Historical Background of Data Compiling Activities*

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

Received October 31, 1966

The work on the collection, critical analysis, and compilation of numerical data for science and technology, that has been carried on in the world, including the International Critical Tables, the "Landolt-Börnstein Tables," the Annual Tables, the establishment of the Office of Critical Tables of the National Research Council under the National Academy of Sciences, the establishment of the National Standard Reference Data Program of the National Bureau of Standards, and the establishment of the International Committee on Data for Science and Technology under the International Council of Scientific Unions, is reviewed.

Data compiling activities are among the most important areas of science today—the marshalling of known facts and extracting from them reliable data for science and technology. Since it is not possible to cover here all the work that has been done in this field, this paper will be concerned mainly with what we have come to call critical tables of standard reference data, covering general fields.

In the very early days of science, it was possible for a person to know substantially all the science that existed. But as science developed, the scientist began to limit his area of concern to one of what we now term separate disciplines, such as chemistry or physics. With the further expansion of science, a scientist in one discipline, such as chemistry, found it necessary to confine his attention to one branch of that discipline, like physical chemistry or organic chemistry, in order to keep abreast of what was going on. Now we are in the midst of an explosion of scientific knowledge, said to be doubling every eight years. Today we find that a chemist concerns himself not with all of chemistry or even one branch of it, such as physical chemistry, but only with one specialized area of this branch, such as thermodynamics. Even in this one specialized area, a chemist must work hard to keep himself fully up to date, because of the great expansion of scientific information.

When an investigator in a specialized area needs to know the present status of a field peripheral to his own, he usually has difficulty on two counts: He will not be able to digest all of the literature relevant to this peripheral field because of its magnitude; and he will usually not possess the expertise necessary to reduce the enormous literature of the given field to some understandable and useful meaning.

ATOMIC WEIGHTS

An example of a specialized area peripheral to most others about which nearly every chemist needs to have up-to-date knowledge is the area of atomic weights. It is, of course, not possible for every chemist to become expert in the field of atomic weights. Long ago, this problem was solved for chemists by having an international group of experts, constituting the Commission on Atomic Weights under the International Union of Pure and Applied Chemistry, maintain a continually up-to-date table of internationally agreed-upon atomic weights.

FUNDAMENTAL CONSTANTS

Let us look at another example: A chemist working in physical chemistry in the specialized area of thermodynamics has need to know the present best values of the fundamental constants, such as the Faraday constant, the velocity of light, the Avogadro number, the gas constant, etc., because he uses the values of one or more of these constants in presenting the results of his own experimental observations or theoretical calculations. Until about 50 years ago, each scientist essentially had to fend for himself in seeking the best values for the fundamental constants. Then, in 1926 appeared Volume I of the International Critical Tables (*1*), which presented a complete list of recommended values of the fundamental constants. But this list quickly became out of date. Meanwhile, Birge, working independently, tackled the problem and published several papers on recommended values of the fundamental constants over the years 1929 to 1945 (*2, 3, 4*). Du Mond and Cohen, also working independently, published several sets of recommended values of the fundamental constants over the years 1948 to 1955 (*5, 6, 7*).

Meanwhile in 1952, the Subcommittee on Fundamental Constants of the Committee on Physical Chemistry of the National Research Council published a recommended set of values of the fundamental constants for chemistry. But in 1955, the absolute temperature of the triple point of water was defined by international agreement (*9*), and in 1961 came the international agreement to unify the scales of atomic weights of the chemists and of the physicists (*10*). Both of these changes made the 1952 list

* Presented before the Division of Chemical Literature, Symposium on Compilations of Data on Chemical and Physical Properties of Substances, 152nd National Meeting of the American Chemical Society, New York, N. Y., Sept. 12, 1966.

of fundamental constants obsolete. Meanwhile, the National Research Council had established a Council-wide Committee on Fundamental Constants which encouraged Du Mond and Cohen to prepare a new list of recommended values of the fundamental constants, so that international agreement might be obtained through the National Academy of Sciences. By 1963, this NRC Committee on Fundamental Constants managed to get international agreement, from the International Union of Pure and Applied Chemistry and the International Union of Pure and Applied Physics, on an up-to-date list of recommended values of the fundamental constants. This problem now is to maintain this set of fundamental constants up to date.

NUMERICAL DATA FOR SCIENCE AND TECHNOLOGY

Now there is a third area of specialized knowledge to which full access must be available for every scientist in accordance with his needs—that is, the area of numerical data for science and technology. The analysis and appraisal of quantitative scientific information is an intellectual task of high order. Though it can be aided by machines, automatic devices, and high-speed computers, such work can be performed effectively and efficiently only by skilled scientists expert in that field. The advance of science, technology, and industry rests heavily upon the quantitative scientific information we identify as critical tables of standard reference data.

About 150 years ago, there were in the world about 50 scientific journals; 100 years ago, about 500; 50 years ago, about 5000; and today, on the order of 50,000. Now, more than one million scientific or technical papers appear in publication each year in journals, bulletins, and related documents. A large percentage of these papers contains numerical data on the properties of substances, which must be culled, analyzed, and correlated, to obtain the selected best values which constitute critical tables of standard reference data.

For quantitative scientific information of this kind, the entire communication process may be pictured in four steps: (a) the generation and publication of the data by the original investigator; (b) the collection, calculation, analysis, and correlation of the data, and the compilation of critical tables of standard reference data; (c) the dissemination of these critical tables in appropriate form; and (d) the use of the tables by the scientific and technical community.

It is impractical and much too costly to have each end-using investigator, or even each laboratory, develop an individual set of tables of selected best values of the physical and thermodynamic properties. The solution is to interpose, between the original literature and the end-using scientist, a system of review, appraisal, and compilation by qualified experts.

These critical tables of standard reference data should be produced by staffs of competent experts adequately supported, so that we can have efficient utilization of our scientific manpower. The scientists in research laboratories in educational, governmental, and industrial organizations will then have a maximum of time for pursuing the work of their respective missions, fortified with the knowledge that they have at their disposal essentially all of the existing quantitative information in the literature in the form of standard reference data, critically prepared by experts. The number of manhours of scientific time that could be saved in our laboratories would be incalculable. But what is more important, the numerical values available would be of much higher quality than can be produced by the sporadic effort of scientists primarily interested in other problems. This latter point is very important in the highly competitive technological world of today, where the precise control of temperature and pressure makes possible the conduct of industrial reactions and processes heretofore considered impossible.

With this statement of the problem, let us now turn our attention to the record and see just what has happened over the years. As might well be expected, the first attempts at compiling tables of numerical scientific data were by individual workers or groups of workers operating privately or with the support of governmental or industrial organizations. Until about 30 years ago, an appreciable amount of such work was done by individual workers in their spare time. The first attempt at a national coordination of work in this field came about 10 years ago with the establishment of the Office of Critical Tables by the National Academy of Sciences–National Research Council. The first real attempt at international coordination on a broad scale came this year with the establishment, under the International Council of Scientific Unions, of an International Committee on Data for Science and Technology.

EARLY TABULATIONS

One of the early important general compilations were the "Landolt-Börnstein Tables" (*11*), the first edition of which appeared in 1883, with 281 pages in one volume. The second, third, and fourth editions were also in one volume, coming out in the years 1894, 1905, and 1912, with 575, 877, and 1330 pages, respectively. The fifth edition comprised eight volumes, appearing over the years 1923 to 1936, containing a total of 7457 pages. The sixth edition was in 11 volumes and appeared in the years 1950 to 1959, containing a total of 8492 pages. It is my understanding that, because of the magnitude of the task of preparing a complete general compilation, the "Landolt-Börnstein Tables," following the sixth complete edition, are issuing only separate volumes covering specialized topics.

Another enterprise related to the problem of quantitative scientific information was the "Annual Tables of Constants and Numerical Data" (*12*), involving about 10 volumes appearing over the years 1910 to 1930, produced at Paris, France. These early "Annual Tables" constituted simply a transcription of the numerical data from the original literature into one central volume or volumes. Under new editorship, the Office of the Annual Tables at Paris is issuing compilations of data in specialized areas, with appropriate appraisal (*26*).

The most notable example of an effort to provide critical tables of standard reference data is the International Critical Tables of Numerical Data of Physics, Chemistry, and Technology (*1*). The International Critical Tables resulted from discussions begun at the 1919 Conference of the International Union of Pure and Applied Chemistry at London. For this project, the United States was assigned the financial and editorial responsibility. The National Academy of Sciences–National Research Council accepted the executive, editorial, and financial responsibilities for the United States, and, with the cooperation of industry, through the American Chemical Society and the American Physical Society, created a Board of Trustees to take charge of the financial and business administration and a Board of Editors, under Edward W. Washburn, to supervise the preparation of the tables proper. The entire enterprise was made possible by a fund of $200,000, contributed largely by approximately 200 industrial companies of the United States. This famous collection of numerical data was the result of cooperative efforts by 408 scientists in 18 different countries. Seven volumes with a total of approximately 3500 pages were published in the years 1926 to 1930, with a separate index volume appearing in 1933. This was the broadest single compilation job in history. The Board of Trustees of the International Critical Tables had hoped that they would become an established institution, with supplements and revisions from time to time, but unfortunately this never came to pass.

API RESEARCH PROJECT 44

Another important general compilation, though limited to one class of substances, is that of the American Petroleum Institute Research Project 44 (*13*) on the physical, thermodynamic, and spectral properties of hydrocarbons and related compounds. This project was authorized in 1941 and began operations in 1942 at the National Bureau of Standards. In 1950, the Project moved with its director to the Carnegie Institute of Technology, and in 1961 from there to Texas A and M University. The Project has published two bound volumes, one in 1947 (*18*) and one in 1953 (*19*). As of January 1, 1966, the number of loose leaf data sheets extant was 2344 on physical and thermodynamic properties and 7226 on infrared, ultraviolet, Raman, mass, and nuclear magnetic resonance spectra.

A companion project to the API Research Project 44 was established in 1955 by the Manufacturing Chemists Association, and operated under the same direction, on the properties of chemical coupounds, until 1966.

THERMODYNAMIC DATA

There is one class of data where critical appraisal and selection of best values is particularly important. This is the class of thermodynamic data. Here, one is required to have values which are not only consistent among themselves but also consistent with the relations of thermodynamics, physics, and chemistry. Because of the particular importance attached to these tables, a summary is presented of what has been done in this area, apart from what has appeared in the International Critical Tables and the "Landolt-Börnstein Tables."

One of the earliest compilations, comprehensive for that time, covered heats of formation. It was put together by Julius Thomsen at Copenhagen and published in four volumes 1882 to 1886 (*14*).

In 1923 appeared the thermodynamic tables of Lewis and Randall, published in the famous text entitled "Thermodynamics and the Free Energy of Chemical Substances" (*15*). In 1936 Bichowsky and Rossini published their book on "Thermochemistry of the Chemical Substances" (*16*). The thermochemical tables of Bichowsky and Rossini were revised and extended in the National Bureau of Standards Circular 500 entitled "Selected Values of Chemical Thermodynamic Properties," by Rossini, Wagman, Evans, Levine, and Jaffe. This compilation was completed in 1950 and appeared in publication in 1952 (*17*). Beginning in the middle 1930's and continuing for several decades, work on compilations of the thermodynamic properties of metallurgically important substances was being done by Kelly (*20*) under the auspices of the U. S. Bureau of Mines at Berkeley, Calif. The results of this work have appeared in the form of a number of monographs published through the U. S. Government Printing Office.

Within the past several years, several compilations of thermodynamic properties of special classes of substances by Hultgren (*21*) and by Stull (*22*) have become available in the open literature. Also within the past year, the first two parts of the total revision of National Bureau of Standards Circular 500 have appeared, with others to follow (*23*).

OFFICE OF CRITICAL TABLES

Meanwhile, by 1955, formal compilation projects were being carried on in several countries, expressing the urgent need of science and technology for more up-to-date numerical data in support of the research programs being accelerated everywhere.

In the United States, which had been the headquarters of the International Critical Tables, special situations took place. The National Research Council established a Committee on Tables of Constants in the early 1940's. Although this Committee considered it desirable to have a revision of the International Critical Tables, it saw no ready solution to that problem.

In 1955, the National Research Council Committee on Tables of Constants concluded that there was no hope of repeating the work of the International Critical Tables, as one large compilation job, for the following reasons:

1. Since the appearance of the International Critical Tables in 1926 to 1930, the fields of chemistry, physics, engineering, technology, and other sciences have expanded in size and created new areas, requiring many times more and newer data than before.
2. The precision of measurement in science and the precision of manufacturing in industry have been pushed to another magnitude, requiring more accurate data of greater precision.
3. A rough estimate indicated that an adequate and complete revision and extension of the International Critical Tables would be a job 100 to 200 times as great as the original job, meaning that, counting the change in the value of the dollar, nearly $50,000,000 would be required.

4. From 1938 to 1955, a number of large data compiling projects operating on a continuing basis had come into existence in the United States, involving total annual expenditures approaching one million dollars. These projects included: the American Petroleum Institute Research Project 44 on hydrocarbons and related compounds; the Manufacturing Chemists Association Research Project on chemical compounds; the Nuclear Data Project of the United States Atomic Energy Commission; the compilation work on metallurgically important compounds of the U. S. Bureau of Mines; the Thermophysical Properties Center at Purdue University; and several compilation projects on thermochemical, thermodynamic, and other properties at the National Bureau of Standards.

It was also important to note that any new undertaking of this kind must provide for one essential element lacking in the International Critical Tables—namely, firm provision for continuity.

On consideration of the foregoing points, the National Academy of Sciences-National Research Council decided that what was needed was a new plan for providing science and industry in the United States with continuing up-to-date critical tables of data. There was then established in the National Research Council an Office of Critical Tables, with the following responsibilities:

1. To survey the needs of science and industry for critical tables of numerical data.
2. To stimulate and encourage, and expand as appropriate, existing critical data compiling projects.
3. To promote uniform editorial policy and procedures, and high standards of quality.
4. To provide a directory of continuing critical data compiling projects.
5. To assist in the establishment of needed critical data compiling projects for new scientific areas.

The Office of Critical Tables of the National Research Council began operation late in 1957 with Dr. Guy Waddington as Director. In the first six years, significant progress was made in carrying on the first four objectives listed above. But the fifth objective required considerable funding and little progress was made in this area.

NATIONAL STANDARD REFERENCE DATA PROGRAM

However, in 1963 a new impetus was given to the last objective stated above when funding was made available by the United States Government to the National Bureau of Standards to take over the operation for the Government of a National Standard Reference Data Program. This program is under way under the direction of Dr. Edward L. Brady. The Office of Critical Tables is working closely with the National Bureau of Standards in this work, and the advances which have been made under this program are reported elsewhere in this symposium (*24*).

INTERNATIONAL COMMITTEE ON DATA FOR SCIENCE AND TECHNOLOGY

Two years ago the work in the United States reached the point where urgent consideration needed to be given to the problem of providing an organization that can coordinate critical data compiling projects on an international level in the same way as did the Office of Critical Tables of the National Research Council in the United States.

Efforts initiated two years ago by the Office of Critical Tables of our National Research Council have resulted in the establishment by the International Council of Scientific Unions of an International Committee on Data for Science and Technology. The organization meeting of this International Committee was held in Paris in June 1966. It has been agreed that the headquarters for this International Committee will be located in Washington, D. C., for two years under the direction of Dr. Guy Waddington, who will continue half-time as Director of our Office of Critical Tables under the National Research Council. Dr. Waddington has reported on the status of the international program (*25*).

CONCLUSIONS

It is appropriate to make some comments about critical data compiling projects in general, as we would like to see them throughout the world:

1. The tables should cover all substances and properties for which any information is available.
2. The tables should have a standard order of arrangement of the compounds, with arrangement, language, and symbols readily understood by workers in all countries, in any language.
3. The tables should be fully self-consistent with all physical relations and with the recommended best values of the fundamental constants.
4. The tables should be easily available at reasonable cost in any part of the world.
5. The tables should be produced by competent scientists of high capabilities, with adequate rewards for them.
6. The tables should be produced on a continuing basis and maintained up-to-date by revision at appropriate intervals.
7. The work should be adequately supported, with resources provided by both government and industry as partners in this important work.

In making assignments of areas of responsibility, we should take advantage of, and utilize to the fullest extent, the capabilities of all workers interested in, and willing to work on, the preparation of critical tables, wherever they happen to be in the world.

Also, in making assignments of areas of responsibility to different projects, we need to avoid significant duplication of effort. At the same time, however, we need enough overlapping to ensure that we get the benefits of competitive comparison of the quality of the work performed by different projects.

A point important in the entire business of producing critical tables of standard reference data is that the data must be ready at the time of need. We cannot wait until there is an important need for a specific lot of data. It will be too late. Many important industrial enterprises have been subject to very costly delays because of the lack of a few data of the right kind at the right time. Similarly, many research investigations have suffered because of a similar lack. In a related vein, unreliable data have been known to lead to great losses in industrial enterprises, and to much wasted effort in research investigations.

LITERATURE CITED

(1) "International Critical Tables," Volumes **I** to **VII**, McGraw-Hill, New York, 1926 to 1930.
(2) Birge, R. T., *Revs. Modern Phys.* **1**, 1 (1929).
(3) *Ibid.*, **13**, 233 (1941).
(4) Birge, R. T., *Am. J. Phys.* **13**, 63 (1945).
(5) Du Mond, J. W. M., Cohen, E. R., *Revs. Modern Phys.* **20**, 82 (1948).
(6) Du Mond, J. W. M., Cohen, E. R., *Phys. Rev.* **82**, 555 (1951).
(7) Cohen, E. R., Du Mond, J. W. M., Layton, T. W., Rollett, J. S., *Revs. Modern Phys.* **27**, 363 (1955).
(8) Rossini, F. D., Gucker, F. T., Johnston, H. L., Pauling, L., Vinal, G. W., *J. Am. Chem. Soc.* **74**, 2699 (1952).
(9) Advisory Committee on Thermometry of the International Committee on Weights and Measures (1955).
(10) Report of the International Commission on Atomic Weights (1961), *Pure Appl. Chem.* **5**, 255 (1962).
(11) "Landolt-Börnstein Tabellen," Springer, Berlin.
(12) "Annual Tables of Constants and Numerical Data," Paris.
(13) American Petroleum Institute Research Project 44, Texas A and M University.
(14) Thomsen, J., "Thermochemische Untersuchungen," Volumes **I, II, III, IV**, Barth, Leipzig, 1882, 1882, 1883, 1886.
(15) Lewis, G. N., Randall, M., "Thermodynamics and the Free Energy of Chemical Substances," McGraw-Hill, New York, 1923.
(16) Bichowsky, F. R., Rossini, F. D., "Thermochemistry of the Chemical Substances," Reinhold Publishing Corp., New York, 1936.
(17) Rossini, F. D., Wagman, D. D., Evans, W. H., Levine, S., Jaffe, I., "Selected Values of Chemical Thermodynamic Properties," National Bureau of Standards Circular **500**, U. S. Government Printing Office, Washington, D. C., 1952.
(18) Rossini, F. D., Pitzer, K. S., Taylor, W. J., Ebert, J. P., Kilpatrick, J. E., Beckett, C. W., Williams, M. G., Werner, H. G., "Selected Values of Properties of Hydrocarbons," National Bureau of Standards Circular **461**, U. S. Government Printing Office, Washington, D. C., 1947.
(19) Rossini, F. D., Pitzer, K. S., Arnett, R. L., Braun, R. M., Pimentel, G. C., "Selected Values of Physical and Thermodynamic Properties of Hydrocarbons and Related Compounds," Carnegie Institute of Technology Press, Pittsburgh, 1953.
(20) Kelley, K. K., Bulletins of the U. S. Bureau of Mines.
(21) Hultgren, R. R., "Selected Values of the Thermodynamic Properties of Metals and Alloys," McGraw-Hill, 1963.
(22) Stull, D. R., "Joint Army, Navy, and Air Force Thermochemical Tables," PB 168370, U. S. Government Clearing House for Publications, Washington, D. C., 1965.
(23) Wagman, D. D., Evans, W. H., Halow, I., Parker, V. B., Bailey, S. M., Schumm, R. H., "Selected Values of Chemical Thermodynamic Properties," Parts I and II, National Bureau of Standards, Technical Notes 270-1, 270-2, Washington, D. C., 1965–66.
(24) Brady, E. L., "The National Standard Reference Data System," J. CHEM. DOC. **7**, 6 (1967).
(25) Waddington, G., "A World System of Evaluated Numerical Data for Science and Technology," *ibid.*, p. 20.
(26) "Tables de Constantes et Données Numeriques," Volumes **1** to **14**, Gauthier-Villars, Paris, 1947–1965.

The National Standard Reference Data System*

EDWARD L. BRADY
Office of Standard Reference Data, Institute for Basic Standards,
National Bureau of Standards
Washington, D. C. 20234

The National Standard Reference Data System is a government-wide effort to give to the technical community of the United States optimum access to quantitative data on the physical and chemical properties of substances and their interactions, critically evaluated and compiled for convenience. The general functions of the system are to coordinate and integrate existing data evaluation and compilation activities into a systematic, comprehensive program, supplementing and expanding technical coverage when necessary, establishing and maintaining standards for the output of the participating groups, and providing mechanisms for the dissemination of the output as required. The plan of operation and the general status of activities initiated by the National Bureau of Standards are described.

"Of course, it's important, but I wouldn't want to do it myself." This is a view of data compilation activities expressed by many, or perhaps most, scientists when their possible interest in undertaking a particular project is discussed with them. It is significant, however, that this is probably not the view of many of the leading and most creative specialists in most fields of physical science.

* Presented before the Division of Chemical Literature, Symposium on Compilations of Data on Chemical and Physical Properties of Substances, 152nd National Meeting of the American Chemical Society, New York, N. Y., Sept. 12, 1966.