The ICSU Committee on Data for Science and Technology (CODATA)*

FREDERICK D. ROSSINI
University of Notre Dame, Notre Dame, Ind.

Received June 24, 1970

Some historical aspects of the problem of numerical data for science and technology are reviewed, and the present status of the international picture, as viewed by the chairman of CODATA, the Committee on Data for Science and Technology of the International Council of Scientific Unions, is discussed.

The problem of numerical data for science and technology has occupied the attention of scientists for many years. The problem is burgeoning. In the past 150 years, the number of scientific journals has increased by a factor near 1000. In the past 50 years, we have seen the quantity of scientific information doubling every eight to 10 years. Each year we have of the order of 1 million scientific papers, reports, and related documents appearing.

Fifty years ago, an individual scientist had little difficulty reading substantially all of the scientific papers in the literature relating to his specialty. But today, this is not possible. If a scientist devotes his time to reading all the papers in his field, he will not have time to do anything else. Now he can, of course, further subdivide his field of interest and hold his own that way. But there is a limit to the continued splintering of science—learning more and more about less and less.

Something different is required: In principle, the solution is simple—one arranges for the original literature to be reviewed and appraised by experts devoting a substantial part of their time continuously to such work. Then we can have the bench scientist or engineer digest these highly condensed appraisals. In this way, by interposing a system of expert appraisal one makes it possible for the bench scientist to assimilate fully all of the new knowledge being generated in his field.

In the field of numerical data of science, this system of appraisal results in the preparation of critically evaluated tables of reference data. This work constitutes an intellectual task of high order. Though machines, automatic devices, and high-speed computers can aid and expedite the work, effective and efficient performance requires management by skilled scientists. To ensure that one will have enough competent scientists engaged in this work, one must see that they are accorded appropriate high status in the scientific community, with corresponding compensation.

*Presented at the Symposium on Communications in Chemistry and Chemical Technology, sponsored by the Division of Chemical Literature, ACS/CIC Joint Meeting, Toronto, Canada, May 25, 1970.

THE PROBLEM

Observation and measurement, which give rise to numerical data, constitute the life-blood of science. The precision and accuracy of measurement have increased enormously with advances in science. Over the past four centuries, for example, the precision of measurement of length has increased by a factor of about 1 million. Today, much of the numerical data appearing in the literature is of very high precision and accuracy. In the appraisal of such data, it is important that all of the accuracy and precision be preserved in the transition from the original record to the final compilation. The appraiser must be alert to identify the fundamental constants used by the investigator, how he calibrated his instruments for measuring length, mass, time, energy, temperature, pressure, etc., and what values he may have used for the reference substances involved. The appraiser must derive from the detailed content of all the reports that weighted value which is most likely to be near the true value for the property under investigation.

This system of appraisal by qualified experts has great advantage over any system involving the efforts of scientists working only sporadically on the literature. The organized effort will normally produce results of much higher quality. This is very important for our industrial technology today, where the precise control of temperature, pressure, and other variables makes possible the conduct of industrial processes heretofore considered impossible.

A brief sketch is given of some of the earlier efforts on the compilation of data for science and technology in various countries, including the U.S.A., so that one can better appreciate the origin of CODATA.

EARLIER WORK

The "Landolt-Bornstein Tabellen," headquartered in Germany, first appeared in 1883, with an edition of 281 pages. In 1894, came the second edition of 575 pages; in 1905, the third edition of 877 pages; and in 1912,

the fourth edition of 1330 pages; from 1923 to 1936, the fifth edition in eight volumes of 7457 pages; from 1950 to 1969, the sixth edition in 26 volumes of about 20,000 pages. Because it was no longer practical to cover simultaneously all areas of science, the Landolt-Bornstein Tables established a new series of volumes, on specialized topics. These include nuclear physics and technology, magnetic properties, astronomy and astrophysics, atomic and molecular properties, crystal and solid-state physics, geophysics, etc.

The "Tables Annuelles de Constantes et Donnees Numeriques," headquartered in France, had 10 volumes appear over the years 1910 to 1930. From 1936 to 1945, there were 40 installments, which covered the literature to 1939. Then a new title was given to this work "Tables de Constantes Selectionees," to cover specialized topics. These include nuclear physics, spectroscopy, oxidation-reduction potentials, optical rotatory power, semiconductors, and terpenoids.

The "Kaye and Laby Tables of Physical and Chemical Constants," headquartered in England at the National Physical Laboratory, have appeared in one volume, running through 13 editions, from 1911 to 1967.

The "International Critical Tables of Numerical Data of Physics, Chemistry, and Technology," headquartered in the U.S.A., at the National Academy of Sciences, came out in the years 1926 to 1933, in one edition of eight volumes and 3819 pages. Contributions to this work came from 408 scientists in 18 countries. The organization of the International Critical Tables was established in 1919 by the International Union of Pure and Applied Chemistry, 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 was the predecessor of our present International Council of Scientific Unions (ICSU), gave its blessing to the project. It was hoped that the International Critical Tables would become a continuing operation, with revisions from time to time, but, unfortunately, the editor-in-chief, Edward W. Washburn, died in 1934, and this continuity never came to pass. The importance of the International Critical Tables to science and technology of the world is evident from the fact that a significant number of sets of these tables are being sold today, 44 years after their original issue.

RECENT WORK IN THE U.S.A.

In the years 1938 to 1957, other data-compiling projects, operating on a continuing basis, came into existence in the U.S.A. These projects together involved total expenditures approaching 1 million dollars per year, and included the following: The American Petroleum Institute Research Project 44 on physical, thermodynamic, and spectral properties of hydrocarbons and related compounds; the Manufacturing Chemists' Association research project on physical, thermodynamic, and spectral properties of chemical compounds; the U.S. Atomic Energy Commission project on nuclear data; the U.S. Bureau of Mines project on thermodynamic data on metallurgically important compounds; the Purdue University Center on thermophysical properties; the Dow Chemical Company project on thermodynamic properties of selected compounds; and several National Bureau of Standards projects on thermodynamic, thermochemical, atomic and other properties of a large variety of compounds. With some change in sponsorship, nearly all these projects are still operating.

In 1957, to help coordinate the increasing number of data-compiling projects in the U.S.A., the National Research Council established the Office of Critical Tables, under Guy Waddington.

In 1963, with the support and encouragement of the Office of Critical Tables, the U.S. established, under the National Bureau of Standards, the National Standard Reference Data Program, under Edward Brady.

ESTABLISHMENT OF CODATA

With this background, let us now discuss the establishment of CODATA, its assignments, and the work it is doing.

In the spring of 1964, it was suggested that the International Council of Scientific Unions, ICSU, could provide international coordination and guidance in this field. In June 1964 ICSU established a working group to study the problem. This working group, under Harrison Brown, met in 1964 and formulated a recommendation that ICSU should establish a committee in this field. In April 1965 the recommendation was approved and the working group was requested to prepare a constitution and seek nomination for membership from unions and countries. In January 1966 the general assembly of ICSU, under the presidency of Sir Harold Thompson, approved the establishment of a Committee on Data for Science and Technology, or CODATA, as it is called, with a consititution and initial membership.

Assignments of CODATA. CODATA was given the following assignments on a worldwide basis: To ascertain, through the unions and appropriate national bodies, what data-compiling work is going on and what the needs are; To achieve coordination among and provide guidance for data-compiling projects; To encourage support for data-compiling projects by appropriate private, governmental, and intergovernmental agencies; To encourage the use of internationally approved constants, units, and symbols, and, when desirable, uniform editorial policy and procedures; To produce a compendium-directory-survey of continuing data-compiling projects and related work; To encourage and coordinate research on new forms for preparing and distributing critically evaluated numerical data.

Meetings of CODATA. Following the organizational meeting in Paris, France, in 1966, CODATA has held three annual meetings; in 1967, at Moscow, U.S.S.R.; 1968, at Frankfurt, Germany; 1969, at Rome, Italy. In 1970, the annual meeting will be held at St. Andrews, Scotland, U.K.

Meetings of the Bureau have been held at the same time and place as the annual meetings and in addition at one other place each year: 1966, at Enumclaw, Washington, U.S.A.; 1967, 1969, and 1970, at Frankfurt, Germany; 1968, at London, U.K.

Membership of CODATA. To accomplish its aims, CODATA has several categories of memberships: union members, to achieve contact with the worldwide scientific community in specific disciplines: national members, to

provide resources of funds and manpower from the technologically advanced countries; coopted members, to provide expert advice in specialized areas; liaison representatives, to maintain contact with other international organizations and other parts of ICSU interested in the work; and invited observers from governmental-funding and appropriate other organizations. At the present time, CODATA has 10 union members, nine national members, and 1 coopted member, together with a number of liaison representatives and observers that are invited to its meetings.

The union members include those from the Unions of Chemistry, Physics, Crystallography, Biological Sciences, Biophysics, Geological Sciences, Geography, Geophysics and Geodesy, Mechanics, and Astronomy.

The national members include the six original countries, France, Germany, Japan, U. K., U. S. A., and U. S. S. R., together with three new member countries, Canada, Italy, and Poland. We have extended invitations for membership also to Sweden, Netherlands, and Belgium. Inquiries have been received from East Germany, Hungary, and Israel.

Beginning July 1968, CODATA has one coopted member, in the person of Guy Waddington, the original executive director of CODATA.

Liaison representatives to CODATA have come from IAB, the ICSU Abstracting Board; UNESCO, the United Nations Educational, Scientific, and Cultural Organization; IAEA, the International Atomic Energy Agency; FID, the International Federation for Documentation; WMO, the World Meteorological Organization; FAGS, the Federation of Astronomical and Geographical Sciences; OECD, the Organization for Economic Cooperation and Development; and WFEO, the World Federation of Engineering Organizations.

Observers have included chairmen and members of CODATA task groups, ICSU officials, members of the national committee for CODATA of the country serving as host for the annual meeting, and representatives of government agencies from the various countries.

Bureau of CODATA. The Bureau of CODATA, until the coming annual meeting in September 1970, is composed as follows: Frederick D. Rossini, U. S. A, president; Sir Gordon Sutherland, U. K., and Boris Vodar, France, vice-presidents; Wilhelm Klemm, Germany, secretary-treasurer; Masao Kotani, Japan, and M. A. Styrikovich, U. S. S. R., additional members. The officers will change after the 1970 annual meeting. Also, at the meeting a seventh member of the Bureau will be elected—probably to represent the smaller countries.

Central Office of CODATA. The Central Office of CODATA was operated for its first two years (1966–68) in Washington, D. C., U. S. A., with Guy Waddington as executive director. Beginning in July 1968, the central office of CODATA was transferred to Frankfurt, Germany, with Christoph Schafer, who had spent one year with Dr. Waddington in Washington, as the new executive director. The staff of the central office has now been expanded to five persons. To maintain the international character of the operation, the five persons in the central office come from four countries: two from Germany, and one each from France, U. K., and U. S. S. R.

Funding of CODATA. The funds for the expenses of CODATA have so far come from two sources—about $5,000 per year from the parent body, the International Council of Scientific Unions, to defray the cost of meetings of the Bureau; $60,000 to $70,000 per year from annual dues from the countries holding national membership, to cover the other expenses. The annual dues of national members are fixed according to a schedule based essentially on the UNESCO scale. By 1972, the cost of operating CODATA will amount to more than $100,000 per year, at which time additional funds, to include the cost of operating task groups, will be sought from a modest increase in the annual dues and from other sources, such as UNESCO.

MISSION OF CODATA

CODATA is concerned with critically evaluated numerical data on a worldwide basis. All such compilations, taken together, should have the following characteristics:

Cover all substances and all properties of interest to all sectors of the scientific and technical community.

Be fully self-consistent with all physical relations and with the internationally approved constants, units, symbols, and nomenclature.

Have an accepted standard order of arrangement, understood and usable at the working level by scientists in all countries.

Be produced in forms needed by the various sectors of the scientific and technical community.

Be produced by scientists of high capabilities, adequately compensated.

Be maintained up-to-date by revision at appropriate intervals.

Be adequately supported from governmental and private industrial sources.

Be readily available at reasonable cost in any part of the world scientific community.

To carry out its over-all mission, CODATA must recognize the existence of data-compiling projects throughout the world. Between and among all these projects and its central office, CODATA must provide communication, linkages, and coordination.

Communication. CODATA provides needed communication among all segments of the scientific community interested in data for science and technology in four ways:

First, with the CODATA Newsletter, containing news of the activities of CODATA and of data-compiling work, issued twice each year, to scientists throughout the world, without charge.

Second, with the CODATA Bulletin, containing reports of CODATA task groups, issued irregularly, to scientists throughout the world, without charge.

Third, with the CODATA International Compendium of Numerical Data Projects, containing information about data-compiling projects throughout the world. The first edition of 295 pages, issued in November 1969, was published by Springer, Germany, at $13.20 U. S. dollars or 48 DM per copy.

Fourth, with the CODATA International Conferences on the Generation, Compilation, Evaluation, and Dissemination of Numerical Data for Science and Technology. The first conference was held in 1968 in Germany; the second will be held in September 1970, in Scotland. These are "Gordon-type" conferences, attended by about 100 scientists from many countries, and provide a worldwide meeting ground for specialists in various scientific areas.

Linkages. CODATA can provide linkages to the data-compiling projects throughout the world by listing them in the CODATA International Compendium and by communication with them by surface mail, air mail, cable, telephone, telecommunication, etc. At some stage, CODATA will be faced with the delicate and sensitive problem of giving some label of recognition to those projects producing acceptable compilations of data. This will be a difficult problem.

Coordination. CODATA can provide coordination among all the data-compiling projects by means of two kinds of task groups: General task groups, each having cognizance over a subject that is of interest and concern to data-compiling projects in all areas of science and technology as, for example, task group on computer use, and task group on fundamental constants; and area task groups, each having cognizance over a given area of science, as for example, task group on thermodynamics, and task group on chemical kinetics.

A word now about the general task groups. Each of these task groups will cover a broad general subject which will concern all data-compiling projects to some degree. The membership of each general task group consists of leading experts of the given subject from different countries, so that their final recommendations will carry appropriate weight with the scientific and technical community of the world. Also, given a limited membership, appropriate rotation of members is desirable to secure reasonable representation throughout the countries.

With regard to area task groups, eventually CODATA may have eight or 10 such task groups to cover sustantially the entire spectrum of science with regard to numerical data. These could include, in addition to the two mentioned, others such as the following: nuclear properties; atomic and molecular properties; spectral properties; ordinary physical properties; special solid-state properties; mechanical and related properties; etc. I believe that CODATA has an obligation to provide an umbrella or cover for numerical data for substantially all areas of science, and to work down in depth in each area, in accordance with its importance, as more and more resources become available to CODATA.

The membership of each area task group should consist of the directors of the principal data-compiling projects of the given area from different countries, with some representation of expert users. With such membership, the final recommendations of the task group will carry appropriate weight among the data-compilers as well as with the total scientific and technical community interested in the given area.

CONCLUSION

One important point to note here is that CODATA is concerned directly with projects that compile numerical data, not directly with projects that create data by laboratory experimentation. It is the responsibility of the data-compiling projects to be totally knowledgeable about the scientific literature of their respective fields and to be somewhat familiar with projects that create data experimentally.

Now for some general comments about the over-all plan. The basic mission of CODATA is to promote, encourage, and coordinate data-compiling projects wherever they exist in the world. The actual data-compiling work is done by each project itself. Each such project has to arrange for local funding for the cost of its work. CODATA itself does not support data-compiling work per se, but can lend its prestige in helping recognized data-compiling projects obtain needed financial support from their respective national sources. In this way, while benefiting from the communication, linking, and coordination provided by CODATA, as well as its prestige, each data-compiling project maintains essentially complete local autonomy for financial responsibility, scientific direction, and advisory services.

A true data-compiling project is one which provides a product useful to 98 or 99% of the scientific and technical community of the given area, rather than a product useful only to the 1 or 2% of that community represented by the top experts in the given field. Working scientists and engineers need reliable numbers and are willing to trust the appraisal and judgment of the qualified compilation experts that provide the critically evaluated tables of standard reference data.

Finally, it is very important for all of us to work toward international coordination and cooperation, with appropriate local autonomy, to achieve the desired goal of providing numerical data of high quality and in adequate quantity, for science and technology through the world.