INFORMATICS FOR PHASE-BASED MATERIALS DATA

CONTACT DETAILS:
PROJECT NAME: MATERIALS GENOME INITIATIVE
CONTACT: JAMES WARREN, URSULA R. KATTNER, CARELYN E. CAMPBELL, ALDEN A. DIMA
ORGANISATION: NIST
EMAIL: JAMES.WARREN [AT]NIST.GOV

Objectives

Current material property databases focus on either the engineering properties of commercial materials or on first principles results of selected material classes. The availability of phase-based property databases is limited, although these data are essential for designing new materials. This work focuses on the development of informatics tools and file repositories for phase-based property data. Experimental and computational phase-based property data, which serve as the basis for CALPHAD assessments, are diverse and semi-structured. In addition, these data sets are often incomplete (i.e. essential metadata are missing). To accommodate these incomplete data sets and the diversity of the data requires extensible representations. The strategy for developing these representations includes defining universal identifiers for materials and phases; a working ontology for phase-based materials; data schemas for specific material properties; and informatics tools.

On-going activities:

Current efforts to develop informatics tools for materials include developing a working-ontology, data-schemas for specific materials properties, and universal identifiers for materials and phases. The working ontology and data schema development efforts are currently focused on diffusion, phase equilibria, and thermochemical data. The ontology will be linked to the developing data schemas. Initial tracer and interdiffusion data schemas were introduced to experts in the community. Feedback from this community is currently being incorporated into revised schemas. Universal phase identifiers are being developed using an InChI-based format.

File repositories, using DSpace, have been established for CALPHAD-based assessments and first-principles based calculations. These repositories provide permanent identifiers, various Creative Common licenses, and metadata for the associated data files. These repositories are currently being tested by the user communities.

Results:

A materials data infrastructure in support of the Materials Genome Initiative is being developed. The initial focus is on phase-based data. Current efforts are focus on the development of a working ontology, data schemas, universal identifiers for materials and phases and crystals structures, and file repositories. The development of this infrastructure also requires engagement of the community to provide data. This is the biggest challenge for the development a materials data infrastructure as the community must change the way it thinks about its data.

URL: http://www.nist.gov/mml/msed/thermodynamics_kinetics/materials-informatics.cfm