The Information Analysis Center—Key to Better Use of the Information Resource*

ODIN WILHELMY, Jr., and PATRICIA L. BROWN
Department of Economics and Information Research, Columbus Laboratories, Battelle Memorial Institute, Columbus, Ohio

Received March 11, 1968

The steadily growing mass of information relating to food and agriculture represents a significant resource in efforts to cope with the increasingly serious world food problem. The effective application of this resource to this problem requires the development of techniques commensurate with the complexity of the problem. The establishment of an "information analysis center," oriented to the needs of government agencies, private industry, and scientists, would permit exchange of essential knowledge that presently cannot be achieved. The mechanism for setting up and operating an information analysis center has been worked out and proved in other fields of science and industry. The types of information that would logically go into a world food and agriculture center are readily identifiable. The variety of outputs available can be geared to the most pressing needs of those who would benefit most.

Over the past 18 months, a number of us at Battelle have been giving a great deal of thought and attention to the question of how the growing mass of information in the field of food and agriculture can be more effectively mobilized to facilitate a solution to the increasingly critical world food problem. This thinking, and the steps that we have taken thus far in this regard, have stemmed primarily from our information scientists and from our agricultural and food economists, working together. The former group has a background of more than 15 years of experience in setting up and operating a variety of information centers to serve the needs of government and industrial sponsors. Out of that experience has come an appreciation of just how useful a properly structured system can be in coping with the information problem. Over a similar period of years, we in agricultural and food economics have become increasingly aware of the need for such a tool, both in our research in this country and in developing countries overseas. It was only natural, therefore, that these two groups would come together in their thinking. It was equally natural that other research groups at Battelle—in food technology, biochemistry, chemical engineering, and mechanical engineering—gave us full support and encouragement in this effort.

Our initial activities in this area, in the summer of 1966, took the form of personal interviews with top management of selected food and agriculture-oriented companies whom we felt might have an interest in a group research program on opportunities in the developing countries. Our visits included such companies as Swift, Armour, Dekalb, Pillsbury, International Harvester, Quaker Oats, and others. We found a rather wide variation in the degree of interest in overseas opportunities on the part of these companies, but great unanimity of concern over the availability and quality of basic information needed to evaluate such opportunities intelligently.

*Presented in Symposium on "The Literature of Agricultural and Food Sciences—Problems and Some Solutions," Divisions of Chemical Literature and Agricultural and Food Chemistry, 155th Meeting, ACS, San Francisco, Calif., April 1968.

We decided at that point to reorient our efforts towards the formulation of a research program aimed at closing this information gap. In the fall of 1966, we structured a proposal for the development of a food and agricultural information analysis center directed at certain aspects of the world food problem. In succeeding months, we presented this proposal to officials of the USDA's Economic Research Service, the Agency for International Development, and the newly formed Agribusiness Council. In every case, there was genuine interest in the concepts that we presented and general agreement that an approach of this type had real potential merit. While, for a variety of reasons, nothing tangible has yet materialized from our efforts, we remain convinced of the need and value of such an information analysis center.

As might be inferred from my remarks thus far, we have attached a world-food-problem dimension to this paper. We have done so for two reasons. On the one hand, there are many people who believe that the coming world food crisis and its potential ramifications will be the most difficult and wide-ranging problem confronting the world between now and the end of the century. By the same token, one can make a rather strong case that a really effective means for coping with all the information bearing on the world food problem could have greater long-range benefits than a system aimed at a narrower or more specialized field of interest. I am sure, however, that most of the thoughts expressed can find equal validity and application to the other agricultural and food problems with which many may be more immediately concerned in their day-to-day work.

THE NEED FOR A BETTER SYSTEM AND THE INGREDIENTS OF THAT SYSTEM

Information (both quantitative and nonquantitative) is becoming recognized as a world resource—a resource growing rapidly in volume and complexity, but still a resource which, properly used, can be instrumental in the solution of world problems. This information takes a multitude of forms—scientific and technical journals, classified and proprietary reports, monographs, news reports, patents, preprints, letters, data, trip reports, abstracts, and foreign literature. The sources of such information—publishers, companies, government agencies, professional societies, universities, and the like—are as numerous as its forms.

On top of all this, the total package of information with which one must be concerned in the food and agricultural field is both multidimensional and interdisciplinary. On the one hand, there are technological, economic, sociological, geographic, and political aspects that must be considered. On the other, one must recognize that essential knowledge may derive from any of a variety of scientific disciplines—agriculture, chemistry, economics, chemical or mechanical engineering, psychology, or demography, to name just a few. It is the sheer enormity of the information problem implicit in these facts that convinces us of the urgent need for a better system than those now in common use to cope with this problem. We are not going to realize the full potential of the information resource until such a system is made available and effectively used.

For those who question the need for improved methods of information handling and analysis, let me cite several strong evidences of need as we see them in our research activity at Battelle. One obvious and familiar example is the need for getting an adequate grounding in a new subject area—be it a product, an industry, or a country—in order to carry out an assignment that has been handed you. Because of the diverse nature of our studies at Battelle, we probably face this problem more frequently than others, but I'm sure only in terms of degree. None of us can hope to be "instant experts" in every area in which we are called upon to function, and any increased efficiency that can be attained in the job of dredging up basic knowledge in these areas can represent significant savings in time, dollars, and energy.

For those who feel this day-to-day need for information is not enough justification for developing a sophisticated information analysis center, let me cite another problem area where it may be justified. We are all aware, at least to some degree, of the far-flung research efforts in food and agriculture being conducted by AID, many universities, various foundations, the USDA, and local government agencies in developing countries around the world. Without question, the better grasp that all these researchers have of pertinent knowledge generated in the past, the better they will be able to do their current tasks. But beyond this obvious need for previously generated information, we submit that a well-conceived information analysis center also could contribute significantly to the exchange of current and future knowledge.

If the benefits of all this on-going research are to be realized as fully and as rapidly as possible, there must be some mechanism for scientists in one part of the world to learn quickly of the progress being made by their counterparts in other countries or continents. An oil palm expert in Central America will not profit much from the findings of an oil palm expert in Africa if it takes ten years for those findings to filter through to him. Similarly, if a university agronomist working in India on an AID contract winds up his two-year tour of duty and comes home, his successor from another university may not get the benefit of his predecessor's hard-won know-how unless there is an adequate means for passing that experience along. One could well make a similar case with respect to the volumes of information flowing out of our state colleges and experiment stations and the USDA in this country.

A third evidence of need that one can cite stems from the growing interest of U. S. food and agriculture-oriented companies in possible overseas production or marketing opportunities. Many of these companies have never ventured outside the U. S.; others may have international operations, but only in the more familiar areas of the world such as Western Europe. In either case, serious consideration of a venture in a developing country of Asia, Africa, or Latin America may well require a very substantial pre-investment effort to determine the product and marketing requirements, the potential profitability, or the simple mechanics of getting the proposed venture off the ground. The existence of an adequate information facility designed to provide such information quickly and at reasonable cost certainly would facilitate the decision-making process in such cases.

The facility that we have in mind and which we offer for serious consideration is an information analysis center designed to serve the needs of workers in agricultural and food science and technology. The primary activities of such a center are illustrated in Figure 1. Like other information-handling systems with which you may already be familiar, it includes the three basic functions of (1) acquisition, (2) storage and retrieval, and (3) service. However, there are a number of very significant ways in which an information analysis center differs from more conventional systems—differences that we believe make the analysis center considerably more useful.

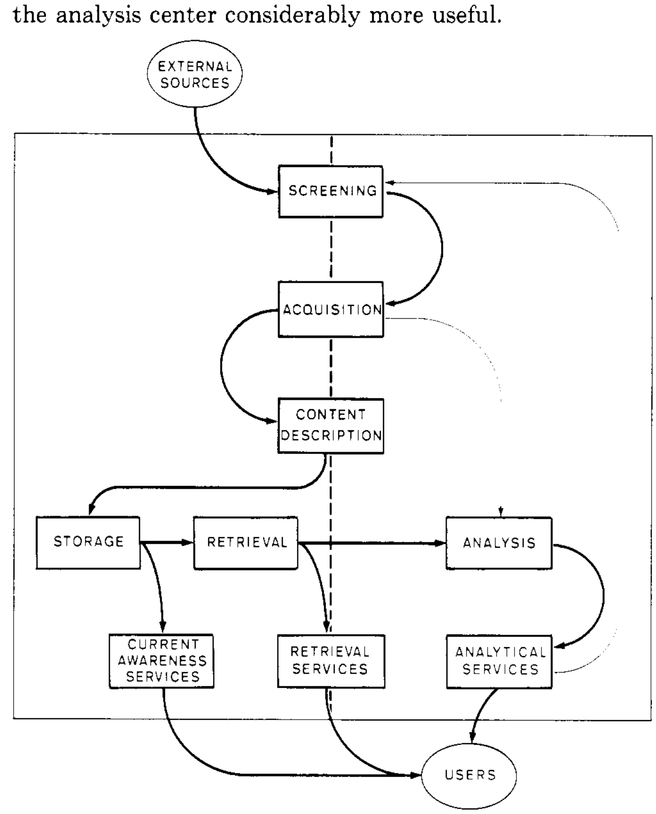

Figure 1. The primary functions of an information analysis center.

The most important of these extra ingredients is the close and continued association of trained scientists and engineers with the organization and operation of the center. Traditionally, functions to the left of the dashed vertical line in Figure 1 are handled by information specialists. Those to the right are handled by subject specialists. A few are best handled by a combination of both. The analysis center emphasizes the role of the practicing subject specialist in the area of his greatest competence—analyzing the information in his field, thus permitting the center to provide authoritative service, technically interpreted for the particular user group.

An analysis center is mission oriented. That is, it is established and run for the specific purpose of serving the information needs of people concerned with a given field or subject-matter area. It is not tied to any particular form of information handling.

As shown in Figure 1, screening is the key operation by which all information items enter the system, including those internally generated in the analysis process. Aggressive, accurate, and continuous screening and acquisition must be conducted for the center to be effective. The utmost care must be taken to monitor and control these two functions. The next three areas—content description, storage, and retrieval—should be treated cautiously. They are the means to an end, not the end itself. Care is needed in content description and choice of storage and retrieval methods, but these functions are significant only if they are geared to facilitate analysis and use of the information—the real payoff of the center. To achieve this payoff, the analysis center blends the competence of the experts in information handling with that of the scientists and engineers in the field—the same type of people who use the products of the center.

The many types of information inputs—ranging from a book to notes on a phone conversation—that can be fed into an analysis center have already been mentioned. The potential outputs of such a center are equally varied in nature, depending upon the needs of the user for the three types of services—current-awareness, retrieval, and analytical—normally offered. Current-awareness services encompass periodic publications sent to all or selected users and include abstract bulletins and bibliographies or indexes to current acquisitions. Many types of retrieval service are possible, both periodic and on demand, with provision for rapid reference service for answering questions from users. Analytical services, however, are perhaps the most significant feature of information analysis centers. These are tasks requiring specialists in the subject areas and using the information store as a base. Out of this can come (1) summaries or state-of-the-art reports, (2) identification of gaps in current knowledge, (3) recommendations for research programs and policy formulations, (4) management information to aid in decision making, (5) syntheses of new ideas based on current knowledge, and (6) research assistance ranging from "quick-fix" solutions to long-range development.

Information analysis centers, serving a number of government agencies and industry groups, already exist in various parts of the U.S. Battelle operates a dozen of them—five on behalf of the Department of Defense, AEC, and NASA. The other seven cover such diverse fields as copper, cobalt, packaging, transducers, and area economics. To our knowledge, there is no information analysis center in the U. S. to date relating specifically to the food and agriculture area.

POTENTIAL STRUCTURE, USE, AND BENEFITS OF A FOOD AND AGRICULTURE INFORMATION ANALYSIS CENTER

Having briefly reviewed the general characteristics, functions, and advantages of an information analysis center, we may consider how such a center might logically be structured and used by agricultural and food scientists and technologists. Let me stress at the outset that the intended inputs, method of operation, and outputs must be carefully worked out in advance by both users and operators of the center to ensure that it will do the job that it is designed to do. Recognizing this, one may still postulate the likely ingredients of a world food information analysis center.

In the proposal that we developed at Battelle, we suggested that information coverage encompass five major areas, as follows:

> Resources and Their Deveolpment—land, climate, natural vegetation, inland water, marine resources, fertilizers, power, agricultural finance, and research and extension services.
>
> Existing Food Supply and Consumption—indigenous foods, imports, exports, analysis and composition, geographic distribution, and accessibility.
>
> Sociological Factors—population characteristics, economic status, diets, employment, education, and government factors.
>
> Methods of Food Production—planting, harvesting, irrigation, fertilization, pest and disease control, conservation practices, animal husbandry, farm management, and fishing practices.
>
> Methods of Food Marketing and Distribution—processing and preservation, handling, storage, marketing system, and transportation.

Other variables to be considered include the extent to which various commodities are covered and the geographic area or areas that should be of primary concern. The information sources that are to be stressed, in terms of number, type, and location, is another decision to be made. Still other variables include acquisition procedures, processing and storage methods, and the products and services to be made available to users.

The potential users and benefits of a world food and agricultural information analysis center already have been spelled out to a considerable degree in the previous discussion of the need for a better system. However, they are worth reiterating. The users would logically include the many government and private agencies seeking solutions, through research and development, to the world food problem; the thousands of individuals involved in these efforts; and those industrial firms with a stake—either existing or potential—in such efforts.

The benefits will take the form of (1) factual bases for the formulation of policy; (2) factual bases for the development of action plans; (3) improved effectiveness in the planning, coordination, and conduct of technical assistance, training, and research programs; and (4) a stable base to provide a chain of continuity for such programs. Based on our experience at Battelle in operating other information analysis centers, we are confident that such benefits are attainable. Anyone who is genuinely concerned with, or involved in, the search for solutions in this area would be hard put to deny that such benefits are desirable.

It would seem that all that remains is for those of us who believe so firmly in the value of this research tool to convince the potential users that an effective world food information analysis center would be an investment well worth making. Thus far, our efforts in this regard have not borne fruit. However, we have not given up and we do not intend to. At some point in time—hopefully soon—sufficient support for such a center will develop on the part of industry government or both to make the dream a reality. When this happens, a lot of us who are trying to do a better job of agricultural and food research will have a powerful new resource to draw upon.

CRYSTAL DATA Editor—Automatic Proofreader for the 1968 Edition

LEONARD B. TAYLOR and GERALD G. JOHNSON, JR.
Materials Research Laboratory & Department of Computer Science,
The Pennsylvania State University, University Park, Pa. 16802

Received March 18, 1968

A computer program in FORTRAN IV on the IBM 360/67 has been developed to proofread the third edition of the crystallographic reference *CRYSTAL DATA*. Proofreading is accomplished in four phases: data preparation, record searching routine, numerical verification preparation, and numeric validation. The merging of the character-processing ability of computers and crystallographic theory yields increased reliability and accuracy as well as a decreased time delay for future editions.

The third edition of *CRYSTAL DATA*, a crystallographic reference publication, is scheduled to be published this year, under the joint sponsorship of the American Crystallographic Association and the National Standard Reference Data System of the National Bureau of Standards.[1] This edition is of particular interest to crystallographers, physicists, and chemists since it will contain over 30,000 entries with information current to January 1968. Previous editions were not as large nor as current as the third edition, owing to the time required for proofreading and constructing indices. It is now possible to reduce the time previously needed to publish large current and accurate editions by exploiting the character-processing ability of electronic computers.

Prior to this edition, proofreading of the text was done manually, consuming many man-hours; the level of accuracy using this method was often questionable. A computer program to perform this proofreading of the more than 30,000 entries of *CRYSTAL DATA* has been developed by the authors at The Pennsylvania State University. This program will enable proofreading of all input records to the third and subsequent editions, utilizing a standard level of accuracy which is compatible with crystallographic theory.

The program currently processes one record per second on an IBM 360/67 at The Pennsylvania State University. The source program was written in FORTRAN IV, LEVEL G, RELEASE 12.

DATA PREPARATION

The input data is subjected to four phases of verification prior to final acceptance of the entry. Each of these phases—translation, record searching routine, numeric verification preparation, and numeric validation—concentrates on specific sections of each record to verify the existence and accuracy of items within each entry. Figure 1 shows a sample page of *CRYSTAL DATA*, unjustified for publication. Each entry begins with the characters *** and may be of any reasonable length. The input to the program, *CRYSTAL DATA* Editor, is in the form of octal tapes prepared at the Government Printing Office by the National Bureau of Standards. To produce a compatible code for the IBM 360, these tapes are subjected to a translation routine, phase one of the program. This phase is simply a character-for-character translator from octal code to hexadecimal code, the internal notation utilized for the IBM 360/system. Figures 2 and 3 show the output of the translation section. Figure 2 was printed with the 64-character upper-case print chain, and Figure 3 utilized the 120-character upper and lower case print chain. Either form of input is acceptable, and

***A,ED,NM,OR,000109
Ammonium nickel cyanide (2·1·4) 3-hydrate
&&&.5828 .4801 8.80 15.10 7.25 ... 4 qual. 1.706, 21° 1.741
92°37′ 95°23′ 88°59′
Diammonium nickel cyanide trihydrate, $(NH_4)_2Ni(CN)_4 \cdot 3H_2O$ (Brasseur & de Rassenfosse, Mem. Soc. Roy. Sci. Liege, 4, 397 1941:15.10, 8.80, 7.25; 95°23′, 92°37′, 88°58.5′; $(a{:}b{:}c)_{morph}$ =1.686 : 1: 0.824)$0\bar{1}0/\bar{1}00/00\bar{1}$. Yellow orange. Twinning. Opt. neg. (546 mμ) =1.473, 1.597; 2V 25°. Pseudo-hexagonal.
***A,ED,NM,OR,000110
Rubidium nickel cyanide (2·1·4) 3-hydrate
&&&.5838 .4824 8.99 15.40 7.43 ... 4 qual. 2.455, 20° 2.480
92°22′ 95°29 89°37′
Dirubidium nickel cyanide trihydrate, $Rb_2Ni(CN)_4 \cdot 3H_2O$ (Brasseur & de Rassenfosse, Mem. Soc. Roy. Sci. Liege, 4, 397, 1941: 15.40, 8.99, 7.43; 95°29′, 92°22′, 89°36.5′. $(a{:}b{:}c)_{morph}$ =1.705 : 1 : 0.8345)$0\bar{1}0/\bar{1}00/00\bar{1}$. Twinning. Opt. neg. Pseudo-hexagonal.
***A,ED,NM,OR,000111

Figure 1. Unjustified page of *CRYSTAL DATA*