The PAULING FILE database – experimental data for inorganic substances

The PAULING FILE project, named after the Nobel Prize laureate Linus Pauling, was launched some 25 years ago by Japan Science and Technology (JST). The aim was to create a tool for materials scientists searching for new inorganic functional materials. One of the main components of the project was the building up of a comprehensive database that would group experimental data published over the years and over the world. A series of tools for data mining and simulation would then be linked to this nucleus. The winning project was presented by Shuishi Iwata, Japan, and Pierre Villars, Switzerland; the latter is since then editor-in-chief of the PAULING FILE database (*http://paulingfile.com/*).

The database contains three parts: phase diagrams, crystal structures, and physical properties. The third part covers a broad range of different intrinsic properties (examples in parentheses) of single-phase samples: electronic and electrical properties (energy gap, resistivity, Hall coefficient, effective mass), ferroelectricity (Curie temperature, permittivity, polarization), magnetic properties (transition temperatures, susceptibility, magnetization, hysteresis, Mossbauer spectra), mechanical properties (hardness, bulk modulus, magnetostriction), optical properties (absorption coefficient, refractive index, luminescence, optical spectra), phase transitions (melting point, temperatures and pressures for structural transitions), superconductivity (T_c, critical field, current density, energy gap), thermal and thermodynamic properties (thermal expansion, enthalpy, entropy, thermoelectric power). All the data are accompanied by details concerning the preparation and experimental procedure. In January 2018, the PAULING FILE database contained 48'629 Phase Diagram entries, 335'016 Crystal Structure entries, 136'838 Physical Property entries.

The link between the different parts of the database is made *via* the *distinct phases* concept. As the basic principle, a *distinct phase* is defined by the chemical system and the crystal structure. This concept will also enable linking to other databases.

In 2001, PAULING FILE Binaries Edition was published as a test version with innovative software offering numerous options. During a few years Japan defined other financing priorities, and the further development of the database was ensured by contracts with different publishers, who include selected parts of the data in their products. Today most of the crystallographic data for inorganic substances in the well-known file for phase identification PDF4+, produced by ICDD, USA, come from the PAULING FILE. The large database SpringerMaterials (Germany), based on the long series of Landolt-Börnstein handbooks, has a special section, called Inorganic Solid Phases, which contains PAULING FILE data: phase diagrams, crystallographic data, and physical properties. ASM, USA, proposes online binary and ternary phase diagrams from the PAULING FILE in the ASM Alloy Phase Diagram Database. The classical off-line database Pearson's Crystal Data follows Pearson's Handbook, which was specialized in intermetallics. The updated, electronic, version contains, in addition, an even larger amount of data for oxides and halides, accessed *via* attractive software. Japan is back as an important partner, represented by the National Institute for Materials Science (NIST). Their free-of-charge database AtomWork, based on PAULING FILE Binaries Edition, will soon be complemented by an advanced, charged version containing also data for multinaries. In accordance with the initial aim to serve as a tool in materials design, the PAULING FILE is also included in products for simulation. The most recent product, Materials Platform for Data Science, produced by MPDS, Switzerland, was launched in 2017 and contains all the data categories (*http://www.mpds.io/*).

Pearson's Crystal Data, which contains all the crystallographic data of the PAULING FILE, is the most comprehensive database for metals and alloys, and intermetallic compounds. In addition to complete refinements, it also contains entries where only the cell parameters were refined and atom coordinates have been assigned based on the structure type, which is particularly useful for this class of compounds. The crystallographic data are presented as they were published, and converted so that all representatives of the same structure type can be directly compared. Search on atomic environment is also possible.

It is tedious work to understand and summarize the content of publications from different journals (250 titles are followed cover-to-cover for the selection of articles), extract relevant data and enter the information into the database frame. This work can only be done by qualified scientists with experience in experimental solid-state chemistry or physics, who acquire proficiency by processing thousands of publications. A set of some 100 checking modules ensures consistency at several levels, control over units and symmetry restrictions, *etc.* About 95% of the editorial work is at present carried out in Lviv, which gives me one more reason to be happy to get the opportunity to publish a longer article presenting the PAULING FILE in this issue of *Chemistry of Metals and Alloys*.

Roman GLADYSHEVSKII
Karin CENZUAL