Technical Note

NATIONAL STANDARD REFERENCE DATA PROGRAM

Background Information

U. S. DEPARTMENT OF COMMERCE
NATIONAL BUREAU OF STANDARDS
THE NATIONAL BUREAU OF STANDARDS

Functions and Activities

The functions of the National Bureau of Standards are set forth in the Act of Congress, March 3, 1901, as amended by Congress in Public Law 619, 1950. These include the development and maintenance of the national standards of measurement and the provision of means and methods for making measurements consistent with these standards; the determination of physical constants and properties of materials; the development of methods and instruments for testing materials, devices, and structures; advisory services to government agencies on scientific and technical problems; invention and development of devices to serve special needs of the Government; and the development of standard practices, codes, and specifications. The work includes basic and applied research, development, engineering, instrumentation, testing, evaluation, calibration services, and various consultation and information services. Research projects are also performed for other government agencies when the work relates to and supplements the basic program of the Bureau or when the Bureau's unique competence is required. The scope of activities is suggested by the listing of divisions and sections on the inside of the back cover.

Publications

The results of the Bureau's research are published either in the Bureau's own series of publications or in the journals of professional and scientific societies. The Bureau publishes three periodicals available from the Government Printing Office: The Journal of Research, published in four separate sections, presents complete scientific and technical papers; the Technical News Bulletin presents summary and preliminary reports on work in progress; and the Central Radio Propagation Laboratory Ionospheric Predictions provides data for determining the best frequencies to use for radio communications throughout the world. There are also five series of nonperiodical publications: Monographs, Applied Mathematics Series, Handbooks, Miscellaneous Publications, and Technical Notes.

A complete listing of the Bureau's publications can be found in National Bureau of Standards Circular 460, Publications of the National Bureau of Standards, 1901 to June 1947 ($1.25), and the Supplement to National Bureau of Standards Circular 460, July 1947 to June 1957 ($1.50), and Miscellaneous Publication 240, July 1957 to June 1960 (includes Titles of Papers Published in Outside Journals 1950 to 1959) ($2.25); available from the Superintendent of Documents, Government Printing Office, Washington 25, D.C.
NATIONAL BUREAU OF STANDARDS

Technical Note 194

ISSUED JUNE 1963

NATIONAL
STANDARD REFERENCE DATA
PROGRAM

Background Information

Prepared by
Standard Reference Data Committee
National Bureau of Standards
U.S. Department of Commerce

NBS Technical Notes are designed to supplement the Bureau's regular publications program. They provide a means for making available scientific data that are of transient or limited interest. Technical Notes may be listed or referred to in the open literature.

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington 25, D.C. Price 25 cents
Excerpt from an Announcement by the
Office of Science and Technology, Executive Office of the President

June 7, 1963

Dr. Jerome B. Wiesner, Director of the Office of Science and Technology in the Executive Office of the President, today announced the establishment of a National Standard Reference Data System. This national effort will be administered by the National Bureau of Standards, which is already compiling standard data. The system integrates to a single point of responsibility the present data-compiling activities of the National Bureau of Standards, Department of Defense, Atomic Energy Commission, National Aeronautics and Space Administrations, the National Science Foundation, and several other agencies.

The announcement follows the action taken by the Federal Council for Science and Technology, a government-wide group of top level agency officials in science and technology, based upon recommendations of its Committee on Scientific Information. According to Adm. Charles B. Martell of the Office of the Secretary of Defense, Chairman of the Committee on Scientific Information, the intent is to provide an articulated system of activities under such coordination and direction as to ensure a compilation of data meeting quality standards, and also to maintain continuous policy guidance of the system at the level of the Executive Office. The need to improve scientific and technical information in the Federal government has been increasingly apparent during recent years.
Every day the working scientist or engineer is finding it harder to locate the specific information he needs when he needs it. More important, even when he does find it the information is often incomplete, the data from separate sources may be incompatible and, unless he is a specialist in the field, he may find it impossible to judge its reliability and authority. The problem arises from the enormous increase in the volume of scientific and technical literature generated each year and the consequent difficulty experienced by the specialist in trying to keep abreast of all the literature relevant to his field. This problem in communication undoubtedly results in increased costs for research and development, waste of technical manpower, duplication of effort, and delay in the industrial application of scientific and technological advances.

The general problem of the transfer of scientific and technical information has been studied by a number of groups, including the President's Science Advisory Committee, the Federal Council on Science and Technology, the Senate Committee on Government Operations, the House of Representatives Committee on Science and Astronautics, the National Science Foundation, and the National Academy of Sciences-National Research Council. In addition, many professional societies have standing committees working on information transfer problems in specific areas.

This report is concerned with a single important aspect of the general science-information problem; namely, the urgent need for compilation of Standard Reference Data. Standard reference data is defined as critically evaluated data on the physical and chemical properties of materials, authoritatively documented as to reliability, accuracy and source. The value of tabulations of such data is far reaching and well recognized: data of known reliability are thereby made conveniently available for use by the scientist or engineer to design an experiment or a piece of hardware; the individual worker is relieved of the necessity of searching the literature and attempting to evaluate data in fields in which he is not expert; areas in which additional work is needed become more clearly defined; and relationships not previously apparent are recognized.

The National Bureau of Standards, as well as other organizations in this country and abroad, has been active in the compilation of standard reference data for many years. However, in view of the great accumulation of data over the past few years, the present rate of production of new data, and the urgent needs of American science and industry, it is apparent that a substantially greater effort is demanded. At present there is no single organization which is planning and coordinating standard reference data activities on a national basis. As a result, the nation's storehouse of such data is not being effectively utilized to assist in the rapid, orderly development of science, technology, and the national economy. This report proposes a program to provide the standard reference data that are needed in the physical sciences and engineering.
Size and Importance of the Problem

To gain an idea of the size of the problem, let us consider some figures. There are about 35,000 recognized scientific and engineering journals in the world today, producing each year about 1,000,000 research papers dealing with the physical sciences and engineering. About 200,000 of these papers contain data of lasting value. Ultimately, data from this smaller group of papers are critically evaluated and compiled for publication in handbooks and other basic sources of standard reference data. The average lag in time, however, between publication of the data in a journal and its ultimate appearance in a compilation is estimated to be about ten years, leaving about 2,000,000 papers in the backlog of un-evaluated data sources. No single institution, relying solely on its own, in-house resources, could hope to solve the problem.

When we break the numbers down on a decentralized working basis, however, the problem begins to assume manageable proportions. For example, if there were 5,000 scientists throughout this country, in universities, in government, and in private laboratories, collecting and evaluating data in specialized subject-matter areas, each individual scientist would need to scan only about 200 papers a year, he would have a backlog of about 400 papers, and even granted the doubling of scientific information output every ten years, his scanning workload would rise to only about 400 per year. This is truly a sizeable job, but not one which must be abandoned at the outset because of its enormity.

Thus, we see that the problem of providing an adequate storehouse of standard reference data is at least capable of solution through the part-time efforts of a considerable number of scientists. But is this massive investment of time and professional manpower justified by the benefits that will accrue? To answer this question, let us look more closely at the problem and the situation it creates.

Three commonly used sources of standard reference data are International Critical Tables, Handbook of Chemistry and Physics, and Perry's Chemical Engineer's Handbook (references 7, 10, and 11). In a recent study by the American Institute of Chemical Engineers (reference 18), these three sources were analyzed in terms of the availability of data on 16 important properties (such as specific heat, viscosity, thermal conductivity, and vapor pressure) for 13,150 compounds. The average number of compounds for which data were available covering any property was 5 percent, and for only one property were as many as 11 percent of the compounds covered.

Undoubtedly, many of the data do exist somewhere in the ten-year backlog of unevaluated papers. But even if there were ten times as many data in the backlog as in the handbooks, the total coverage would be no more than 55 percent, with the remaining properties yet to be determined experimentally. This situation points up two important aspects of the problem; first, the backlog of unevaluated data must be brought under
control. Second, a sound, comprehensive experimental program must be a part of any program to compile and evaluate the existing data in the literature.

In view of the poor coverage in data compilations, some companies have established a cut-off point, beyond which they will not search the literature for data they need. In other words, if they can perform the experimental measurements to obtain the data at a cost less than the cut-off point, which may be as high as $100,000, they will not bother to consult the literature. As the total cost or value to the company of a top research scientist may be as much as $200 per day, it is often less expensive to make the necessary measurements than to have him spend his time in a library seeking data which may not be available.* Although the individual company may have no choice, the resulting duplication of experimental work cannot but have a serious effect on our total national technological effort.

Lack of adequate reference data can also lead to erroneous conclusions on the economic feasibility of developing new products and new processes, and can thus retard the introduction of technological change into industrial production. The use of erroneous values or incomplete data in the design of production facilities can have serious results. For example, decisions based upon erroneous and incomplete data on the thermodynamic properties of boron compounds led to the premature construction and eventual shut-down of two chemical plants estimated at 38 million dollars each. Many factors played a role in these decisions, but lack of adequate scientific data was a major one.

In the recent Venus space shot it was found that the vehicle overheated, apparently because of errors in the thermal data used in its design. Efforts are now being made to obtain better thermal emittance data, and the Aerospace Industries Association has recommended initiation of a multimillion-dollar program devoted solely to emittance data.

Lack of data on heats of polymerization was a problem in the design of the early synthetic rubber plants, making it necessary to overdesign cooling systems. The result was a great increase in costs. Today, although progress is being made in a number of fields relevant to the technology of organic substances, there are no up-to-date reviews of polymer properties which the polymer specialist requires in developing a new material or improving the efficiency of an industrial process.

On the other hand, the petroleum industry has benefited significantly from the program to compile tables of hydrocarbon properties which has been supported for many years by the American Petroleum Institute, the

*This statement applies to simple measurements possible with equipment and staff ordinarily available in general-purpose laboratories. It does not apply to many of the more complex measurements.
National Bureau of Standards, the Bureau of Mines, and several universities. The relatively small sums that have been devoted to this research during the last three decades have led directly to such technical developments as the shipment of liquefied methane in cargo vessels, the improvements in the efficiency of the production of petroleum, and the perfection of many chemical processing and refining techniques. A detailed knowledge of the phase behavior of hydrocarbons, which was summarized in a review by Sage and Lacy as part of the API program on hydrocarbon properties, led to the realization that retrograde condensation phenomena occur in certain types of oil and gas fields. This finding, in turn, led to the development of a method of production which avoided losses of millions of barrels of light hydrocarbon condensates.

Current research directed toward the further development and application of the laser is also benefiting from compilations of such standard reference data as atomic energy levels and transition probabilities of gases and inorganic crystals. Many of the same properties are of broad interest in other areas of research and development, such as semiconductors and thermoelectric and nuclear materials.

Proposed Solution

What is needed, then, is a comprehensive program which will result in (1) the accumulation, in all fields of the physical sciences, of those quantitative data which have been determined by careful evaluation to be useful to the nation's scientists and engineers; and (2) the critical evaluation of both the significance and reliability of the accumulated data. The results of such evaluations might, in some cases, be simple statements in footnotes to tables of data. In others, they might be extensive monographs surveying the state of knowledge in a particular field. Indeed, in some cases authoritative statements as to the limits of our present knowledge will be of more value than tabulation of what is known. In each field the effort should be to evaluate and present what is known in a form that will be most useful to those who need to apply the knowledge in science and technology.

A third important objective of the program should be the creation of new knowledge in the form of correlations between existing data, interpolations and extrapolations to obtain new data, and determinations of the specific areas where further work is required. These results could then be used to stimulate research groups to produce data to fill in gaps or to confirm newly developed relationships.

We believe such a program can best be carried out by centralized coordination of a decentralized operation across the country. We envision this program as consisting of three parts: an input from scientists in many different locations, a central source of the evaluated data, and an output system geared to the needs of the nation's scientists and engineers (Fig. 1).
FIG. 1

DATA EVALUATION

WORLD LITERATURE IN PHYSICAL SCIENCES

STANDARD REFERENCE DATA CENTER

DATA OUTPUT

PERIODICAL SERVICE
SUBSCRIPTION SERVICE
REFERRAL SERVICE
CORRELATION & PREDICTION SERVICE
ANALYSIS & DATA PROCESSING SERVICE
APERIODICAL PRODUCTS
SUMMARY REVIEWS
Input: The input will come from scientists who are comprehensively reviewing the literature in their fields of specialization and critically evaluating the data for ultimate inclusion in the storehouse of standard reference data. These scientists may be in universities or in industrial or government laboratories. Some will be at NBS, others will be under contract to the Bureau. They will work singly or in small groups oriented to the traditional scientific disciplines. At the same time other scientists will be engaged in experimentally determining the standard reference data that do not exist in the literature. Again, these scientists may be at NBS, or they may be performing their work under contracts from the Bureau. Clearly, the interplay between the two groups must be close and continuous.

Central Core: The central core (Fig. 2) will consist of a Standard Reference Data Center at NBS where the evaluated data will be located, in punched cards, on magnetic tape, in notebooks, in many other forms, all mechanized for storage and retrieval. A review and control office will label the incoming data as to relative quality and reliability. The SRD Center will classify the data into as many major and minor categories as are required by the needs of the data users.

Output: The output will take the form of a series of services aimed at different technical levels and tailored to the needs of various segments of industry. In general, it will be oriented toward the application of the data, rather than toward a field of science. The output services will include:

(1) Periodical Service

designed to keep the user up to date on new data acquisitions in the SRD Center. It will provide information on the data available in the Center (but will not provide the data themselves) by means of a monthly news letter and by annual and semiannual reviews of data acquisitions.

(2) Subscription Service

in which the user pays to receive all available data on a specific subject on a continuing basis. These data packages will be designed to meet the needs of specific industries, industry groups, or Government research and development programs.

(3) Referral Service

which will handle narrow, one-time requests for data by referral to the files of the SRD Center. In general, this service will take care of needs that are not met by the other output services.
STANDARD REFERENCE DATA CENTER

ATOMIC PHYSICS
NUCLEAR PHYSICS
THERMO DYNAMICS
ETC.

QUALITY CONTROL
QUALITY CONTROL
QUALITY CONTROL
QUALITY CONTROL

FILE
FILE
FILE
FILE

DATA
PROCESSING
AND
DISSEMINATION

FIG. 2
(4) Correlation and Prediction Service
for computing values in areas where some data exist, but where requests come in for specific information not contained in the SRD Center. Values will be computed by making use of correlations based on molecular structure and the properties of related compounds.

(5) Mathematical and Statistical Service
which will offer mathematical and computer techniques to customers for evaluating new data for subsequent inclusion in the files of the SRD Center or for individual use. It will also provide techniques to assist in the Prediction and Correlation Service.

(6) Aperiodical Products
including tabulations, review monographs, review papers, computer card decks, and computer tapes. These will constitute the formal and products of the SRD Center.

(7) Summary Reviews
to provide a rapid assessment of the state-of-the-art in fields where there are few data but which must suddenly be explored because of scientific breakthroughs or crash programs.

In planning the details of the program, the needs of American industry, academic scientists, and Government laboratories must be ascertained and taken into account. Undoubtedly, limitations in funds and manpower will require establishment of a priority system of some kind. The choice of work to be undertaken from such a vast field must necessarily be made by NBS, but the Bureau will seek the cooperation and advice of the Office of Critical Tables of the National Academy of Sciences-National Research Council, the NAS-NRC Advisory Committees to the Bureau, interagency panels, expert consultants in the subject-matter areas, and working committees of the scientific and engineering societies and industry associations that are active in the field of critical data.

The report of the Weinberg Committee (reference 20) pointed out the informational cross-referencing required to take into account the fact that some information users are "mission-oriented" while others are "discipline-oriented." In planning priorities, content, and format of each of the activities that comprise the program, this "mission-discipline duality" will be given consideration. In particular, activities related to output services will cut across disciplinary lines to provide information in a form directly related to users' needs. Thus the data content of the SRD Center will be packaged in many ways, depending on the interests of the customers.
Management of the Program

In order to manage the Standard Reference Data Program, we propose to set up an Office of Standard Reference Data at NBS under a Program Director who will have over-all responsibility for the conduct of the program and who will report directly to the Director of NBS. The Program Director will be assisted by a General Advisory Committee consisting of distinguished members of the scientific community, by a Steering Committee drawn from the NBS staff, and by discipline-oriented technical coordination groups headed by NBS staff members. Initially these groups will cover the fields of atomic and molecular physics, nuclear physics, thermodynamics and transport properties, kinetics, and mechanical properties. Later other coordination groups will be formed as considered necessary.

As a means of enlisting the advice and aid of the scientific community, advisory committees in these disciplines will also be formed, comprised of representatives of several of the important data producer and user groups. Early in the planning in each major field, we plan to call a meeting of a group of experts, larger in number than the advisory committee for that field, who will be informed of the Bureau's plans and requested to make recommendations for activities in their specialty. These recommendations would then be evaluated with the aid of the advisory committee.

The Program Director's Office will include the review and control office for labeling data, and a liaison staff. However, the actual data compilation activities will not be done under the direct supervision of the central office, but will be conducted in the technical operating divisions of the Bureau or, if the necessary competence does not exist within the Bureau, by a suitable outside expert, or group of experts, under contract with NBS, as previously mentioned. Similarly, the data processing and data dissemination functions of the program will be carried out by existing activities within the Bureau, rather than by the central office.

By its very nature such a program must involve the enthusiastic participation of NBS scientists who are actively engaged in conducting research and development activities; that is, the program cannot be carried out by a separate staff having standard reference data as its sole responsibility. It is expected that ultimately a large fraction of the senior scientists at the Bureau will participate in the work. In addition, the Bureau plans to invite distinguished scientists to spend some months at the Bureau, using its technical, administrative, and information retrieval services for the purpose of producing critical reviews and compilations.

**Proposed Organizational Structure**

To carry out this program, we propose an organizational structure consisting of the following major components (see Fig. 3):
1. The Program Director's Office will consist of:

(a) The Program Director, who will be a member of the NBS staff, together with his immediate office staff.

(b) The General Advisory Committee, which will review operations of the entire program, make recommendations to the Program Director, and report to the Director of NBS. This Advisory Committee will consist of the Area Program Chairmen, as in 2 (a).

(c) The Steering Committee, drawn from the NBS staff to serve as advisors to the Program Director as required.

(d) A liaison Staff to facilitate cooperation with the various NAS-NRC organizations, and with university and industrial laboratories, scientific and engineering societies, and other Government agencies.

(e) Area Directors with their supporting staffs. The Area Directors will be NBS staff members responsible for implementation of Scientific Area Programs, based on recommendations of the Area Program Chairman, under the general supervision of the Program Director.

(f) A Product and Evaluation Officer, with supporting staff as required. This Officer will be responsible for the certification that each output of the Program (table, graph, review article, etc.) has been adequately reviewed, evaluated, and clearly labeled for accuracy and dependability.

2. Each Scientific Area Operating Program will consist of:

(a) An Area Program Chairman, who should be a scientist of high international prestige (a member of the National Academy of Sciences or of comparable stature) drawn from the community to serve on a part-time (approximately 10 percent) basis on contract. He will be responsible for program formulation, development of requirement priorities, and review of all operations in his area.

(b) One or more Area Steering or Advisory Committees, Sub-committees, or Panels, drawn from the scientific community, as required to assist and advise the Area Program Chairman.
(c) The Area Director. As noted in 1(e), he will be a member of the NBS staff working in the Program Director's Office, who is responsible for operation of the Area Program.

(d) Data Evaluation Projects operating either within NBS or outside NBS by contract between NBS and the participating laboratory. These projects will survey the literature in specific areas and will produce data analyses, tabulations, review articles, and other information services as directed by the Area Program. They will also conduct a limited number of experimental investigations when necessary to fill critical data gaps.

Each Scientific Area Operating Program will be responsible for reviewing and evaluating existing data projects and information centers in the subject matter area on a continuing basis to determine the extent of existing coverage and its adequacy. Where significant gaps in information coverage are found, it will assess their importance and priority, and will attempt to find ways of filling them by locating competent centers with available manpower and contracting for their services. The Scientific Area Operating Program will monitor the performance of the contractual services and will ensure that their output is made available to the scientific and technical community promptly and effectively.

3. The Centralized Services will be carried out initially by existing units within NBS, as projects financed by the Program, with the project leaders responsible to the Program Director for project policy and completion of assigned tasks. The following Centralized Services are planned:

(a) Data Processing Systems Service. This Service will advise the Scientific Area Programs on their operations, provide or develop equipment for them as required, perform the same functions for the Data Dissemination Service (below), advise the Program Director on foreseeable technical needs and on desirable data processing developments, and carry out such developments as directed.

(b) Data Dissemination Service. This Service will involve the publication and circulation of written, tabulated, and graphical material. It will also include inquiry answering and referral services. A variety of techniques will be employed to reach the scientific and technical public and to respond to its requests.

Funding

It is apparent the program as outlined will require a substantial amount of money to operate when fully developed. The exact amount can be
Summary of Proposed Funding

(Thousands of dollars)

<table>
<thead>
<tr>
<th></th>
<th>FY 63</th>
<th>FY 64</th>
<th>FY 65</th>
</tr>
</thead>
<tbody>
<tr>
<td>Program Director's Office</td>
<td>---</td>
<td>$60</td>
<td>$180</td>
</tr>
<tr>
<td>Centralized Services</td>
<td>---</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>Data Handling R&amp;D</td>
<td>---</td>
<td>50</td>
<td>150</td>
</tr>
<tr>
<td><strong>Area Scientific Programs</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internal to NBS</td>
<td>$810</td>
<td>1,330</td>
<td>1,600</td>
</tr>
<tr>
<td>External Contract</td>
<td>---</td>
<td>310</td>
<td>1,000</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td>$810</td>
<td>$1,800</td>
<td>$4,030</td>
</tr>
</tbody>
</table>
determined only after some years of operating experience, but it is estimated that an ultimate level of perhaps $10 million per year will be required. No estimate is available of the amount now being devoted to the scattered activities of this type in progress throughout the world.

Because the products of the program will be of broad, general service to the entire scientific and technical public, it is appropriate that the bulk of the funding be obtained as part of the regular budget of NBS. However, when certain critical reviews and compilations are of unique interest to a particular agency, it would seem proper to seek financial support from the agency concerned.

It is recommended that the program be conducted on a rather modest budget until many of the lessons to be learned from a pilot-scale operation have been assimilated. The Bureau's fiscal 1963 budget includes $810,000 for data compilation activities, all of which are now being conducted in house. We propose that this figure be increased to $2,280,000 in FY 1964 to provide support for the new program, and to $4,087,000 in FY 1965. Details of the proposed funding are given in the table.

If the program as outlined above seems ambitious in conception, it must be remembered that the problem itself is truly monumental and that only a massive attack can be expected to be successful. Even so, it must be borne in mind that we are addressing ourselves to only one part of the gigantic problem of communication of scientific information in this country. Perhaps this part represents no more than one-fifth of the total problem, but in view of the critical role of standard reference data in science and industry, this part cannot be ignored until a grand solution is found to be the entire science information problem.
References


29. U. S. House of Representatives, Committee on Government Operations, Subcommittee on Government Information. (John E. Moss, Chairman.) A series of reports over the period from 1956 to present on Availability of Information from Federal Departments and Agencies.


THE NATIONAL BUREAU OF STANDARDS

The scope of activities of the National Bureau of Standards at its major laboratories in Washington, D.C., and Boulder, Colorado, is suggested in the following listing of the divisions and sections engaged in technical work. In general, each section carries out specialized research, development, and engineering in the field indicated by its title. A brief description of the activities, and of the resultant publications, appears on the inside of the front cover.

WASHINGTON, D.C.


Office of Weights and Measures.

BOULDER, COLO.

CRYOGENIC ENGINEERING LABORATORY


CENTRAL RADIO PROPAGATION LABORATORY


RADIO STANDARDS LABORATORY


Joint Institute for Laboratory Astrophysics-NBS Group (Univ. of Colo.).