NIST and The Materials Genome Initiative

James A Warren
Technical Program Director for Materials Genomics
Material Measurement Laboratory
National Institute of Standards and Technology

Science advances one funeral at a time - Max Planck
NIST and The Materials Genome Initiative

James A Warren
Technical Program Director for Materials Genomics
Material Measurement Laboratory
National Institute of Standards and Technology

Science advances one funeral at a time - Max Planck

Take time to stop and smell the data - Nate Silver
Outline/Goals

- 18 month perspective about the National effort
- Data, Culture and what’s coming
- Share info about the NIST effort
  - Pilots
  - Supporting Activities
  - External Engagement
To help businesses discover, develop, and deploy new materials twice as fast, we’re launching what we call the Materials Genome Initiative. The invention of silicon circuits and lithium ion batteries made computers and iPods and iPads possible, but it took years to get those technologies from the drawing board to the market place. We can do it faster.

-President Obama
Carnegie Mellon University, June 2011
... 2X faster & 2X cheaper
The Materials Genome Initiative

Goal: to decrease the cost and time-to-market by 50%

1. Develop a Materials Innovation Infrastructure

2. Achieve National goals in energy, security, and human welfare with advanced materials

2. Equip the next generation materials workforce

http://www.whitehouse.gov/sites/default/files/microsites/ostp/materials_genome_initiative-final.pdf

The MGI is a direct effort to enable the creation of new materials by leveraging the growth in computational power with a commensurate focus on modeling and simulation.
The Materials Innovation Infrastructure
Three-pronged approach to advance materials design

“The Materials Genome Initiative will develop new integrated computational, experimental, and data informatics tools. These software and integration tools will span the entire materials continuum, be developed using an open platform, improve best-in-class predictive capabilities, and adhere to newly created standards for quick integration of digital information across the materials innovation infrastructure.”

Materials Genome Initiative for Global Competitiveness
Accelerating the development of advanced materials is critical for achieving global competitiveness.

*From McDowell and Olson, Sci Model Simul (2008) 15:207–240*
Why Now?

Materials Are Complicated Systems
Modeling is a Challenge

• Advanced materials are complex: multi-component and multi-phase
• Without adequate modeling, informatics and data exchange, the development of next generation materials using empirical approaches is bogged down by their complexity
• The Materials Genome Initiative seeks to advance materials design capabilities to promote faster, cheaper

Alloy cooled from 300 °C

Alloy cooled from 800 °C

• Composition and processing affect properties
• Phases change as a function of processing
• Microstructures consist of mixtures of multiple material phases
• Finer microstructure results in a much stronger alloy
Why Materials Design?
Materials development is a “needle in haystack” problem

• If an empirical approach is taken to develop new materials, number of possible compositions to evaluate is overwhelming.
  – e.g., superalloys, which often contain 12 components, if one only considers the current composition ranges the number of compositions to investigate exceeds $1 \times 10^8$
Why Materials Design?
Materials development is a “needle in haystack” problem

• If an empirical approach is taken to develop new materials, number of possible compositions to evaluate is overwhelming.
  – e.g., superalloys, which often contain 12 components, if one only considers the current composition ranges the number of compositions to investigate exceeds $1 \times 10^8$

The right tools:
MGI?
Why an MGI?

Figure 1. Schematic of how the design criteria for a given material dictate the needed material properties and thus define the needed experiments, models and data.
Multi-Agency Process

• 2010-11: NIST, DOE, NSF, and DOD work with OSTP to prepare MGI Whitepaper
• 2011: NIST develops FY12 Initiative on Advanced Materials for Industry
• WH Submits budget in Feb 2011, with NIST Initiative (Request at 14 M)
Federal Programs

The federal government has announced several programs and funding opportunities related to the Materials Genome Initiative. Learn more about these activities below:

Federal Funding Announcements

FY 2012
Department of Energy:

- Funding announcement for the development of lightweight materials and awardees.
- Funding announcement for predictive theory and modeling for advanced materials.
- Funding announcement for Scientific Discovery through Advanced Computing (SciDAC) partnerships.

Department of Defense:

- Funding announcement for the Enterprise for Multiscale Research of Materials program, led by the Army Research Laboratory (ARL).
- Funding announcement for the Air Force Research Laboratory (AFRL) Center of Excellence for Integrated Computational Material Science and Engineering of Structural Materials
- Funding announcement for the Air Force Research Lab (AFRL) Center of Excellence in advanced organic composites
- Beta release of the Ab-Initio Electronic Structure Library (AFLOWLIB), maintained by Duke University in partnership with the Office of Naval Research, with open access to over 17,000 compounds derived from the Inorganic Crystal Structure Database and over 160,000 binary alloys

National Institute of Standards and Technology:

- Funding announcement for the NIST Advanced Materials for Industry program.

National Science Foundation:

- Funding announcement for the FY 2012 NSF Designing Materials to Revolutionize and Engineer our Future (DMREF) program.
FY12 NIST Budget Initiative: Advanced Materials For Industry (14.2M Requested)

NIST will further enable advanced materials by working with our partners in other government agencies, academia and industry to develop

• Computational tools, validated databases, data assessment tools and techniques, and standards
• Reference models and simulations
• Mechanisms for exchange of information
• Consortia to determine consensus standards for data exchange
• Creation of a Center of Excellence
NIST’s Mission

To promote U.S. innovation and industrial competitiveness by advancing measurement science, standards, and technology in ways that enhance economic security and improve our quality of life.
NIST at a Glance

Major Assets

- ~2800 employees ~(50/50 technical/admin)
- ~2600 associates and facilities users
- ~1600 field staff in partner organizations
  (Manufacturing Extension Partnership)

Major Programs

- NIST Laboratories
- Baldrige Performance Excellence Program
- Hollings Manufacturing Extension Partnership
NIST Products and Services

Measurement Research
- ~ 2,200 publications per year

Standard Reference Data
- ~ 100 different types
- ~ 6,000 units sold per year
- ~ 226 million data downloads per year

Standard Reference Materials
- ~ 1,300 products available
- ~ 30,000 units sold per year

Calibration Tests
- ~ 18,000 tests per year

Laboratory Accreditation
- ~ 800 accreditations of testing and calibration laboratories
Multi-Agency Efforts

• NSTC Subcommittee (CoT) on the Materials Genome Initiative formed (First meeting 4/12)
• Membership now includes NIST, DOE, DOD, NSF, NASA, NIH, USGS, NNSA, DARPA, NSA, and OMB.
• Coordination with the NNCO and new Nanotechnology Knowledge Infrastructure (NKI) Signature Initiative.
• Coordination with NITRD and OCI (NSF)
• MGI White House Kickoff Event (Discussed Later)
• New focus teams: Data, Code Curation and Technical Gaps, Industry Outreach
Nanoinformatics, the NKI-NSI and the MGI

“The Nanotechnology Knowledge Infrastructure developed by ten agencies will stimulate the development of models, simulation tools, and databases that enable the prediction of specific properties and characteristics of nanoscale materials.

Also approaches, protocols, and standards developed through MGI activities may be initially explored, tested, or evaluated specifically for nanomaterials under NKI efforts.

The cross fertilization between NNI and MGI will yield broader knowledge dissemination and can be facilitated by the NKI effort.”

Fact Sheet on Progress on the Materials Genome Initiative
Executive Office of the President, May 14, 2012
The Problem with Materials
Science Data
A simple example to illustrate the key questions, gaps, obstacles to realizing materials by design
The Problem with Materials Science Data

A simple example to illustrate the key questions, gaps, obstacles to realizing materials by design

Jim states the obvious? Or causes trouble
Q: What is Data?
A: Standard Answer

- **Data**: result of measurement
  - **Physical**: experimental result, uncertainty
  - **Virtual**: simulation result, uncertainty
- **Metadata**: information describing the measurement process
  - **Physical**: experimental setup, settings
  - **Virtual**: explicit underlying model, software*, input parameters
Q: What is Data?
A: Standard Answer

- Data: result of measurement
  - Physical: experimental result, uncertainty
  - Virtual: simulation result, uncertainty
- Metadata: information describing the measurement process
  - Physical: experimental setup, settings
  - Virtual: explicit underlying model, software*, input parameters

* simulation software is data!
Traditional Approach Example

- Measure a diffusion coefficient (perhaps by tracking the root mean displacement of tracer particles)
- record values is some sort of table (perhaps Excel)
- perhaps publish
- publication contains metadata
What the #$&%^ is a diffusion coefficient?

- It’s only a model (not reality)
- It’s an equation
- What you think you’re measuring is NOT the diffusion coefficient
- You are measuring the (maybe) root mean displacement of particles and assuming the diffusion equation is true.

\[
\frac{\partial c}{\partial t} = D \nabla^2 c
\]

All Models are Wrong, but some are Useful -- George Box
Old School, Computation

- Write down Equations
- “Solve them” (verify please)
- Analyze, make graphs, compute numbers of interest (results)
- Publish results

We can do much much much better!
Better Models = Less Data

- Take the LHC as an example
- Data produced at 1PB/sec!
- Reduced data saved: 300 MB/s
- That’s a darn good model
- Other end: Biology?
- Materials: In the middle

- But you need to know the model to make sense of the data!
What is Data “2.0”?  

Science is characterized by the iteration of experiment and models, yielding higher fidelity with lower uncertainty.

The measurement or computation of a quantity (data) is generally meaningless without the associated quantifying model that defines both the data and its uncertainties.

Thus dissemination of data is ideally the dissemination of the following information:

1. (1) Measured quantities,
2. (2) Associated quantifying models, and
3. (3) Raw data, including the protocols (and equipment) by which it was obtained

Without these descriptions (expressible through software), the data becomes divorced from meaning, making interpretation “difficult.”

see R. Peng Science 2 December 2011: 1226-1227
A Comment on Data Driven Discovery for MGI

• Biology community comfortable with this
• Physics community finds this approach “weird”
• Materials Science is in the middle
• Presents unique challenges for Computer Science and informatics research
It’s all about mapping structural information

Synchrotron diffraction set up at Beamline 7-2 (SSRL)

The entire 3” wafer (300 spots) can now be measured in 2 hrs
Consequences of the traditional approach

- High barrier to adoption of methods and results
- Extra expense due to duplication of effort
- Lost opportunities to enable discovery & further science

So Now What?
NIST Role

Provide the Measurement and Standards Infrastructure Needed to Realize the MGI

Implementation

NIST will work with stakeholders in industry, academia, and government to develop standards, tools and techniques for the

• The acquisition, representation and interoperability of materials data, whether from simulation or experiment
• The interoperation, across multiple length and time scales, of modeling systems
• The quality assessment of models, simulations, and the materials data generated from them, using
  • validation through experiment
  • quantification of uncertainty
  • verification of mathematical and physical consistency
  • determination of data pedigree

Expected Outcomes

• Improved access to data/models/simulations
• Easier assembly of models operating at differing length and time scales
• Improved model reliability and confidence in results
• Reduced barriers to adopting state-of-the-art methods and techniques
NIST will enable the widespread adoption of ICME by improving materials data and model exchange, and providing the tools to assess materials data and model quality.