems. The data are needed in areas including prediction of the operating parameters of electrical and non-electrical plants, and in modeling reservoirs. ERDA and USGS laboratories are currently actively engaged in these studies on basic scientific information.

(2) Compilations of information on known geothermal resource areas within the United States. Urgently needed are numerical data on surface water chemical analyses, reservoir fluid physical and chemical properties, and geophysical measurements. These data are valuable for many uses including: correlating theoretical models with actual reservoirs; indicating corrosion, erosion, and scaling problems which might arise; providing a basis for research and development of materials (e.g., piping) which come in contact with geothermal fluids. Currently, such U.S. data are not readily available for general dissemination.

(3) Evaluation of the economically viable geothermal energy resource of the U. S. The information is important in assessing the potential for geothermal energy to produce electricity, and for other uses (e.g., space heating). The U. S. Geological Survey recently issued a report on Assessment of Geothermal Resources of the United States - 1975. (Ref. 6).

(4) Measurements on environmental aspects of geothermal energy exploration and utilization, particularly centered around background data. This will permit an accurate differentiation of environmental effects due to utilization of geothermal energy from the naturally occurring effects. This is currently an active area, for example, the subsidence network in Imperial Valley, California, (Ref. 16).

(5) Methods and models for evaluating data from KGRA's, facilitating future predictions of pollutant distributions and alteration of land subsurfaces. For example, currently at LBL, there is a project sponsored by ERDA utilizing computer generated plots to update geodetic maps.

Acknowledgment

Thanks are given to S. R. Schwartz and J. J. Herr for their generous advice and review.

References

1. White, H. J., Jr, in Conference on Thermodynamics and National Energy Problems, National Academy of Sciences-National Academy of Engineering, June 10-12, 1974, p.372. Washington, DC.

2. A Recommended Research Program in Geothermal Chemistry, Lyon, R. N. and Kolstad, G. A., (Co-chairmen), WASH-1344, U. S. Energy Research and Development Administration, Div. of Physical Research, Oct. 1974. Available from National Technical Information Service, U. S. Dept. of Commerce, Springfield, VA 22161. $5.45.

3. Conference on Thermodynamics and National Energy Problems, National Academy of Sciences-National Academy of Engineering, June 10-12, 1974. Washington, DC.

4. Presser, T. S. and Barnes, I., Special Techniques for Determining Chemical Properties of Geothermal Water, U. S. Geological Survey, Water-Resources Investigations 22-74, August 1974. Available from Water Resources Div., U. S. Geological Survey, 345 Middlefield Rd., Menlo Park, CA 94025.

5. Wollenberg, H. A.; Asaro, F.; Bowman, H.; McEvilly, T.; Morrison, F.; and Witherspoon, P., Geothermal Energy Resource Assessment, UCID-3762, Lawrence Berkeley Laboratory, July 1975. Available from Technical Information Department, Lawrence Berkeley Laboratory, Berkeley, CA 94720.

6. Assessment of Geothermal Resources of the United States - 1975, White, D. E. and Williams, D. L. (Eds.), U. S. Geological Survey Circular 726 (1975). Available from U. S. Geological Survey, National Center Reston, VA 22092.

7. Banwell, C. J., "Geophysical Methods in Geothermal Exploration," in Geothermal Energy - Review of Research and Development, Armstead, H. G. H. (Ed.), p. 41-48 (1973). The UNESCO Press, 7 Place de Fontenoy, 75700 Paris, France.

8. Geothermics Special Issue 2, Proc. U. N. Symp. on Dev. and Utilization of Geothermal Resources, Pisa, Italy, Sept. 22 - Oct. 1, 1970, Vol. 1 (1974). Available from Maxwell Scientific International, Inc., Fairview Park, Elmsford, NY 10523. Price $60.00.

9. Axtmann, R. C., An Environmental Study of the Wairakei Power Plant, P. E. L. Report No. 445, Department of Scientific and Industrial Research, Lower Hutt, New Zealand, Dec. 1974. Available from Physics and Engineering Lab., DSIR, Private Bag, Lower Hutt, New Zealand.

10. "A Bibliography: Geothermal Resources, Exploration and Exploitation", TID-3354, U. S. Energy Research and Development Administration, Office of Public Affairs, March 1975. Available from National Technical Information Service, U. S. Dept. of Commerce, Springfield, VA 22161. Price $10.60.

11. Bowen, R. G., "Environmental Impact of Geothermal Development," Chapter 10 in Geothermal Energy, Kruger and Otte (Eds.), Stanford University Press, (1973), Stanford, CA 93405. Price $17.50

12. Proc. Conf. on Research for the Development of Geothermal Energy Resources, NSF-RA-N-74-159, Jet Propulsion Laboratory, Pasadena, CA, Sept. 23-25, 1974. (Dec. 31, 1974),

13. Hickel, W. F., "Geothermal Energy: A National Proposal for Geothermal Resources Research, NSF-RA-N-73-003, University of Alaska, College, AK, (1972).

14. Preliminary Report: "National Plan for the Development of Geothermal Energy," 5040-6, Jet Propulsion Laboratory, Pasadena, CA, July 21, 1975.

15. "GEODOC - The Document File for the National Geothermal Information Resource," Trippe, T.; White, V.; Phillips, S; Henderson, F.; and Schwartz, S., Draft copy, Oct. 1975. Available from National Geothermal Information Resource, Lawrence Berkeley Laboratory, University of California, Berkeley, CA 94720.

16. "Geothermal Subsidence," LBL-3220. Available from National Geothermal Information Resource, Lawrence Berkeley Laboratory, University of California, Berkeley, CA 94720. Subscription price $20.

Data Base for Surface Mining and Land Reclamation

Ronald R. MacDonald
Geographic Information Systems
Energy and Environmental
Systems Division
Argonne National Laboratory
Argonne, Illinois

Abstract

Based on past experience in both data use and data management, it is felt that the distributed data base concept is the only feasible method of implementing the very large national data bases needed by the ERDA/DBER Assessment and Evaluation programs. These national data bases are needed to minimize wasteful duplication of effort by the various ERDA laboratories as well as maximize access to this data by the researchers working on assessment and evaluation projects.

In order to get moving in this direction, the following objectives are established:

1. Promote cooperation between researchers, data managers both within and among the ERDA laboratories as well as other agencies. Provide the coordination necessary to achieve this goal.
2. Establish standards and specifications for data compatability.
3. Identification of and coordination with other ERDA committees concerned with data utilization.

Data Base for Surface Mining and Land Reclamation

Recent efforts in reclamation and land use planning have indicated the need to develop and extend current baseline data management systems such as the existing Argonne Mine Land Reclamation information system and the Oak Ridge National Laboratory information system. In addition, there is a need to identify information gaps and develop data-acquisition techniques to complete baseline information. It is important that data and techniques developed in the aforementioned systems be extended to provide information relating to physical and chemical characteristics of overburden, baseline exchange capacity, species abundance, composition, and spatial organization, current land use and potential methods of mining. It will be necessary to identify areas where significant data gaps do exist so that primary information activities can be instituted.

Information systems are needed to develop the base for land reclamation and land-use planning. A significant feature of such a program will be the establishment of an active two-way exchange of information with the public sector, other federal agencies, the professional community, and the coal industry. Characterization of strip-mined land requires the collection, analysis, and interpretation of a large amount of primary and secondary data. The major effort of this information storage, retrieval, analysis, and brokerage system will involve the identification of all available biogeophysical, biogeochemical, socioeconomic, and land-use data. The system must include the adaptability for merging the spatial nature of this data into its temporal nature; along with the merging of the geocoded earth science, and ecological information to the engineering, socioeconomic, and land-use information (shown in Fig. 1).

The nature and the large amount of data that must be statistically handled by this information system creates the need for the project to be divided into programs.

First, a library program will be developed to locate data for the user that is available in the main information storage system and as to how that information can be obtained (shown in Fig. 2).

Second, the main information system will be enlisted, and the sorting of user data completed.

Third, the information will be prepared in either graphic or tabled form for output to users. This will include use of remote graphic terminals, microfilm plotters, and microfilm output either in addition or in place of the standard line printer output and paper graphics (shown in Fig. 3).

Establish Standards and Specifications for Data Compatability

Define standard building blocks -- the data system needs a geographic unit on which the regions can be built for the spatial studies. Principal criteria for the definition of the Primary Unit are as follows:

1. The unit must have time independent boundary.
2. The unit should be large enough to be statistically nondisclosing.
3. The unit should be small enough to allow sensible definition of a "region."

The methods by which these units can be put together into regions or regions taken apart, needs to be studies.

Define the Spatial and Temporal Filter Function of the System --

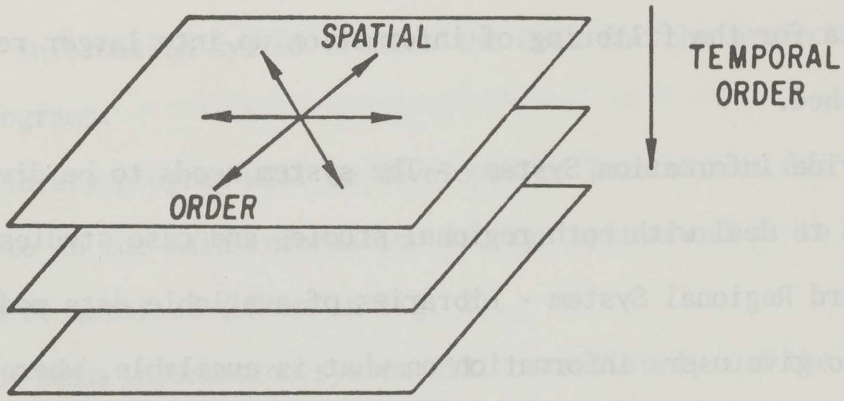
Standard methods need to be derived by which information can be transformed to uniform spatial and/or temporal systems. The criteria for the filtering of information down into smaller regions needs to be established (ZOOM-IN). The criteria for the filtering of information up into larger regions needs to be established.

Subdivide Information System -- The system needs to be divided into major parts to deal with both regional studies and case studies.

Standard Regional System - Libraries of available data modules need to be set up to give users information on what is available, where these data reside, and its quality characteristics. The Regional Data Modules should be available for area studies on a stand-by basis to the management system.

Case Study System - Libraries of available raw data on associated projects need to be established. This raw data must include quality characteristics information of that data. In order for this system to work with maximum coordination between participants, the system must maintain internal consistency and the production of a glossary of terms within the system will be undertaken.

GEOGRAPHIC INFORMATION SYSTEMS



THE ROLE OF DATA BASE MANAGEMENT
IN ENVIRONMENTAL PLANNING

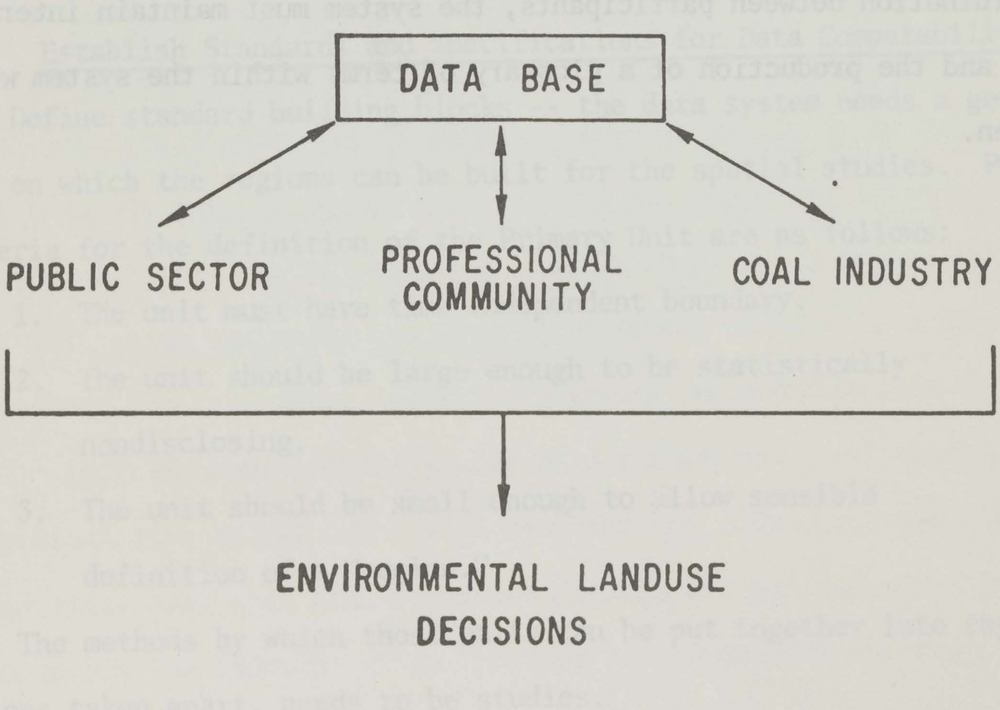


FIG. 1

FLOW CHART

REGIONAL INFORMATION SYSTEM

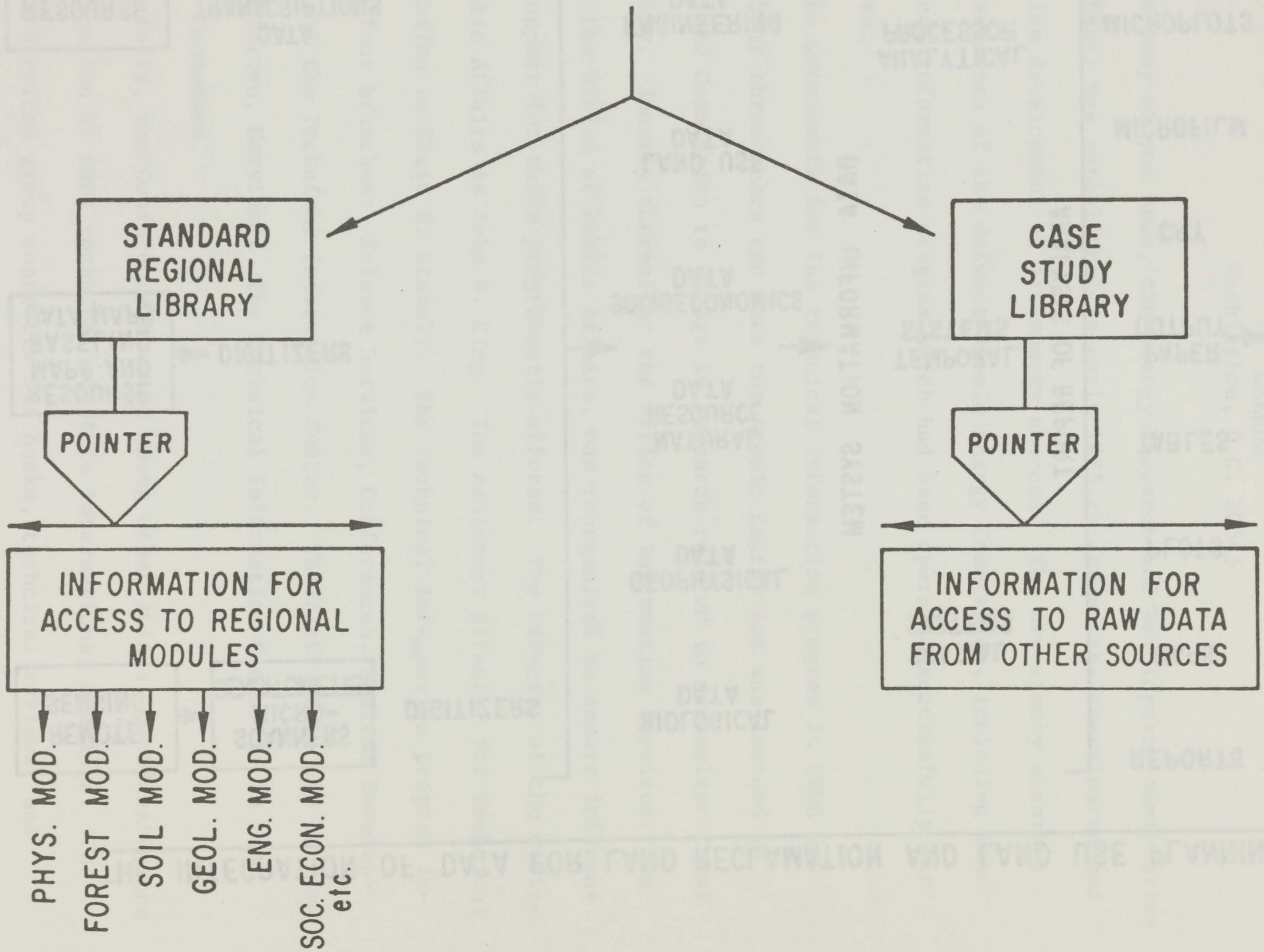


FIG. 2

THE INTEGRATION OF DATA FOR LAND RECLAMATION AND LAND USE PLANNING

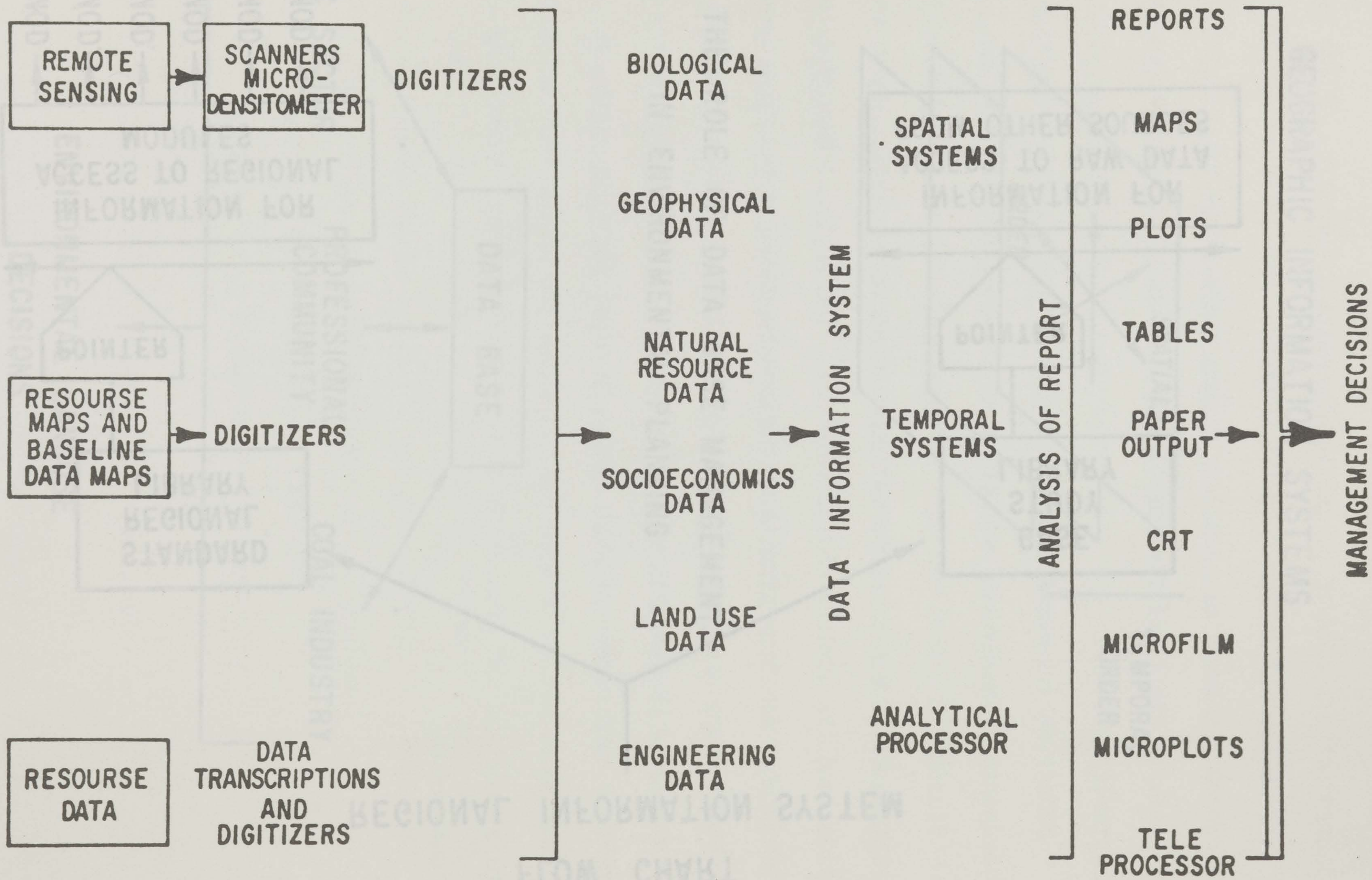


FIG. 3

The ERDA Technical Information Program

Carol G. Alexander
Office of Public Affairs
USERDA
Washington, D.C. 20545

As many of you know, the Energy Research and Development Administration (ERDA) was established in early 1975 to assure the coordinated and effective development of all energy sources. The new agency absorbed major segments of the defunct Atomic Energy Commission, including its technical information program which had been operated successfully for 28 years.

The groundwork for the technical information program in ERDA was laid about three years ago when the Atomic Energy Act was amended to permit the Commission to engage in research related to nonnuclear forms of energy. Shortly thereafter the Office of Information Services, now called the Office of Public Affairs, was reorganized to assure information support for these programmatic efforts. The Director of the Office of Public Affairs is John W. King. The Assistant Director for Technical Information is Edwin E. Stokely. The technical information program includes four branches: Science Services, Conferences, Systems Development, and the Technical Information Center. The first three are located in Germantown, Maryland. The Technical Information Center is in Oak Ridge, Tennessee.

Briefly, the Conference Branch, among other things, coordinates the participation of ERDA technical staff in international meetings. The Science Services group contracts for books, technical reviews, and newsletters to be written by experts in all energy fields outside of the Agency and published by the private sector. The Technical Information Center gathers and processes documents and bibliographic information from worldwide sources and enters them into ERDA's Energy Information Data Base. The Systems Development Branch develops machine-based information retrieval services (such as ERDA/RECON) and improved methods of access to original documents. We also develop new data bases for ERDA-wide use, including bibliographic data bases, information on energy research and development activities worldwide, referral and directory data bases, and data bases to provide access to scientific and technical numerical information. In addition this branch coordinates technical information programs in ERDA contractor libraries in the field.

At the time that the office was reorganized a recommendation was made to the Commission that the Office of Public Affairs begin to assemble a comprehensive bibliographic data base covering the entire field of energy. That plan was approved in 1973 and we have been implementing it since then.

As part of the implementation we have followed certain policies, which are:

1. to develop a mission-oriented system designed to meet the research and development needs of the Agency;
2. to build this system by making maximum use of information already available both in this country and abroad, acquiring it both by purchase and exchange; and

3. to seek the assistance, as appropriate, of the private sector in adapting information materials to meet ERDA's needs.

We determined that this data base would be compiled in machine-readable form for exploitation in a variety of ways. Bibliographic information added by ERDA will be made available to the private sector for further use through the sale of magnetic tapes. Our thinking is that ERDA will limit its publication efforts to issuing, for our own use, an abstract service to be called ERDA Research Abstracts, which will be a product of the data base and will cover only ERDA-originated, energy-related scientific and technical reports, patents, journal articles, conference papers, theses, books, and critical reviews.

Subscriptions for this abstract service will be sold by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. One-year subscription prices are as follows: Energy Research Abstracts and an annual index, \$119.00 (domestic), \$148.75 (foreign); index only, \$30.50 (domestic), \$38.15 (foreign); and individual copies, \$7.10 (domestic), \$8.90 (foreign).

With 26 years of experience in publishing Nuclear Science Abstracts we have decided not to publish an abstract journal with broad coverage of the energy field, but feel that such a service should be published by the private sector. Nuclear Science Abstracts will be discontinued after June 1976 and replaced by an expanded version of INIS Atomindex, published by the International Atomic Energy Agency in Vienna. We will issue specialized publications such as bibliographies drawn from the data base, and we will continue to publish the abstract service Energy Abstracts for Policy Analysis. You may remember that this one is sponsored jointly by ERDA, the National Science Foundation, and the Federal Energy Administration.

To test the coverage of nonnuclear subjects in our data base, we have published three bibliographies this year. They are:

Coal Processing, report number TID 3349, 7741 entries;

Geothermal Resources, report number TID 3354, 3890 entries;

Solar Energy, report number TID 3351, 3545 entries.

These are sold through the National Technical Information Service, Springfield, Virginia.

In addition to the subjects of these three bibliographies, our data base includes bibliographic records in support of the following program efforts:

1. Selected citations in support of general research in physics, chemistry, materials, and instrumentation;
2. Energy conservation and management;
3. Energy engineering;
4. Fossils and synthetic fuels;
5. Nuclear energy; and
6. Environment and safety.

Table 1 shows the composition of the data base and the approximate number of entries within each class as of October 1975.

I would like to mention one more publication connected with the data base, ERDA Subject Indexing and Retrieval Thesaurus, which is being developed, structured, and maintained by the staff at the Technical Information Center in Oak Ridge. The thesaurus is available for sale from the National Technical Information Service as report number TID 7000. It incorporates more than 15,000 terms identified through a variety of sources already in use in energy organizations. An updated edition will be issued in the spring of 1976, with regular quarterly updates scheduled after that.

We are pleased to announce two other publications which may interest you. The first is the Directory of Librarians and Technical Information Specialists in ERDA and Its Contractor Organizations. It is a fifth revision as of May 1975. The second publication is the Directory of USERDA Information Centers, which is a listing and description of those groups within ERDA who provide analyses of technical information or data services for all members of the scientific, technical, and managerial communities. Both are available free of charge from:

USERDA
Technical Information Center
P. O. Box 62
Oak Ridge, Tennessee 37830

Publications represent one type of output from the data base. Other products and services are available through ERDA special libraries, information analysis centers and the ERDA/RECON system. The acronym RECON stands for REmote CONsole and refers to a computerized on-line, interactive storage and retrieval system designed to permit rapid access to bibliographic records stored in a large file. ERDA/RECON is operated for us by the Computer Sciences Division of Union Carbide Corporation, Nuclear Division, in Oak Ridge. The system runs on IBM 360 equipment and is fully described in reports published by Carbide. A listing of these citations as well as those needed to use RECON may be obtained by writing Thomas E. Hughes, Chief, Systems Development Branch, USERDA, Washington, D.C. 20545.

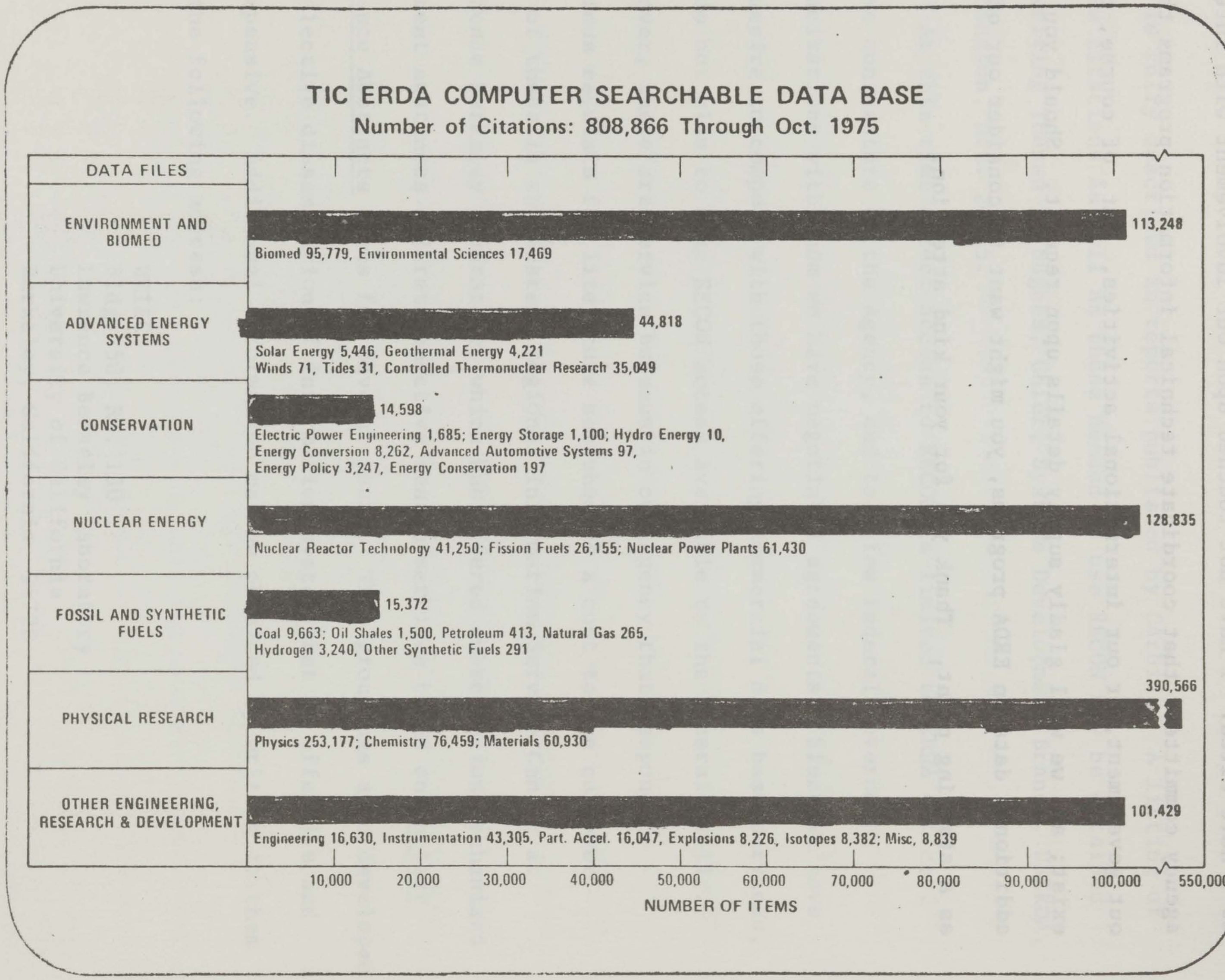
At this time direct access to RECON is limited to ERDA offices, prime contractors of the Agency, and to a few Federal government organizations with whom we have negotiated agreements. Since we have no desire to compete with those offering commercial data base services, we do not plan to make RECON access available to the general public. However, there are service bureaus in our agency that respond to serious requests for literature searches at a cost to the customer. One of these is the Western Regional Information Service Center at Lawrence Berkeley Laboratory, which has offered subscription mechanized current awareness and retrospective search services based on <u>Nuclear Science Abstracts</u> tapes for several years. This group has also developed a selective dissemination of information system that is effective and inexpensive. Additional information may be obtained by writing to them at the following address:

WRISC
Bldg. 50, Rm. 130
Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720

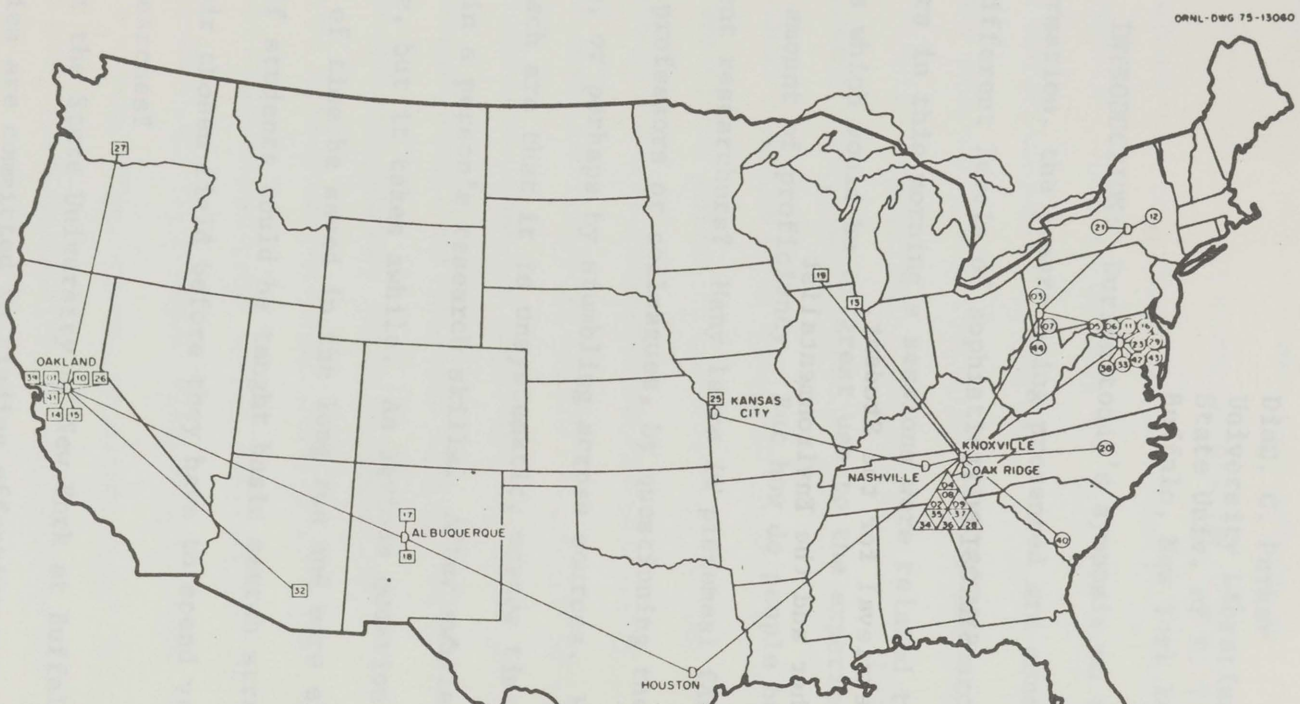
Other library and information facilities are listed in the two directories mentioned earlier.

I have mentioned briefly most of the activities now operational. Additional information on all of these is available from our office or as I have cited. I have not touched upon our involvement with interagency committees that coordinate technical information programs throughout government, or our international activities. But, of course, these exist; and we will gladly supply details upon request. Should you need additional data on ERDA programs, you might want to consider our office as a starting point. Thank you for your kind attention.

Table 1


TIC ERDA COMPUTER SEARCHABLE DATA BASE
Number of Citations: 808,866 Through Oct. 1975
DATA FILES
ENVIRONMENT AND BIOMED
113,248
Biomed 95,779, Environmental Sciences 17,469
ADVANCED ENERGY SYSTEMS
44,818
Solar Energy 5,446, Geothermal Energy 4,221
Winds 71, Tides 31, Controlled Thermonuclear Research 35,049
CONSERVATION
14,598
Electric Power Engineering 1,685; Energy Storage 1,100; Hydro Energy 10,
Energy Conversion 8,262, Advanced Automotive Systems 97,
Energy Policy 3,247, Energy Conservation 197
NUCLEAR ENERGY
128,835
Nuclear Reactor Technology 41,250; Fission Fuels 26,155; Nuclear Power Plants 61,430
FOSSIL AND SYNTHETIC FUELS
15,372
Coal 9,663; Oil Shales 1,500, Petroleum 413, Natural Gas 265,
Hydrogen 3,240, Other Synthetic Fuels 291
PHYSICAL RESEARCH
390,566
Physics 253,177; Chemistry 76,459; Materials 60,930
OTHER ENGINEERING, RESEARCH & DEVELOPMENT
101,429
Engineering 16,630, Instrumentation 43,305, Part. Accel. 16,047, Explosions 8,226, Isotopes 8,382; Misc, 8,839
10,000
20,000
30,000
40,000
50,000
60,000
70,000
80,000
90,000
100,000
550,000
NUMBER OF ITEMS

ORNL-DWG 75-13060



□ = GDA6765 (WESTERN CIRCUIT)

01 = L.B.L.,	BERKELEY, CA.
10 = L.L.L.,	LIVERMORE, CA.
13 = ARGONNE,	CHICAGO, ILL.
14 = G.E.B.R.O.,	SUNNYVALE, CA.
15 = G.E.N.E.D.,	SAN JOSE, CA.
17 = L.A.S.L.,	LOS ALAMOS, N.M.
18 = SANDIA,	ALBUQUERQUE, N.M.
19 = WISC. U.,	MADISON, WISC.
25 = BENDIX,	KANSAS CITY, MO.
26 = SANDIA,	LIVERMORE, CA.
27 = BATTELLE N.W.,	RICHLAND, WASH.
32 = ARIZONA U.,	TUCSON, ARIZ.
39 = U.S.F.S.,	BERKELEY, CA.
41 = E.P.R.I.,	PALO ALTO, CA.

D = TELCO

△ = 88GD1067 (OAK RIDGE CIRCUIT)

04 = O.R.N.L. MASTER,	OAK RIDGE, TN.
02 = T.I.C.,	OAK RIDGE, TN.
08 = N.S.I.C.,	OAK RIDGE, TN.
09 = C.R.L.,	OAK RIDGE, TN.
28 = K-25 LIBRARY,	OAK RIDGE, TN.
34 = T.I.C.,	OAK RIDGE, TN.
35 = T.I.C.,	OAK RIDGE, TN.
36 = T.I.C.,	OAK RIDGE, TN.
37 = T.I.C.,	OAK RIDGE, TN.

○ = GDA24803 (EASTERN CIRCUIT)

03 = BETTIS,	PITTSBURGH, PA.
05 = U.S.E.R.D.A.,	GERMANTOWN, MD.
06 = B.T.L.,	BETHESDA, MD.
07 = W.N.E.S.,	PITTSBURGH, PA.
11 = N.O.A.A.,	WASHINGTON, D.C.
12 = K.A.P.L.,	SCHENECTADY, N.Y.
16 = W.R.S.I.C.,	WASHINGTON, D.C.
20 = N.C. ST. U.,	RALEIGH, N.C.
21 = CORNELL U.,	ITHACA, N.Y.
23 = L.C.,	WASHINGTON, D.C.
29 = F.E.A.,	WASHINGTON, D.C.
33 = U.S.F.S.,	ARLINGTON, VA.
38 = V.P.I.,	BLACKSBURG, VA.
40 = S.R.L.,	AIKEN, S.C.
42 = B.L.M.	WASHINGTON, D.C.
43 = U.S.N.R.C.,	WASHINGTON, D.C.
44 = E.R.C.,	MORGANTOWN, W.VA.

ERDA/RECON NETWORK
NOVEMBER, 1975

SYMPOSIUM II:

Data Retrieval for the Student,

Researcher and the Environmentalist

TEACHING LIBRARY AND LITERATURE SEARCH STRATEGY TO GEOLOGY STUDENTS

Diane C. Parker
University Libraries
State Univ. of N. Y., Buffalo
Buffalo, New York 14214

INTRODUCTION: During today's symposia on retrieval of geoscience information, the papers being presented are concerned with data retrieval at different levels of sophistication or research expertise. Most of the papers in this morning's sessions were related to the development of data banks which would be of great use to the experienced researcher with a certain amount of proficiency. But how do people become experienced and proficient researchers? Many learn in piecemeal fashion through recommendations from professors or colleagues, by questioning the staff in research institutions, or perhaps by stumbling across sources. Major characteristics of this approach are that it is unsystematic, wastes time, and can cause some odd gaps in a person's research skills. After awhile a system usually does emerge, but it takes awhile. An obvious question comes to mind. Wouldn't a lot of time be saved in the long run and more effective, complete research done if students could be taught basic search strategies and the literature of their chosen field before they have to spend very much time doing literature searches?

At the State University of New York at Buffalo, the University Libraries are committed to finding effective ways to teach students the skills of library and literature search strategy. In a survey done in February, 1974, we discovered that with very few exceptions, the academic teaching departments did not see it as one of their functions to teach students how to use the libraries to retrieve bits of data or bibliographic information. At the same time, the Libraries were offering very little to students whose last exposure to research methods usually was an enforced library lesson or two in junior or senior high schools. The general kind of information offered in those kinds of situations is soon forgotten and, in any case, grossly inadequate for people who are training for careers in a specific field. In order to fill this gap for geology students, a one unit course called "Introduction to Geological Literature" was developed to help the students learn to use the library as a basic research tool. The course is called Geology 307, Introduction to Geological Literature.

INITIATING THE COURSE: I said earlier that the University Libraries at SUNY, Buffalo are committed to finding effective ways to teach library research. This means that my boss was sympathetic to having the Science & Engineering Library's geology bibliographer (me) spending time to do it. Hence, the Libraries funded the instructor. While support from the Libraries was important, support from the academic department was also very important.

Some of the librarians at SUNY, Buffalo offer instruction with no assignments and no credit, but I wanted the students to have supervised practice in actually doing research on a topic of their own choice. Such assignments would entail a lot of work, and I felt the students should receive some credit for this work. Fortunately, members of the Geology Department's undergraduate curriculum committee agreed. After a discussion of possible course content, the only change they suggested was adding practice in writing abstracts and annotations.

And so Geology 307 was born ... almost. A "Course Approval Form" with the Department's approval still had to be sent to the University's Division of Undergraduate Education for approval. The form included a course description, abstract for the catalog, a rationale, and a duplication check. Presumably the D.U.E. found that the course seemed rational enough and also did not duplicate any course already offered within the University, because they did approve it in due time. Geology 307 was first given in fall, 1974.

Support from the parties concerned and lots of paperwork seem to be the necessities of starting a course. Anyone thinking of teaching one would be well-advised to start early. If necessary, be prepared to nurse the paperwork along every step of the way. You're the instructor, and you have the primary responsibility for seeing that everything comes together.

PLANNING THE COURSE: A prerequisite to planning is definition. For whom is the course designed? What is its scope? What should students taking the course be able to do after they have finished?

The course was designed initially for undergraduate students majoring in geology. Other interested undergraduate students would have been welcomed if they had some interest and/or commitment to the geosciences. Graduate students were welcome to audit the course, but there was at that time no provision for giving them graduate credit.

Scope included the logical procedures followed for doing library research in most if not all science disciplines. (See Exhibits 1 and 2.) In addition to these general search strategies, geological literature and distinct search strategies applicable to geological literature were emphasized. Also, because geology is itself so interdisciplinary and allied to chemistry, physics, engineering, geography, mathematics and computer science, the major publications in those fields were included. Because geology students need to use maps and government publications, these also were included too, along with a demonstration of computer searching. Patent literature was outside the scope of the course, although one student even requested it. There simply was no time to include it within the fifteen weeks the course ran. A lecture which included style manuals and writing was reinforced with assignments on writing annotations and an abstract.

After they have finished the course, students should be able to define their information need more precisely, select the most appropriate materials to solve their problem, use indexes and cumulated volumes to save time, locate peripheral sources when the most likely sources fail to help, and turn to the literature of related disciplines for problems which overlap with other fields. As D.N. Wood points out in his book called Use of Earth Sciences Literature (1973, page 28) at least 16.5% and possibly up to 48% of periodical titles used by geologists are from fields outside the earth sciences. For survival's sake, students need to learn methods applicable to other disciplines.

Once the purpose, scope and audience have been determined, the next step is to write a course outline and plan the method for teaching.

METHODOLOGY: The course was given in a fifteen week period with a formal, one-hour lecture given each week, either in a classroom or in one of the University Libraries. Informal consultation on weekly assignments or the semester project was available whenever the librarian was in the library. Thus, the students did in effect have lab sessions available.

A "Questionnaire/Pretest" was given during the first class session to determine the students' interests and level of ability in literature searching and library methodology. A "Post-test" or final exam given at the end of the course repeated many of the pretest questions but was much more elaborate in that it included more of the materials covered during the semester.

Assignments consisted of five worksheets and a semester project. The worksheets were designed to point out important features of the bibliographic materials being used. The worksheets were 'interactive' in several ways. For the major portion of the worksheet, students were asked to use the materials examined to find bibliographic citations for their semester projects. They had to list subjects or other entries tried, say whether the entry was useful or not, and if the entry was not successful, explain why they thought it wasn't. This process was intended to make the students critical about the approaches they used, flexiable enough to think of several approaches and persistent enough to keep trying when the first approach didn't work. Finally, while correcting the assignment, the instructor often did the same search and suggested approaches or subjects the students hadn't thought to try. With this method, the students had to make the first effort, and as in most learning situations, this is the critical factor. Expertise comes only with practice. The worksheets given were as listed below:

1. Glossary of Geology; Geosaurus; L.C. Subject Headings
2. Card catalogs; Catalog of the U.S. Geological Survey Library
3. Geoscience bibliographies, indexes, abstracting services
4. General science indexing services useful for geology
5. Monthly Catalog of U.S. Government Publications; Publications of the U.S. Geological Survey; Government Report Index and Government Reports Announcements
6. Using the map collection (not used; ran out of time)

In addition to the worksheets, there was a semester project. At the beginning of the semester, students were asked to choose a topic of interest to them. Some of the topics chosen were subjects of personal interest, some were related to other course work, and one was anticipated as a topic for graduate work. The students were asked to compile a bibliography of at least twenty-five items, four of which had to be annotated and a fifth abstracted. Remember, a significant part of the worksheets was to try to find citations which could be used in the semester project. Students were asked to assign subjects to each citation in a manner that could be useful to them if they were building a personal index file of citations on a subject they were researching in depth. As a guide to selecting proper subject headings, they were to use the thesaurus and dictionaries covered in the first worksheet and their own common sense. In correcting the projects, the librarian was looking for choice of citations appropriate for the topic, correctness and consistency of bibliographic form, appropriate choice and consistency in format for subject headings, annotations and abstracts written clearly and in proper style and coverage, and good choice and variety in sources for citations. It is not enough to take all of one's citations from one good bibliography!

Finally, the students were asked to evaluate the course for its usefulness to them. Formal evaluation forms were devised (see Exhibit 3); the students filled them out and turned them in to a secretary who then compiled the results. The students rated all aspects of the course high except the final exam. (I am inclined to agree that this particular exam was more appropriate for library school students and would not use it again.) The students also were unanimous in their opinion that the course should be worth two credits instead of one.

Grades for the course were based primarily on the semester project. However, the worksheets had to be turned in for the course to be completed, and a high score in the final exam was sufficient grounds for raising a grade.

RESULTS: The measurable results of the committed students' work were very good. Six students took the course for credit, and others audited from time to time. One class is too small to generalize from. Nevertheless, I'll risk generalizing from this one sample to this extent: the students who did have a strong interest in the earth sciences as a careet spent a lot of time on this course and were enthusiastic about the results. Often I would find them in the library helping other students who were not taking the course. This was wonderful from my point of view, since I was assigned to work in the geology library only part time. A couple of the students found the coursework more than they cared to handle. By the end of the semester, faculty in the Geology Department had heard enough good comments about the course that they were willing to increase the number of credits for the course to two and also crosslist it as Geology 307/511, so that it could be taken for credit by graduate students. The course is being given in fall, 1975 for two credits. Somehow during a move of the Geology Department and the geology library to new quarters last spring, the paperwork for listing the course with a graduate number got lost. So, that change will have to wait for next year.

Another change that has taken place is a change in instructors. Last spring I was offered a promotion, and another librarian was hired to take my place. Notice that another librarian was hired. The Libraries are trying very hard to maintain commitments started. Of course there's no law that says courses have to be offered every year or even ever again. However, when we start something like this, there should be intent to make it good and commitment to try hard to keep it going.

TIME COMMITMENT: How much time does it take for a librarian to plan and give a course like the one I have described? A conservative estimate would be 150 hours or about 10 hours a week. A breakdown is as follows:

Preparation of:

course approval form	1
course outline	4

	14 lectures @ 2 hours each	28
	5 worksheets @ 3 hours each	15
	9 handouts	9
	pretest & post-test	4
	project design	1
	project correction sheet	1
	flier announcing course	1
	distributing flier, etc.	2
	evaluation	2
		68
Contact hours:		
	class lectures	14
	work in library "lab"	28
		42
Correcting papers, etc.:		
	5 worksheets (10 min. - 2 hrs. ea.)	21
	Pretest	.5
	Post-test	.5
	projects (3 hours each)	18
		40

Summary:

Preparation	68
Contact	42
Correction	40
	150 hours

If preparation time were cut to 28 hours (2 hours for each lecture) the second time a course was given, the time spent on the course would be 110 hours or about 8 hours per week. This two hours preparation time might include enough leeway for revising handouts, but I doubt it. About .25 of a full time employee would be a more reasonable estimate of time needed. If all this time is going to be spent preparing the course, it might as well be given for as many students as is feasible. (I would consider ten a maximum without changing the structure of the course assignments.) You may have to advertise the course in order to have as high an enrollment as you would like.

ADVERTISING: One of the most effective ways to advertise a course the first time it is given is to have it included in the college catalog and the semester class schedule. If, however, you decide to give the course after deadlines have passed for including it in those publications, you'll have to resort to other means. You can announce the course in the school paper. You can write a flier to be distributed to students at registration lines. And don't forget to give your flier to student advisors and members of the geology club if your school has one. Post your flier in the department coffee room too. I also posted a copy of my Pretest in the coffee room. One student tried it, couldn't answer any of the questions, signed up, and became one of my best students. Use your imagination, especially the first time the course is offered. When you do give the course, work hard to make it good, and above all, enjoy yourself. After that, let's hope you've done such a good job that you'll have a core of enthusiastic past students, as well as supportive members of the academic department, who will advertise the course for you.

READINGS & TESTS:

1. Cochran, Wendell, Fenner, Peter and Hill Mary. Geowriting; a guide to writing, editing, and printing in earth science. 2nd ed. American Geological Institute, Falls Church, Virginia, 1974.

2. "Geoscience information and user needs." Geoscience Information Society Newsletter. No. 30. June, 1974. pp. 3-8.

3. Ward, Dederick C. and Wheeler, Marjorie W. Geologic reference sources. Scarecrow, Metuchen, N.Y., 1972.

4. Wood, D. N. Use of earth sciences literature. Archon Books, Shoestring Press, Hamden, Conn., 1973.

EXHIBIT 1

ABRIDGED COURSE OUTLINE

Geology 307
Introduction to Geological Literature

Week:	Topic:
1	Course objectives; outline of literature and information search procedures; pretest.
2	Characteristics of geological literature; bibliographic citations, annotations, format for term project; abstracts; personal index-files.
3	Subject approach to literature searches; thesauri, dictionaries, lists of subject headings.
4	Retrospective literature searches; card catalogs; card catalog reproductions; U.S.G.S. Library catalog.
5	Geoscience indexing and abstracting services.
6	Indexing & abstracting services for disciplines related to geosciences.
7	Organization of academic libraries -- access to and control of materials; classifications systems; types of public services; union lists & catalogs.
8	Citation and technical report literature.
9	Map collections.
10	Government documents.
11	Reference materials -- geology.
12	Reference materials -- geology.
13	Reference materials -- general.
14	Semester review; projects returned and discussed.
15	Final exam; course evaluation.

Diane C. Parker — Fall, 1974
State University of New York, Buffalo

LITERATURE and INFORMATION SEARCH PROCEDURES

Step I. Locating Bibliographic Information

A. Begin with your knowledge of:

1. Author's Name
2. Specific Title
3. General Subject

Expand or modify with:

Guides to Subject Headings
1. Library of Congress subject heading guide
2. Thesauri or dictionaries for your subject field

B. Find bibliographic citations in:

CARD CATALOG for
1. Material by a known author
2. Specific titles
3. Background information in your subject - retrospective searches

SPECIAL INDEXES for
1. Journal articles
2. Abstracts or summaries of articles, reports, dissertations
3. Technical report literature
4. Patent literature
5. Government documents

PRINTED CATALOGS of holdings in other large libraries

BIBLIOGRAPHIES of work in your subject field

C. Locate materials in:

THIS LIBRARY
1. Check shelves using location number found in card catalog
2. Ask library staff to help locate or recall material not found on shelf

OTHER S.U.N.Y.A.B. LIBRARIES
1. Check under main entry (usually the author's name) in the card catalog at Lockwood Library
2. Use serial printout to see if a journal is in a S.U.N.Y.A.B. library

LIBRARIES OUTSIDE S.U.N.Y.A.B.

Inquire about interlibrary loan procedures

Step II. Sources of Information

Use GUIDES TO THE LITERATURE For a general description of seach logic reference tools, and sources of information for specific subject fields

For:

- **Handbooks** — For summaries of basic theories, applications, and methods
- **Tables** — For compilations of frequently used data
- **Encyclopedias** — For general orientation to specific topics
- **Dictionaries** — For definitions of technical or foreign language terms
- **Biographical Dictionaries** — Science Citation Index — For checking a person's credentials and contributions to a field
- **Style Manuals** — For help in writing papers

Step III. Continuing Research

- **Directories of Periodicals** — For lists of journals published in your field
- **Current Awareness Services** — For continuous updating — Abstracts; Current Contents
- **Annual Reviews and Yearbooks** — For summaries of significant research
- **Information Analysis Centers, Research Centers** — e.g. Cryogenic Data Center, etc. For dissemination of hard data
- **Trade Literature** — For suppliers of materials and equipment
- **"Invisible College"** — 1. Correspondence with colleagues 2. Professional society meetings

Diane Parker + Cheryl Carlson

March, 1974

EXHIBIT 2

EXHIBIT 3

COURSE EVALUATION

Geology 307
Fall, 1974

I. Considering a) the amount of work you were asked to do for this course
b) the amount of work you actually did for this course
c) the amount of work you do for other credit courses

How many units of credit do you think you should receive for a course like Geology 307? _____

II. Please rate the usefulness of the following parts of this course: (scale of 1 - 10; 10 is highest)

Pretest — 0 2 4 6 8 10

Lectures — 0 2 4 6 8 10

Handouts — 0 2 4 6 8 10

Worksheets
- a) General questions — 0 2 4 6 8 10
- b) Project-related sections — 0 2 4 6 8 10

Project
- a) Bibliographies — 0 2 4 6 8 10
- b) Annotations — 0 2 4 6 8 10
- c) Abstract — 0 2 4 6 8 10
- d) Subject headings — 0 2 4 6 8 10
- e) Sources — 0 2 4 6 8 10

Final exam/Post-test — 0 2 4 6 8 10

III. Once it's all over, do you think what you've learned in this course will be useful to you in the future?

0 2 4 6 10

IV. Comments, if any

Water Resources Literature Cited in Wisconsin
Department of Natural Resources Publications, 1964-1973

by

Richard D. Walker
and
Gregory J. Zuck
Library School, University of Wisconsin, Madison 53706

Introduction

With the enactment in 1964 of the Water Resources Research Act (Public Law 88-379) Congress intended that there should be a substantial increase in research into the several aspects of water resources. Furthermore, Congress stipulated that the results of such research should be widely disseminated among potential users who would in turn apply the research findings to their areas of interest.

The investigation of the problems of the transference of information from the researchers who produce the research in the area of water resources into the hands of "applied" water resources personnel is seen as necessary. In general, the effective transfer of technical and scientific information is perceived as a major problem whose solution is both desired and needed. In 1967 this desire and need was acknowledged with the establishment of the Water Resources Scientific Information Center (WRSIC) which was authorized to promote water resources research, training, and information dissemination as provided for by the 1964 Act. In addition, the Federal Council for Science and Technology (FCST) designated WRSIC as the national center for scientific and technical information in water resources; WRSIC, therefore, attempts to disseminate scientific and technical information to the national water resources community. In 1968 the Committee on Water Resources Research (COWRR) of the FCST called attention to the need for more effective dissemination of water resources research results.

This study represents one phase of a more comprehensive research study of the transfer of information in the area of water resources. It deals with the use of literature in the Wisconsin Department of Natural Resources (WDNR) publications on topics of water resources, water quality and quantity, water supply, the water cycle, water planning, and other water subjects. Literature cited in all WDNR publications issued during the period 1964-1973, is analyzed.

The Problem

The purpose of this study was to examine the water resources publications of the Wisconsin Department of Natural Resources (WDNR) in order to ascertain what literature is used by the persons employed by the WDNR in preparing their publications. It was assumed the literature is a principal means of communication and is a representative record of the knowledge and activity of the WDNR. The problem can be stated as follows: What literature is cited by the WDNR personnel when writing on and reporting results of investigations concerned with water-related subjects?

An objective of the study was to create an array of descriptive statistics including a list of most frequently cited sources. An examination of the references cited by the WDNR authors of water-related publications issued by WDNR generates descriptive statistics as to the subject, age, form, and publisher of each citation. An additional objective was to determine how relevant the research funded by the OWRT and other federal agencies is to the water resources personnel of the WDNR.

More specifically, this study was designed to examine the following aspects of the literature cited in the water resources publications of the WDNR:

1. The subject coverage of the cited literature.

2. The age of the material as an indication of the useful life of the literature used or the obsolescence of the literature.
3. The several publication forms (or literature types) used by WDNR personnel, i.e., journals, books, conference proceedings, technical reports, bulletins, unpublished and internal reports, personal communciations, and theses, to name a few of those frequently cited.
4. The publishers of the literature cited, i.e., the several government agencies [OWRT, EPA, U.S. Geological Survey (USGS), and others], universities, non-profit organizations, societies and commercial publishers.
5. The dispersion of the journal titles cited, including a list of the most frequently cited journals.
6. The dispersion of the technical reports and bulletins as to publisher.
7. An analysis of all federally supported literature cited in order to determine how much use is made of the water resources research funded by OWRT, EPA and other federal agencies.
8. Comparative analysis of this study's findings with the results of other similar research studies.

<u>Why the Department of Natural Resources</u>

The Department of Natural Resources is the central unit of state government for quality management and protection of all waters of the state, ground and surface, public or private. The Department must develop a comprehensive plan to mobilize governmental effort and resources at state, federal, and local levels to accomplish the greatest result for the people of the state as a whole [<u>Wis</u>. <u>Stat</u>., Secs. 144.025 (1) and (2) (a)].

Method of Investigation

The WDNR Publications

All publications originating in the Wisconsin Department of Natural Resources (WDNR) from January 1964 through the end of December 1973, dealing with any aspect of water were used in this study. The references cited by the authors, either as a list or bibliography appearing at the end of the publication or notes throughout the publication were analyzed so that each publication cited is recorded and tabulated for the following characteristics: 1. subject content, 2. year of publication, 3. form or type of publication, 4. publisher, 5. journal title dispersion.

All WDNR publications were examined to determine which of those published during the ten-year period under investigation dealt with some aspect of water resources. The following series were initially examined: Research Reports, Technical Bulletins, Surface Water Resources, Bureau of Fish Management-Management Reports, Bureau of Fish Management-Miscellaneous Reports, Lake Use Reports, Mini Reports, Pollution Investigation Reports, Special Publications, and Wisconsin Wetland Inventories; all other non serial publications were also checked for water-related topics.

The Analysis

To determine what literature was considered relevant and useful to the water resources interests of WDNR the technique known as citation analysis was employed. This research method is a well-established bibliometric tool that can be used to determine what literature was used by researchers in a given field or to study the cited literature of a particular group of researchers.

There is at least one shortcoming to this method of determing the significant writings. The field of water resources is relatively new and the support of the EPA and OWRT is only quite recent and, therefore, conditions may be changing too rapidly for such a study to be valid for an extended period of time. However, it does give a true representation of the present and recent past. An analysis was made of all cited literature (not only the journal or other serial literature). Additionally the study was designed to determine if federally funded research in water resources is considered relevant and useful by the writers of publications issued by the WDNR in much the same way Burton and Green (4) determined the use made of technical reports by physicists in their published research.

The Gross and Gross study (6) was designed to ascertain which journal titles would best serve the information needs of chemists in a small liberal arts college and it contained the following assumptions: The usefulness of a journal in any scientific field may be measured indirectly and objectively by determination of the number of times the journal is cited in the literature of that field, that is, the greater the number of citations, the greater the usefulness of the journal (6). We are interested in ascertaining what literature, regardless of form, WDNR water resources personnel consider relevant and useful as reflected in their citation of that literature, we therefore, modify the above assumption to encompass all forms of literature and do not restrict it to journals.

If a citation is a measure of the cited work's usefulness and relevance then the overall usefulness and relevance of a literature can be determined by an analysis of the cited literature. Margolis has made the observation that, "the value of a scientific paper can be measured by the influence it has on others ... It is reasonable to expect that the best contributions would have been those cited most, while relatively unimportant papers would have attracted few, if any citations" (7). Furthermore, he adds, "authors or original contributions are probably the best qualified critics of the literature in their field" (7). It follows that papers thought to be useful and relevant are cited, and trivial or marginal ones are ignored. If a literature is not cited, it can be assumed to be either not relevant or unknown to the authors.

One additional point is made by Margolis: An author, he adds, "knows that his article may be overlooked if it does not cite pertinent literature ..." (7). This study assumes that care is taken in the author's choice of cited references and that all or nearly all references cited by an author were used by him and that he cited all the relevant/useful material he used in the preparation of his work.

Subjects

The subject covered by a cited publication was determined by the title of the publication as given by the citing author. Each cited publication was assigned one subject.

Publisher

The source of publication by type of publisher was recorded for all references by assigning each citation one of the following:

Federal.

State.

Other Government.

Non-profit Sources, Societies and Associations. Includes universities, University presses, international organizations and committees, and academies of science.

Commercial Publishers.

Form or Type of Literature

For each cited reference the form or type of literature was recorded using the following categories of material:

1. Journals (achival, research, trade and society)
2. Technical or research reports
3. Bulletins (circulars, pamphlets, etc.)

4. Books
5. Conference and symposium proceedings
6. Personal communications and unpublished materials
7. Internal WDNR reports ("open file" reports)
8. Abstracting and indexing services
9. Congressional publications
10. Theses and dissertations
11. Other government documents
12. Reviews (both annuals and others)
13. Maps

Descriptive Statistics

The general characteristics of the literature cited are described in the following order: 1. subject, 2. age, 3. form, and 4. publisher. Under these broad categories the three different groups of WDNR publications, Research Reports, Technical Bulletins, and Surface Water Resources and miscellaneous publications, are analyzed separately and in combination.

Subject

The single most frequently occurring subject in the Research Reports and Technical Bulletins taken separately or in combination is "Water quality management and protection (pollution and pollutants)." Of the 564 references cited in the twenty-one Research Reports, 184 are to this subject, resulting in 33 percent of the total number of cited references. The subject "Water quality" in Technical Bulletins accounts for 163 cited references or 57 percent of the total. In combination this subject ranks first with 347 items or almost 41 percent of the total number of cited references.

Two of the next four subjects in the Research Reports and Technical Bulletins are identical, although their rank order differs. The second most frequently cited subject is "Plants" in Research Reports with 90 references, or 16 percent and ranks third in Technical Bulletins with 17 "Water cycle" with 56 items or approximately 10 percent and ranks second in Technical Bulletins, having only 22 cited references at 7.7 percent of the total.

When combining the subject order of Research Reports with a total of 544 cited references with the ordering of Technical Bulletins having approximately half the total number, 285 citations, the rank order of the combined total closely parallels that of the Research Reports with the following rank ordering as listed according to number of times cited and percent of the total. First is "Water quality" at 347 references or nearly 41 percent; second is "Plants" having 107 references at 12.6 percent; third in rank is "Water cycle" with 78 references cited; or 9.19 percent; fourth "Water resources planning" contains 63 items at 7.42 percent; and fifth is "Animals" with 62 citations or 7.3 percent of the total number of cited references.

Surface Water Resources and miscellaneous publications have a subject assignment that contrasts markedly with those of the Research Reports and Technical Bulletins. For Surface Water Resources "Water cycle" is first with the largest number of cited references, 131, or nearly 30 percent of the total number of all cited references in that publication series. The subject category "Other nonwater"* contains 114 cited references or nearly 26 percent of the total number of cited and appears second. This subject differs appreciably in its ranking when compared to the order in Research Reports and Technical Bulletins where "Other nonwater" is, respectively, ninth place with 18 cited references or 3 percent and sixth place with 12 references or 4 percent of the total number of cited references. The subjects "Plants" and "Animals" tie for third in ranking. Each subject receives 37 citations, or 8.3 percent of the total.

---

*The category "Other nonwater" includes a number of individual subjects all of which are nonwater in nature.

When combining all WDNR publications "Water quality" retains its position as the subject with the most frequently cited references. The ranking of the second subject, "Water cycle" and the third subject, "Other nonwater" with 209 references at 16 percent and 144 or 11 percent, respectively, is influenced, in part, by the concentration of a large number of cited references in these subjects by the Surface Water Resources publications. Tied for third with the same number of cited references is "Plants" and "Animals" ranks fifth with 99 references or 7.64 percent of the total. Placed sixth is "Water resources planning" with 83 items cited or 6.41 percent. The next nine subjects rank from approximately 5 percent of the total to 1/15 of one percent of the total number of citations. Of marginal interest are "Resources data," "Nature of water," and "Manpower" which have less than 1/2 of one percent of the total number of cited references. "Water supply" and "Engineering work" are each less than 1 1/2 percent of the total.

Agé of Cited Literature

The age of the cited literature is given as a measure of the length of time the WDNR authors consider their literature useful to their research. The evidence indicates that the sciences with severe time pressures have literatures with a short useful life and in the more stable sciences the useful life span is longer. The rate of obsolescence of the literature of a field of science can be measured by the age of the cited literature and therefore provide, if not a true measure, at least an indication of the probable life expectancy of the literature of the future.

The age of a cited reference is recorded as the difference between the date of the cited reference and the date of the citing publication.

As in the case of the ranking of subjects, Research Reports and Technical Bulletins show similarities in age distribution. The range in ages of cited references in Research Reports is 0-53 years, for Technical Bulletins the range is 0-77 years. Spreads in the distribution by age result in a mean age of 11 years for Research Reports, 8.13 years for Technical Bulletins and a combined mean age of 8.37 years. Median ages reflect a similar pattern with Research Reports at a median of 6 years, Technical Bulletins at 4 years, and in combination 6 years. The mode for Research Reports and Technical Bulletins individually and in combination is, in each case, 3 years.

Surface Water Resources and miscellaneous publications again vary in comparison to Research Reports and Technical Bulletins. The range of ages of cited references is 0-121 years with items dating 116, 103 and 97 years in age. The mean age is nearly 13 years; the median is 7 years. The mode however, remains at 3 years. It should be noted that the second most frequently occurring subject category, "Other nonwater" accounts for some of the oldest materials cited. Many of the references are to early geological surveys, maps, county histories, census figures, and early agricultural reports.

Combining all WDNR publications studied produces composite figures that reflect the concentration of the older materials cited in Surface Water Resources publications. The mode, of course, continues to be 3 years but the mean age is nearly 10 years and the median is 6 years.

Form

For both Research Reports and Technical Bulletins publications, journal articles are the single most frequently occurring form. Research Reports cite 221 references to journals or 37 percent and Technical Bulletins have 127 cited references for nearly 45 percent of the total number of references cited in that publication. Research Reports and Technical Bulletins both have technical reports as the second most cited form. Technical reports represent 189 cited references in Research Reports for 22 percent of the total and the same form produced 66 references for approximately 23 percent of all cited references in the Technical Bulletins. In Research Reports bulletins and books follow in frequency of citation with 103 cited items or 13.12 percent and 102 references for 12.59 percent, respectively. Technical Bulletins reverse the order with books having 31 cited references for 11 percent and bulletins with 29 references at 10 percent of the total.

Having nearly twice the number of citations the rank order of the first five forms as listed in Research Reports is replicated in the listing that results from the combination of Research Reports and Technical Bulletins. First, journal articles dominate with 328 cited references, representing 40 percent of the total; second, technical reports at 189 items or 22 percent; third, bulletins with 103 cited references or 12.73 percent, fourth is books with nearly the same figures of 102 citations and 12.01 percent; fifth, conference proceedings at 72 cited items or 8.48 percent.

In contrast, Surface Water Resources and miscellaneous publications rank technical reports as most cited with 176 citations or slightly more than 39 percent of the total. Journals appear in fifth place as producing but 9 percent of the total, or 40 cited references. In second place comes bulletins at 76 items or 17 percent. Third place is books with 62 citations at nearly 14 percent of the total number of citations. In fourth place other government documents appear for the first time in the listing of the five most frequently cited forms of publications with 47 citations or 10.5 percent of the total.

In the combination of all the WDNR publications the rank order of the first five most cited forms follows the same listing patterns recorded in the Research Reports and the combined Research Reports-Technical Bulletins with the following alternations as to number of cited references and percent of the total: First, journals with 378 citations or 29 percent; second, technical reports at 365 cited references at 28 percent; third, bulletins having 179 references cited or nearly 14 percent; fourth is books with 164 items or 12.6 percent; and fifth place is conference proceedings with only 76 cited items for approximately 6 percent.

Books seem to play an important role in the literature of water resources literature, the WDNR authors cited 164 references to books, almost 13 percent. Other forms are roughly at the level one finds for other sciences. Reviews and abstracting and indexing are, no doubt, more heavily used than the 6 references cited might suggest. These forms are seldom cited, as well they should not, but are used to locate other literature which may then be cited.

Internal reports (including what are sometimes called "open file reports") and personal communication with 32 and 23 citations, respectively, may suggest a healthy communication among researchers both within and without the WDNR, but may also suggest provicialism, not enough is known to make a determination. In addition, it should be noted there were 150 references to WDNR-produced material. This self-citation is almost 12 percent of all the material cited.

As in most fields, theses play a relatively unimportant role in the water resources literature, at least in the thesis format, no attempt has been made to determine if other cited work was also reported in thesis format elsewhere.

Maps account for a surprisingly low percentage of the references cited, less than 1 percent, with only 10 maps cited. This low figure may reflect more the citing habits of the authors than the contribution maps make to the literature

Not one citation to a patent is recorded, suggesting that patent specifications are not a source of information among researchers at the WDNR interested in water.

Publisher

Non-profit organizations are responsible for more of the literature cited than any other source of publication in both Research Reports and Technical Bulletins. For Research Reports 225 cited references or 45 percent of the total are accounted for by non-profit publishers. Fifty-one percent or 146 citations result from this same type of publishing source for Technical Bulletins. In combination, non-profit publishers represent 47% or 401 items cited.

Close agreement continues for the second ranked publisher, government, with Research Reports having 41 percent or 232 cited references resulting from this source, and Technical Bulletins record nearly 40 percent or 113 items. The figure of 345 citations or 40.6 percent for government publishers is the result of the combination of Research Reports and Technical Bulletins.

Commercial publishers are poorly represented as a source of publication. In Research Reports only 77 items or approximately 13 percent are the result of commercial publication. Similarly, Technical Bulletins list 25 cited references or 9 percent from this type of publisher. The combination of Research Reports and Technical Bulletins produce 103 citations or 12 percent of the total attributable to commercial publishers.

The WDNR Surface Water Resources and miscellaneous publications contrast sharply with this ranking of publishers. Government as publisher is strongly represented in first place with a total of 342 cited references that account for nearly 77 percent of the total number of all references cited. Non-profit publishing is in second position with only 69 items or 15 percent and commercial publishers are third in rank at 35 references or nearly 8 percent.

When combining all WDNR publications, the high concentration of cited references of Surface Water Resources in government publishers places this source of publication into the dominant position with 687 cited references or 53 percent of the total. Non-profit organizations are second with 470 items at 36 percent and commercial publishers account for only 138 citations or approximately 11 percent. Figure 1 shows sources of publication of technical reports and bulletins cited.

Dispersion of Cited Journal Literature

The dispersion of titles has long been of interest to researchers using citation analysis. Stevens defined title dispersion in his review of citation studies as

> "... the degree to which the useful literature of a given subject area is scattered through a number of different books and journals. If there is comparatively much scattering, the title dispersion is high; if a large portion of the literature is contained in a few journals, the title dispersion is low." (8, p.12)

As will be seen below, this is an inadequate description of dispersion because it says nothing about the subject of the journal titles. The authors also prefer the term "concentrated" to Stevens' term "low."

Gross and Gross in the first citation study were only interested in ascertaining which titles a small liberal arts college should acquire and maintain in order to best serve the research needs of chemists with the smallest possible list (6). The distribution of citations among journal titles is of great interest to research librarians although there has been some serious criticism of citation analysis when utilized to determine library collections based on the most cited journals. Estelle Brodman's study of physiology journals covers this criticism (3). As Durrance points out, this criticism is only valid if the literature is interpreted to mean library use. (5).

This criticism notwithstanding, it has been well established that a citation analysis can generate a list of the journals which account for a high percentage of all references. These journals have become known as the core journals of a field, area, or discipline. Four titles (Journal of Water Pollution Control Federation, Journal of Wildlife Management, Journal of American Water Works Association, and Ecology) account for 84 of 375 references or over 22 percent of all journal references. Sixteen titles account for just over 50 percent of all references to journal articles. See Figure 2.

R. E. Stevens drew the conclusion that, "Title dispersion is greater for literature of technologies than it is for those of pure sciences," and also, "Title dispersion is greater for literatures of new science than it is for those of old sciences" (8, p. 13). Water resources as a separate field is relatively new, and displays many characteristics of a technology when viewed from the perspective of the WDNR, therefore, one might expect a relatively high dispersion of the literature cited in WDNR publications. If all serials are considered, which is the most common practice in citation analysis, then the title dispersion is concentrated, if not low.

The 378 cited references to journal articles are from 127 different journal titles, and the thirty-six most cited journal titles are widely dispersed among many disciplines. Among these titles are found journals dealing with water, sewage, water pollution, water resources research, limnology, and oceanography. In addition, the following subjects are represented by journal titles: Wildlife, ecology, fish and fisheries, general science, law, civil engineering, bacteriology, botany, chemical engineering, geology, agriculture and soil science. This high subject dispersion would suggest high dependence on the work of other fields as is expected of a newly developed applied discipline.

Bradford Distribution

To address our attention first to the citation of scientific and research journals we can see the dispersion of titles more clearly when the cited titles are arranged in descending order of frequency of citation.

When the cumulative number of cited journal titles is plotted against the cumulative frequency of citations for each title or the cumulative percentage of frequency of citation the curve in Graph 1 results. This characteristic has become known as a "Bradford distribution" or as S. C. Bradford first described this phenomenon, which he called, "scattering" (1), the extent to which articles occur in journals devoted to other subjects than that of the core journals. A few journals in each subject field include the majority of the articles on the subject and then a large number of articles on the subject are scattered throughout a large number of journals. Bradford's verbal formulation of the law reads:

> If scientific journals are arranged in order of decreasing productivity of articles on a given subject, they may be divided into a nucleus of periodicals more particularly devoted to the subject, and several groups or zones containing the same number of articles as the nucleus, when the number of periodicals in the nucleus and succeeding zones will be $1:n:n^2$ ... (2, p.154)

where n is a constant, being the ratio of p/p_1, where p, p_1, p_2... are the number of periodicals in the nucleus and succeeding zones respectively. This constant is also known as the Bradford multiplier.

In an article published in 1972, Elizabeth Wilkinson points out there is a difference between Bradford's own presentation of the law: one by a graph and a second by words (10). She concludes the verbal formulation expressed Bradford's theory and the graphic formulation expressed his observation. We take Bradford's theory as he expressed it in words and have applied it to determine if our data coincide with his theory of scatter.

The journal literature of our WDNR water resources references follows the Bradford law of scattering almost precisely. Even the journals contributing two references each and those contributing one reference each follow the Bradford scatter. This would suggest that the subject of water resources, as cited by the water personnel of the WDNR, is not restricted to water resources literature but dips into many fields (as noted earlier) and is not self-contained. As stated above, 16 journals contribute 51 percent of all the cited journal literature, thirty-six titles contribute almost 71 percent of all journal literature.

Foreign Literature

A total of twenty citations (1.6 percent) are to foreign journal titles, of these only two are not in the English language, one is in German and the other is in Dutch. Although this section deals only with journal citations, we shall discuss all foreign literature here. Even if the eight international conferences cited are counted as foreign (all of which are in the English language), the total number of citations to materials published outside the United States is only 43. Of these, forty-one are in the English language. Nineteen of these references are published in Canada and ten others are from Great Britain. Sweden and Germany account for 5 citations each, Italy two, and one each from The Netherlands and New Zealand. In addition to the 20 journals, there are six reports, six books, three bulletins and the eight international conference proceedings previously mentioned.

General Conclusions and Discussion

The purpose of this chapter is to bring together the several conclusions and to complement them with a brief interpretation and discussion.

Characteristics of the WDNR References

Journal Literature.

1. Over 35 percent of all the literature cited is in the form of the journal.
2. English is the predominant language of the works cited.
3. The major source of journal literature cited is the United States.
4. Concentrated use is made of a few journals.

5. A few additional journals are little used; contrary to the findings of many previous citation analysis where concentrated use of a few journals is complemented by a large number of references spread over a large number of journals.
6. The total number of different journals cited is small.
7. The journals cited are dispersed among many subject disciplines. Among the top 36 journals, in addition to water and water-related journals, are journals devoted to subjects as varied as wildlife, fish and fisheries, ecology, general science, law, limnology, civil engineering, bacteriology, botany, chemical engineering, geology, agriculture and soil science.

Technical Report Literature.

8. Over 28 percent of all references in the publications studied were to technical reports.
9. The Office of Water Research and Technology is not a major producer of technical reports cited.
10. The Environmental Protection Agency is not a major producer of technical reports cited.
11. State and local governments are the major publishers of the technical reports cited.
12. Heavy use (56 percent) is made of reports published by Wisconsin Department of Natural Resources and comparable departments in the several states.

Bulletin Literature.

13. Heavy use is made of bulletins published by the state of Wisconsin (exclusive of WDNR).
14. Heavy use is made of WDNR Technical Bulletins and Research Reports (self-citation).
15. Federal agencies are relatively unimportant sources of bulletins cited. The most frequently cited agencies are the U.S. Geological Survey, the U.S. Department of Health, Education and Welfare, and the U.S. Department of Agriculture.

All Forms.

16. The federal government is not a major source of the cited literature.
17. Less than one percent of all references cited is to publications of the Office of Water Research and Technology.
18. There is little reliance upon serial publications, which is a typical of scientific literature in general.
19. There is heavy use made of recent materials (this may reflect the newness of the field).

Discussion

The literature cited in the publications studied is in many ways typical of other literatures cited by similar groups of authors. It is no surprise that English language journals, produced in the United States, are the most frequently cited materials. Nor is the fact there is a concentrated use made of a few journals unexpected. Citation studies often identify a list of "core" journals that represent a large percentage of the journal citations, this "core" often consists of but a few titles. The "core" journals are usually followed by a long list of journals that are referred to a few times each, often only once or twice. What was found in our sample is a small number of additional journals referred to a few times each. Of course, this means there is a small total list of journals cited. This may be due to the lack of available materials within the WDNR; the follow-up study to this one of the awareness and availability of relevant research should yield more information and hopefully, additional evidence from which to make a judgment about this phenomenon.

The most revealing finding is the widely scattered subject coverage of the journals cited within this relatively small number of journal titles. The small number of titles suggests low scattering (or concentrated use) but the high number of quite different subject disciplines suggests high scattering (or unconcentrated use) of materials in this form. Again we do not have enough data but the follow-up study may produce additional evidence.

The high number of references to reports is not unexpected but the high percentage (56) to nonfederal reports is. The low use of OWRT and EPA reports suggests either the material in report form does not reach the personnel of the WDNR or it does not address itself to WDNR needs in a way they recognize as useful.

Self-citation is expected of such an organization, as is the heavy use of local (state of Wisconsin) publications. Also, heavy use of material produced by agencies with comparable missions from other states is expected. There is not enough evidence to know whether the citation of WDNR internal reports and personal communications is a healthy indication of informal communication or an indication of provincialism, only more information will reveal what these facts signify.

There is a low reliance upon serials in general and this is atypical of scientific literatures.

References

1. Bradford, S. C. 1934. "Sources of information on scientific subjects." Engineering 137:85-6 (January 26).

2. Bradford, S. C. 1948. Documentation. Crosby, London.

3. Brodman, E. 1964. "Choosing physiology journals." Bull. Medical Lib. Assoc. 32:479-83.

4. Burton, R. E. and B. A. Green, Jr. 1961. "Technical reports in physics literature." Physics Today 14, No. 10:35-7.

5. Durrance, R. 1975. A Comparison of Characteristics of Journal Articles of Laboratory Science and Environmental Science. Ph.D. thesis. University of Wisconsin, Madison, Wis.

6. Gross, P. L. K. and E. M. Gross. 1927. "College libraries and chemical education." Science 66:385-9.

7. Margolis., J. P. 1967. "Citation indexing and evaluation of scientific papers." Science 155:1213-9.

8. Stevens, R. E. 1953. Characteristics of subject literatures. (ACRL Monograph No. 6.) American Library Association, Association of College and Research Libraries, Chicago, Ill. pp. 10-21.

9. Walker, R. D. and G. J. Zuck. 1975. The Literature Cited in the Wisconsin Department of Natural Resources Publications on Water Related Subjects, 1964-1973. Water Resources Center, Madison, Wisconsin, (Technical Report WIS WRC 75-03).

10. Wilkinson, E. A. 1972. "The ambiguity of Bradford's law." J. Doc. 28:122-30.

Figure 1 Sources of publication of technical reports and bulletins cited in WDNR Research Reports, Technical Bulletins, Surface Water Resources series and miscellaneous publications published during the period, 1964-73

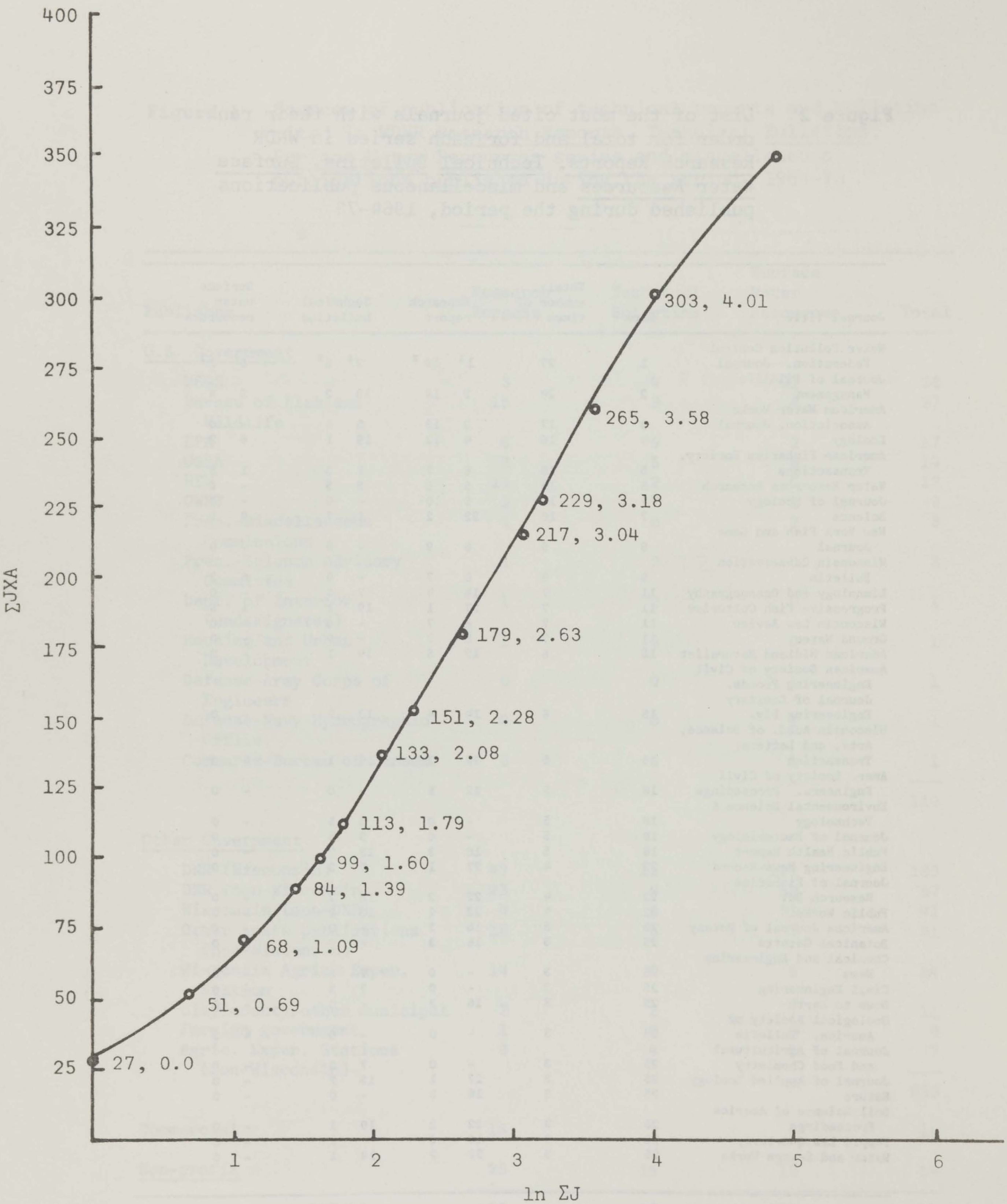
Publisher	Research Reports	Technical Bulletins	Surface Water Resources	Total
U.S. Government				
USGS	3	4	21	28
Bureau of Fish and Wildlife	16	3	8	27
EPA	6	9	2	17
USDA	3	5	6	14
HEW	11	2	0	13
OWRT	1	7	1	9
Pres.-Miscellaneous Commissions	1	0	2	3
Pres.-Science Advisory Committee	1	2	0	3
Dept. of Interior (undesignated)	1	0	0	1
Housing and Urban Development	1	0	0	1
Defense-Army Corps of Engineers	0	0	1	1
Defense-Navy Hydrographic Office	0	0	1	1
Commerce-Bureau of Census	0	0	1	1
				119
Other Government				
DNR (Wisconsin)	47	15	121	183
DNR (non-Wisconsin)	23	4	30	57
Wisconsin (non-DNR)	4	6	31	41
Other state publications (non-Wisconsin)	18	6	7	31
Wisconsin Agric. Exper. Station	14	0	0	14
City/county/other municipal	2	2	7	11
Foreign government	1	8	0	9
Agric. Exper. Stations (non-Wisconsin)	3	4	0	7
				353
Commercial	16	2	0	18
Non-profit	25	16	13	54
TOTALS	197	95	252	544

Figure 2 List of the most cited journals with their rank order for total and for each series in WNDR <u>Research Reports</u>, <u>Technical Bulletins</u>, <u>Surface Water Resources</u> and miscellaneous publications published during the period, 1964-73

Journal Title	Rank	Total number of times cited	Research report		Technical bulletins		Surface water resources	
			[1]	[2]	[1]	[2]	[1]	[2]
Water Pollution Control Federation. Journal	1	27	1	19	2	6	6	2
Journal of Wildlife Management	2	24	2	18	13	2	3	4
American Water Works Association. Journal	3	17	3	13	6	4	–	0
Ecology	4	16	4	12	19	1	6	2
American Fisheries Society. Transactions	5	15	8	7	7	3	1	5
Water Resources Research	6	14	6	9	3	5	–	0
Journal of Ecology	7	10	5	10	–	0	–	0
Science	7	10	22	2	1	7	9	1
New York Fish and Game Journal	9	9	6	9	–	0	–	0
Wisconsin Conservation Bulletin	9	9	8	7	–	0	6	2
Limnology and Oceanography	11	7	14	4	7	3	–	0
Progressive Fish Culturist	11	7	27	1	19	1	1	5
Wisconsin Law Review	11	7	8	7	–	0	–	0
Ground Water	11	7	8	7	–	0	–	0
American Midland Naturalist	15	6	12	5	19	1	–	0
American Society of Civil Engineering Procds. Journal of Sanitary Engineering Div.	15	6	14	4	13	2	–	0
Wisconsin Acad. of Science, Arts, and Letters. Transaction	15	6	22	2	19	1	4	3
Amer. Society of Civil Engineers. Proceedings	18	5	12	5	–	0	–	0
Environmental Science & Technology	18	5	–	0	3	5	–	0
Journal of Bacteriology	18	5	–	0	3	5	–	0
Public Health Report	18	5	16	3	13	2	–	0
Engineering News-Record	22	4	27	1	7	3	–	0
Journal of Fisheries Research Bd.	22	4	22	2	13	2	–	0
Public Works	22	4	22	2	13	2	–	0
American Journal of Botany	25	3	16	3	–	0	–	0
Botanical Gazette	25	3	16	3	–	0	–	0
Chemical and Engineering News	25	3	–	0	7	3	–	0
Civil Engineering	25	3	–	0	7	3	–	0
Down to Earth	25	3	16	3	–	0	–	0
Geological Society of America. Bulletin	25	3	–	0	–	0	4	3
Journal of Agricultural and Food Chemistry	25	3	–	0	7	3	–	0
Journal of Applied Ecology	25	3	27	1	13	2	–	0
Nature	25	3	16	3	–	0	–	0
Soil Science of America Proceedings	25	3	22	2	19	1	–	0
Temple Law Quarterly	25	3	16	3	–	0	–	0
Water and Sewage Works	25	3	22	2	19	1	–	0

[1] Rank order

[2] Number of times cited



Graph 1. Bradford distribution. Cumulative frequency of citations (ΣJXA) and Log of cumulative number of cited journal titles (lnΣJ).

GeoRef Developments

John G. Mulvihill
Manager, GeoRef
American Geological Institute
5205 Leesburg Pike
Falls Church, Va., 22041

GeoRef is a machine-readable bibliographic data base which covers all aspects of geology, worldwide, produced by the American Geological Institute. The principle printed product of GeoRef is the Bibliography and Index of Geology, which is published by the Geological Society of America.

This is a report on three related developments at GeoRef: (1) adoption of the UNISIST Standard for bibliographic citations, (2) use of intelligent, CRT terminals for input, and (3) use of Samantha for data base maintenance and photocomposition. These were implemented concurrently, with the characteristics of each in mind as they affected the other two. An unrelated development, conversion of USGS tapes to GeoRef format,is also discussed.

UNISIST Standard--This is set forth in the Reference Manual for Machine-Readable Bibliographic Descriptions, prepared by the UNISIST/ICSU-AB Working Group on Bibliographic Descriptions, compiled by M. D. Martin; UNESCO, Paris, 1974, 71 pages, Pub. No. SC.74/WS/20.

Included therein are a basic set of data elements to be used for bibliographic citations, together with detailed specifications for the format of each element in machine-readable form.

Adopting the Standard has given AGI a carefully worked out, internally consistant document on which to base its input to GeoRef. Furthermore, use of the Standard will facilitate exchange of new citations between GeoRef and other bibliographic data bases, which overlap in coverage. Economic pressures for such exchanges, at least of bibliographic information if not indexing, is mounting. The National Science Foundation is sponsoring research on overlap which may increase this pressure. This exchange of citations would not have been feasible with the old GeoRef tape format. Also, the Standard, by tagging the information in GeoRef in greater detail, has given us more precision in arranging information for publication in the various printed products of GeoRef.

A key feature of the UNISIST Standard is to require categorization of each bibliographic citation by document type e.g. book or serial, and bibliographic level. The bibliographic levels are collective, monographic, and analytic. To illustrate the categorization, a book cited as a unit would be categorized "book-monographic." Each of the chapters of a book, if cited individually, would be categorized "book-analytic." A multivolume book (more than one physical unit) taken as a whole, would be categorized "book-collective." See Appendix 1, Document Type-Bibliographic Level Combinations.

The categorization clarifies the nature of each citation. It is the basic criterion for computer validation of each GeoRef citation, and it is the chief means of determining the format of a citation printed from GeoRef. A wide variety of types of documents and bibliographic levels are common in GeoRef. Serials, books, reports, theses, and maps are regularly cited. Separately authored chapters of books and reports are cited. The Standard forces us to decide immediately, as the first step for inputting a document, that document's type and the level of the citation. The indexing form, data elements, validation checks, and printed citation format follow from that decision. Prior to the Standard, much was fuzzy concerning input and format which is now clarified by the type-level categorization.

For each document type-bibliographic level combination, certain data elements are appropriate--some required, some optional. Other data elements are never used for that type-level combination. For the author, corporate author, title, and affiliation elements, the data has a different tag depending on whether it pertains to the analytic, monographic or collective level. This type-level differentiation necessarily results in many tags. In the case of GeoRef, adoption of the UNISIST Standard increased the number of data elements from 21 to 80. See Appendix 3, "GeoRef Data Elements."

To accomodate this increase in data elements, we designed a different Citation Input Form for each type-level combination. One of these forms, "Book-Analytic" is included as Appendix 4.

Intelligent Terminals--Datapoint 1100 diskette terminals are micro-computers with keyboards and CRT screens. We use three of these as input devices for GeoRef citations. In addition to making the citations computer-readable, these terminals: (1) for each type-level combination, display only that series of screens containing the appropriate data elements (see Appendix 2, "Screen Chains"); (2) save data common to a series of citations and append that data to each document in the series; (3) substitute index terms for letters representing those terms to form term sets, with the result that each term need be keyed only once for a document while being repeated in several term sets; (4) force the entry of certain required data elements; (5) verify CODEN and ISSN by means of check digits; (6) limit the type of characters permitted in certain data elements, e.g. restrict the Category Codes to the digits 01-29; and (7) eliminate the typing of data element tags by attaching the appropriate tag to each data element.

A Diablo printer is connected to one of the terminals to provide hard copy. This terminal, a two diskette machine, also functions as a work station for remote job entry. Citations are transmitted from this machine over local telephone lines to a computer center in Washington, D. C. Notice of computer job status and of job results are received on the terminal from the computer center.

Computer Processing--in February, 1975 AGI purchased Samatha, a package of computer programs for data base maintenance and for photocomposition. Samantha was purchased to give AGI ownership of its own processing programs, flexibility in choosing computer centers for processing, and direct control of its computer processing schedule. Cost reduction was another important consideration.

Prior to the use of Samantha, AGI had sent out its raw input data for processing to the company which did its photocomposition. AGI was billed for each new citation processed, for each record corrected, etc. With Samantha AGI in effect handles its own computer processing. Now AGI pays by the hour for computer time used. And because we are a steady customer and can use any of several IBM equipped computer centers in the D.C. area, we pay competitive prices.

Samantha has all the capabilities offered by the company formerly used by AGI, plus additional features not available previously, notably more edit checks, proofs of citations which include embedded error messages, more easily correctable proofs of subject and author indexes, and more flexibility in formatting photocomposed pages.

AGI contracted with Information and Publishing Systems, Inc. (I&PS) which developed Samantha, to have each routine computer job in the GeoRef production cycle set up, tested, and made operational before being handed over to the GeoRef production staff. This includes the photocomposition runs for each type of publication from GeoRef. Following this startup phase which is not yet completed, I&PS, which is based in Bethesda, Maryland, is to be retained on call to handle any unusual processing problems which may occur.

The existing AGI staff, which does not include a programmer, has been able to handle the normal computer processing associated with GeoRef.

AGI production jobs, submitted at AGI from the Datapoint terminal, are normally run overnight and delivered before eight o'clock the following morning. The jobs are set up in such a way that the job control language need not be altered each time they are run. Tape rotation and tape backup for each GeoRef update are provided for in the job control language.

Samantha (1) produces a proof sheet for each new or changed citation (see Appendix 5); (2) produces proofs of each index for editing prior to photocomposition; (3) provides extensive correction capabilities; (4) sorts and cumulates new accessions as added, to produce the GeoRef Master File tapes; (5) produces tapes which drive the Videocomp photocomposer for each type of publication made from GeoRef; and (6) produces GeoRef search tapes.

In the course of processing a new or changed citation, Samantha (1) checks for duplicate citation numbers; (2) flags missing, multiple or unauthorized data elements, based on document type, bibliographic level, and the presence of other data elements; (3) adds text and punctuation to citations before photocomposition, e.g. inserts "Report No." before each report number in a citation; (4) looks up and inserts short titles and full titles for serials into citations on tape, based on input of CODEN or ISSN, and inserts full text for languages and countries, based on input of standard language and country codes; and (5) verifies level one and level two index terms, both for spelling and capitalization.

Using Samantha a 458 page special bibliography and index of Colorado geology has been produced. The October, 1975 issue of the <u>GSA Bibliography and Index of Geology</u> will be the first issue done by the new programs.

Samantha accomplishes all the major computer processing needs associated with GeoRef. It is also our plan to produce by Samantha an information retrieval thesaurus and an update to AGI's Glossary of Geology.

Conclusion--The processing changes described above reduce substantially the number of keystrokes needed to input GeoRef, extend the number of editing functions performed by the computer, and make the GeoRef data base citations compatable with other machine-readable files. The changes will save on editorial and proofreading time. Computer processing costs for maintenance of the GeoRef data base and for preparation of camera-ready copy will be reduced. The turnaround time to process a citation will be shortened. The last three statements are in the future tense, since we are still in the throes of converting from the old to the new systems.

The effect of these changes on on-line searching of GeoRef at Systems Development Corporation (SDC) is yet to be determined. It is important to keep the SDC on-line GeoRef search file unified. For this reason we may decide to map the new GeoRef tags into the existing SDC fields. If this were done there would be no change in the searching of GeoRef before and after October, 1975, with the possible exception of adding a new field for author's affiliation which information was not included in the GeoRef file until October, 1975.

USGS Tape Conversion--By a complex conversion program, Jack Wolfire of I&PS was able to analyze the citations of the 1961-1968 photocomposition tapes for the Bibliography and Index of North American Geology (NAB) into the old data elements used for GeoRef. A total of 65,122 citations were converted. This is not the total number of citations for the 1961-1968 NAB. Some citations were missed due to gaps in the tapes from USGS.

These converted citations are suitable for photocomposition. Some citations were selected for and appear in the Colorado bibliography which AGI recently completed. We plan to add these 1961-1968 citations to the SDC on-line GeoRef file, which will extend our coverage of North American geology back eight years.

Appendix 1
Document Type-Bibliographic Level Combinations*

	Analytic	Monographic	Collective
Serial	X	X	X
Book	X	X	X
Report	X	X	
Thesis		X	
Map		X	X

*This matrix differs from the UNISIST Standard in that Map is used and Patent is omitted. Also, the Standard does not include Serial-Collective as a valid combination.

Appendix 2
Screen Chains

DOC TYPE BIBLIOGRAPHIC LEVEL	COMMON SCREENS**	NON COMMON SCREENS
Serial Analytic	1,[2,3],4,5	11-26
Book Analytic	1,[2,3],6-9	11-26
Report Analytic	1,[2,3],7,10	11-26
Serial Monographic		1,[2,3],4-6,12-14,7,16-26
Book Monographic		1,[2,3],6,14,7-9,16-26
Report Monographic		1,[2,3],6,14,7,10,16-26
Thesis Monographic		1,[2,3],6,7,10,16-26
Map Monographic		1,[2,3],6,14,7-9,16-26
Book Collective		1,[2,3],6,14,8,9,16-26
Map Collective		1,[2,3],6,14,8,9,16-26

** The data on these screens is common to all chapters of a book or report, and all articles of a journal issue. It is keyed once and added to each part at the Datapoint conversion step.

Appendix 3
GeoRef Data Elements

TAG	FIELD NAME	TAG	FIELD NAME
A17@	Address, corp auth-A*	Z21	Index term, primary
A18@	Address, corp auth-M*	Z22	Index term, secondary
A19@	Address, corp auth-C*	Z23	Index term, tertiary
A14@	Address, prim. affil.-A	A06	Issue or part of serial
A15@	Address, prim. affil.-M	A06	Issue-collection
A25@	Address, publisher	Z41	Language (spelled out)
Z37	Affiliation, other	A24	Language code, summary
A14	Affiliation, primary-A	A23	Language code, text
A15	Affiliation, primary-M	Z04	Lit Type (S,B,R,T or M)**
A11@@	Alternate spelling, Auth-A	Z34	Map Type
A12@@	Alternate spelling, Auth-M	Z35	Medium of original document
A13@@	Alternate spelling, Auth-C	A31	Meeting address or loc
A43	Availability of report	A32	Meeting begin date
Z05	Bib level (A, M or C)	A32	Meeting inclusive dates
A02	CODEN	A30	Meeting name
Z03	Category code	A99	Miscellaneous (unused)
Z29	Conference? Yes or No.	A29	Number of pages
A17	Corporate author-A	A28	Number of parts-C
A18	Corporate author-M	Z31	Operator
A19	Corporate author-C	A07	Other issue/pt ID
Z40	Country (spelled out)	A20	Page numbers
A17@@	Country code, corp auth-A	A11	Person-analytic
A18@@	Country code, corp auth-M	A13	Person-collective
A19@@	Country code, corp auth-C	A12	Person-monographic
A31@	Country code, meeting	A25	Publisher
A14@@	Country code, prim. affil.-A	A39	Report number
A15@@	Country code, prim. affil.-M	A11@	Role, person-A
A25@@	Country code, publisher	A13@	Role, person-C
Z26	County	A12@	Role, person-M
A21	Date of publication (YYYYMMDD)	Z33	Scale
A41	Degree granting institution	A03	Short title of serial
A42	Degree level	Z38	Source note
A27	Edition	Z27	Special bib code
Z30	Editor/Indexer	A46	Summary only? Yes or No
Z25	Formation, geologic	Z24	Title annotation
Z12	Full title of serial	A08	Title-A (incl. type code)
Z36	Grid location (unused)	A10	Title-C (incl. type code)
Z01	ID Number	A09	Title-M (incl. type code)
A26	ISBN	A05	Volume of serial
A01	ISSN	A05	Volume-collection
Z32	Illustrations		

* @=a subfield of the tag, A=analytic, M=monographic,C=collective
** S,B,R,T or M=serial, book, report, thesis, or map

01 ID 30055

BOOK-ANALYTIC(1)

ED/IND JC

PUB DATE 1974

06 PER-M/C Surange,

ROLE editor

ALT

07 TITLE-MON (O M L T X)

Aspects and appraisal of Indian palaeobotany

08 TITLE-COL (O M L T X)

VOL-C

ISSUE-C

NUMB PTS-C

09 PUBLISHER Birbal Sahni Inst. Palaeobot.

LOC Lucknow

COUNTRY IND

EDN

ISBN

GeoRef, July 1975

01 ID BOOK-ANALYTIC(2) ED/IND

11 PER-A ROLE ALT

Gowda, S. Sambe

[Note - author's names are at ends of articles]

12 PRIME AFFIL Bangalore Univ., Dep. Geol.

LOC Bangalore COUNTRY IND

13 OTHER AFFIL

15 TITLE-A (OMLTX)

Precambrian and Cambrian ——

16 CATEGORY Ø9

LANG OF TEXT LANG OF SUMM

PG NUMBERS 1-8

ILLUSTRATIONS tables

MAP TYPE SCALE

GRID

SPEC BIB CODE M12 SUMMARY ONLY

MEDIUM

GeoRef, July 1975

17 TITLE ANNOT Algae, palynomorphs, India
18 COUNTY

FORMATION

SOURCE NOTE

19	A	Asia	S	biostratigraphy	22
	B	India	T		
	C	paleobotany	U		23
	D	algae	V		
	E	palynomorphs	W		
20	F	Plantae	X		
	G	Precambrian	Y		
	H	Cambrian	Z		24
	I	stromatolites	1		
	J	floral studies	2		
21	K	Cyanophyta	3		
	L	Dasycladaceae	4		
	M	fungi	5		25
	N	acritarchs	6		
	O	fossil wood	7		
22	P	megaspores	8		
	Q	miospores	9		
	R	occurrence			

26	SET	B C F D E M G H	SET	E J G H B R S N P Q
	SET	G A B S F D E M O	SET	
	SET	H A B S F D E M O	SET	
	SET	D J G H B R S I K L	SET	

Appendix 5
Computer Proof from SAMANTHA*

DATE: 75343 DATA SET NAME: GE075MF

FIELD	WORD		FIELD DATA	WORD
2	1	Z01	75-30055	2
100	1	Z04	B (book)	2
200	1	Z05	A (analytic)	2
300	1	Z29	N (not a meeting)	2
400	1	A21	1974	2
500	1	Z30	JC (indexer)	2
600	1	Z31	PP (operator)	2
700	1	A12	Surange, K. R. @editor	5
800	1	A12	Lakhanpal, R. N. @editor	5
900	1	A12	Bharadwaj, D. C. @editor	5
1000	1	A09	0 Aspects and appraisal of Indian palaeobotany**	8
1100	1	A25	Birbal Sahnj Inst. Palaeobot. @Lucknow @IND	7
1101	1	Z40	India	2
1200	1	A11	Gowda, S. Sambe	4
1300	1	A14	Bangalore Univ., Dep. Geol. @Bangalore @IND	7
1301	1	Z40	India	2
1400	1	A08	0 Precambrian and Cambrian plants**	6
1500	1	Z03	09	2
1600	1	A20	1-8	2
1700	1	Z32	tables	2
1800	1	Z27	M12 (a micro-paleontology bib code)	2
1900	1	A46	N (not a summary)	2
2000	1	Z24	Algae, palynomorphs, India	4
2100	1	Z21	India	2
2200	1	Z22	paleobotany	2
2300	1	Z23	Plantae, algae, palynomorphs, fungi, Precambrian,	6
	7		Cambrian	7
2400	1	Z21	Precambrian	2
2500	1	Z22	Asia	2
2600	1	Z23	India, biostratigraphy, Plantae, algae, palynomorphs,	6
	7		fungi, fossil wood	9
2700	1	Z21	Cambrian	2
2800	1	Z22	Asia	2
2900	1	Z23	India, biostratigraphy, Plantae, algae, palynomorphs,	6
	7		fungi, fossil wood	9
3000	1	Z21	Algae	2
3100	1	Z22	floral studies	3
3200	1	Z23	Precambrian, Cambrian, India, occurrence,	5
	6		biostratigraphy, stromatolites, Cyanophyta,	8
	9		Dasycladaceae, Tasmanites	10
3300	1	Z21	palynomorphs	2
3400	1	Z22	floral studies	3
3500	1	Z23	Precambrian, Cambrian, India, occurrence,	5
	6		biostratigraphy, acritarchs	7

* Typed reproduction

** The 0 before the title indicates "original title."

The Awareness of Relevant
Water Resources Literature*

by

Richard D. Walker
Library School, University of Wisconsin, Madison 53706

Introduction

The Problem

What recently published literature in the field of water resources is known to the Wisconsin Department of Natural Resources (WDNR) personnel interested in any of the several aspects of water resources.

In addition, a number of complementary aspects of the relevant water resources literature were dealt with during the investigation: 1) sources of financial support, 2) availability of publications in the WDNR, 3) physical characteristics, 4) publishers and geographical sources, 5) subject concentration, and 6) journal title dispersion. Differences between the known and the unknown literature were as investigated.

This study, then, investigated the level of awareness of water resources literature within one state agency, the Wisconsin Department of Natural Resources (WDNR), charged with seeing to it that the latest available information is used to solve water resources problems within the state. It was assumed, there is a logical relation between information and the application of knowledge. It was also assumed there exists an interest in finding new information and applying new knowledge. Further, it was assumed relevant and useful information does exist and it is contained in the WRSIC file and is published in Selected Water Resources Abstracts (SWRA).

The Office of Water Research and Technology (OWRT) of the Department of the Interior operates the Water Resources Scientific Information Center (WRSIC), which seeks to collect and disseminate scientific and technical information to the water resources community. See the description of the information retrieval program at WRSIC which appears in the Proceedings of the Ninth Annual Meeting of the Geoscience Information Society, November 18, 1974, Miami Beach, Florida.

The Office of Water Research and Technology has long been aware of the importance of supporting scientific and technical research and development in the area of water resources and has promoted it in many ways. It has recognized the need to improve the availability of water resources information and has exercised leadership in this area. Much research into information systems has been supported by OWRT; the present study is an example of their interest in learning more about the use of the water resources literature. The enormous volume and complexity of water resources information, the multi-disciplinary aspect of much of the work in water, and the urgency of solving far reaching problems make it difficult to fully appreciate, identify, or acquire for use.

The importance of the OWRT to exercise leadership in the field of water resources information research, its creative dissemination, and information exploitation is well understood. If the research in the area of water resources is to be of any appreciable value to society it is necessary that the Water Resources Research Act of 1964, as amended (P.L. 88-379), be fully implemented, including the requirement to disseminate the literature. There can be no viable policy for water resources research without a policy for processing and disseminating the literature produced as a result of research. It must be a coherent, comprehensive information policy, covering all aspects related to water resources.

Users of Water Resources Information

Water resources information is used by many groups within society for many different purposes. It is used by the scientist to add to scientific knowledge; the scientist may be a biologist, a chemist, geologist, limnologist, or other specialist. It is used by the engineer to aid in the interpretation of other information or data newly recorded, or in the design of a new piece of equipment, or to help solve a technological problem in some area of water resources. Administrators and bureau chiefs in all levels of government use it to organize and interpret; to help make decisions in matters of water quality, water quantity, water facilities, water supply, and other water matters. The citizen also must use the information to gain understanding and make appropriate choices in water matters.

Availability of Water Resources Information

In water resources, as in any field, relevant and useful information may be produced in any part of the world, in one of many languages. Many formats are used to communicate research findings and technological advances; formal professional literature appears in books, journals, technical reports, conference proceedings and many other forms. Much new information is exchanged at scientific meetings and during formal and informal visits between researchers. An individual interested in some aspect of water resources may have a difficult time attempting to learn what is already "known," where it was published, or who distributed it, or how to obtain a copy of a desired piece of information.

Bibliographic Control of Water Resources Literature

Bibliographic tools exist to aid the individual scientist in his search for water resources information. Many of the disciplines contributing to water resources have abstracting and indexing services covering much of the literature of their respective disciplines.

Many now take advantage of the capabilities of computers to help create and manipulate bibliographic records which were initially created to produce the printed bibliographic tools. The OWRT's own Selected Water Resources Abstracts with its computer searchable file is of prime interest to any scientist, technologist or other researchers in the area of water resources.

WRSIC Thesaurus and File

WRSIC published The Water Resources Thesaurus, second edition, 1971 as an aid to the indexing and retrieving the literature of water resources research. The thesaurus was developed under contract with the Science Information Exchange (SIE) of the Smithsonian Institute, which has a computerized file of all ongoing federal research, including that in the area of water resources. Just as WRSIC depends on its centers of competence to abstract reports of completed research it depends on SIE for its information about ongoing research. It must be admitted that WRSIC relies on incomplete sources for its data, however, its Selected Water Resources Abstracts is the most comprehensive abstracting service available and its file is the most complete file available for computerized searching even if it is far from comprehensive.

The WRSIC file is designed to serve the scientific and technical information needs of the water resources community; scientists, engineers, technologists, planners, and managers. Subject coverage includes published literature dealing with the water related aspects of the life, physical, and social sciences as well as related engineering and legal aspects of the characteristics, conservation, control, use, or management of water (4, p.iii).

As the present investigation was completed the file contained approximately 80,000 abstracts to monographs, journal articles, reports, and other publication forms; adding 28,500 during the period the investigation covered, 1973-1974. Each entry (see example in Appendix) includes a full bibliographical citation, abstract as it appears in SWRA, and a full set of descriptors or identifiers used to index and retrieve the entry. The descriptor word control is maintained through the use of the Water Resources Thesaurus. Each entry also indicates the subject area categories into which it is classified. There are ten fields and sixty groups similar to the water resources research categories established by the Committee on Water Resources Research of the Federal Council for Science and Technology (4, p.iii).

Access to the computerized file is by remote, interactive terminal via leased/line to the Holifield National Laboratory at Oak Ridge, Tennessee. Using the AEC/RECON software retrieval program, the WRSIC file can be searched by author, descriptor key words, identifier key words, and field and group codes. The present study was designed to serve: 1) as a way of identifying a potential user group of water resources information, 2) as a means of alerting that group to the available information, and 3) to demonstrate a felt, but up until now, unarticulated need for water resource information. In addition, information was obtained about the awareness of the several forms in which water resources information appears and their relative level of accessability and awareness.

The Hypotheses

Specifically these two primary research hypotheses were tested:

1. The water resources personnel at the Wisconsin Department of Natural Resources are not aware of the current useful and relevant literature published in the area of water resources.
2. Of the literature determined to be relevant and useful to Wisconsin Department of Natural Resources water resources

personnel more is supported by federal funds than either not supported or supported by non-federal funds.

And six secondary null hypotheses were tested:

1. More research appearing in technical reports is known than is research reported in journal articles.
2. The journal article is more frequently used to report federally sponsored research considered relevant and useful than is the technical report.
3. The literature relevant and useful to WDNR personnel is not available within the WDNR.
4. There is more relevant and useful OWRT research known than unknown.
5. There is more unknown OWRT supported research than unknown research not OWRT supported.
6. There is more unknown EPA supported research than unknown research not EPA supported.

A number of selected characteristics of the literature were described in this study; they pertained to the relevant and useful literature selected by the OWRT personnel.

<u>Characteristics of Relevant Literature</u>

Characteristics of scientific literature help to provide a theoretical basis for understanding the nature, the use, the production, and the consumption of scientific literature; this basis necessarily precedes better design, implementation, and improvement of information transfer systems intended to be used by scientists and others. There is ample research precedent for selection of certain characteristics for study. The earliest major study of literature characterization was conducted by Fussler in his study of the literature of chemistry and physics (1).

Studies have frequently included five characteristics: 1) form 2) subject dispersion 3) title dispersion 4) language and 5) national origin of publication. The present study includes each of these. Another characteristic which has been included frequently, age, has not been included in the present study since all of the material under examination is recent. In addition to the traditional five characteristics named above, five other characteristics have been included in the present study: 1) type of publisher 2) source of support (funding) 3) number of references 4) number of pages and 5) number of authors per publication.

The composite profile of the literature then consists of these ten separate characteristics used herewith to characterize all analyzed publications. For purposes of discussion these ten may be separated into two categories: Those characteristics which reflect the production of the documents and those which reflect the form and topics of the documents themselves. Those reflecting production are source of publisher, sources of support, country of publication, language, number of authors; those reflecting the physical document are, journal title dispersion, pages, references, and subject.

Methodology and Research Design

Data related to the relevance, awareness, availability, subject coverage, source, and physical characteristics of the water resources literature used in this study were collected from computer printouts resulting from subject searches of the WRSIC file for retrospective and current publications. The searches represented the subject interest profiles for a random sample of the water resources personnel employed by the Wisconsin Department of Natural Resources (WDNR).

The Universe - Water Resources Personnel

The term "water resources personnel" was broadly interpreted, and covered all professional persons in the WDNR engaged in an activity that could be described as water related, except those activities dealing exclusively with water law and water regulation enforcement.

The Sample

The water resources personnel group then consisted of some 74 persons as described above. A sample of twenty-three persons was randomly drawn from this group. Three of this sample elected not to participate and a fourth dropped out after some weeks. The data analyzed were gathered from the remaining 19 participants. Every individual in the sample had professional responsibilities within the Department requiring some use of technical or scientific information and had expressed an interest in the literature.

The Collection of Data

The method used to identify relevant research material and publications of professional interest to the WDNR water resources personnel retrieved from the WRSIC file was a straight forward one and is described briefly here. First, a refined interest profile was developed and then employed to identify a body of relevant documents as determined by the sample of WDNR participants in both a retrospective search printout and a series of printouts as a part of a demonstration Selective Dissemination of Information and Document Delivery Service conducted over a six month period of time. Because of the unique application of these interest profiles in this study, special care was taken to create profiles of high refinement in order to yield a high recall of relevant documents.

Research Design

A great deal of research has been conducted where precision is used as one of the measures of success of an information retrieval, indexing, or SDI system. The present research used the refined profile design to produce as many citations to relevant publications with as high a recall as possible while providing the participant (user) with few non-relevant citations. The purpose was to create a relevant body of both retrospective and current water resources literature for the purposes of determining the participants' level of awareness of that literature, its availability in the WDNR, and to analyse and describe certain of its characteristics. The composite literature profile consisted of ten separate characteristics (described earlier) which were used in this study to characterize all publications designed as relevant by the WDNR participants. Since the time involved in location, acquisition and examination of each relevant document would have been prohibitive, the printout citation/abstract served as a surrogate source of information in lieu of examination of the actual documents. The use of surrogate data of course means that interpretation was dependent upon the accuracy and completeness of the information provided in the bibliographic descriptors on file in the data base. The SDI/delivery service provided during the demonstration period was selected as a part of the research design, in part, as a "cover" for the collection of the data to be analysed in this study. However, the SDI/delivery demonstration was incorporated into the research design also because the researcher thought such a service had potential application with this group of users.

The researcher was interested in gathering information that might be used to improve the level of awareness of the participants in the future and to gain an indication of the kind of information used and needed by this potentially heavy user group of the Wisconsin Water Resources Reading Room collection, the resources available in the several libraries on the Madison campus of the University of Wisconsin, and in state agency libraries. The information gained will also be used to complement earlier findings concerning the informational sources used by WDNR personnel (8). This information is considered of value as evidence the University of Wisconsin is interested and capable of providing complementary library and information services to state agencies on a continuing basis.

The method selected and described above served still another important function. The researcher believed, and the experience sustained his belief, it would be a means of alerting this group of potential users of water resources literature of the range, amount, sources, characteristics, and availability of the literature needed by them to better perform their mission within the WDNR. Alerting this group (larger than just the sample participants, because others interested in water resources were cognizant of the demonstration and what it produced) not only to the individual citations listed on the series of six printouts received by each participant, but also alerting them to the source of the printouts, the WRSIC file and the Selected Water Resources Abstracts as valuable services. Many of the persons employed in the WDNR displayed ignorance of both these services as this investigation got underway.

Rationale

As with most studies into the efficiency of an information retrieval system or SDI system, high relevance with high recall was considered desirable. In this study the researcher wished to provide citations to literature that could produce as many "General Relevant" (GR) and "Specific Relevant" (SR) citations to publications as possible so as to generate a large body of such citations to be used to determine if the WDNR personnel interested in water related matters were aware of the literature relevant to their professional interests and their job oriented requirements. He also sought to determine how this group became aware of the relevant materials when they indicated prior knowledge of the cited work. Knowledge of the availability of material within the WDNR was also sought.

It was reasoned, a) if materials identified as relevant (GR and SR) do represent the literature in the WRSIC file relevant to the needs of WDNR personnel, and b) if the WRSIC file contains a high percentage of the literature of water resources, then c) an analysis of the material marked relevant represents a substantial portion of the recent and current water resources literature being produced from all sources including the federal government and, in particular, that produced by the Office of Water Research and Technology (OWRT), the Environmental Protection Agency (EPA), the U.S. Geological Survey (USGS), and the U. S. Department of Agriculture (USDA), and therefore d) such an analysis identifies the literature most needed by this group of water resources personnel, and e) further identifies how much of it is known by them, and f) what the major sources of knowledge of the literature are.

Report of Results

Each participant was asked to note the following information on labels attached to the printouts adjacent to each citation/abstract.

I. Is this publication

a. of General Relevance. If yes, "GR".
b. of Specific Relevance. If yes, "SR".
c. not relevant? If so, "X".

II. If relevant, that is either "GR" or "SR", were you aware of the publication represented by the citation/abstract prior to this listing? yes_____ no_____.

III. If relevant and if you were aware of it, indicate how you became aware of it.

a. Scanning of the literature in general?
b. Personal subscription to the literature?
c. By means of written communication?
d. By means of oral communication?
e. Other means of awareness?

IV. is there a copy of this publication available to you within the DNR? yes_____ no_____

V. If you should like to receive a copy of this publication, please so indicate

Primary Hypotheses

The first primary hypothesis tested was: the water resources personnel of the WDNR are not aware of the current useful and relevant literature published in the area of water resources.

The hypothesis was accepted. The 19 participants in our sample were aware of only 140 of the 1267 relevant publications identified during the study. This represented only 11 percent of the relevant literature. Stated another way, this is a novelty ratio of 89 percent. Novelty is a measure of the participants awareness of relevant literature of his or her field.

The second primary hypothesis was: the literature determined to be relevant and useful to WDNR water resources personnel is not supported by federal fundings.

This second hypothesis was rejected. Of the 1267 publications designated as relevant by the 19 participants in the sample, 743 were known to have been supported and 68 per cent of these had been supported by federal funds.

Secondary Hypotheses

There are six null hypotheses tested. The first two were: 1) more research appearing in technical reports is known than is research reported in journal articles, and 2) the journal article is more frequently used to report federally sponsored research considered relevant and useful than is the technical report.

Both of these null hypotheses were rejected. More research appearing in technical reports was not known than was research in journal articles. In fact, a greater percentage of the relevant journal literature was known than the percentage of the relevant technical report literature. Also the journal article was not more often used to report federally sponsored research considered relevant and useful literature than was the technical report used for the same literature.

A third null hypothesis was: the literature relevant and useful to WDNR personnel is not available within the WDNR.

This hypothesis was accepted. With less than 6 percent of the relevant literature available to the water resources personnel at the WDNR there was no question that this was a major factor having effect on other aspects of this study. There was and is no central source of publication within the organization; there was and is no systematic program of acquisitions of material and no bibliographic apparatus of any kind to make known the acquisition or existence of publications within and between the several organizations divisions of the WDNR.

The fourth null hypothesis: there is more relevant and useful OWRT research known than unknown is rejected. Only 5 percent of the 121 OWRT supported research publications designated as relevant by the participants were known to them at the time of their appearance on one of the printouts distributed as a part of this study. This too is a reflection of the lack of any systematic and concerted effort within the organization to acquire and maintain the literature as it is published or otherwise made available for distribution.

The remaining two null hypotheses: 1) there is more unknown OWRT supported research than unknown research not OWRT supported and 2) there is more unknown EPA supported research than unknown research not EPA supported. Both of these hypotheses were rejected. The data supports the hypothesis that EPA and OWRT research is as well known as research from other sources; there is a small, but significant degree of association between awareness and source.

Availability

Only 71 of the 1267 relevant publications were available within the WDNR, less than 6 percent of the total relevant literature. Secondary hypothesis number 3 is supported by these data, the relevant and useful literature to the WDNR personnel interested in water resources is not available within the WDNR.

Subject

Water Quality Management and Protection is the subject field most frequently represented by publications relevant to the mission of the participants in our study; 907 of the 1267 relevant publications cover this subject field, nearly 72 percent. The next most frequently identified subject field is Water Resources Planning, with 158 publications or approximately 13 percent. Figure 1 shows the complete breakdown by subject field and subject groups within the larger subject fields.

Within the broad subject field Water Quality Management and Protection, the most frequently occurring subject is Water Treatment Process, with 382 or slightly more than 30 percent of all relevant publications; this is followed by 161 publications on Sources of Pollution (13 percent). In fourth position is Identification of Pollutants, with 95 publications nearly 8 percent. Only Water Law and Institutions, with 92 publications (7 percent) came close to these four Water Quality topics, tying with still another Water Quality topic, Water Quality Control. All other fields or groups are in much less demand.

Source of Publication

Almost 23 percent (22.49) of all publications designated by the WDNR personnel as having either general or specific relevance are published or distributed by societies; 285 of the 1267. Commercial publishers account for 16 percent (178 publications), following as the next most important sources of publication of relevant literature. If, however, all domestic government publications from all levels of government are grouped together they become the type of publisher producing the greatest number of publications relevant to our sample member's professional and mission interests; this group accounts for 492 publications or almost 39 percent. The government agency represented by the highest number of publications is the Environmental Protection Agency (EPA) with 33; closely following is the Office of Water Research and Technology (OWRT) with 29. The U.S. Geological Survey (USGS) contributed 16 and "county and city" governments contributed 11 more, while "other U.S." accounted for 76 publications. If university publications are grouped with societies, together they account for 421 publications or almost exactly one-third of all publications.

Country of Publication Origin

The United States produced 1000 of the relevant publications, 79 percent. Of the remaining 267 publications, 65 were published in the United Kingdom, 43 in Canada, 37 in the U.S.S.R., and 26 in West Germany; these four countries accounted for another 13.5 percent and were followed by France (15), Poland (14), Czechoslavakia (9), Japan (7), and Italy; another 4 percent. The seventeen remaining countries account for the other 3.5 percent.

Language

English is the predominant language, 1183 publications (93.4 percent). Of the remaining 84 relevant publications, 71 are journal articles, 5 are conference papers, 3 are technical reports, 4 are bulletins, and one is a book. It is obvious from these and the above figures that many of the foreign publications are in the English language; of the 267 foreign sources only 83 are in languages other than English or less than one-third. Almost 100 are from English speaking countries, but many are written in English while being produced in non-English speaking countries.

Forms of Publication

Two secondary hypotheses tested are concerned with the form or format of the relevant literature as evaluated by WDNR personnel. The first of these hypotheses is simply: more research appearing in technical reports is known than is research reported in journal articles. The second of these hypotheses is: the journal article is more frequently used to report federally sponsored research considered relevant and useful than is the technical report.

The journal article is the most frequently occurring form among the 1267 relevant publications with 495 or 39 percent; while only 301 technical reports (24 percent) are recorded. It is clear that the journal article is still the form more often containing relevant information as seen by the WDNR personnel. The first hypothesis above is rejected. The probability that the number of technical reports is greater than or equal to the number of journal articles is less than 0.05 (X^2=4.42). Only 32 of the 301 technical reports were known to our sample participants, while 79 of the 495 journal articles were known. It is of particular importance however, that the level of awareness of the journal article and the technical report was almost the same for those publications appearing on the retrospective search printouts. In those printouts our participants knew or were aware of 31 of the 228 journal articles (13.66 percent) and 15 of the 114 technical reports (13.16 percent), suggesting, at least, the technical report format is not a deterrent to awareness. This also suggests journal articles come to the attention of our participants sooner than do technical reports.

The second hypothesis is also rejected; the journal article is not more often used to report federally sponsored research considered relevant and useful literature than is the technical report. Of the 68 OWRT publications considered here all were technical reports, as were all 67 EPA publications. Only one of the four USGS publications appeared as a journal article and only 12 of the 112 "other U.S. Government" publications were in the journal form.

As noted earlier, the overall awareness is but 11 percent or 140 known publications of the 1267 considered here. Books and conference papers appear to be two of the more elusive of the forms represented; of 18 books only 1 is known and only 16 of the 271 conference papers, being approximately 6 percent of these formats. Patent specifications appear to have some relevance to the interest of the WDNR mission, but are unknown to them if our sample is a true reflection of the lack awareness of the form, none of the eight designated patent specifications are known to our participants.

Journal Title Dispersion

There are 495 journal articles cited in 224 different journal titles. Forty-four journal titles accounts for 283 of these articles, less than 20 percent of the titles account for over 57 percent of the citations. One title, the Journal of the Water Pollution Control Federation, had 45 citations to it, being 9 percent of all journal citations. The top eleven journals have a total of 156 citations, almost one third (31.51 percent). Of these eleven titles, ten are of U.S. origin and one is a publication of an international society, published in the U.S. and elsewhere (Water Research, the Journal of the International Association of Water Pollution Research, published by Pergamon Press, Oxford, London, New York, and Paris); three are commercial publications and seven are publications of societies. Thirty of the top 44 are U.S. publications--16 societies, 12 commercial, and 2 governmental publications. One of the remaining 14 top representatives is an international journal, 12 are foreign (3 Canadian, 3 British, 2 each from U.S.S.R. and West Germany, and one each from Czechosolovkia and Poland), and one is of unknown origin. See Figure 2.

Conclusions

There is no question the WDNR water resources personnel are unaware of a great proportion of the literature designed for their use, regardless of the form or the source of publication. Their best source of awareness is their individual personal subscriptions to a few journals of interest; next best is scanning of a few additional journals available for this purpose. The Federal government supports more of the relevant research than does any other type of publisher. Federally supported research was not better known than the non-federally supported research; society and commercially produced publications were better known than government publications.

Perhaps the most telling findings of all, and one that may influence all other findings more than any other single fact, is the low availability of the relevant literature as defined by our sample participants. "Low availability" may be an understatement --"unavilability" may be a better term. The 6 percent of the relevant material available within the WDNR is a surprisingly low figure for an organization such as this; an organization necessarily relying so heavily upon current information in such a rapidly changing field.

Recommendations for further Study

This study and the previous reported citation study of WDNR publications have investigated the level of awareness and level of need of water resources literature of this special group of users. Additional research into the use of the water resources literature must be conducted to establish a better understanding as to how it might be transferred to persons whose responsibility it is to apply it. Other groups need to be investigated in much the same way so to continue to gain understanding of the reasons literature is known and unknown, understood and not understood, applied and not applied. Consulting engineers and environmental scientists working in the areas of planning and design of water works, sewerage, and industrial waste water treatment or preparing environmental impact reports are clearly persons needing information dealing with water resources. This group does work in areas of vital concern to governments at all levels, from the federal government to local municipalities.

There are still other potential users of water resources literature--these groups' needs also warrant study. Each group studied will make it clear what literature is known and available, will make the gaps in the several literatures more obvious, and will demonstrate the needs for areas where efforts need to be made to transfer information to the appropriate groups in the appropriate forms and levels of presentation. This much is clear--much of what has been produced and "made available" is not known by the persons for whom it was intended; any effort that improves the communication of information is a benefit that can be neglected only with some serious risk of loss to the water resources community and to society at large.

REFERENCES

1. Fussler, H. H. 1948. Characteristics of the research literature used by chemists and physicists in the United States. Ph.D. Thesis. University of Chicago, Chicago, Ill.

2. Kahle, R., et al. 1974. A Q-methodological study of attitudes toward water resources and implications for using mass media in dissemination of water research results. University of Missouri, Columbia, Water Resources Research Center. 349 pp.

3. Lamb, J. C. III. 1975. Methods for transferring water resources research findings to practicing engineers. University of North Carolina, Water Resources Research Institute, 1974. (Report No. 96) 48 pp.

4. Selected Water Resources Abstracts. 1968- . Water Resources Scientific Information Center. Semi-monthly.

5. Stewart, J. M. and D. H. Howells. 1972. Perception of water resources research dissemination, and utilization of research findings. University of North Carolina, Water Resources Research Institute. 95 pp. (Report No. 58).

6. Walker, R. D. and G. J. Zuck. 1975. The Literature cited in the Wisconsin Department of Natural Resources publications on Water related subjects, 1964-1973. Water Resources Center, Madison, Wisconsin. (Technical Report WIS WRC 75-03).

7. Walker, R. D., et al. 1975. The Awarness of the Reveland Water Resources Literature by the Personnel of the Wisconsin Department of Natural Resources. Water Resources Center, Madison, Wisconsin. (Technical Report WIS WRC 75-09).

8. Zweifel, L. G. and J. R. Luedtke. 1973. U.W. Water Resources Information Program interface with Department of Natural Resources. Final Report. Madison, Wisconsin. 47 pp.

Figure 1 Subject coverage of the publications designated as either of general relevance or specific relevance by WDNR Water Resources Personnel and the breakdown of the WRSIC file

Subject Field and Subject Group	Number of Publications	Percentage: Field	Percentage: Group
1. Nature of Water	0	0.0	
2. Water Cycle	89	7.02	--
a. General	3		0.24
b. Precipitation	0		0.0
c. Snow, Ice, Frost	0		0.0
d. Evaporation and Transportation	0		0.0
e. Streamflow and Runoff	2		0.16
f. Groundwater	7		0.55
g. Water in Soils	7		0.55
h. Lakes	17		1.34
i. Water in Plants	8		0.63
j. Erosion and Sedimentation	10		0.79
k. Chemical Processes	5		0.39
l. Estuaries	30		2.37
3. Water Supply Augmentation and Conservation	7	0.55	--
a. Saline Water Conversion	0		0.0
b. Water Yield Improvement	0		0.0
c. Use of Water of Impaired Quality	0		0.0
d. Conservation in Domestic and Municipal Use	2		0.16
e. Conservation in Industry	3		0.24
f. Conservation in Agriculture	2		0.16
4. Water Quantity Management and Control	52	4.10	--
a. Control of Water on the Surface	33		2.60
b. Groundwater Management	8		0.63
c. Effects on Water of Man's Non-water Activities	7		0.55
d. Watershed Protection	4		0.32
5. Water Quality Management and Protection	907*	71.60	0.08
a. Identification of Pollutants	95		7.50
b. Sources of Pollution	161		12.71
c. Effects of Pollution	133		10.50
d. Waste Treatment Processes	382		30.15
e. Ultimate Disposal of Wastes	21		1.66
f. Water Treatment and Quality Alteration	22		1.74
g. Water Quality Control	92		7.26
6. Water Resources Planning	158	12.47	--
a. Techniques of Planning	12		0.95
b. Evaluation Process	27		2.13
c. Cost Allocation, Cost Sharing, Pricing/Repayment	1		0.08
d. Water Demand	7		0.55
e. Water Laws and Institutions	92		7.26
f. Nonstructural Alternatives	3		0.24
g. Ecologic Impact of Water Development	16		1.26
7. Resources Data	33	2.60	--
a. Network Design	1		0.08
b. Data Acquisition	22		1.74
c. Evaluation, Processing and Publication	10		0.79
8. Engineering Works	21	1.66	--
a. Structures	8		0.63
b. Hydraulics	2		0.16
c. Hydraulic Machinery	4		0.32
d. Soil Mechanics	1		0.08
e. Rock Mechanics and Geology	1		0.08
f. Concrete	0		0.0
g. Materials	4		0.32
h. Rapid Excavation	0		0.0
i. Fisheries Engineering	1		0.08
9. Manpower, Grants, and Facilities	0	0.0	
10. Scientific and Technical Information	0	0.0	
Totals	1,267	100.00	100.00

* One publication assigned the Subject Field 05, with no Subject Group assigned.

Figure 2 List of the most cited journals with their rank order, showing number of citations percentage of citation, and percentage of total publications designated relevant, country of origin, and type of publisher.

Rank	Title	Number of citations n = 495	Percent of citations	Percent of relevant publications	Country of origin	Type of publisher[1]
1.	Water Pollution Control Federation Journal	45	9.09	3.55	U.S.	Soc.
2.	Water Research	22	4.44	1.74	I[2]	Soc.
3.	American Water Works Assoc. Journal	19	3.84	1.50	U.S.	Soc.
4.	Water and Wastes Engineering	12	2.42	.95	U.S.	Comm.
5.	Water Resources Bulletin	11	2.22	.87	U.S.	Soc.
6.	Water Resources Research	9	1.82	.71	U.S.	Soc.
7.	American Society of Civil Engineers J. of the Environmental Eng. Div.	8	1.62	.63	U.S.	Soc.
7.	The American City	8	1.62	.63	U.S.	Comm.
7.	Public Works	8	1.62	.63	U.S.	Comm.
10.	J. of Environmental Quality	7	1.41	.55	U.S.	Soc.
10.	TAPPI. J. of the Technical Assoc. of the Pulp and Paper Industry	7	1.41	.55	U.S.	Soc.
12.	Civil Engineering (ASCE)	6	1.21	.47	U.S.	Soc.
12.	Pulp and Paper Magazine of Canada	6	1.21	.47	Canada	Soc.
12.	Water Pollution Control	6	1.21	.47	U.K.	Soc.
12.	Bumazhnaya Promyshlennost	6	1.21	.47	USSR	Govt.
16.	Effluent and Water Treatment Journal	5	1.01	.39	U.K.	Comm.
16.	Environmental Journal	5	1.01	.39	Unk.[3]	Unk.[3]
16.	Water and Pollution Control	5	1.01	.39	Canada	Comm.
16.	The Minnesota Volunteer	5	1.01	.39	U.S.	Govt.
20.	Science (AAAS)	4	0.80	.32	U.S.	Soc.
20.	Chemical Engineering	4	0.80	.32	U.S.	Comm.
20.	Water and Sewage Works	4	0.80	.32	U.S.	Comm.
20.	Paper Trade Journal	4	0.80	.32	U.S.	Comm.
20.	Verhandlungen. Internationale Ver. Theor. Ange. Lim.	4	0.80	.32	W. Germany	Soc.
20.	Water and Water Engineering	4	0.80	.32	U.S.	Comm.
20.	Pulp and Paper	4	0.80	.32	U.S.	Comm.
20.	Human Ecology	4	0.80	.32	U.S.	Comm.
28.	Limnology and Oceanography	3	0.61	.24	U.S.	Soc.
28.	Ground Water	3	0.61	.24	U.S.	Soc.
28.	Environmental Science & Technology	3	0.61	.24	U.S.	Soc.
28.	Soil Science Society of Amer. Proc.	3	0.61	.24	U.S.	Soc.
28.	Przeylad Papierniczy	3	0.61	.24	Poland	Govt.
28.	Beef	3	0.61	.24	U.S.	Comm.
28.	Boston University Review	3	0.61	.24	U.S.	Soc.
28.	American Forests	3	0.61	.24	U.S.	Soc.
28.	Remote Sensing of the Environment	3	0.61	.24	U.S.	Comm.
28.	Hydrobiological Studies	3	0.61	.24	Czech.	Govt.
28.	Health Physics	3	0.61	.24	U.S.	Comm.
28.	Arkansas Out-of-Doors	3	0.61	.24	U.S.	Govt.
28.	Central Atlantic Environmental News	3	0.61	.24	U.S.	Soc.
28.	Das Papier	3	0.61	.24	W. Germany	Comm.
28.	Gidrobiologicheskii Zhurnal	3	0.61	.24	USSR	Govt.
28.	Canadian Chemical Processing	3	0.61	.24	Canada	Comm.
28.	Paper	3	0.61	.24	U.K.	Comm.
		283[4]	57.18[5]	21.43[6]		

[1] Type of publisher: Soc. = not-for-profit organization (Society)
Comm. = profit organization (Commercial)
Govt. = government agency

[2] I = International
[3] Unk. = Unknown
[4] 283 of 495
[5] % of 495 journal occurrences;770 were to forms other than Journals.
[6] % of 1267

Figure 3. Examples of citation/abstracts as they appeared on the computer printouts received by the WDNR participants. Examples of the abbreviated response label are included

44/6/00401-00522// 518 CONTINUATION 1 PAGE 477

FARM WASTES / ACID MINE WATER/WATER QUALITY/WATER QUALITY CONTROL/LEGISLATION/WASTE DISPOSAL/TERTIARY TREATMENT/WATER MANAGEMENT(APPLIED)/

44/6/00401-00522// 519
69R05753 WRA-W2-14 05.DO (W69-05753)
A FEASIBILITY STUDY OF A LIVESTOCK WASTE DISPOSAL SYSTEM INVOLVING THE REUSE OF WATER/
PRATT,GEORGE L./(NORTH DAKOTA STATE UNIV.,FARGO.)
RESEARCH PROJECT TECHNICAL COMPLETION REPORT TO OFFICE OF WATER RESOURCES RESEARCH,DECEMBER 1968,WASHINGTON,D.C.24 P, 11 TAB.2 REF.OWRR PROJECT A-001-NDAK./
05D/05G/
SEPARATING SOLID MATERIALS FROM LIQUID WASTES THAT HAVE BEEN REMOVED FROM A LIVESTOCK BARN CAN BE ACCOMPLISHED IN SEVERAL WAYS. IN TRIALS AT THE NORTH DAKOTA AGRICULTURAL EXPERIMENT STATION SETTLING TANKS AND SAND FILTERS WERE EVALUATED. IN ONE PHASE OF THE WORK TREATMENT OF OVERFLOW FROM A SETTLING TANK THAT COLLECTED BEEF WASTE WAS EMPHASIZED. NO TREATMENT,AERATION,AND TREATMENT WITH ALUMINUM SULFATE (ALUM)WERE TESTED FOR THE SETTLING TANK OVERFLOW.THE SETTLING TANK REMOVED THE BULK OF THE SOLIDS.ALUM TREATMENT OF THE OVERFLOW WAS EFFECTIVE IN REDUCING TOTAL SOLIDS TO THE POINT WHERE IT WAS NEARLY FEASIBLE TO PROVIDE FINAL TREATMENT.IN THE SECOND PHASE OF THE WORK A SLOW SAND FILTER WAS TESTED UNDER CONTROLLED CONDITIONS TO DETERMINE ITS ADAPTABILITY FOR FINAL TREATMENT OF RECLAIMED WATER.TESTS WERE RUN AT THREE TEMPERATURES, 45 DEG F,70 DEG F,AND 100 DEG F.CONTROLLED AMOUNTS OF ORGANIC MATERIAL WERE ADDED TO THE INFLUENT OF THE FILTER. THE FILTER WAS EASILY MANAGED AND ADAPTABLE TO A WIDE RANGE OF CONDITIONS.ITS PERFORMANCE WAS BETTER AT 70 DEG F THAN AT 45 DEG F OR 100 DEG F./
WATER REUSE/STOCK WATER/DOMESTIC ANIMALS/FARM WASTES/POTABLE WATER/SOLID WASTES/URINE/WATER POLLUTION/AERATION/AEROBIC TREATMENT / BIOLOGICAL TREATMENT / COAGULATION/FILTERING SYSTEMS/FILTRATION/FLOCCULATION/OXIDATION/WASTE STORAGES/WASTE WATER TREATMENT/WATER PURIFICATION/WATER TREATMENT/

Rel.___
Aw. ___y___n
Scan.______
Per.sub.____
Per.com.____
oral___
writ___
Other_______
DNR av._____
Want copy?
___y___n

44/6/00401-00522// 520
69R05417 WRA-W2-13 05.DO (W69-05417)
WATER QUALITY CHANGES IN CONFINED HOG WASTE TREATMENT/
MCKINNEY,ROSS E./BELLA,ROBERT/(KANSAS UNIV.,LAWRENCE.)
CONTRIB NO 24,KANS WATER RESOURCES RES INST PROJ COMPLETION REP,1968.88 P,21 FIG,12 TAB,13 REF.OWRR PROJECT A-011-KAN. /
HOG FARM WASTES/
05D/
TREATMENT METHODS FOR THE WASTES OF CONFINED HOGS ARE EXAMINED.CONFINED-ANIMAL WASTES ARE A MAJOR POLLUTION PROBLEM. THERE ARE ABOUT 55 MILLION HOGS IN THE U.S.,AND THEIR WASTES ARE LARGELY UNTREATED.WHEN THEY OCCUPY A LARGE AREA, FORAGING FOR FOOD,THIS IS NO PROBLEM.BUT UNDER CONFINEMENT THE PROBLEM IS SIMILAR TO THAT OF MUNICIPAL WASTE DISPOSAL.A FACILITY CONSISTING OF AN AEROBIC BIOLOGICAL TREATMENT SYSTEM FOR 10,000 HOGS WAS STUDIED.AN IMPORTANT CONSIDERATION WAS ODOR CONTROL BECAUSE OF A NEARBY METROPOLITAN AREA. RECYCLED EFFLUENT WAS USED FOR RAW-WASTE INPUT DILUTION,AND PADDLE-WHEEL AERATORS WERE SELECTED FOR HIGH OXYGEN TRANSFER EFFICIENCY AND LOW OPERATING COST.EACH BUILDING HAD AN AERATION DITCH UNDER ITS SLOTTED FLOOR FOR AEROBIC PROCESSING,AND FINAL DISPOSAL WAS BY A 2-STAGE INFILTRATION POND SYSTEM AND FIELD SPREADING OF SOLIDS. MEASUREMENTS WERE MADE OF DO,BOD,COD,PH,SUSPENDED SOLIDS,N,AND TRACE METALS. CHEMICAL ANALYSIS RESULTS AND COST DATA ARE TABULATED.OPERATION EXPENSE IS ABOUT 1-2 CENTS PER HOG PER DAY.STREAM POLLUTION IS PREVENTED OR GREATLY REDUCED AND ODORS ARE NO PROBLEM.(KNAPP-USGS)/
FARM WASTES/WASTE TREATMENT/AEROBIC TREATMENT/HOGS/FARM LAGOONS/SETTLING BASINS/SOIL DISPOSAL FIELDS/SEWAGE TREATMENT/ DISPOSAL/ENVIRONMENTAL ENGINEERING/SLUDGE DISPOSAL/WASTE STORAGE/

Rel.___
Aw. ___y___n
Scan.______
Per.sub.____
Per.com.____
oral___
writ___
Other_______
DNR av._____
Want copy?
___y___n

44/6/00401-00522// 521
69R04527 WRA-W2-11 05.DO (W69-04527)
ANAEROBIC DIGESTION OF PINEAPPLE MILL WASTES/
MCMORROW,MARTIN J.K./YOUNG,REGINALD H.F./BURBANK,NATHAN C.,JR./LAU,L.STEPHEN/KLEMMER,HOWARD W./(HAWAII UNIV.,HONOLULU.)
17/
TECH RPT NO 27 ... 19... 23 P, ... FIG, ... B,2 ... F.O ... PROJ ... A-O ... I./
ANAEROBIC DIGES...N/CANNERY WASTES/

ENVIRONMENTAL POLICY MAKING AND INFORMATION

by

Marta L. Dosa

School of Information Studies
Syracuse University
Syracuse, New York 13210

Introduction

In recent years many countries, especially in the industrialized areas of the world, were jolted into recognizing that, "the environment could not be compartmentalized, for it is constituted by a network of interacting relationships extending through all sectors of human activity."[1] The ever-changing pattern of these relationships is not yet reflected by integrated policies for environmental planning. However, governments are beginning to move in this direction as they develop national environmental policies within the framework of their own developmental needs and perceptions about the quality of life.

This paper focusses on the United Nations Environment Programme (UNEP), located in Nairobi, which endeavors to coordinate national planning for environmental priorities. UNEP was established by the General Assembly on December 15, 1972. The Governing Council of 58 national representatives is responsible for the coordination of the manifold on-going environmental programs in the United Nations System (e.g. FAO's Soil Data Processing System, the International Geological Correlation Program, the Intergovernmental Oceanographic Commission, UNESCO's Man and the Biosphere Program, etc.). UNEP's role is basically a catalytic one. It establishes a systematic framework, draws attention to problem areas and urges collaboration between governments. The Environmental Fund makes a modest $20 million available for programs yearly, most of which are carried out by other UN agencies and national governments. Examples of problem areas include: human settlements, arid land and grazing land ecosystems, rational uses of national resources and substitutions, oceans, energy, natural disasters and the law of shared natural resources.[2]

All programs of UNEP emphasize the need for a keener understanding of environmental relationships and consequences through the utilization of available knowledge and the generation of new information. Geoscience data, along with information produced and disseminated by other multidisciplinary fields, play a vital role in the International Referral System (IRS) sponsored by UNEP. The IRS was developed as part of Earthwatch to assure that the knowledge of experts and the resources of data systems are available to governments in their environmental problem solving. Currently, the IRS can be used by governments only, but the coordinators of the system are committed to the notion that eventually the non-government sectors should have access to this vast resource. In the United States the Environmental Protection Agency acts as the IRS focal point. A federal Inter-Agency Committee on the IRS acts as advisor and coordinator.

This global network of national focal points of environmental data and information is conceived as a mechanism to interlink information users with sources by a central computer-based directory of addresses. A user, upon turning to the IRS system with a question, would be given the addresses of sources which would most likely supply the answer.[3] Thus, the concept of information referral, already familiar to the information community in other contexts, is central to the world-wide planning of UNEP.

Since each member nation of UNEP is responsible for designating a national focal point (about forty-four had been established by the summer of 1975), and for inventorying environmental data resources, various professions and disciplines in each country have a unique opportunity to critically review their priorities in planning data and information systems. An analysis of national information policies would be in order to identify the points of leverage that can optimize the processes (e.g. planning, systems design, evaluation and education) that are needed to support environmental research and programs. Questions should be faced, for instance: Are various institutions-government, the information industry, agencies in the private sector- capable and ready to accept information referral responsibilities? Will governmental funds go into program development only or also into research that would identify problems related to referral? Does the IRS hold real potential for information users in the private sector? What do we know about the individual expert as an information source and his role in a global network? Meetings of professional organizations can serve as forums for the discussion of such questions that need the attention of researchers. A review of information policies may be seen against the background of two trends that converged in the 1970's to create a new role for the professions: (1) the development of national environmental policies, and (2) participation of the nongovernmental sectors in the policy making process.

Environmental Policies

Whether at the local, regional, national or international levels, policies may be of two kinds: (a) those made in response to a suddenly felt need or crisis; (b) those that result from systematic long-range planning. Traditionally, governments have followed an uncoordinated and piecemeal approach to environmental problems. Conservationists, long aware of the need to protect finite natural resources, had their own battles to fight. In the United States, for instance, we have seen separate legislative achievements attained by the conservation movement and by its relatively newer ally, the anti-pollution movement. The enactment of the National Environmental Policy Act of 1969 was the first sign of a unifying effort for national environmental policy formation. In the 1970's legislative attempts to establish a National Environmental Policy Institute (92 Congress, (Senate) S.1216) and a Council on Energy Policy (93d Congress, (Senate) S.70) signalled intensifying public attention given to priorities. However, the prevailing dichotomy of the environmental and energy issues as perceived by many in government and the private sector, still impedes a true comprehensive approach.

The theme of environmental planning winds itself through the conference documentation of the 1970's. At a Symposium of the United Nations Economic Commission for Europe in 1971, it was agreed that, "...the main task for governments..[was] to attain an optimal allocation of resources, including common property resources, such as air, water, etc., for the various alternative uses..."[4] In 1972, the United Nations Conference on the Human Environment in Stockholm expanded the policy concern to the global level. In their pre-conference report, prepared with the assistance of a multi-national committee, Barbara Ward and René Dubos insisted that, "...a deeper and more widely shared knowledge of our environmental unity" must transcend nationalistic and local concerns.[5]

Some of the most compelling policy issues facing any joint effort by governments were discussed in 1973, at the Ninth World

Congress of the International Political Science Association. The questions raised indicated that the role of information was seen as a crucial one:

-What systems are developing for storing and retrieving data applicable to the measuring of the quality of life?

-How do environmental attitudes of the Third World countries differ from those of industrialized countries?

-What is the relationship of economic development and international technology transfer?

The political implications of information policies were identified: "...information is a crucial variable in any system's guidance capacity. As long as the political system has no information lead -- i.e....devices for prewarning, priority-setting, output control, and so on, -- it will react asthmatically."[6]

Scientists have perceived the significance of integrated policies governing vast quantities of research data and systems for their storage, retrieval and utilization long before other disciplines turned their attention to information as a powerful resource. This long-standing concern was reflected by Andre Hubaux: "The very future of the geological profession may depend on its capability to set a policy for the presentation to proper authorities of the basic geological data to estimate the reserves of commodities."[7]

Participation of the Non-governmental Sector

Environmental organizations at the national and grass roots levels have moved into political activism with increasing force.

Their impact may be traced to a combination of critical studies of environmental problems and government policy, scientific and legal information interpreted and packaged for the public, and mobilization of citizen participation relative to pending legislation. These organizations have been accepted as a social feedback mechanism by governments because they represent a wide array of disciplines. For instance, professional groups concerned with environmental education are becoming visible at national and international conferences and the mass media. Professionals concerned with environmental information have not yet gained such visibility. At the Stockholm Conference in 1972, only the International Federation for Documentation (FID) and the Special Libraries Association (SLA) were represented. The latter, through its Special Committee on Environmental Information, has undertaken during the past two years the development of some models for the integrated environmental information activities of a professional group. Since professional groups are "non-governmental organizations" (NGOs), I will briefly discuss UNEP-NGO relations.

The United Nations and the specialized agencies have, over the years, evolved an intricate structure of relationships with non-governmental organizations (NGOs). The very definition of NGOs has been a elusive matter. Sometimes they have been described as voluntary non-profit organizations, professional, humanitarian or political in nature. It has been observed that "no clear-cut typology of such organizations on the basis of legal or administrative criteria is possible, since legislation relating to associations varies from country to country..."

Traditionally, there have been two kinds of NGOs recognized by the U.N. agencies: (a) international NGOs with consultative status, accredited by the U. N. Economic and Social Council's Committee on Non-governmental Organizations, and (b) NGOs registered with the U.N. Office of Public Information (OPI) as observers. The process of accreditation is long, involved and inflexible. The NGOs approach to environmental problems and, consequently, to UNEP's work is innovative, politically aggressive and impatient. They are guided by the notion that UNEP's very existence is a result of the Stockholm conference which was characterized by an unusual and dramatic participation of NGOs - international, national and local - from most parts of the world. After Stockholm, these organizations expected a continued interactive relationship with UNEP and a liberalization of conventional requirements for accreditation. As one critic of the system expressed it, "It really is quite extraordinary that in this time of increasing social crisis there should be this U.N. impulse to say 'that body is not good enough to help.'"[9] Although the official status of NGOs clustering around UNEP has never been formally clarified, based on the "open policy" of its executive director, Maurice Strong, UNEP facilitated several unprecedented arrangements that brought NGOs into direct interaction with the organization.

Many of the environment-oriented NGOs are public interest groups. They have become sophisticated in the use of well-marshalled information as a source of political participation.[10] In spite of perennial frustrations with bureaucratic hurdles, they continue to be alert to the policy-setting opportunities of UNEP-related meetings, and they systematically inform their membership of UNEP-sponsored studies and official documentation. In the following I will summarize information on some of the current channels for NGO participation which are utilized by a number of professional organizations:

1. An Environment Liaison Board (ELB), established by the International Assembly of Non-Governmental Organizations Concerned with the Environment in 1973 and reconstituted in 1975, acts as a facilitator for cooperative work by NGOs, develops NGO profiles for a directory and is in charge of organizing NGO assemblies. (Address: c/o International Council of Voluntary Agencies, 7 Avenue de la Paix, 1202 Geneva, Switzerland).

2. The ELB directs an NGO Environment Service Centre in Nairobi. This center serves as a clearinghouse for information and as contact point for NGOs. Services are based on regular annual contributions by NGOs. (Address: P. O. Box 72461, Nairobi, Kenya).

3. In the United States, several umbrella organizations of NGOs provide contact points, and assistance in developing strategies and recommendations for international meetings. Two examples are:

Environment Forum, Inc.
3021 Cambridge Place, N.W.
Washington, D.C. 20007

North American Committee of NGOs
Concerned with the Human Environment
c/o Community Development Foundation
345 East 46th Street
New York, N.Y. 10017

4. NGOs, whether formally accredited by ECOSOC or not, have been invited to send observers to a number of international conferences. The degree of "open policy" for participation fluctuates, dependent on

the various organizing committees. The mechanism of participation has been the NGO "assembly", "tribune", or "forum", held immediately before or simultaneously with the official sessions of delegates. Such assemblies were held in relation to UNEP in Geneva (1973) and Nairobi (1974 and 1975), The Law of the Sea Conference, Caracas (1974), The International Population Conference, Bucharest (1974), the World Food Conference, Rome (1974), and the World Conference of the International Women's Year, Mexico City (1975). NGOs interested in HABITAT: The United Nations Conference on Human Settlements to be held May 31 to June 11, 1976, in Vancouver, and the related NGO meeting to be called Habitat Forum, may contact the NGO Committee for Habitat (c/o ACSOH, Box 48360, Bentall Postal Station, Vancouver, B.C. Canada.)[11]

5. "Hotline International" makes long-distance participation in international conferences possible. Recommendations of NGOs were transmitted to the NGO linkage centers at the conferences in Nairobi and Mexico City by an IBM 370/168 computer, operated by the National CSS Corporation with the Community Development Foundation. Access to this computer via leased lines was available from many locations in the United States and Canada. For instance, the University of Hawaii transmitted to UNEP in Nairobi technical advice related to research on non-polluting energy sources. The system also stores data on the activities and programs of NGOs and coming conferences. (54 Riverside Drive, New York, N. Y. 10024).

6. The Center for International Environment Information is a non-governmental service partially sponsored by UNEP and established by the United Nations Association of the United States of America.

International environmental news and announcements are carried by the bi-weekly World Environment Report. (Address: 345 East 46th Street, New York, N.Y. 10017.)

A Potential Role of the GEoscience Information Society

In a remarkable article on geological data systems and their implications nationally and internationally, Daniel F. Merriam focused on a number of interesting policy questions.[12] For example, he suggested establishment of a U.S. Index of Geoscience Data that would include computer-based files and unpublished open-file reports in addition to traditional forms of literature. Perceptions such as this should become the basis of systematic policy formulation concerning scientific data and information. National resources such as the Index of Geoscience Data could then be linked with other resources in UNEP's International Referral System.

The Geoscience Information Society may consider the following approaches:

1. The society could explore its potential role as an "NGO" having specialized expert knowledge in the transfer of geoscience data and information. A committee or task force could investigate possible linkages with UNEP and ways to contribute to environmental information policy formation.

2. Geoscientists and information specialists in government, industry and the non-governmental sector, should develop stronger links of communications with each other in order to discuss policy alternatives concerning data resources. Regional workshops could be jointly sponsored by various institutions to facilitate

a) the on-going review of current policies at the national and international levels and the potential impact of the International Referral System (IRS) on the coordination of geoscience data systems, the standization of data recording and the exchange of information,

b) recommendations to appropriate policy setting groups in the public and private sectors,

c) dissemination of information to the scientific community.

Specific professional areas which merit consideration by a policy-oriented scientific group include:

a) representation of the geoscience disciplines and their data needs in the planning process of UNEP and the further development of the national components of the International Referral System (IRS),

b) contacts with the U.S. Federal Inter-Agency Committee on the IRS and the national focal point, the Environmental Protection Agency, to assure that information about new IRS developments will reach the membership of professional societies,

c) recommendations to appropriate funding sources that funds should be allocated for research on innovative information sharing mechanisms such as referral and for seminars and workshops in relevant data retrieval and utilization techniques,

d) systematic attention to the educational implications of the relationship between geoscience data resources and information policies.

References

1. United Nations Environment Programme. "Action plan for the Human Environment: Programme Development and Priorities, Report of the Executive Director," Geneva, April 2, 1973, UNEP/GC/5, §10.

2. United Nations Environment Programme, Governing Council. "Proposed Fund Programme Activities, Revised 1975, Proposed 1976-1977. Note by the Executive Director, Nairobi, March 7, 1975 UNEP/GC/32.

3. United Nations Environment Programme. "The International Referral System for Sources of Environmental Information (IRS). Report by the Executive Director," Nairobi, Feb. 24, 1975, UNEP/GC/31/Add. 3.

4. United Nations. Economic Commission for Europe/ Symposium on Problems Relating to Environment: Proceedings and Documentation, New York, United Nations, 1971. p. 52 (ST/ECE/ENV/1).

5. Ward, Barbara and René Dubos/ Only One Earth. New York, W. W. Norton, 1972. p. 213.

6. Zapf, Wolfgang/ The Policy as a Monitor of the Quality of Life. American Behavioral Scientist, 17 (No. 5); 652 (May/Je 1974).

7. Hubaux, Andre, "A New Geological Tool - the Data," Earth Science Review, 9 (2): 160, 1973. Cited in Merriam, D. F., "Resource and Environmental Data Analysis," in Earth Science in the Public Service, Symposium presented during dedication ceremonies, U. S. Geological Survey National Center, Reston, Virginia, July 10-13, 1974. Washington, D.C. G.P.O. 1974, pp. 37-45.

8. United Nations, Economic Commission for Europe/ op. cit. p. 317.

9. Anthony Judge/ Mobilization for Alienation vs. Catalysis for Participation, the Critical Choices for the United Nations System, International Associations 25 (No. 8/9): 410 (1973).

10. Henderson, Hazel / Information and the New Movements of Citizen Participation, American Academy of Political and Social Science Annals, 412:34 (March 1974).

11. "The Involvement of NGOs in the UN Conference/Exposition on Human Settlements," International Associations, 26 (No. 6/7): 358-359 (1974).

12. Merriam, D. F. op.cit.

A Comparison of Journal Articles in Environmental and Laboratory Science

Raymond E. Durrance
School of Library Science
University of Michigan
Ann Arbor, Michigan

INTRODUCTION

The purpose of this investigation is, through journal article analysis, to statistically compare the characteristics of the literature of representative scientific disciplines in Environmental and Laboratory Science. In recent years increased emphasis upon interdisciplinary activity in science has revealed a need for better bibliographic control of interdisciplinary literatures. A review of research in this area shows that for several decades citation analysis has been used to measure the characteristics of literature used by scientists with the result that we have a better understanding of the bibliographic structures of the literatures of individual disciplines. However, previous investigators have not compared interdisciplinary groups of literatures such as Environmental Science to Laboratory Science, or Life Science to Physical Science. (Stevens 1949) Perhaps more importantly, only a few earlier researchers report statistically tested comparisons between the characteristics of two or more disciplines. Fussler (1949) is a notable exception.

Statistical validation is desirable since substantive apparent differences are reported between characteristics of the literatures of the various scientific disciplines that have been examined. For example, there seems to be a striking 55% difference in the age of literature used by geologists and that used by physicists: Laosunthara (1956) reported that 67% of the literature used by geologists was more than

10 years old while Fussler (1949) found that only 12% of the literature used by physicists fell within that age span. Similarly, according to Craig (1969) journals account for about 77% of the literature used by geologists, while according to Brown (1956) journals account for 94% of the literature used by chemists. It seems probable that if such marked differences between the literatures used by scientists could be statistically validated, then this knowledge could be exploited with greater confidence in improving the design and management of information retrieval systems intended to meet their needs. However, dissimilarities in the methods used tend to hamper post hoc statistical comparisons between the various investigations. Citation analysis, the basis for most of these previous studies, usually has been conducted in one of two ways; (1) either by randomly selecting individual citations for analysis from large populations of citations in source journals; (2) or by pooling and analyzing all the citations within a smaller sample of journal articles. The first method does not hinder statistical analysis but it does destroy the integrity of the selected lists of references chosen by the authors, while the second method does hinder statistical analysis since it ignores the statistical interdependence between the elements of these lists of citations.

METHOD

With the above considerations in mind a method was devised for this investigation to statistically compare characteristics of journal articles of the interdisciplinary groups, Environmental Science, and Laboratory Science. The journal article was chosen as a basic unit of analysis in preference to the citation for three reasons:

(1) The journal article is the single most important bibliogra-

phic unit in scientific literature. It serves the scientist as a vehicle for reporting his research and the information specialists as the primary unit for indexing.

(2) It permits expansion of traditional citation analysis by including more characteristics than would be possible through citation analysis alone. This is especially true of those characteristics representing the physical dimensions of the articles and their texts.

(3) Analysis of the entire article preserves the integrity of the lists of references selected by the authors and at the same time provides a means for dealing with the statistical inter-dependence of the references.

The profile of characteristics analyzed for each journal article involved the measurement of the physical dimensions and the text of the article as well as citations to other materials.

The profile is as follows:

(1) The number of authors per article;

(2) The number of references per article;

(3) The number of pages per article;

(4) The language of the article;

(5) The median age of material cited by each article;

(6) The percentage of journals cited by each article, or the journal citation rate; and

(7) The percentage of unique journal titles cited in each article, or journal title distribution.

A stratified random sample (n=600) of journal articles was selected from Biological Abstracts; Meteorological and Geoastrophysical Abstracts; Deep Sea Research Oceanographic Abstracts; Bibliography and Index of Geology; Science Abstracts, Part A Physics Abstracts; and Chemical

Abstracts. Descriptive statistics were compiled for each of the seven variables; authors, pages, references, language, age, journal citation rate, and journal title dispersion, for the inter-disciplinary groups: Environmental Science, Laboratory Science, Life Science and Physical Science, and for each of their selected representative component disciplines: ecology, meteorology, physical oceanography, geophysics, biochemistry, genetics, chemistry and physics. Then, using the factorial design shown in Figure 1, a Mann Whitney U-test procedure (Conover 1971) was used to test hypotheses ($p < 0.05$) regarding differences between the characteristics of the test groups, Environmental and Laboratory Science and between the control groups, Life and Physical Science.

RESULTS

It may be seen from Table 1 that for four of the seven variables, authors, pages, language, and journal citation rate, significantly greater differences were found between the test groups than between the control groups. There were no statistically greater differences between the test and control groups in the number of references per article, the age of cited material or in journal title dispersion. Examination of each component discipline is also instructive. Figures 2 through 8 are bargraphs showing the observed mean values within 95% confidence intervals for each of the seven variables and for each of the eight separate disciplines. By comparing these to Table 1 it can be seen that marked differences in the values for characteristics of the articles of the component disciplines often were masked by blending them into the interdisciplinary groups and also that there appears to be considerable difference in variance from one discipline

TABLE 1

Statistical Comparison of Journal Article

Profiles in Environmental and Laboratory Science

	Median Age of Cited Material in Years	Number of Authors per Article	Number of Pages per Article	Number of References per Article	Percentage of English Language Articles	Journal Citation Rate	Journal Title Dispersion
Environmental Science	8.2	1.66	12.6	21.3	74	73.5	67.9
Laboratory Science	6.0	2.23	7.6	16.8	84	82.2	62.6
Difference	2.2	0.57	5.0	4.5	10	8.7	5.3
Life Science	7.9	1.92	10.4	21.6	82	80.0	68.5
Physical Science	6.4	1.97	9.8	16.5	76	75.8	61.9
Difference	1.5	.05	0.6	5.1	6	4.2	6.6
Test Groups Show Statistically Greater Difference than Control Groups	No	Yes	Yes	No	Yes	Yes	No

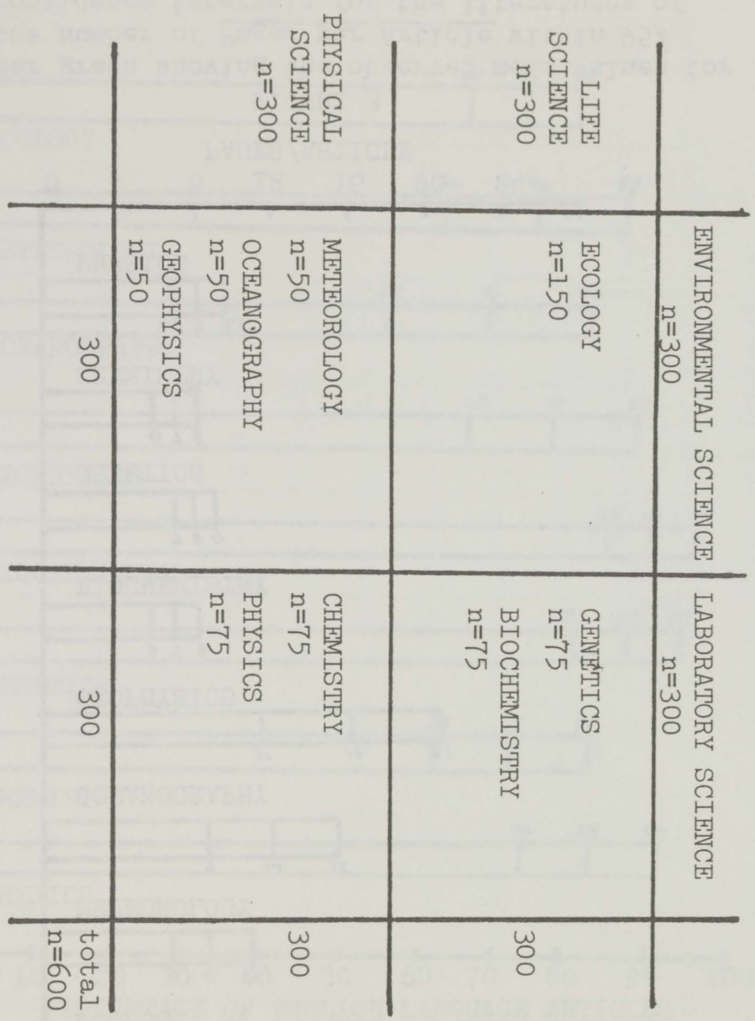


FIGURE 1

Diagram showing the factorial design used for testing for differences between characteristics of eight literatures representing the test groups Environmental Science and Laboratory Science and the control groups Life Science and Physical Science.

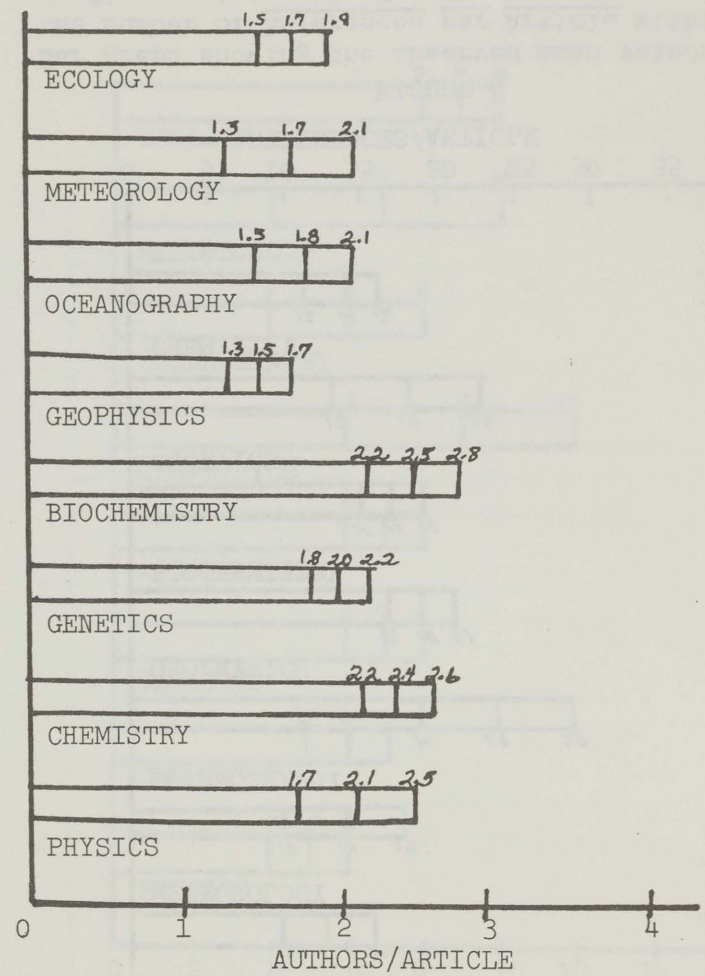


FIGURE 2

Bar graph showing the observed mean number of Authors per Article within the 95% confidence intervals for the literatures of eight disciplines.

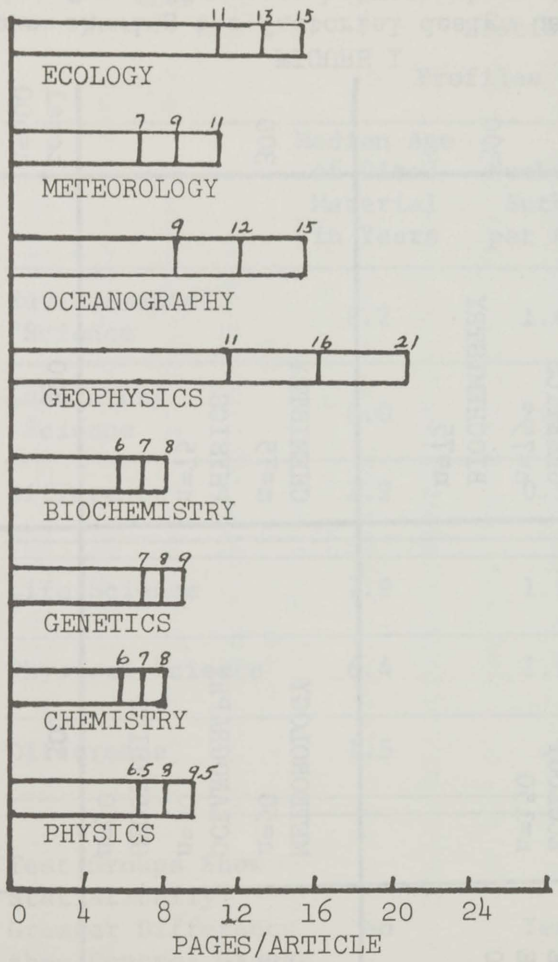


FIGURE 3

Bar graph showing the observed mean values for the number of Pages per Article within 95% confidence intervals for the literatures of eight disciplines.

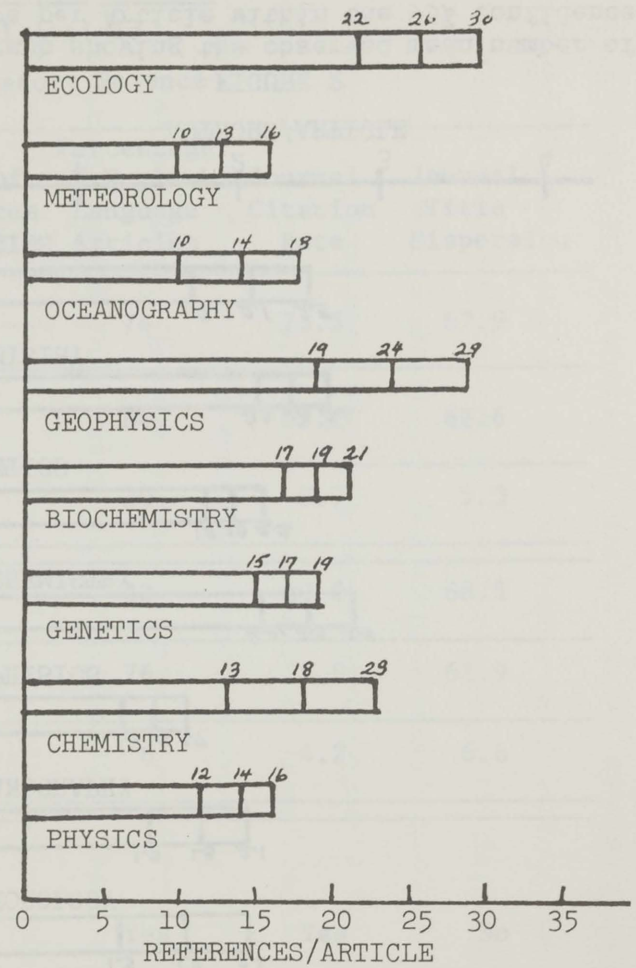


FIGURE 4

Bar graph showing the observed mean values for the number of References per Article within 95% confidence intervals for the literatures of eight disciplines.

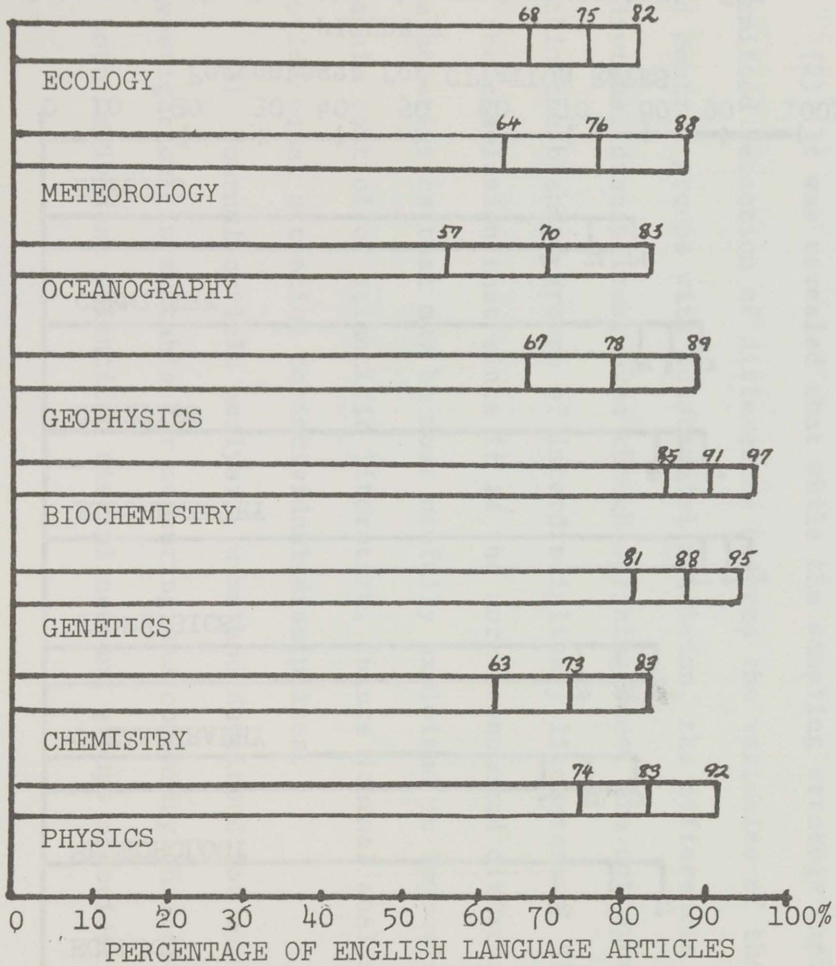


FIGURE 5

Bar graph showing the observed percentages of <u>English</u> <u>Language</u> articles within 95% confidence intervals for the literatures of eight disciplines.

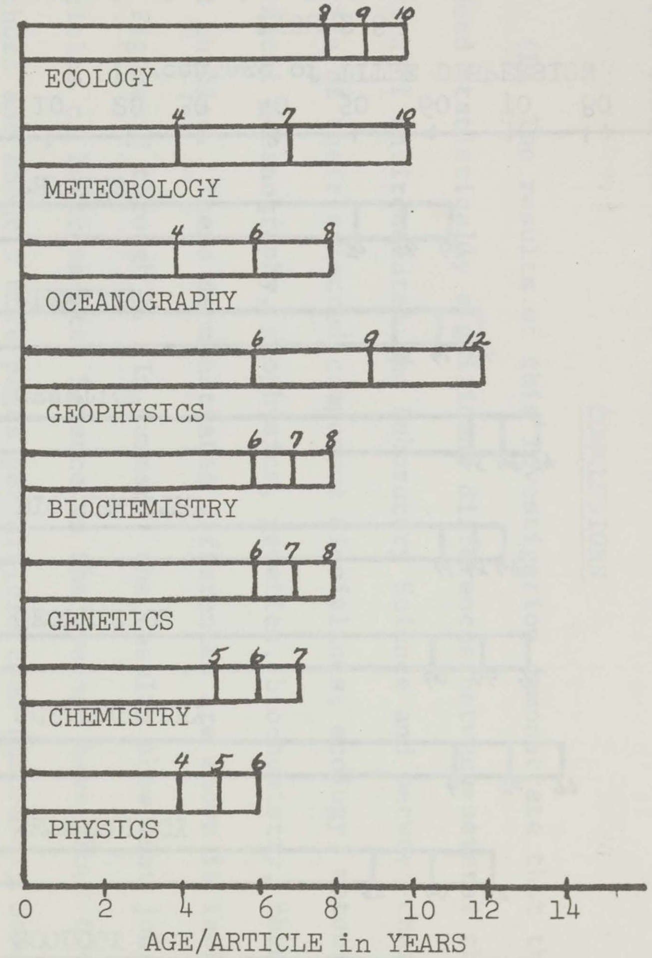


FIGURE 6

Bar graph showing the observed mean values for typical <u>Age</u> <u>per</u> <u>Article</u> within 95% confidence intervals for references cited in the literatures of eight disciplines.

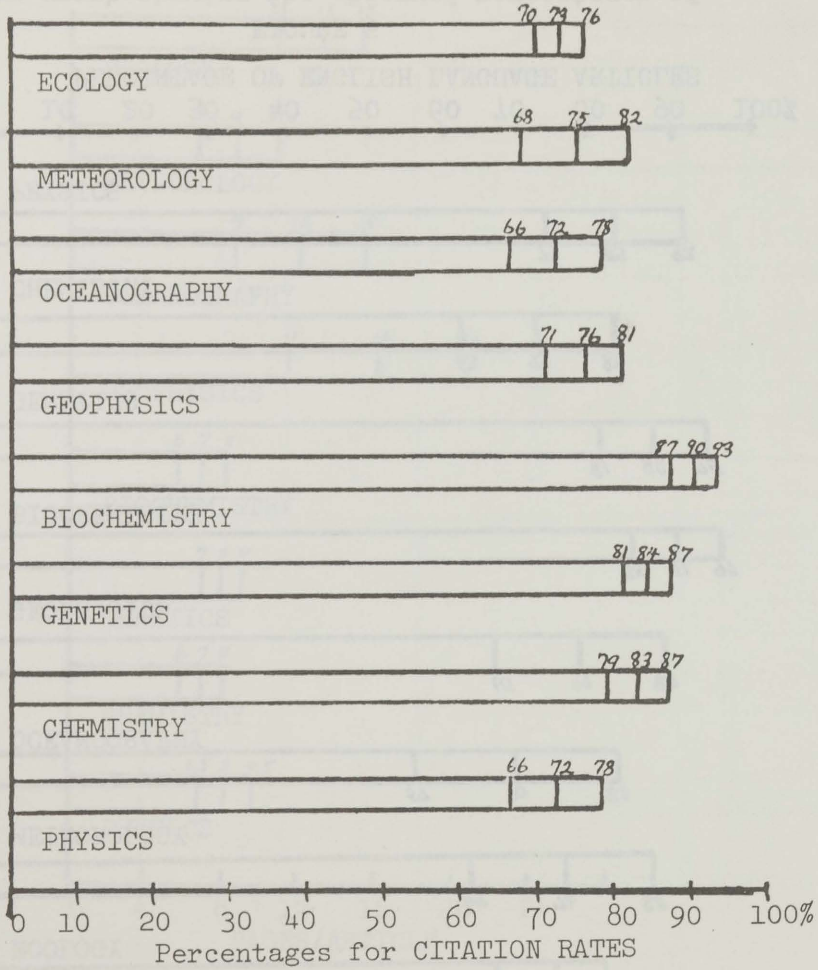


FIGURE 7

Bar graph showing the observed percentages for Journal Citation Rates per article within 95% confidence intervals for the literatures of eight disciplines.

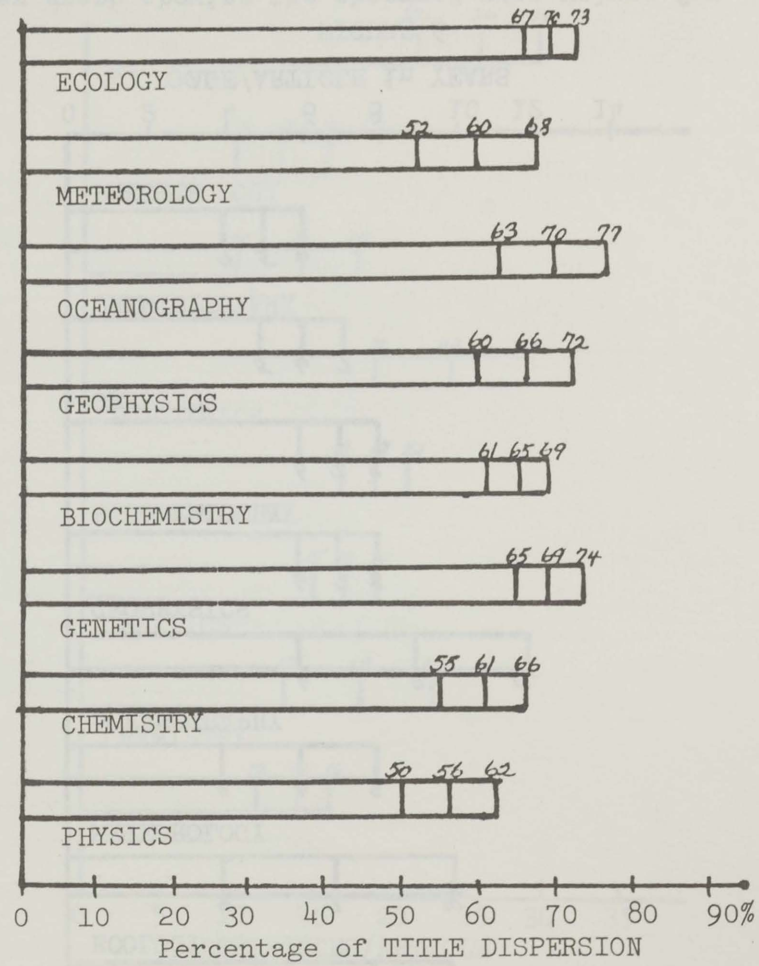


FIGURE 8

Bar graph showing the observed percentages for Journal Title Dispersion per article within 95% confidence intervals for the literatures of eight disciplines.

to the next and from one characteristic to the next.

CONCLUSIONS

(1) The results of this investigation demonstrate that there are indeed statistically significant differences between several characteristics of Environmental and Laboratory Science and between characteristics of their selected component disciplines; ecology, meteorology, physical oceanography, geophysics, genetics, biochemistry, chemistry, and physics. These deomonstrated differences are shown in Table 1 and in Figures 2 through 8. In summary, the results show that journal articles of Environmental Science on the average have about 0.6 fewer authors and about 5 more pages per article than Laboratory Science. They also show that authors of journal articles in Environmental Science cite journals about 9% less often and publish in non-English languages about 10% more often than those in Laboratory Science.

(2) It was revealed that while the sampling strategy employed permitted detection of differences between the variables of the test and control groups with substantial precision, the differences between individual disciplines, even though striking ones, are not apparent within the blended groups of interdisciplinary literatures. This leads to the conclusion that since it is the more pronounced differences in characteristics that may be most usefully exploited in improving bibliographic control of scientific literature, future journal analyses should provide equal attention to individual disciplines.

(3) Journal article analysis, a method developed for use in this investigation, is suitable for measuring and comparing characteristics of both individual scientific disciplines and groups of scientific

disciplines. It is superior to citation analysis for three basic reasons:

(a) It deals with the journal article, a bibliographic unit which is of primary importance.

(b) It maintains the integrity of the list of the references brought together by each author, and it facilitates statistical comparisons.

(c) It incorporates characteristics of the journal articles that cannot be included in traditional citation analysis alone.

BIBLIOGRAPHY

Brown, Charles H. Scientific Serials. ACRL Monograph No. 16. Chicago: American Library Association. Association of College and Reference Libraries, 1956.

Conover, W.J. Practical Nonparametric Statistics. New York: John Wiley and Sons, Inc., 1971.

Craig, J.E.G., Jr. "Characteristics of Use of Geology Literature." College and Research Libraries, 30:230-236 (1969)

Fussler, Herman H. "Characteristics of the Research Literature Used by Chemists and Physicists in the United States." Library Quarterly 19: 19-35 (1949)

Laosunthara, Maria. "Some Bibliographical Characteristics of Serial Literature in the Field of Geology." Unpublished Ph.D. dissertation, Indiana University, 1956.

Stevens, Rolland E. Characteristics of Subject Literatures. ACRL Monograph No. 9. Chicago: American Library Association. Association of College and Reference Libraries, 1953.

THE MINT SYSTEM: A SCHEME FOR ORGANIZING A SMALL MAP COLLECTION

Nancy Jones Pruett
Edwin Jay Foscue Map Library
Science/Engineering Library
Southern Methodist University
Dallas, Texas 75275

When you need a particular map that you've used before, do you end up rolling and unrolling all the ones you have? And do you still sometimes have to redo one or reorder it because you can't find it even though it has to be there somewhere? If so, you're not unusual. Maps, because they're cumbersome, don't fit easily into most company filing systems. They require a system all their own. But often no one has the time to create one, or the time to keep it up afterwards.

The MINT (Maps In No Time) System was designed for a collection of about 500 maps in a petroleum consultant's office. These consultants had tried marking the maps with state and county names but found they still could not find a particular map they wanted when they wanted it. Valuable time was being wasted as they rolled and unrolled maps looking for one they remembered. They needed a system by which they could quickly find one or all the maps on a particular area or subject and which would allow new maps to be included easily in the system.

The MINT System is a coordinate indexing scheme designed to do just that. It consists of three files: the maps, a set of Descriptive Cards and a set of Keyword Cards.

<u>The Maps</u>: Each map is assigned a number, which can be simply an accession number. For instance, the first map you pick up is assigned Map Number 1, the second is Map Number 2, etc. The maps are kept in order

by this number, which should be clearly written on all four corners of the map if it's to be rolled. If you must keep your maps rolled and not in the more convenient horizontal map cases, you can roll any number of maps together as long as you label the roll as to which numbers it contains.

<u>The Descriptive Cards</u>: For each map fill out a Descriptive Card (see Figure 1). This card has the Map Number on it and all the information about the map. Keep these in order by Map Number and if nothing else you can save time when you need to search by searching the Descriptive Cards, which are less cumbersome than the maps themselves.

<u>The Keyword Cards</u>: A Keyword Card is shown in Figure 2. A keyword for the subject or area of a map is written at the top. The numbers of every map which has that subject or covers that area are listed below it. For instance, if you wanted to know what maps you have on Winkler County, Texas, you could pull the Keyword Card shown in Figure 2 and find that maps 1, 29 and 302 do so. Now what? Either go pull maps 1, 29 and 302 or go to the Descriptive Cards and pull cards 1, 29 and 302 to help you decide which map you really want to see.

It is also possible to narrow down your search for a specific map by pulling another Keyword Card. For example, assume you want a structure map of Bee County, Texas. If you pull the BEE COUNTY, TEXAS Keyword Card (Figure 3) you will find there are 16 maps listed. All of them aren't structure maps. If you don't want to look through all 16 maps until you find a structure map, pull the STRUCTURE Keyword Card (Figure 4). Any map numbers on both the BEE COUNTY, TEXAS card and the STRUCTURE card (in this case Maps 119 and 298) represent maps which are both structure maps and of Bee County. So without unrolling a single map, you have narrowed your search from nearly 500 maps to 16 to 2. Practically "maps in no time", isn't it?

The procedure for entering a map into the MINT System consists of filling out a Descriptive Card completely for each map and then posting the number of the map to the appropriate Keyword Cards. With a little practice it can be done very quickly. The System also allows easy withdrawal and reorganization.

ENTERING A MAP IN THE SYSTEM

1) Fill out a Descriptive Card, beginning with the Map Number taken from the corner of the map. Fill in all the areas of the Descriptive Card.

AREA. Examples: Henderson County, Texas (always specify counties if there are not more than three); Montana; Gulf Coast; North Sea.

SUBJECT. Examples: Base map; structure; geology; well log; cross section; topography; bathymetry.

TITLE. This is the title as it is given on the map. Examples: Atlantic Ocean north of Cape Hatteras; Jacko Bay, Louisiana; Petroleum Exploration and Development Map--Permian Basin.

DATE. Fill in the month and year or just the year, whatever is given on the map.

SCALE. Horizontal scale can be indicated in any convenient way. Examples: 1" = 4000'; 1:20,000; 1" = 2000 feet. Put contour intervals and other indications of vertical scale in the form of notes in the SERIES area of the card.

JOB NUMBER. This is a space for an internal job number.

PUBLISHER. This is the authority most responsible for a map's existence. Examples: USGS (U.S. Geological Survey); C&GS (U.S. Coast and Geodetic Survey, now the National Ocean Survey); Heydrick; Midland Map Company.

SERIES. Examples: USGS 7.5' topographic series; USGS 1:250,000 series; Bathymetric maps, Eastern Continental margin, U.S.A., Sheet 3 of 3; C&GS 1274.

Describing the series serves two purposes. First, maps in certain series are very similar (e.g. the 7.5' quadrangles) and stating that it is in that series is a quick way of giving information about it. Secondly, sometimes a map will be referred to by a series number or description without its title (e.g. the Coast and Geodetic Survey Chart 1274). Entries for series could be placed in the Keyword Card file at any time as long as the information is recorded on the Descriptive Card.

SIZE. Size categories serve two purposes. First, they help you know what kind of map you're looking for when you go to the map rolls. Second, it is easier to roll maps that are approximately the same size together.

TYPE. "Col." means a published or printed map like those in the USGS topographic series. "Bl." indicates a blue-line map or one which has been copied from an original but is not easily reproducible itself. "Orig." indicates a map or drawing on acetate or sepia which could be reproduced.

OVER. Circle the word OVER if you make any notes on the reverse of the card.

KEYWORDS. In this section of the card enter the keywords with which the map is associated. You will use this section to pull appropriate Keyword Cards and post the map numbers onto them. The form of the keywords should be correct. For instance, under AREA you may have written "Texas--Henderson County." But if you look in the Keyword Card file you will find the card is actually entitled HENDERSON COUNTY, TEXAS. HENDERSON COUNTY, TEXAS is what should appear in the KEYWORD section of the Descriptive Card.

Most often the keywords to be posted are the AREA and SUBJECT as written at the top of the Descriptive Card. If the form is correct, you may simply write "A" for Area and "S" for Subject in the KEYWORDS space, and just look at the top of the card when it is time to post.

Some thought must be given ahead of time to what keywords to use in the Keyword Card file. If the first time you have a map of Dallas you put it under TEXAS--DALLAS COUNTY and the next one you put under DALLAS COUNTY, TEXAS, when you go to the file under TEXAS--DALLAS COUNTY you will find one map and miss the other. There are two ways to minimize the possibility of this kind of mistake. One is to make a basic decision at the beginning on the form of area entries. Either pick what corresponds to your regular filing system or pick what makes sense for the overall composition of your collection. But whatever you decide, stick with it. Another way is to be alert anytime you add a new Keyword Card--what cross references should be put in? What other keyword might someone look for this map under?

<u>2) Post the Map Number onto the appropriate Keyword Cards.</u>

Take a Descriptive Card and look at the first KEYWORD at the bottom. Pull the corresponding Keyword Card from the file, list the map number on it and refile. Back on the Descriptive Card, put a check in front of each keyword you post so you will know when you look at any Descriptive Card that the keywords have been posted.

If one of the keywords you've decided on is not already in the file, make a new Keyword Card for it. But check first to make sure it isn't there in some other form.

Also make any cross-references that seem necessary. For instance, if you decide to enter everything on Phoenix Field under the county and not in both places, put a card in the Keyword Card file that says, "PHOENIX FIELD see MARTIN COUNTY, TEXAS."

Or if you make a decision to use a certain keyword in the Keyword Card file it is helpful to make cross references from all the other places people might look. For instance, if you decide to use ATLANTIC COAST for maps of the east coast of the U.S., make a card saying "EAST COAST, U.S. see ATLANTIC COAST, U.S."

Any notes clarifying what the keywords refer to can also go right on the card. For instance, GULF COAST is used for maps of the Gulf of Mexico covering more than two states. A map covering the coast of only one or two states is entered under "TEXAS COAST" or "FLORIDA COAST". So there is a note on the GULF COAST Keyword Card: "(covers more than 2 states); see also (STATE) COAST."

3) Notes.

With the USGS topographic series, always make a Keyword Card for the title (the name of the quadrangle). Also make entries for the relevant counties if possible.

It is possible and sensible to give a particular series of maps one number and note on the descriptive card that it is a series and where the index is located.

WITHDRAWING A MAP FROM THE COLLECTION

1) Pull the Descriptive Card that goes with the map.

2) Look at the bottom to see what Keyword Cards the map number was recorded on. Pull each of those cards and cross off the map number.

3) Destroy the Descriptive Card and the map.

The map number may now be assigned to another map and it is useful to keep a note in the front of the file of numbers which are available.

REORGANIZING:

If at some time it seems desirable to bring together physically maps which are separated from one another this can be easily done by assigning them new numbers closer together. One must be very careful in the record-keeping, however.

The first step is to do the first two things listed above for withdrawing a map: pull the Descriptive Card and cancel the map number from all Keyword Cards. Then, instead of destroying the map and the Descriptive Card, simply cross out or erase the old number and assign a new one. Then post the new number to the appropriate Keyword Cards. Be careful to do this, since the keywords at the bottom will already have checks in front of them from from having been posted before and it is easy to forget. If the new number is not posted there will be no way to find the maps, except to flip through all the Descriptive Cards--always a last resort.

SUMMARY

The MINT System is a coordinate indexing system for organizing a small map collection. It is simple to enter maps in the system, easy to find them once they are entered and possible to remove them or reorganize them.

FIGURE 1 A DESCRIPTIVE CARD

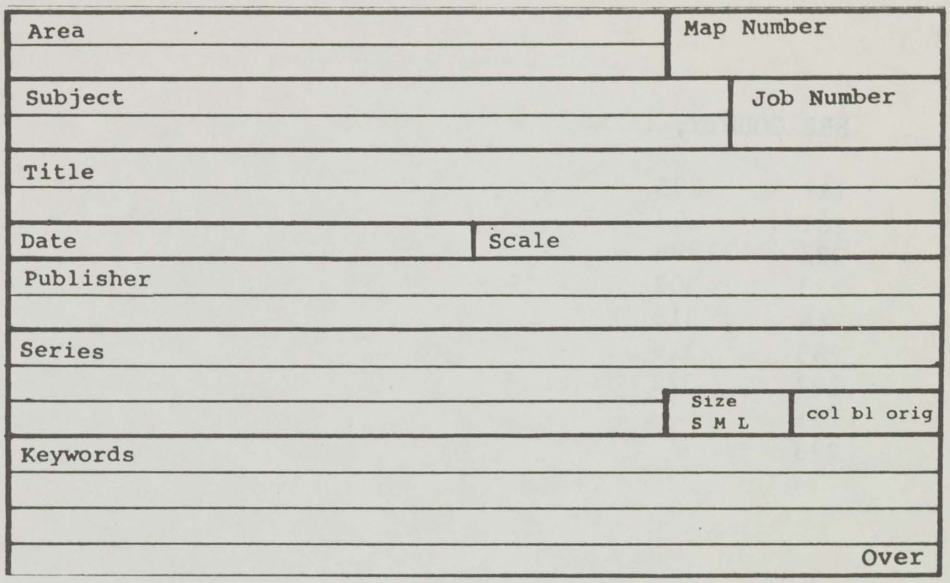


FIGURE 2 A KEYWORD CARD

WINKLER COUNTY, TX.

1
29
302

FIGURE 3

BEE COUNTY, TEXAS

119	295
281	297
282	298
283	307
284	318
285	319
287	322
291	
293	

FIGURE 4

STRUCTURE

11	79	135	271	302
21	80	176	273	303
22	112	219	276	306
23	→119	241	277	312
45	120	242	285	
51	131	243	287	
66	132	246	296	
77	133	248	→298	
78	134	250	299	