Physicochemical Property Data Treatment*

B. S. SCHLESSINGER**
Technical Information Services Department, Olin Mathieson Chemical Corp., 275 Winchester Ave., New Haven, Conn. 06504

Received August 26, 1968

The need for data by laboratory scientists is reviewed. Physicochemical property data coverage in Chemical Abstracts indexes and four major compilations is discussed. Suggestions are presented for three possible alternatives to present coverage: (1) measurement and collection of data by a central laboratory, (2) analysis and indexing of all property data literature by some center, possibly the present NSDRS, and (3) an express service of Chemical Abstracts Service. It is strongly recommended that the criterion for approaches (2) and (3) be the existence of the data, and that no critical analysis be attempted, in the interest of time.

One of the basic tasks of the chemist is determining or locating physicochemical values of materials studied. These values provide essential information for decisions in the further study and use of new materials.

The location of physicochemical property data has always been a difficult task for both the laboratory and the literature scientist. Several investigators have published papers on searching for physicochemical property data. Maizell[1] grouped sources of property data into four categories to provide for systematic search with maximum speed. Chase[2] described an organized approach to searching the open literature for property data in an attempt to avoid duplication of effort. Ross[3] surveyed data sources available for the calculation of free energy and heat of reaction.

With the proliferation of technical literature, the problem of finding one numerical value in the forest of words has become almost insurmountable. The difficulties prompted the founding in 1963 of the National Standard Reference Data Program, under the direction of E. L. Brady. Plans for this program were presented in a symposium at the ACS 1966 meeting in New York. The papers[4] were published in the *Journal of Chemical Documentation* in 1967.

Two polls have been conducted in the past few years by the National Standard Reference Data System. The first, in conjunction with the American Chemical Society, in August 1965, showed "that present compilations of data satisfied poorly, or at best only moderately, the requirements of the membership of ACS."[5] In the more recent poll, *Industrial Research*[6,7] questioned scientists on their materials properties data requirements. The answers indicated that 75% of scientists and engineers have problems in obtaining materials properties data, 55% look up property data from one to five times in a week; 22% do this from six to 10 times and 19% must perform the task over nine times in the average week. The majority located necessary information in less than 1 hour; however, 24% indicated search times of up to 8 hours and 8% reported search times of days or weeks.

INTERNAL HANDLING OF PROPERTY DATA

The problems of finding physicochemical property data reflected in the above figures from the *Industrial Research* poll are apparent in our own technical information center at Olin. Many of the reference questions we answer concern location of such data. Our literature searches frequently deal with location of papers on determination of certain properties or in some cases a value for a specific property of a compound.

The general search mechanism used by the laboratory scientist in this area usually begins with a consultation of the handbooks and the library book collection. His next source is the journal literature, most often through *Chemical Abstracts* indexes. Frequently he also consults the company data file, if there is one, and his colleagues in the field. If unsuccessful, he may, at this point, turn the searching over to a technical information specialist, if the company for which he works provides such a person. Many times, more often than we would like to think, the researcher returns to the laboratory early in his literature search and determines the desired properties himself.

In our own organization we have tried to provide access to internal property data from two sources. First, in the indexing of our internal reports we emphasize the inclusion of physicochemical properties, which are named insofar as possible in accordance with *Chemical Abstracts* index nomenclature. This provides for a uniform index access link between internal and external literature. Similar kinds of properties are grouped together. For instance, spectral properties all begin with "spectra," followed by type of spectra, including nuclear magnetic resonance. Table I lists 202 properties included in the last volume of our internal reports index, which covered 2600 documents. It indicates the wide range of physicochemical properties with which an industrial information center can be concerned.

* Presented before the Division of Chemical Literature, Symposium on Redesign of Chemical Literature, 156th Meeting, ACS, Atlantic City, N. J., September 11, 1968.
Editor's note: Other papers from this symposium were published in the November 1968 issue.

** Present address, Department of Library Sciences, Southern Connecticut State College, 501 Crescent St., New Haven, Conn. 06515

Table I. Physicochemical Properties and Processes in the Olin Reports Index

Abrasion Resistance	Dielectric Loss	Green Strength	Printability
Azeotropy	Diffusion	Hardness	Pyrolysis
Acid Number	Dimensional Stability	Haze	Reduction Resistance
Acidity	Dissociation Constant	Heat Capacity	Refractive Index
Activation Energy	Distillation	Heat Content	Ripening
Activity Coefficients	Drying	Heat of Chlorination	Rise
Adsorption	Dyeability	Heat of Combustion	Saponification Number
Ageing	Elasticity	Heat of Dilution	Scorch Resistance
Alkalinity	Electric Conductivity	Heat of Distortion	Shear Strength
Aromaticity	Electric Current Density	Heat of Evaporation	Shrinkage
Bleaching Power	Electric Current Efficiency	Heat of Formation	Softening Point
Boiling Point	Electric Decomposition Potential	Heat of Fusion	Solidification Point
Bond Energies	Electric Discharge	Heat of Hydration	Solubility
Branching	Electric Insulation	Heat of Nitration	Solubilization
Breakage Modulus	Electric Moments	Heat of Reaction	Specific Gravity
Brightness	Electric Polarization	Heat of Solution	Spectra, Fluorescence
Buckling	Electric Potential	Heat of Vaporization	Spectra, IR
Caking	Electric Properties	Humidity	Spectra, Mass
Castability	Elongation	Hydration	Spectra, NMR
Castor Oil Gel Strength	Embrittlement	Hydrolysis	Spectra, UV
Catalytic Activity	Entropy	Hydrophilicity	Spectra, Visible
Chemical Oxygen Demand	Equilibrium	Hydrophobicity	Spectra, X-Ray
Chlorine Availability	Ester Content	Hydroxyl Number	Spectroscopy
Chlorine Demand	Evaporation	Impact Strength	Stability
Chromatography	Explosibility	Impact Sensitivity	Stearic Hindrance
Cleaning Potential	Extraction	Ion Exchange	Stereoregularity
Cloud Point	Extrusion	Isomer Ratio	Stiffness
Coagulability	Fatigue	K-Factor	Storage
Color	Film Durability	Kinetics	Strength
Colorimetry	Film Properties	Light Scattering	Stress
Compaction	Flaking	Load Deflection	Surface Tension
Compressibility	Flash Distillation	Lubricity	Tear Strength
Corrosion	Flash Point	Magnetic Susceptibility	Tensile Strength
Cracking	Flex Fatigue	Mechanical Properties	Thermal Conductivity
Critical Constants	Flexibility	Melting Point	Thermal Decomposition
Crystal Form	Flow	Microscopy	Thermal Expansion
Crystal Growth	Fluidity	Molecular Structure	Thermal Insulation
Crystal Size	Foamability	Molecular Weight	Thermodynamics
Crystal Structure	Foaming	Odor	Titrimetry
Crystallinity	Foam Height	Oil Absorptivity	Torch Test
Crystallization	Foam Shrinkage	Oil Resistance	Toxicity
Curing	Foam Strength	Oscillometry	Transition Temperature
Decolorization	Freezing Point	Overvoltage	Tunnel Test
Defoaming	Freeze-Thaw Stability	Oxidation Resistance	Turbidity
Demulsification	Free Energy of Formation	Permeability	Unsaturation
Density	Free Energy	Permeation	Vapor Pressure
Detergency	Gel Strength	Physical Properties	Viscosity
Dew Point	Gelability	Polarography	Viscosity, Intrinsic
Dielectric Constant	Graining	Potentiometry	Viscosity Index
Dielectric Dissipation Factor	Granulation	Pour Point	Volatility
			Weathering

In addition to the report indexes, as noted above, physicochemical property data can be obtained through the Olin Compound Registry System which provides a unique number for each compound synthesized at Olin and for other compounds of special internal interest. Data sheets on preparation and physicochemical properties are maintained for each of these compounds and are accessible to laboratory staff.

MAJOR SOURCES OF INFORMATION

There are many handbook and older literature compilations which are available to the data searcher. Among these are Gmelin, Beilstein, Mellor, Landolt-Bornstein, the International Critical Tables, and countless others. However, the largest basic compilations of data which are used in searching for chemical compound property information (outside of spectral-type data), and which are ongoing projects in our own country, are (1) the Thermophysical Properties Research Center Data Book and the Retrieval Guide to Thermophysical Properties Research Literature (edited at Purdue University by Y. S. Touloukian), (2) the Texas A & M Thermodynamics Research Center publications (embracing efforts of the American Petroleum Institute and the Manufacturing Chemists Association and edited by B. J. Zwolinski), (3) the National Bureau of Standards (NBS) Circular **500** (now being reissued as the NBS Technical Note **270** by the National Standards Reference Data System (NSRDS), and (4) the Joint Army, Navy, and Air Force (JANAF) Thermochemical Tables (edited by D. R. Stull at Dow Chemical Co.). The accompanying Table II affords descriptive data for the four services in the areas of properties covered, materials covered, funding period covered, total chemical substances in compilation, and materials arrangement. Data for this table were taken from knowledge through use of the published compilations and from the recent survey of data compilations sponsored by the National Science Foundation.[8]

Several points can be made about the data in the table. First, the figures in the column, "Total Chemical Substances in Compilation," reveal the basic difficulty in the use of these sources: very few of the total number of chemical compounds (estimated at about 3 million) are represented by any of the services. The number of compounds included range from 625 in the JANAF tables to 9030 in the Touloukian project. Those compounds which are included are generally the most common compounds, the values for which can be found in most handbook compilations. In fact, it has been our experience that it is relatively infrequent that a value for a specific property not found in the handbook literature is found in any one of these four services.

The table also reveals a duplication of effort by these major compilations. The JANAF tables, the NBS Circular **500**, and the Texas A & M publications all publish thermodynamic and thermochemical data, in many cases for the same compounds. In view of the government support in 1967 of all work at JANAF, Purdue, and NBS and 35% of work at Texas A & M it is understandable that one of the objectives of the National Standards Reference Data System was the coordination of this work.

In addition to the duplication of thermochemical and thermodynamic data coverage, the lack of data for some basic physical properies is disturbing, although the Texas A&M service *does* cover refractive index, viscosity, boiling point, vapor pressure, and density for the compounds

Table II. Comparison of Major Compilations

	Properties Covered	Materials Covered	Funding	Period Covered	Total Chemical Substances in Compilation	Arrangement
Joint Army, Navy, Air Force Tables	Thermochemical properties	All substances of interest in fuel and other selected applications	100% Government	1925 to present 26,000 data sources	625	Hill formula arrangement alphabetically by element
Thermophysical Properties Research	Thermal conductivity Thermal Diffusivity Heat capacity Thermal radiative properties Emissivity Mass diffusivity Prandtl number	All matter	100% Government	1920 to 1964 33,700 data sources	9030	By name
National Bureau of Standards Circular **500**	Thermodynamic properties	All inorganic substances and organic substances with molecules containing no more than 2 C atoms	100% Government	1940 to present 55,000 data sources	4500	Standard order of arrangement given at beginning of circular, based on element-within-formula approach
Texas A&M	Thermal and thermodynamic data Viscosity Refraction Boiling point Freezing point Density Vapor pressure	**Primary** Hydrocarbons and related compounds **Secondary** All substances of interest to chemical industry **Classes of Compounds** Alcohols C_{1-20} Acids C_{1-5} Esters of acids up to C_{20} Haloalkanes and alkenes up to C_{20}	35% Government	1848 to present 30,000 data sources	5300	Formula-oriented in tables listed at front of volume

included, and The Thermophysical Research Center publications cover viscosity and mass diffusivity, as well as the thermal properties of inorganic oxides.

The above difficulties are further complicated by the different arrangements of materials in the compilations. Laboratory and literature scientists faced with such problems often resort to more basic sources of physicochemical property data references, usually the *Chemical Abstracts* indexes. Here, one can approach the problem of a specific value for a compound, either by compound or by specific property. No matter which route is chosen, the searcher must be alert to the difficulties in the use of the indexes.

In searching under the compound, the physicochemical property is not necessarily found conveniently alphabetized at the beginning of the modification. This is particularly true if several properties have been studied in the paper and a catch-all modification like "physicochemical properties of" is used. Neither the specific physical property nor the words "physical property" may be mentioned in a modification at all, as in the cases where (1) the paper is a review on the compound; (2) the paper covers synthesis with properties included; and (3) the paper covers so many topics that an all-inclusive modification could not be synthesized by the indexer.

Problems are also experienced when the searcher uses the property approach and looks for the specific compound, as many of you are aware. The search for a specific compound may be hampered by: (1) the possibility of a number of names for the compound; (2) the presence in the paper of values for a number of compounds, with the modification referring to a class of compounds or an even broader classification; (3) the presence of the compound name as a second, or later entry in an alphabetical list of compounds; and (4) the use of the formula rather than name in the modification.

Added to these difficulties in the use of the *Chemical Abstracts* indexes is the fact that a significant number of common property headings are general in the sense that only new techniques and broad studies of interest are indexed there. A list of 111 general properties from *CA* index Vol. **58** is in Table III.

Comparing the *CA* list with the previous list of properties in the Olin reports index, one finds that 44 properties are included on both lists, i.e., there are 44 properties presumably of industrial interest for which the researcher would not be able to find the newer literature data in *CA*, even if such data did exist in the literature, because these are deemed properties worth indexing only for broad interest studies, not for a study of the specific property for a specific compound or class of compounds.

RECOMMENDATIONS

Overall, it would seem that in the area of physicochemical property data the information sources for thermodynamic and thermochemical properties are adequate, but only for a very small portion of the organic compounds and for an even smaller portion of the inorganic compounds, except for certain petroleum industry chemicals and chemicals of interest to the military establishment. As for the basic physical properties (properties like density, viscosity, melting point, boiling point), these have not received the attention they warrant by any of the major compilations. In our own experience, this is substantiated by the fact that we conduct the majority of searches of this type through the older volumes of *Chemical Abstracts* and the old, but still useful, International Critical Tables.

I have previously mentioned the duplication of effort by the various services in this area. In this era, when so much information floods us all, there should be a con-

Table III. General Properties and Processes in the CA Indexes

Abrasion	Coulometry	Iodine Number	Specific Volume
Acetyl Number	Decomposition Point	Kappa Number	Spin, Electronic
Acid Number	Deformation	Melting Point	Spin, Nuclear
Ageing	Density	Microscopy	Stiffness
Aniline Point	Dew Point	Molecular Weight	Stoichiometry
Asymmetry	Ductility	Nucleophilic Strength	Strength
Atomic Number	Elastic Deformation	Octane Number	Stress (Mechanical)
Atomic Volume	Elasticity	Oiliness	Supersaturation
Atomic Weight	Equilibrium	Optical Rotation	Swelling (Physical)
Azeotropy	Equivalent Weight	Oscillator Strength	Symmetry
Bending Strength	Ester Number	Oxidation	Tackiness
Boiling Point	Expansion	Parity	Taste
Breaking Strength	Extinction Coefficient	Permeability	Tear Strength
Brittleness	Extraction	Peroxide Number	Tensile Strength
Bromine Number	Fatigue	Plastic Deformation	Tension
Bromometry	Fire Point	Plasticity	Thickness
Bursting Strength	Flash Point	Polymorphism	Toughness
Carbonyl Number	Flexibility	Porosity	Toxicity
Cetane Number	Flow Point	Pyrolysis	Turbidity
Chemical Oxygen Demand	Fluidity	Reduction	Unsaturation
Chloramine Value	Fracture	Relaxation	Viscosity
Chlorine Number	Gamma Number	Rigidity	Viscosity Index
Chlorometry	Hardness	Ripening	Volume
Chromatography	Heterogeneity	Saponification Number	Wear
Chromometry	Hydrophilicity	Softening Point	Weathering
Cloud Point	Hydroxyl Number	Solidification Point	Whiteness
Colorimetry	Hexachromic Effect	Solubility	Yield Point
Compressibility	Impact Strength	Specific Impulse	

certed effort made to avoid the intellectual scrutiny of the same sources by at least five different organizations to decide whether a particular piece of data should be included in their compilation. The duplication is compounded by the philosophy nowadays that (1) only critically evaluated data should be allowed to exist in certain compilations, and (2) in the case of *Chemical Abstracts* indexing, that only certain property headings in the index should include all reported measurement.

There are at least three encouraging elements in this picture today in this country: (1) the aforementioned NSRDS system, (2) the properties available (but relatively unindexed) in the *Journal of Chemical and Engineering Data,* and (3) the property calculation program undertaken by the American Institute of Chemical Engineers (an interesting and ambitious, although expensive, undertaking).[9] But more is needed. Time is a vital factor. With each passing year, the volume of data in the literature grows enormously.

I would like to suggest that what is needed is a compendium of physicochemical properties containing, in one printed volume and one central file, measured properties for specific compounds. Most important to industry are the data of economic value, related more directly to sales, customer service, and process economics. The center doing the work should be located in a major city with easy WATS (Wide Area Telephone Service) access (an important consideration in these times). The work and aims of the National Reference Data System are a step in this direction, and their hard work is indeed to be commended, but it is essential that this work move forward with all possible speed to provide the necessary data.

There are two ways to achieve this goal: (1) Ideally, the compendium recommended above should be set up on the basis of standards taken in a laboratory equipped with all modern methods and established for such purpose. Some of this type of activity is already under way at the Purdue Center and more should be encouraged. Property measurement of this sort at first sounds prohibitively expensive but when one considers the amounts expended in duplicate acquisition and indexing, the small numbers of properties and compounds covered by the various compilations and the amounts expended in the country by duplicate searches for the same pieces of information, such a national laboratory might be overall the cheapest source of useable physicochemical property data. (2) Alternatively, a center for analysis and indexing of all property data literature should be seriously considered. This could conceivably utilize input from any of the major compilation sources in a standard format on punched cards, magnetic tape, or disk to produce a merged output. To this body of data could be added other properties not presently covered. Rapid telephone or computer access could be arranged.

The criterion for inclusion of data in the file should be its very existence; many laboratory researchers do not require evaluated or best values so much as they require some data to help in their work. The decision as to value can be made by reference to the original paper. The 1927 measurement of a common property is usually as valid today as the 1968 measurement, and certainly is useful to the scientist if it avoids costly laboratory setup and research time.

A compendium based on the criterion of data existence would avoid the consideration of a critical review and should be much more easily attained than presently planned projects, which may be equally necessary. In such a program, it would be possible to involve individual chemists on a grant basis, especially those in small universities and colleges, to provide the screening of literature sources.

The design of a physicochemical property compendium in a printed form should probably follow the order for chemical compounds most recognizable to the scientific community, i.e., the standard Formula Index Hill alphabetical approach or possibly the Wiswesser Line Notation which is being increasingly used by industrial information centers. The properties to be included would fall in the electrical, magnetic, molecular, spectral, and physical areas, and probably should, as a first step, cover all the properties presently included in the *Chemical Abstracts* indexes.

In theory at least, another possibility for the handling of physicochemical property data in a compendium form is an express service through the Chemical Abstracts Service (CAS). *CA* personnel already analyze the literature for property data as they abstract and index. Perhaps this is the most economical method of establishing such a compendium, especially in view of the criterion of existence previously mentioned. One attractive aspect of CAS involvement would be the opportunity to tie the Compound Registry System into a file such that the search of the CAS Registry would establish more than the mere existence of a compound. I realize, however, that CAS is very busy with present activities and may well find this too large an undertaking.

I would predict the use of such a compendium would grow rapidly in the coming years. The researcher has become more and more involved in seeking information from company files or the total literature about compounds having certain properties and specific ranges making them candidates useful in a given application. The availability of compiled values for properties for a large number of compounds would undoubtedly reveal many new uses for both old and new compounds. This would be still another economic argument for the establishment of such a file accessible to the individual scientist.

ACKNOWLEDGMENT

The author thanks Robert E. Maizell, Technical Information Services, Olin Mathieson Chemical Corp., for stimulating and constructive suggestions that materially aided in the preparation of this paper.

LITERATURE CITED

(1) Maizell, Robert E., *J. Chem. Educ.* **32** (6), 309–11 (1956).
(2) Chase, J. David, *Chem. Eng.* **73** (19), 190–6 (1966).
(3) Ross, Lawrence W., *Chem. Eng.* **73** (21), 205–8 (1966).
(4) Brady, E. L., J. CHEM. DOC. **7** (1), 6–9 (1967).
(5) Weisman, H. M., J. CHEM. DOC. **7** (1), 9–14 (1967).
(6) *Industrial Research* **9** (5), 107–8 (1967).
(7) *Ibid.*, **9** (9), 99 (1967).
(8) Brunner, R. G., *et al.*, **AD 652**, 742 March 20, 1967.
(9) *Chem. Eng. Progress* **63** (7), 37–8 (1967).