is of great value. Many fields are well covered, including extraction, reduction, corrosion, casting, heat transfer, physical metallurgy, alloy properties based on composition, surface chemistry, electrochemistry, petroleum products, metallography, commercial aluminas, and toxicology; this coverage, of course, applies to all published literature. Coverage in the following fields is naturally of less concern to *Chemical Abstracts*, and they are not as well covered: mechanical working, machining, joining and welding, plant equipment, instrumentation, and product design. It is in these areas, particularly in product design, that we have a real problem in patent searching.

In reference to computer-based searching, we should make a comparison here with manual searching. A full-time technically qualified searcher should make 60 to 70 searches per year. These searches will have an average of 15 to 20 years' coverage and will include four or five major sources. This material will be evaluated and a lot of irrelevant material thrown out. There are no computer facilities available, to the best of my knowledge, which can give this combination of broad coverage, time period, and evaluation. If there were such a facility, the costs would probably be in the area of $100 to $200 per year of the literature searched, based on prices now being used by various agencies; in other words, $2000 to $4000 per search. This enormous cost can be reduced by combining several searches into one broad search, but our experience with this approach has been poor. In general, the program would not exclude a large drop-out of irrelevant material, and a lot of pertinent material would be lost. In other words, commercial computer searching is not yet economically sound, except in certain restricted areas—e.g., current awareness programs.

In conclusion, an attempt has been made to present the problems of the literature searcher in the aluminum industry, and possibly assist those who have to deal with them, particularly the scientist or engineer who occasionally does some searching.

CODATA—Its Organization, Activities, and Goals*

GUY WADDINGTON
Office of Critical Tables, National Research Council,
2101 Constitution Ave., Washington, D.C. 20418

Received April 3, 1969

In 1966, the International Council of Scientific Unions (ICSU) organized the Committee on Data for Science and Technology (CODATA) to achieve informal coordination among and provide guidance for numerical data compiling projects on a worldwide basis and encourage support for data compiling projects by appropriate private, governmental, and intergovernmental agencies. The committee now has members from eight countries and ten international unions. Each country and some of the unions have established internal mechanisms for carrying out the CODATA aims. Concerned groups in the various countries are linked by informal communication channels. The central office of CODATA, initially in Washington, D.C., but now in Frankfurt, Germany, will soon issue a compendium analyzing the output of principal data evaluation and compilation centers. Task groups have been established to cope with special problems—i.e., fundamental constants, application of computers to numerical data processing, standardizing "key" input data used by compilers, etc. A trend is evident toward cooperation among numerical data compilers on a worldwide basis.

In September 1966, at the New York ACS Symposium, I reported on the then recently formed Committee on Data for Science and Technology (CODATA) of the International Council of Scientific Unions (ICSU).

Before I bring the story of CODATA up to date, I shall quickly review the background of the problem, the mission, and the framework of the organization.

BACKGROUND

I trace the background, as viewed from the U.S.A. The National Academy of Sciences began its formal interest in the problem of evaluated data for science and technology in 1919. In that year, the International Union of Pure and Applied Chemistry, meeting in London, established the International Critical Tables, with the National Academy of Sciences of the U.S.A. being assigned the financial and editorial responsibility for the work. In 1923, the International Research Council, which later became the International Council of Scientific Unions, gave its patronage to the project. With Edward W. Washburn as Editor-in-Chief, and with the cooperation of 408 scientists in 18 countries, the complete work of seven volumes was issued in the years 1926 to 1930, with a separate index volume following in 1933.[1]

Between 1933 and 1957, a limited number of excellent numerical data compilation centers were organized spontaneously in the U.S., but with no coordination among them and no coverage of many areas of science. The Landolt-Börnstein Tabellen[2] in Germany and the Tables

* Presented before the Division of Chemical Documentation, 4th Middle Atlantic Regional Meeting, ACS, Washington, D.C., February 1969.

of Constants[3] in France continued to function, but had more and more difficulty in providing complete coverage. After World War II, it gradually became apparent that the compilation lag was getting out of hand. In 1955, the National Academy of Sciences, through its Committee on Tables of Constants, reviewed the entire situation and decided that a new plan was needed to provide science and technology in the United States with continuing up-to-date tables of critically evaluated numerical data.

In 1957, there was established in the National Academy of Sciences the Office of Critical Tables, with the following responsibilities: to survey the needs of science and industry for critical tables of numerical data; to stimulate and encourage existing critical-data-compiling projects; to promote uniform editorial policy and procedures, and high standards of quality; to provide an inventory-directory of continuing critical-data-compiling projects; to assist in the establishment of needed critical-data-compiling projects for new scientific areas with financial support. In the first six years of the Office of Critical Tables, significant progress was made in carrying out all the objectives except the last one, which required funds in relatively large amounts.

In 1963, new life was given to the funding situation when the United States Government officially recognized the problem of data for science and technology as an important one and made funds available to the National Bureau of Standards to operate a National Standard Reference Data Program. This program, initially directed by Dr. Edward L. Brady, but now by Dr. D. R. Lide, Jr., is well under way, with close liaison with the Office of Critical Tables.

Soon after the start of the NBS program, it became increasingly clear that the problem of data for science and technology is international. No single country has the resources of money and manpower to support the massive effort that is needed. Accordingly, in the spring of 1964, prominent scientists from major countries proposed to the International Council of Scientific Unions (ICSU) that it should take the lead in providing international coordination and guidance in the important matter of evaluated numerical data for science and technology. In June 1964, the Executive Committee of ISCU established a "Working Group" to examine the problem. The Working Group met late in 1964, formulated a recommendation that ICSU establish a Committee on Data for Science and Technology, and defined its purposes and work. In June 1966, the committee held its organizational meeting in Paris, France, and approved the general nature of its functions.

The Mission of CODATA. The following are the tasks stated in the constitution of CODATA:

1. To ascertain on a world-wide basis, through the Unions and appropriate National bodies, what work on evaluation and publication of numerical data is being carried on in each country, what work is being sponsored by each scientific Union or by other international groups, and what the needs of science and industry are for additional compilations of evaluated data.
2. To achieve coordination among, and strengthening of, existing programs, in such a way as to maximize their effectiveness, to minimize unintentional or undesirable overlap, and to recommend new compilation programs when necessary.
3. To encourage the support of needed work by appropriate private, governmental, and intergovernmental agencies, and to encourage needed experimental work.
4. To encourage the use of nomenclature, symbols, and constants advocated by the responsible Unions, and, when desirable, uniform editorial policy and procedures for presentation of information.
5. On a world-wide basis, to stimulate wider distribution of compilations of high quality, to maintain and distribute an inventory-directory of continuing data compilation projects and related publications, and to encourage adequate indexing of the substances and properties covered by all such compendia.
6. To encourage and coordinate research on new methods for the preparation and dissemination of critically evaluated tables, generally expressed in numerical form.

Organization. The categories of persons who participate in the work of CODATA are members representing International Scientific Unions federated with ICSU, that desire to participate; members representing countries participating in ICSU, and in which there is activity in the data compilation field; and members co-opted for their particular skills.

Upon invitation, meetings of CODATA may also be attended by the following persons: liaison representatives of interested international organizations, such as the International Federation for Documentation, the International Atomic Energy Agency, the World Meteorological Organization, the ICSU Abstracting Board, and other ICSU Committees and Boards, and observers representing governmental funding agencies.

CODATA has an executive committee, called the Bureau, consisting of the president (F. D. Rossini, U.S.A.), two vice-presidents (Gordon Sutherland, U. K., and B. Vodar, France) and a secretary-treasurer (W. Klemm, Germany), M. Kotani (Japan) and Academician Styrikovich (U.S.S.R.). The Central Office of CODATA in Frankfurt, Germany, is headed by an executive director, C. Schafer, with other salaried staff as necessary. The executive director is a nonvoting member of CODATA and its Bureau.

At present, the union members of CODATA are the following: IAU, the International Astronomical Union; IGU, the International Geographical Union; IUBS, the International Union of Biological Sciences; IUCr, the International Union of Crystallography; IUGG, the International Union of Geodesy and Geophysics; IUGS, the International Union of Geological Sciences; IUPAB, the International Union of Pure and Applied Biophysics; IUPAC, the International Union of Pure and Applied Chemistry; IUPAP, the International Union of Pure and Applied Physics; and IUTAM, the International Union of Theoretical and Applied Mechanics.

National members at present are from the following seven countries: Canada, France, Germany (BRD), Japan, United Kingdom, U.S.A., and U.S.S.R. Membership for Italy and Poland is pending.

The Central Office of CODATA, housed in the U.S. National Academy of Sciences during 1966–1968, but now in Frankfurt, Germany, together with the Bureau of CODATA, implement the program generated by the committee. The $60,000 budget of the Central Office, which currently has a staff of four, is provided by dues from member nations, namely, Canada, France, Germany (BRD), Japan, U.K., U.S.A., and U.S.S.R.

THE CURRENT TASKS OF CODATA

The present tasks of CODATA are designed to implement the several aims referred to earlier. Most of these aims involve such words as encourage, stimulate, recommend, informally coordinate, etc. That is, CODATA is not seen as an operating agency which will itself compile tables. Its function is primarily to promote and improve international communication in the realm of data compilation with the aim of stimulating work where needed, suggesting coordination where indicated, developing higher standards of presentation and evaluation, and improving distribution of the products of the various centers. CODATA proposes to accomplish these purposes in various ways.

Inventory of Work in Progress. A first task of the Central Office has been to determine what is going on in the world. Without an international inventory of current data compilation activity, no intelligent action is possible. Thus, the highest priority has been given to the preparation of an International Compendium of Numerical Data Projects,[4] which is now in press. The Compendium will appear about mid-1969. Revisions are planned at intervals of two or three years.

It is interesting to compare quantitatively the amount of activity discovered in 1968, as compared with similar surveys in 1961 and 1965 by the National Research Council, Office of Critical Tables. In 1961, 1965, and 1968, the centers identified in the U.S.A. numbered 35, 53, and 92, respectively. In the world at large, the corresponding totals for the same years were 43, 73, and 153. The rate of increase is gratifying, but it is estimated that currency of coverage for all fields of the physical sciences and engineering will require a three- or four-fold increase of effort.

Communications. Improved communications among compilers are brought about by CODATA by several methods. The following procedures are working effectively.

Union and national members of CODATA have exerted a stimulative influence in their unions or countries. Most of the countries now have a National Committee for CODATA attached to the national body adhering to CODATA—e.g., in the U.S.A., the National Academy of Sciences; in the U.K., the Royal Society; in Japan, the Science Council of Japan; and in the U.S.S.R., the Academia Nauk. These national committees not only serve as advisory bodies to their governments, but in several instances have organized national conferences to focus attention on the numerical data problem. The Royal Society has held such a conference, and in late August of 1969 the Polish Academy of Sciences in Warsaw will hold an International Symposium on Numerical Data for Science and Technology.

A second kind of communication resulting from the existence of CODATA is that managers of national numerical data programs are regularly invited to attend CODATA meetings as observers, and thus have the opportunity to exchange views and to work toward coordination and more uniform standards. In this way, representatives of the Office of Standard Reference Data of the National Bureau of Standards, of the numerical data program of the Office for Scientific and Technical Information of the U.K., and of the State System of Standard Reference Data of the U.S.S.R. have become acquainted and are in continuing communication. This process has also highlighted the fact that some major countries do not have effective governmental focal points for exchange of views in this area, but that the situation in several countries is improving.

A third type of communication program is the holding of relatively large international conferences. The First International CODATA Conference was held in the Taunus Hills at Arnoldshain, Germany, near Frankfurt, in the first week of July 1968. Attended by about 100 experts from 13 countries and aided by the "Gordon Conference" type of meeting arrangements, including a snack bar for evening discussions, it was a lively and stimulating meeting. German governmental interest was expressed by generous financial support to help foreign attendees. The success of this meeting has resulted in planning for a Second International CODATA Conference, which will be held at St. Andrews, Scotland, in September 1970.

Task Groups. A mission of CODATA is to stimulate better and more uniform presentation of numerical data in compilations and in the primary literature; its constitution allows for the organization of special purpose task groups. Hence, the following task groups with the identified functions have been or are being organized.

The Task Group on Key Values for Thermodynamics, with membership covering five countries and two international unions, is preparing a list of key substances and associated key thermodynamic properties, the values of which are used repetitively by thermodynamic compilers the world over. The types of substance-property relationships to be codified include: the standard entropies of the elements; the standard heats of formation of water, carbon dioxide, and other important oxides; the heats of solution of hydrogen halides. The list produced will hopefully be approved by CODATA and will have a status similar to that of internationally approved fundamental constants and atomic masses. Other task groups formed or in process of formation are: the Task Group on Computer Use; the Task Group on Fundamental Constants; and the Task Group on Data for Chemical Kinetics. The Task Group on Computer Use is currently completing a survey in several countries of computer applications to numerical data generation, storage, and retrieval.

Worldwide Network of Numerical Data Centers. Coordination of effort among compilers in each discipline in all countries is desirable. CODATA is currently studying the possibility of designating a given center or several centers to act as the focal coordinating point(s) on an international basis for a given discipline. For example, one center might be considered the tie-point for all crystallographic compilation work; and two centers—one in the U.S. and one in the U.S.S.R.—might take the responsibility in basic data for chemical thermodynamics. With the designation of a number of such centers for various fields, the CODATA Central Office would become the point where information about all the centers, subsidiary centers, and their publication output would be available. This plan may or may not materialize, but CODATA is giving it serious consideration.

Newsletter. The Central Office of CODATA will issue a newsletter about twice a year. It will provide a forum for information on new developments, for dissemination

of approved task group reports, and other types of information of interest to the world community of compilers. The distribution list will include CODATA members, national committee members, data project directors, and other interested persons and groups. CODATA Newsletter No. 1 was distributed in October 1968 from the ICSU-CODATA Central Office, Westendstrasse 19, 6 Frankfurt/Main, Germany-BRD.

The foregoing review tells us that CODATA is a going concern; that it is playing a significant role in giving coherence to the needed worldwide effort to compress, evaluate, and compile the numerical data of science and technology; and that it is taking appropriate steps to improve the quality of data so compiled.

LITERATURE CITED

(1) National Research Council, "International Critical Tables of Numerical Data: Physics, Chemistry and Technology," McGraw-Hill, New York, N. Y., 1926–1933.
(2) Landolt-Börnstein, "Zahlenwerte und Funktionen aus Physik-Chemie-Astronomie-Geophysik und Technik," 6th ed. 1950–).
(3) "Tables de Constantes et Données Numériques," fondées par Ch. Marie (1909): "Constantes Sélectionnées," Pergamon, Paris, current series (1947–).
(4) International Compendium of Numerical Data Projects, Springer-Verlag, Berlin-Göttingin-Heidelberg, 1969.

A Biologically Oriented Data Retrieval System*

THEODORE LEGATT, ROBERT P. GRANDY, and SAMUEL X. DELORENZO
Schering Corp., Bloomfield, N. J.

Received November 21, 1968

A computer-based storage and retrieval system for biological data has been in use at Schering for several years. The system is intended primarily for storage and retrieval of biological data of interest to managers and laboratory scientists, and permits select printouts of cumulative information on aspects of the research program.

Before the computerized file of biological data was developed, a unit record system had been used for search requests. Although this system had been adequate for many queries, it was limited by hardware and input-data restrictions.[1] The need for a comprehensive and accessible information system led to a study to define objectives, examine available hardware systems, and reconsider input data requirements. The conclusion of the study pointed towards a computer system and a reorganization of input data as the most feasible way to accomplish the objectives.

OBJECTIVES

1. To provide means to handle effectively a growing file of biological and chemical information.
2. To provide a foundation for centralizing all information on preclinical studies.
3. To permit the evaluation of current work and assist research management in long-range planning.
4. To provide monthly and quarterly summary reports on the progress of compounds through testing.
5. To identify compounds submitted for screening in which test results had not been reported. This would permit rapid determination of the progress of the compound in the screening stage.

OPERATIONAL FLOW

Samples of compounds synthesized in the laboratories are submitted for testing through the chemical distribution center. This center acts as a central repository for samples of all intermediate and final products synthesized. Compounds are indexed and filed and are readily available. Sample loss has been virtually eliminated by this method.

The distribution center weighs out samples for specific screens and forwards them to the technical information center for processing. The technical information center prepares a "Request for Laboratory Investigation" form (RLI) which accompanies the compound to the screening area. After biological screening has been completed, each investigator forwards a completed test report to the technical information center, where it is duplicated and distributed to research personnel. All sample delivery and test report data are incorporated in the central file system.

INPUT

The sample submission form (RLI) contains standard information that identifies the sample source, the compound, and the tests requested. The chemical structure on this form is not stored on magnetic tape; instead, it is coded and incorporated in a separate optical coincidence file of structures to be reported on later.

The test report returning from the laboratories contains a maximum amount of information in a format that facilitates key punching. The information scientist and laboratory investigator collaborate in the test report design.

Memoranda, letters, and other pertinent reports are also included in the input information. They are indexed in a format similar to that used for the basic test report and can be retrieved by compound number, author, organization, date, or subject.

TECHNICAL CENTER DATA FLOW

The flow of data within the technical information center is shown in Figure 1. The various laboratory forms and

* Presented before the Division of Chemical Literature, 156th Meeting, ACS, Atlantic City, N. J., September 13, 1968.