Dagobert Soergel
University of Maryland, USA

An Automated Encyclopedia — a Solution of the Information Problem? (Part I, Sections 1–4)

Soergel, D.: An **automated encyclopedia – a solution of the information problem?** (Pt.1, Sect. 1–4)
In: Intern. Classificat. 4 (1977) No. 1, p. 4–10

Due to redundancy and lack of specific access, information transfer through the literature is costly in time to the user; it is also costly to the author who must repeat much context information before coming to his new contribution. A non-redundant data store or automated encyclopedia could alleviate these problems. The structure of such a data store and procedures for its creation from the literature by controlled removal of redundancy are described. However, an automated encyclopedia would create problems of its own. Its structure would be enormously complex and it would be very expensive to establish. Its acceptance by users and authors is by no means a given. Pilot projects in high-use areas such as medicine and drugs, or statistical methods, and in numerical data compilation are suggested to test the feasibility of an automated encyclopedia and determine a realistic scope for it. (Author)

Outline:

1. Introduction: Improving the formal information transfer system

1.1 The problem: shortcomings of the present system

Fig. 1 shows a model of the formal information transfer system. At present emphasis in the formal information transfer system is on the left half of the diagram, that is, information transfer through literature. Reference storage and retrieval systems provide the researcher with references to documents, either through SDI or through answering retrospective search requests. The researcher then obtains the documents themselves from the same or another institution. The right half of the diagram is much less developed. There are a few data stores (information analysis centers; social science data archives in a limited way) and data publications, but they are almost exclusively concerned with numerical data. Moreover, the data to be incorporated are taken from the literature and there are no direct contributions to the data store. Note that the term "data" is used in a broad sense, to include not only empirical data, but also theorems, hypotheses, speculations, methods, etc. However, retrieval and dissemination of references and documents solve only half of the problems of the researcher facing the literature, and probably the less difficult ones. This is aptly put by Bohnert and Kochen als follows:

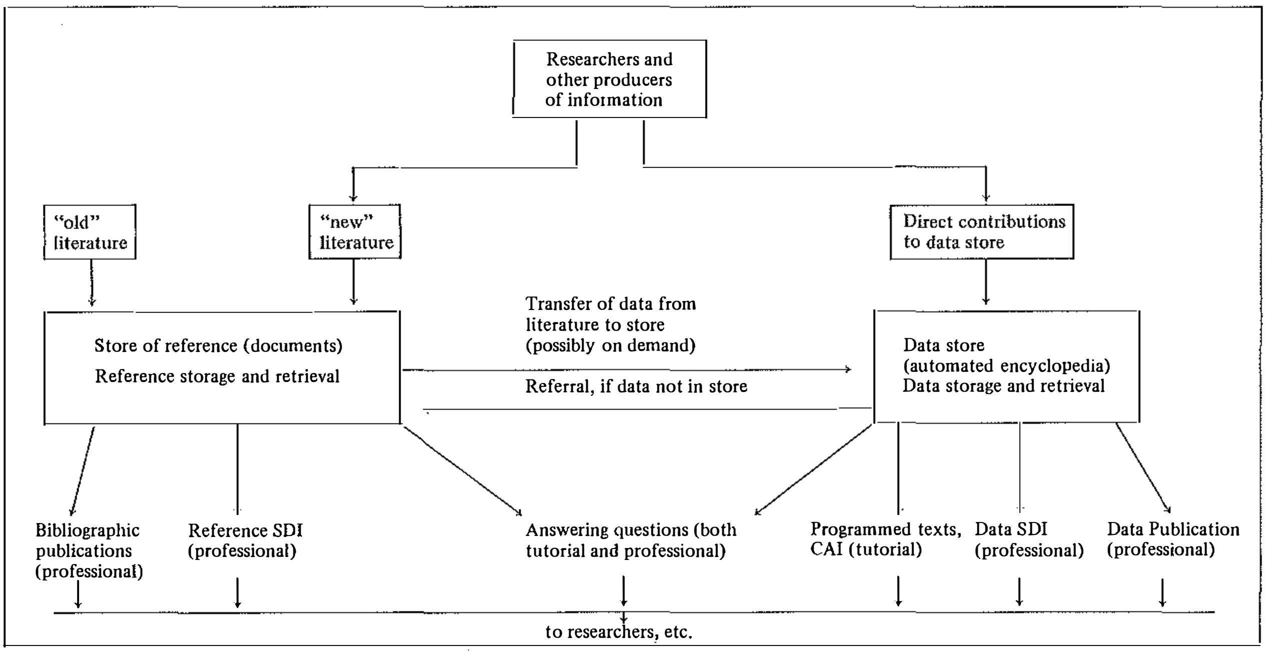

Fig. 1: The Formal information transfer system

"Most of the 'information crisis' literature stresses the difficulty of access to the growing flood of technical literature. But the problem that more often confronts the researcher is that of absorbing the mountain of material already accessible. An increase in accessibility without a corresponding increase in human assimilation rate will be self-defeating. Short of discovering a harmless superbenzedrine, the only way such a thing can be done is by making better use of the natural rate. At present this natural rate is wastefully expended, not only in literature search (the target of most current documentation efforts), but also in wading through series of "self-contained," hence highly overlapping but terminologically idiosyncratic, articles in order to build a consistent, single picture of the state of development in a given field. Often one needs to know only a central idea, result, theorem, or the methods employed, with bibliographical information for later reference, but finds that a short course in unneeded detail is required to get to it." (Bohnert et al 1963. 10, p. 269)

We can analyse this problem into its components: having obtained all relevant documents the researcher still faces the following problems:

(1) *Scatter* – In order to assemble the information necessary for the solution of a problem, it is often necessary to put together bits and pieces from a large number of documents, each of which contains much information not relevant to the problem.

For example, one paper may have raised a question, another paper may have formulated a hypothesis, a third paper reports on an empirical study that, among other things, confirms the hypothesis, but the actual empirical data of that study appear in an unpublished report, and then appears a fourth paper reporting on an empirical study and concluding that the hypothesis should be rejected.

(2) *Redundancy* – The same information element is repeated in many documents. (This is partly a consequence of scatter.)

(3) *Lack of specific access* – In order to retrieve individual data elements, one has to look through much irrelevant stuff.

(4) *Lack of needed detail* – Details of data collection procedures, experimental apparatus, and data collected are often omitted from published documents.

(5) *Lack of coherent framework* – By use of a coherent framework it is often possible to present the same information in a much shorter and more easily comprehensible way. For example, a law in physics enables one to represent a huge mass of empirical data by that law and a small number of base data. In other fields, such as meteorology or in most of the social sciences, the situation is more difficult, but still a reduction in the amount of data being stored could be achieved.

(6) *Time lag* – It takes a long time for new data to be published in journals, let alone textbooks or handbooks. Partly this is due to the simple fact that it takes time to prepare a publication, and partly to the fact that new editions of textbooks or handbooks can only be issued in fairly large intervals.

(7) *There is much bad material* – Objective criteria for evaluation must be developed and applied consistently.

In brief, accessibility of information is low. Furthermore, information density is low, i.e., the number of characters used to represent information in the present system is much higher than would be necessary in an ideal system, resulting in higher costs for storage, transfer, and assimilation of information. It follows that *one key element in any proposal to improve the information transfer system must be the removal of redundancy.*

1.2 A solution: the automated encyclopedia

In order to solve these problems, *Bohnert* and *Kochen* have suggested the creation of an automated encyclopedia:

"At such times, one thinks fondly of the occasions when a quick look-up in an encyclopedia or handbook has yielded an authoritative summary, but reflects with resignation that encyclopedias are necessarily limited in size, directed to the general reader rather than the researcher, and usually more than ten years out of date. So they have been. The computer revolution, however, can change all that. In fact, a reexamination of the encyclopedia concept in the light of current technology suggests that it may serve as the basic concept for a new system of scientific communication, taking in and integrating the latest advances, giving out what information is required, not only as much but also as little, tutorial or technical, elementary or advanced. Since technology offers remote access, it follows that location, and hence size, need no longer be limitations. The master encyclopedia need exist at only one center. It should have a word content equal to thousands of volumes, though of course it need have no resemblance to book form. Technology offers a broad variety of access and storage modes, so that less frequently used, or less valuable, records could be stored more cheaply at the cost of slower access.

The master encyclopedia, then, would be an integrated system of storage systems, discs, tapes, cores, cards, microforms. It would contain not only summaries, but articles on many levels of detail, maps, diagrams, tables, and finally the masses of basic data of science, law, business, medicine and government. It could approach, ultimately, a representation of the total state of knowledge, rather than provide a reference summary only."

"Obsolescence, the chronic defect of encyclopedias in the past, can be overcome once the basic text to be updated exists only in one location and not in printed form. A form of change which permits any number of insertions and deletions yet can serve as the master program for the printing of inexpensive condensations is within the technological horizon ..." (Bohnert et al 1963.10, p. 269)

To the degree that such a plan would be implemented, the information transfer system would change, and emphasis would shift to the right half of the diagram. All the problems mentioned above would be eliminated and thus the information contained in the automated encyclopedia could be obtained much more easily, either from a printed extract or through on-line access to the master file. It is interesting to note that, by overcoming the shortcomings of information transfer through the literature, an automated encyclopedia could take on many of the functions now performed through informal information transfer.

In case a question cannot be answered from the automated encyclopedia, two possibilities are open: either the data are transferred from the literature to the automated encyclopedia and the question is answered then, or the user is referred to the literature. Ideally, each question should be directed to the automated encyclopedia first; only if it cannot be answered from the data available there should the user be referred to the appropriate references or to the reference storage and retrieval system. However, at the beginning the automated encyclopedia will be small and there may be questions which clearly can only be answered by references; in this case, the question might be directed immediately to the store of references. As to the contribution scientists and others have to make, they would no longer be in the form of publications but in the form of direct contributions to the data store, thus eliminating redundancy from the outset.

A partial solution would be a "computer-program system to facilitate the study of technical documents" (Bobrow et al. 1966.10). Such a system assumes that the full text of documents is accessible in machine-readable form; it would retrieve text segments, relevant to a search request.

To the extent that such retrieval based on natural language is feasible such a system would solve the problem of specific access. It would not solve the problem of redundancy.

The remainder of this article will be devoted to a more detailed explanation of the idea of an automated encyclopedia and its feasibility. The reasons for the present state of affairs will be discussed first because this discussion will shed light on the problems to be solved in establishing an automated encyclopedia.

2. Reasons that lead to redundancy in publications in the present system

The same data elements are repeated in different documents for any one of the reasons given in the following. With each of these reasons an indication is given how the creation of redundancy would be avoided in an information transfer system revolving around an automated encyclopedia.

(1) The author of a document did not know that a data element is already contained in another document (be it an element of empirical data or an idea that the author assumed to be new).

With the automated encyclopedia the author would either have found out that the data element existed already or he would be so notified at the time he makes his direct contribution.

(2) New data presented in a document cannot be understood without a context of data that have already been published. Frequently, the data forming this context are scattered over numerous documents and/or formulated in another terminology or for other reasons difficult to access for the reader of the document in question. In this case a repetition of data is unavoidable in the present system in order to spare the reader considerable inconvenience.

Note, however, that different readers with different backgrounds have different requirements as to the background data presented. Thus many, if not most, readers will find some of the context data superfluous and annoyingly repetitious, whereas they will find other context data missing.

With the automated encyclopedia the author would simply mention the context knowledge necessary to understand his contribution and each reader could then get just the context knowledge he needs from the automated encyclopedia in a presentation best suited to his background. The author has to formulate only his own new contribution.

(3) An author wants to present empirical data or ideas that have already been published from another viewpoint. This is of particular importance in the social sciences.

In this case, different presentations of the same data are necessary. However, the user would be made aware of the fact that they are different presentations and that they deal with basically the same data.

(4) Data are repeated to reach different groups of readers. Examples are

– A scientist describes a discovery both for a scholarly journal and for a popular science journal.
– An economist working on the "Economics of information" may report parts of his work in economics

journals and part in information science journals, with considerable overlap.

- Translations.
- The same news item is published in numerous newspapers.
- State of-the-art reports, textbooks, handbooks, etc., contain by their very purpose mainly data that have already been published at other places.
- A closely related phenomenon is the publication of journal articles that are abridged or otherwise revised versions of reports or theses.

In all these cases repetition is useful, and in the present system necessary, for the dissemination of information, primarily the current awareness function. However, for storage and retrospective searching, repetition of information has all the disadvantages discussed in Section 1.1. The basic difficulty is that dissemination of information on the one hand and storage an the other pose different requirements. If, as in the present system, the same documents serve both purposes (if sometimes in varying degrees), difficulties are bound to arise.

With the automated encyclopedia, redundant documents for dissemination purposes would still be produced. However, each data element could be traced back to the non-redundant data store. For example, somebody having read a textbook on 'Statistics for Psychologists' and wanting to broaden coverage to sociology applications does not need to go through a textbook 'Statistics for Sociologists' in its entirety, which would be terribly repetitious. Rather, he would simply go the non-redundant data store and ask for those relevant data elements that are *not* contained in the textbook he has read already.

(5) Status in the scientific community and career opportunities are still highly correlated with the number of publications by an author. This induces many authors to write much and to publish the same data at many different places. Many useless publications result which have only the effect of making access to relevant data more difficult. (The impact of the automated encyclopedia on this problem is beyond the scope of our discussion here.)

3. Use of an automated encyclopedia. Its role in infonnation transfer

In its ideal form an automated encyclopedia would be a data storage and retrieval system based on a non-redundant store of all human knowledge. The uses of such a system, discussed in this section, would revolutionize information transfer. The structure of such a system, discussed in Section 4, must be designed to support these uses.

3.1 Selective dissemination of data

Selective dissemination could be provided, not as is done now by sending references to documents but by sending relevant data. Each user would receive at regular intervals (e.g., once a week) a compilation of data corresponding to his interest profile and his background knowledge. The elements contained in such a compilation may be of widely varying nature, from the treatment of larger contexts similar to conventional articles to short segments conveying an isolated data element of interest.

3.2 Retrospective retrieval of data

In answering a request, the system could be much more specific and adapted to the user's background than any publication. In this respect, the automated encyclopedia would have the same advantages as informal communication. To achieve this, the system would interact with the user in answering a request. In its simplest form the process could be as follows:

The system would retrieve a data element or set of data elements answering the request (the answer set).

Next, it would present to the user a list of the background knowledge necessary to understand the answer and ask what items of the list he already knows and what items he does not know. (This entails questions on the user's background in general and questions on specific prerequisites for understanding or using the answers.) The system would then add to the answer set the necessary context data missing from the user's background. At this point, the data elements in the answer set are in the internal representation of the system. They must now be translated into an external representation that can be understood by the particular user. This external representation can take the form of natural language discourse, mathematical formulae, tables, etc. The external representation may vary from one user to the next according to background and cognitive style. The result is a representation that is appropriate for the information needs of the user on the one hand and to his background knowledge on the other. It is also possible to arrange for the system to keep a record for each user listing his background knowledge, both general (according to education) and special (e.g., through a continuously updated listing of the requests put in by that user and the answers given for these requests). (Of course, these records would not be accessible for other users.) The system would then be in a position to deliver immediately the answer in the right form, together with the necessary background information, without first asking the user.

Often much more complex interactions will be necessary to match the information needs of the user. Of particular interest are methods of computer-assisted instruction. (After all, education is a form of information transfer.)

3.3 Types of requests

The system must be capable of dealing with two kinds of questions:

(1) *Questions for specific data elements.* Among these one can distinguish closed questions having one specific answer (e.g., Who was the foreign minister of India in 1955?) or open-ended, requiring several data elements as answers (e.g., What considerations are important in formulating provisions for international inspection in the nuclear non-proliferation treaty?).

(2) *Questions for general orientation* or in-depth study in an entire field or topic.

This classification of questions corresponds roughly to Kochen's distinction between (1) information dispensing systems and (2) encyclopedic, or tutorial, information systems. The different requirements are highlighted in the following quotation:

"More generally, this system, with its variety of access-currency-detail trade-offs, should provide the effect of a variablefocus, or "zoomar" lens on reality, from the big picture to the ultra-microscopic. Beginning with the briefest summaries, the researcher should be able to follow down more and more detailed or advanced discussions, culminating, when desired, in the most basic data available to society: laboratory reports, patents, case histories and the like. It should also permit him to study his way into disciplines, indicating prerequisites, suggesting a course of study, perhaps through teaching machine consoles, providing detailed instructions." (Bohnert et al. 1963.10, p. 269.)

In the context of an automated encyclopedia specific questions are probably easier to handle than general orientation questions. Another important distinction arises in those cases where information is sought specifically to solve a problem:

(1) Questions where the user knows (or thinks he knows) what data will be useful for solving the problem, and asks for these data.
(2) Questions where the user simply states the problem and expects the system to come up with data useful for solving the problem.

In either case, the answer might be of the specific type or the general orientation type, or a mix of both.

3.4 Data analysis and inference

Often the answer to a specific question is not contained in the system explicitly but can be derived from the data stored by suitable data processing and inference techniques. An automated encyclopedia system must thus have sophisticated data analysis and inference capabilities which are invoked either automatically in the course of a search or through the user.

3.5 Forms of output

In the pure form of selective dissemination from the automated encyclopedia, each individual user would have printed a tailor-made data compilation in regular intervals on his terminal, or printed on a high-speed printer and sent by mail. Alternatively, the user could periodically sit down at a terminal and request an overview of data falling in his interest profile that have accrued since his last current awareness terminal session. He could then probe in more detail points of interest and request hard copy where appropriate.

Individual SDI is expensive. Clusters of users can be detected based on the interest profiles; one data compilation can be printed and a copy sent to each member of the cluster. This would in effect be a low circulation high-specialty journal reflecting user needs. A combination of cluster SDI and individual SDI is also possible.

In retrospective searching it would be expensive to deal with all requests in an on-line mode. One could print texts and handbooks for study and desk reference. "Finally one could print general extracts in different degrees of condensation (e.g., varying from one volume to thirty volumes) which would take the role of the present universal encyclopedias." (Bohnert et al. 1963. 10, p. 269.) All of these extracts could be tailor-made for sepcific audiences and easily updated at regular intervals, especially if they are printed on cheap paper. Only if a request could not be easily answered through one of these printed tools would the central store be accessed on-line.

3.6 Contributing data to the store

An automated encyclopedia would not only make it much easier to find information, it would also make it easier to contribute information. This aspect will be discussed in Section 5. In many interactions users will both contribute information to the system and receive information from the system.

4. Structure of an automated encyclopedia

4.1 Logical structure of the data store

This section gives a general introduction. Some specific problems will be discussed in Section 6.

4.1.1 Representation of data elements

Individual data elements would most likely have a basic representation in the index language of the system. This representation serves for retrieval and processing. In some cases it will be possible to use one of several standard patterns to translate the internal representation into an external one (text, formulae, tables). Otherwise, one or more external representations must also be stored. If sameness-of-meaning of two natural language statements is difficult to determine, both statements may have to be stored.

A set of data elements that together provide introduction to a topic or other coherent context are more difficult to handle. It will probably be necessary to store a coherent natural language text for the entire set. However, to the extent possible the representation of the data elements in natural language must be formulated in such a way that they can be put together to represent larger contexts (modular design). It should be possible to achieve different degrees of detail and/or transmission of necessary background knowledge by elimination or addition of appropriate paragraphs. Very often, it will be necessary to have different natural language representations of the same data elemnt (e.g. popular and scientific representation, "physics for students of medicine," etc.).

The internal representation of data elements should be in a form that achieves high density of information. One might, for example, present all data on loans to developing countries published in *The New York Times* in the form of a table without losing information. This representation would need much less space than the whole of the corresponding entries in *The New York Times.*

Often it will be necessary to introduce context-dependency to achieve high density of information. To give a very simple example: in reporting results from information needs studies for chemists in academic institutions, it would be a waste of space to prefix every result by "For chemists in academic institutions, ..." Rather, this should be a general prefix for the whole set of data elements.

4.1.2 Data reduction through generalization

It is often possible to represent large numbers of empirical data adequately through more or less generalized statements or laws, possibly adding a few empirical base data (this is possible, for example, for many empirical data in physics). This offers obvious savings in storage space. These must be weighed against the costs of deriving specific data when they are needed. (These costs are, in turn, dependent on available data processing technology. The obsolescence of huge volumes of function tables due to the capabilities of pocket calculators is a case in point.) The generalizations arrived at in the process of reducing data may be more useful than the raw empirical data themselves. A related approach is the use of formal similarities to represent data elements more succinctly. Especially, one should try to find formal schemes that may be used for the more transparent presentation of different empirical contents and which would lead thereby to a "formal unification of knowledge".

One could sum up these considerations by saying that theory building has a thorough impact on the amount of data to be stored and transferred. It is claimed here that a systematically organized data store would facilitate theory building. This, in turn, could lead to the systematic formulation of hypotheses and the testing of these hypotheses. In other words, a systematization of the present knowledge could lead to a systematization of the acquisition of new knowledge.

The importance of these considerations for education cannot be overemphasized.

4.1.3 Context indexing of data elements

(1) With each data element or, more precisely, with each representation of that data element, the background knowledge necessary to understand that data element in that representation must be indicated. This will establish a prerequisite structure. Such a prerequisite structure is particularly important in education, for example, in the development of curricula. This can be refined by indicating degree of difficulty, cognitive style for which a particular representation is suited, and similar information to be considered in finding the optimal presentation for a given use.

(2) Individual data elements referring to historical events can often be understood only by considering the role of these events in a complex network of historical processes. Therefore, for each event pointers should be given to earlier events that had an effect on contemporary events that are related to, and later events that were affected by, the event being indexed. The user should be able to ask for this context, possibly in form of a graphical display linking the events.

4.1.4 Reliability indexing

For each data element, the degree of reliability should be determined or at least the information necessary to estimate the degree of reliability should be given. The source indication may serve this purpose if the sources have been subjected to an objective evaluation. A reasonable scale of degrees of reliability of statements might be as follows:

Generally accepted as valid
Probably
Possibly
Probably not
Generally accepted as not valid

A similar scale can be established to indicate the degree of accuracy of numerical values.

4.1.5 Indexing of potential uses/applications

Each data element should be indexed by its potential uses applications, problems that might be solved better by using the data element sets. This type of indexing will make problem-oriented retrieval possible.

4.1.6 Source indications

For each and every data element all sources are indicated precisely, so that it is possible to go back to the original sources whenever this is felt advisable, or even to reconstruct the original sources from representation in the automated encyclopedia. (Compare section 5 Building and updating the data store.)

4.2 Remarks on physical storage organizations

A further aspect of the organization of an automated encyclopedia—very important for the practical implementation of the idea—is the introduction of a hierarchy of stores. Two aspects must be considered in this connection:

(a) Information which is needed frequently is stored in storage devices with fast access; information which is needed seldom is stored in slow access storage devices.

(b) Information which is frequently to be modified is stored in storage devices in which modifications are easy to perform, whereas information which is fixed once and for all can be stored in non-modifiable storage devices.

The costs for a storage device increase as the access time decreases and as the ease with which modification can be performed increases.

The completely processed original documents would be kept, too, so that one could always go back to them. They would be kept in storage devices lowest in the hierarchy (e.g. microfilm or read-only mass storage devices).

(Sections 5–10 in the forthcoming issue 77–2).

Notes (Sections 1–4)

Notes are given by section number

1. General references on the idea of an automated encyclopedia, or, more gencral, a universal encyclopedia not necessarily using computers.
   Wells 1938
   Pollard 1938
   Bush 1945.7 (However, Bush is concerned with the structure of an individuals "external memory," not with a public data store.)
   Manly 1960
   Doren 1962.7
   Bohnert et al. 1963.10 (excellent brief overview; extensively quoted in this article.)
   Milbrath et al. 1964.9, p. 1–2
   Licklider 1965
   Davis 1965.5
   Kochen ed. 1967
   Soergel 1971, Section B6.2, p. 214–234 (This section is superseded by the present article, except for some specifics given in footnotes.)
   Pager 1972.2 (covers data structure, interrogation of, and input to an automated encyclopedia in mathematics.)
   Kochen 1972.12

1.1 The problems facing the researcher vis a vis the literature arc discussed by Manly 1961.7, p. 204 and Deutsch et al. 1963.2, p. 6–7 (Deutsch et al. 1966, p. 83–86) "The need for a data program". The reader is referred to these sources for further illustration of the problem.
An excellent description of the steps needed to extract data from the literature in answering a reference request is given in Alexander 1936.9.

1.2 The idea of retrieving specific text segments was proposed by Mooers 1959.3, p. 85. A program for this purpose is sketched in Licklider 1965, p. 177–179 and Bobrow et al. 1966.10. Menzel 1968 reports how scientists and engineers resort to informal methods of information transfer to overcome the problems listed in Section 1.1.

2. Some of the reasons leading to redundancy are discussed in London 1968.4; for examples and further references see there. The problem is also discussed in ICSU/UNESCO 1968.7. Specific points were made in these references: (1) Author didn't know: te Nuyl 1966.6, p. 97; Manly 1961.7, p. 204; (2) Context knowledge must be repeated: Bohnert et al. 1963.10, p. 269.
3. The type of organization where various overlapping information products are created from a non-redundant store was suggested on the level of reference retrieval by Taeuber 1951 and Smith 1950.4. This idea is implemented in VINITI and is being approached in the envisioned cooperation between Chemical Abstracts, Biological Abstracts, and Engineering Indes. Proposals for a new journal mode (e.g. Shank 1966.1) propound a similar model for dissemination of documents.
3.2 The excellent article by O'Connor 1967.7 discusses the kind of inferences needed to determine whether a data element can be helpful in solving a problem at hand. Thus, this article is relevant in this context as well as for Section 4.1.
3.3 Kochen ed. 1967, p. 1–5. Mooers 1959.3, p. 85, suggested the use of CAI methods.
4.1 The sketch of the logical structure of the data store draws heavily on the more detailed list in Licklider 1965, p. 36–39. All the work done in data storage and retrieval and question-answering is, of course, relevant here.
4.1.1 Mooers 1959.3, p. 85–86, speculates about the possibility of translating from index language to natural language.
4.1.3 Soergel 1972.4 discusses the prerequisite structure in detail. See also Pager 1972.2.
4.1.4 Reliability indexing is suggested in Travis 1963, p. 332–334.
4.2 The multi-level storage organization is envisioned by Bohnert et al. 1963.10, p. 269.

References

Section numbers refer to the section where the item is cited.

(1) Alexander, Carter 1936.9: The technique of library searching. In: Spec. Libr. 27 (1936.9) No.7, p. 230–238.– Section 1.1.
(2) Bernier, Charles L. 1970.9: Terse literatures: 1. terse conclusions. In: J. Amer. Soc. Inform. Sci. 21 (1970.9) No.5, p. 316–319.– Section 6.7.
(3) Bobrow, D. G. 1966.10; Kain, R. Y., Raphael, B., Licklider, J. C. R.: A computer-program system to facilitate the study of technical documents. In: Amer. Doc. 17 (1966.10) No.4, p. 186–189. Also: Cambridge, Mass.: Bolt, Beranek and Newman 1963.11. – Section 1.2.
(4) Boehm, Eric H. 1963.5: Dissemination of knowledge in the humanities and social sciences. In: ACLS Newsletter 14 (1963.5) No.5, p. 3–12.– Section 7: p.11, 5; 8.3: p.6–7.
(5) Bohnert, Herbert G. 1963.10; Kochen, Manfred: The automated multilevel encyclopedia as a new mode of scientific communication. In: ADI Proc. 1963.10, p.269–70.– Section 1.1: p. 269; 1.2: p. 269; 2(2): p. 269; 3.3: p.269; 4.2: p. 269;5.1: p. 269;8.2: p. 269.
(6) Brandejs, Jan F. 1975.5: The C.M.A. information base–a beginning of operational systems in Canada. In: Amer. Federat. of Inform. Process. Soc. AFIPS Conf. Proc. 44 (1975.5) p. 725–731.– Section 9.2.
(7) Bush, Vannevar 1945.7: As we may think. In: Atlantic Monthly 176 (1945.7) No.1, p. 101–108. Repr. in: Kochen ed. 1967, p. 23–35.– Section 1.
(8) Chun, Ki-Taek 1973.7; Cobb, S.; French, J.R.P., Jr.; Seashore, S.: Storage and retrieval of information on psychological measures. In: Amer. Psychol. 28 (1973.7) No.7, p. 592–599.– Section 9.
(9) Chafetz, Morris E. 1976.5: Toward increased quality consciousness for alcoholism literature. In: J. Amer. Soc. Inform. Sci. 27 (1976.5) No. 3, p. 162–170.– Section 7.
(10) Davis, Watson 1965.5: The Universal Brain: (Is centralized storage and retrieval of all knowledge possible, feasible, or desirable?) Paper, National Microfilm Assoc., Cleveland, Ohio, 1965.5.13. Repr. in: Kochen ed. 1967, p. 60–65.– Section 1.
(11) Deutsch, Karl W. 1963.2; Lasswell, H. D.; Merrit, R. L.; Russet, B. M.: The Yale Political Data Program. New Haven, C.: Yale Univ., Political Science Research Libr. 1963.2. 17 p. Abbr. repr. in: Merrit et al. ed. 1966. p. 81–94.– Section 1.1: p. 6–7.
(12) Doren, Charles v. 1962.7: The idea of an encyclopedia. In: Amer. Behav. Sci. 6 (1962.7) No. 1, p. 23–26. Repr. in: Kochen ed. 1967. p. 66–71.– Section 1.
(13) Gervin, Paul L. ed. 1963: Natural language and the computer. New York: McGraw Hill 1963. 398 p.– Source.
(14) ICSU/UNESCO 1968.7: Sichtung, Verdichtung und Ordnung wissenschaftlicher Infonnationen. Ber. d. Arb. gr. 4 d. gemeinsamen Komitees.– Section 2.
(15) Janda, Kenneth 1968.4: Data quality control and library research on political parties. Evanston, Ill.: Northwestern Univ./Int. Comp. Polit. Parties Project 1968.4 7 p. = ICPP Reports No.4.– Section 7.
(16) Kochen, Manfred, ed. 1967: The growth of knowledge; readings on organization and retrieval of information. New York: Wiley 1967. 394 p.– Section 1; 3.3: p. 1–5. Source.
(17) Kochen, Manfred 1972.12: WISE: a world information synthesis and encyclopedia. In: J. Doc. 28 (1972.12) No.4, p. 322–343.– Section 1.
(18) Licklider, J. C. R. 1965: Libraries of the future. Cambridge, Mass.: MIT Press 1965. 219 p.– Section 1, 1.2: p.177–179; 4.1: p. 36–39; 5.2: p. 53, p. 111–118; 9.1.3: p. 33.
(19) London, Gertrude 1968.4: The publication inflation. In: Amer. Doc. 19 (1968.4) No.2, p. 137–141.– Section 2.
(20) Manly, R. 1960: The encyclopedia approach to the storage and utilization of scientific and technical information. Gardena, Calif.: Systems Engineer. Inst.: Intelligence Process. Lab. 1960.– Section 1.
(21) Manly, R. 1961.7: The inadequacy of varying the depth of indexing and other document collection approaches to information retrieval for researchers. In: Amer. Doc. 12 (1961.7) No.3, p. 204–5.– Section 1.1: p. 204; 2(1): p. 204.
(22) Menzel, Herbert 1968: Information communications in science: its advantages and its formal analogues. In: Montgomery, E. B. ed.: The foundations of access to knowledge. Syracuse University, Syracuse, N.Y. 1968. p. 153–163.– Section 1.2.
(23) Merritt, R.L. ed 1966; Rokkan, S., ed.: Comparing nations.– The use of quantitative data in cross-national research. New Haven and London: Yale Univ. Press 1966. 584 p.– Source.
(24) Milbrath, Lester W. 1964.9; Janda, Kenneth: Computer applications to abstraction, storage and recovery of propositions from political science literature. Evanston, Ill.: Northwestern Univ. n. d. 11 p., app. Amer. Political Sci. Assoc. Meeting 1964.9.9–12. Paper. Section 1; p. 1–2.
(25) Mooers, Calvin N. 1959.3: The next twenty years in information retrieval.– Goals and predictions. In: Proc. Western Joint Comp. Conf., San Francisco, Calif. 1959.3.3–5, p. 81–86.– Section 1.2: p. 85; 3.3: p. 85; 4.1.1: p. 85–86; 5.2: p. 86.
(26) Naroll, Raoul 1962: Data quality control: a new research technique. New York: The Free Press 1962.– Section 7.
(27) te Nuyl, T. W. 1966.6: Die Beurteilung von Systemen für das Wiederauffinden von Informationen in organisatorischer Sicht. In: Nachr. Dok. 17 (1966.6) No.3, p. 93–97. Section 2(1): p. 97.
(28) O'Connor, John 1967.7: Relevance disagreements and unclear request forms. In: Amer. Doc. 18 (1967.7) No.3, p. 165–177.– Section 3.2.
(29) Pager, David 1972.2: A proposal for a computer-based interactive scientific community. In: Comm. ACM 15 (1972.2) No.2, p. 71–75.– Section 1; 4.1.3.
(30) Pollard, A.F.C. 1938: The mobilization of knowledge and the "permanent world encyclopedia" of Mr. H. G. Wells. In: Intern. Fed. Doc. XIVth Conf. (1938) p. 161–167.– Section 1.
(31) Ruggles, Richard 1967.8; Ruggles, Nancy: Data files for a generalized economic information system. In: Soc. Sci. Inform. 6 (1967.8) No.1, p. 187–196. Paper Amer. Sociol. Assoc. Reg. Meeting, Harvard Univ. 1966. 3.17.– Section 5.2.
(32) Shank, Russell 1966.1: Letter to the editor. In: Amer. Doc. 17 (1966.1) No.1, p. 45.– Section 3.

References (33) to (40) please see pt. 2 in forthcoming issue.