

# Managing Standards and Critical Evaluation in a World of Big Data

by D. Brynn Hibbert, David Shaw and M. Clara F. Magalhães

IUPAC is very interested in data, big or small. Its web site opens with the statement, “The International Union of Pure and Applied Chemistry (IUPAC) is the world authority on ... many other critically-evaluated data.” While the ‘...’ covers compelling and widely popular topics, such as naming new elements, the mission of IUPAC to give its imprimatur for chemical data is of great importance to health, security, and trade in the world. In this article, after a review of present activities, we will contemplate how a comprehensive approach might be structured under IUPAC project rules and then look to the future in a world of ‘big data’ and ‘smart instruments’.

## History

IUPAC does not perform or sponsor experimental work that produces chemical data. Rather, it encourages the formation of international teams of qualified experts to compile and critically evaluate data gathered by others. Teams have formed within various Divisions to critically evaluate data ranging from fundamental chemical data (e.g., atomic weights) to a wide range of chemical data used by non-chemists as well as chemists in efforts to satisfy human needs and help resolve industrial, environmental, health-related, and other problems, both local and global. Prior to 2001, when IUPAC activities were organized around Commissions within the Divisions, a number of Commissions and Subcommittees within those Commissions focused on critical evaluation, constituting an ongoing pool of expertise in critical evaluation of chemical data. In 2001, [1] groups with ongoing interests in critical evaluation of data included the Subcommittee on Thermodynamic Data (within the Commission on Thermodynamics, I.2), the Subcommittee on Gas Kinetic Data Evaluation for Atmospheric Chemistry (within the Commission on Chemical Kinetics, I.4), the Commission on Atomic Weights and Isotopic abundances, II.1, the Commission on Equilibrium Data, V.6, and the Commission on Solubility Data, V.8.

Since the reorganization of IUPAC to a project-driven system, only three bodies have existed with a continuing focus on the compilation and critical evaluation of data—the Commission on Isotopic Abundances and Atomic Weights (II.1), within the Inorganic Chemistry Division; the Subcommittee on Modeling of Polymerization Kinetics and Processes, within the Polymer Division; and the Subcommittee on Solubility and Equilibrium Data (SSED), within the Analytical Chemistry Division. Based on a review of the titles of active

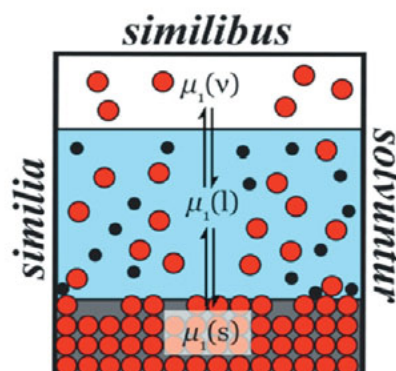
projects listed on the IUPAC web site in November 2016, it appears that three projects involving the critical evaluation of data are underway in the Inorganic Chemistry Division, five in the Polymer Division, and 16 in the Analytical Chemistry Division.

A project, “Interdivisional Discussion of Critical Evaluation of Chemical Data” (project 2016-043-1-500), has recently been started to bring a ‘whole-of-IUPAC’ approach to critical evaluation of data. The Task Group carrying out this project is composed of the authors of this article. We believe that the current organization of IUPAC, in which critical evaluation is organized along disciplinary lines, is appropriate and should be maintained. We also believe that an exchange of information and experience across disciplinary lines can be valuable and should be encouraged. The focus of this project is to convene a working discussion, open to all interested persons, on **Tuesday, 11 July, during the 2017 IUPAC General Assembly**. To begin this discussion, the organizers have posed four questions for consideration:

- How can IUPAC produce critical evaluations that are more useful to chemists and to non-chemist users of chemical data?
- How can IUPAC adjust presentation formats and dissemination channels to make critically evaluated data more accessible to potential users?
- How can groups of critical evaluators within IUPAC better learn from one another’s experience?
- How can IUPAC identify overlooked data categories of high societal value for critical evaluation and organize efforts in response?

As an example of how critical evaluation operates at present, we will describe the workings of the Subcommittee on Solubility and Equilibrium Data (SSED).

Figure 1: ‘Like dissolves like’, the logo of the Subcommittee on Solubility and Equilibrium Data



## The Subcommittee on Solubility and Equilibrium Data (SSED)

The SSED, chaired by Clara Magalhães, coordinates projects whose objectives are:

- the comprehensive compilation and critical evaluation of published experimental data on the chemical solubility and other related thermodynamic data of well-defined substances (solubilities of gases, liquid and solids in liquids and solids) and other equilibrium systems (stability constants for homogeneous reactions); and
- the dissemination of the evaluated data on solubility and stability constants for homogeneous reactions through traditional (journal) and electronic (internet-accessible database) means.

At present the outcomes from the projects directly related to solubility are published in the *Journal of Physical and Chemical Reference Data* and the outcomes of the projects related to the Critical Evaluation of homogeneous systems are published in *Pure and Applied Chemistry*, usually as an IUPAC Technical Report.

The first annual meeting of the Commission on Solubility Data was held at McGill University in Montreal in 1974. Various topics were discussed, including guidelines for the creation of data sheets and evaluations; methods and the preparation of useful evaluations; and the process of recruitment of compilers, evaluators, and editors. By 2016, more than 100 volumes were published on the solubility of well defined systems. In 2001, after the restructuring of IUPAC, the Commission on Equilibrium Data merged with the Commission on Solubility Data, originating the present SSED. Experts in the fields under analysis have done this work, and the recruitment of compilers and other useful evaluators has been a constant effort for more than forty years. The SSED has also taken up work on the Stability Constants Database, a long-standing effort initiated by the Equilibrium Data Commission, and begun work on additional projects in critical evaluation.

To discuss the progress on the various projects, the SSED organizes annual meetings, either in conjunction with the IUPAC General Assembly and the World Chemistry Congress, or at other related international conferences. The SSED is also responsible, through the organization of biannual scientific conferences (International Symposia on Solubility Phenomena), for promoting research in all areas related to solubility and other equilibrium phenomena.

## Approaches to Data Evaluation

Whether data are big or small, their usefulness is in proportion to their accuracy. To a metrologist the problem is easy: data that can be demonstrated to be metrologically traceable to a reference of some standing (such as a realization of a unit in the SI), with an appropriate estimate of measurement uncertainty, are evaluated as fit for some intended use if the measurement uncertainty is within a target range. Even now, when we have a better understanding of the requirements for metrological traceability, [2] published data do not always have clear-cut evidence of traceability or measurement uncertainty estimated by a method compatible with the *Guide to the Evaluation of Uncertainty of Measurement* (GUM). [3] Real-world evaluation is complex and very dependent on both the field of interest and the data themselves.

From its beginnings in the mid-1970s to the end of 2016, the IUPAC-NIST Solubility Data Series has published 103 volumes of compiled and evaluated solubility data and wrestled with many details of the critical evaluation process. Present thinking about best practices in the evaluation of solubility data are described in a guide for preparers and users of evaluated solubility data. [4] According to this guide, "Where two or more compiled independent measurements of solubility for a given system at similar conditions of tempera-

ture and pressure exist, an evaluation is prepared. ... The evaluator checks that the compiled data are correct, assesses their reliability and quality, estimates errors where

*If the knowledge is undigested or simply wrong, more is not better.*

necessary, and recommends numerical values based on all the published data (including theses, reports, and patents) for each given system. Thus, the evaluator reviews the merits or shortcomings of the various data. Only published data are considered." While the Solubility Data Series uses a consistent format to present compiled and evaluated data, evaluators are given considerable latitude to use expert professional judgment to suit their evaluations to the systems with which they are dealing, provided that the process of reasoning leading to an evaluation is made clear. Fitting and smoothing equations as well as graphical representations of the data may be used, but the focus is always on experimentally obtained data rather than modeled values. Evaluators must be aware that their work may be used for purposes with a range of needs for data quality and by individuals with varying levels of knowledge of chemistry. Stability constants are evaluated in a similar manner.

## Data Evaluation in a World of Big Data

'Big data', by its nature, is not amenable to individual scrutiny, and only when condensed, analyzed, and transformed in some way can it be evaluated by traditional methods. Although chemistry is not immediately seen as a producer of big data, we can produce gigabytes of raw signal from NMR or LC-MS/MS, particularly in the biological -omics. Such data can be processed and is often visualized in the form of colorful connection maps expressing a maximum amount of information. It is interesting to speculate how such output might be evaluated if it is to be offered as definitive information. Back to the instrument—which will have been validated and calibrated for the task at hand at each stage of data collection, storage and processing—some trail must be generated that can give a green light to the evaluator. Taking the steps outlined above at each up-scaling of the process, there will need to be evidence that the software and algorithms have done what they purport to do. Coverage intervals on inferences will be necessary to understand how confident we might be in drawing conclusions.

A problem with evaluating big data will be the reliance on the generator and processor of those data to adequately describe and provide evidence for the evaluator. We cannot rely on the 'wisdom of crowds', especially when those crowds are rather small and self-supporting. With small data, we believe we know quite a lot about analytical techniques, for example pH titrations to obtain acid dissociation constants, their limitations and when results are likely to be reliable. This is unlikely to be the case with big data.

Even with modern approaches that might challenge traditional methods (for example "The death of the Job plot, ..." [5]) the attention to careful assessment of uncertainties arising both from the measurements and the modelling of data leads to results in which we can have greater confidence. To conclude, we recall the words of A. S. Kertes in the forwards of volumes 1 to 53 of the Solubility Data Series (1979 – 1993), "If the knowledge is undigested or simply wrong, more is not better." [6]

The authors hope, through the Interdivisional project, that IUPAC will continue to lead in the critical evaluation of chemical data – no matter how 'big'! 🏆

## References

1. IUPAC. *IUPAC Handbook 2000-01*, 2001.
2. P. De Bièvre, R. Dybkaer, A. Fajgelj, and D. B. Hibbert. Metrological traceability of measurement results in chemistry: Concepts and implementation (IUPAC Technical Report) *Pure Appl. Chem.* **83**(10):1873, 2011.
3. Joint Committee for Guides in Metrology. *Evaluation of measurement data—Guide to the expression of uncertainty in measurement*, 100, 2008, BIPM: Sèvres.
4. H. Gamsjäger, J. W. Lorimer, M. Salomon, D. G. Shaw, and R. Tomkins. The IUPAC-NIST Solubility Data Series: A guide to preparation and use of compilations and evaluations (IUPAC Technical Report) *Pure Appl. Chem.* **82**(5):1137, 2010.
5. D. B. Hibbert and P. Thordarson. The death of the Job plot, transparency, open science and online tools, uncertainty estimation methods and other developments in supramolecular chemistry data analysis. *Chem. Commun.* **52**(87):12792, 2016.

Brynn Hibbert <b.hibbert@unsw.edu.au> is Emeritus Professor of Analytical Chemistry, UNSW Sydney. He is Past-President of Division V and editor of the revision of the Orange Book (Compendium of Terminology in Analytical Chemistry). ORCID.org/0000-0001-9210-2941

M. Clara F. Magalhães <mclara@ua.pt> is a professor in the Department of Chemistry and CICECO-Aveiro Institute of Materials, at the University of Aveiro, Aveiro, Portugal; Visiting Fellow in the School of Biological, Earth and Environmental Sciences, Faculty of Science, at UNSW Sydney, Australia, and Chair of the Subcommittee on Solubility and Equilibrium Data of the IUPAC Analytical Chemistry Division. ORCID.org/0000-0002-3359-850X

David Shaw <davidshaw@post.harvard.edu> is Emeritus Professor of Chemistry and Marine Science, University of Alaska, Fairbanks, USA. He is a Titular Member of the Division V and a former Chair of its Subcommittee on Solubility and Equilibrium Data. ORCID.org/0000-0001-8914-2862

