Materials Informatics: Theory and Application

Krishna Rajan

Department of Materials Science and Engineering
NSF Intl. Materials Institute Combinatorial Sciences &
Materials Informatics Collaboratory
Iowa State University
1:00  Introduction to workshop  
      T. Marechaux and L. Peurrung

1:15:  Materials Informatics: Theory and Application  
       K. Rajan, Iowa State University

1:30  NSDL Materials Digital Library Pathway  
      L. Bartolo, Kent State University

Processing information

2:15  Thermo-Chemistry to Phase Diagrams and More  
      P. Mason, ThermoCal

3:00  Break

3:20  Data Patterns and Links to Materials Theory  
      K. Ferris, Pacific Northwest National Laboratory

4:00  Data Mining and Materials Informatics  
      K. Rajan, Iowa State University

4:40  Conclusion and Discussion
Materials Informatics: Growing from the Bio World

Warren H. Hunt, Jr.

Materials science and engineering in our community

or we as a community will not be able to produce new advances in materials rapidly or take full advantage of existing materials capabilities in design.

Materials informatics, is emerging as a new approach for addressing these needs in our field.

Materials informatics fuses collections of annotated materials data. The fused data is integrated in a structured learning framework (e.g., with intelligent data mining tools that include physical models and get "smarter" as the search proceeds). In this way, not just data, but knowledge is extracted. Combined with high-throughput experimental techniques, materials informatics offers a more efficient manner to develop new materials with optimal properties.

On December 8-9, 2005, a group of individuals gathered in Washington, D.C., to discuss materials informatics. The participants developed a working definition, identified and prioritized needs, and outlined actions to advance this new field of materials science. The workshop was sponsored by Pacific Northwest National Laboratory, ASM International and TMS. Outcomes of the meeting are presented in this sidebar.

Since the December workshop, a steering committee has been meeting biweekly via conference call. The group worked to move forward on selected recommendations from the workshop seeking to build the materials informatics community. A key upcoming community event is the Materials informatics workshop scheduled for Sunday, October 15 from 1-5 p.m. in conjunction with the Materials Science & Technology 2006 conference in Cincinnati, Ohio. Further information is available at www.matstech.org or through organizer Krishna Rajan, a professor at Iowa State University, at krishana@iastate.edu. More information is also available from the TMS Technical Services Department at techservices@tms.org.

Warren H. Hunt is executive director of TMS.
THE “OMICS” OF MATERIALS SCIENCE

DATA AND DATA INTEGRATION

In-silico MATERIALS SCIENCE: A Biological Analogue

INFORMATICS: Tools for Materials Discovery

IT platform:

- Ultra large scale databases of appropriate information and data (data warehousing)
- Scientific visualization
- Data mining
- Remote access / distributed computing

Materials Science platform:

- Material theory
- Materials processing
- Property measurement
- Computational techniques
- Materials characterization
- Discrete / atomistic modeling
- Continuum modeling

Data "fusion" through Informatics
INFORMATICS

- **Data acquisition:** The barriers to data acquisition include finding sources of data, locating data within the sources, extracting sufficient metadata to make the property data meaningful, harmonizing the information from different sources, and collecting the results in a fully documented, searchable, and accessible data system. The group identified several means by which the materials community as a whole could contribute to improving data access and acquisition.

- **Data standards:** Specific needs included a listing and taxonomy of the elements of an information framework to describe materials as well as both de-facto and formalized standards for data formats.

- **Dissemination/access:** A requirement to define access policies was identified. It was envisaged that different tiers will probably be required to control access to the data depending on the status of who is trying to access it.

- **Analysis tools:** Knowledge extraction from the fused data sets through the detection of knowledge gaps, identification of anomalies, model development, pattern recognition, and mapping (forward and inverse) techniques is needed. These capabilities need to handle the various forms of information, conform to the likely data standards, and must be joined by a system architecture, as needed by the combinatorial community.

- **Visibility/awareness:** The practitioners of materials science and related fields need to become educated about informatics approaches, even to the point of developing materials informatics as a recognized new sub-discipline within materials science. The group identified a number of activities to raise awareness and visibility of materials informatics broadly and to foster a materials informatics community.

- **Resources:** As with any endeavor of this nature, financial and management support greatly help further progress on resolving the issues and needs described. There may also be a need to develop standard reference materials that can be used to support data validation.