The Materials Commons
A Novel Information Repository and Collaboration Platform for the Materials Community

Staff: Glenn Tarcea, Sravya Tamma,
Domain Scientists: Brian Puchala, Emmanuelle Marquis and John Allison
Information Scientists: Margaret Hedstrom (SI) and H. Jagadish (CSE)
The University of Michigan
Center for PRedictive Integrated Structural Materials Science (PRISMS)

- A 5-year grant to accelerate the development of predictive materials science

- Involves:
  - 11 faculty
  - 5 staff scientists
  - 16 students & postdocs

- DOE resources leveraged by significant UM cost share
PRISMS Overarching Vision
Enable accelerated predictive materials science.

- Linking Experiments & Simulations
- Advanced Quantitative Experiments
- Open Source
- Integrated Hierarchical Multi-Scale
- Computationally Efficient
- Extensible

Leading Edge Metals Science

Use Cases
- Microstructural Evolution
- Mechanical Behavior

PRISMS Computational Tools

The Materials Commons

Collaboration
- Experimental & Simulation Information
- Seamless, Continuous
- Workflow
- Provenance Tracking
PRISMS Overarching Vision
Enable accelerated predictive materials science.

- Linking Experiments & Simulations
- Advanced Quantitative Experiments
- Open Source
- Integrated Hierarchical Multi-Scale
- Computationally Efficient
- Extensible

Leading Edge Metals Science

Use Cases
- Microstructural Evolution
- Mechanical Behavior

PRISMS Computational Tools

The Materials Commons

Collaboration
- Experimental & Simulation Information
- Seamless, Continuous
- Workflow
- Provenance Tracking
PRISMS Overarching Vision
Enable accelerated predictive materials science.

- Linking Experiments & Simulations
- Advanced Quantitative Experiments
- Use Cases
  - Microstructural Evolution
  - Mechanical Behavior
- Open Source
- Integrated Hierarchical Multi-Scale
- Computationally Efficient
- Extensible
- Collaboration
- Experimental & Simulation Information
- Seamless, Continuous Workflow
- Provenance Tracking
- Open Source
PRISMS Integrated Multi-Scale Modeling Framework & Software

- Integrated, open source toolset
- Primarily hierarchical
- Mathematical descriptions appropriate for the scale
- Parameter passing to higher scales
Experiments are key for:
- Mechanisms
- Filling gaps in theory
- Quantitative inputs
- Validation

Tightly Coupled Models & Experiment

- Atomistic Constitutive Laws
- Dislocation Dynamics
- Crystal Plasticity w/ FEM
- Continuum Elastoplasticity with FEM
- Phase Field

Density Functional Theory

Statistical Mechanics

CALPHAD

Experiments, e.g. phase nucleation, growth kinetics, recrystallization

Emmanuelle Marquis

Samantha Daly

Wayne Jones

John Allison

---

PrismS
• Open source software
Materials Innovation Infrastructure

- Open source software
- Team Collaboration
- Integrated framework
- Analysis
- Design

![Diagram](diagram.png)
Materials Innovation Infrastructure

- Open source software
- Team Collaboration
- Integrated framework
- Analysis
- Design
- Community Collaboration
- Federation
Materials Commons

Goals:
• Store materials data (and provenance)
• Collaborate and share data
• Search and use data
• Be a seamless part of the scientific workflow

Our vision of success:
• To be a premier data repository for the metals community in PRISMS technical emphasis areas
• Contribute ideas and software to the wider materials infrastructure community
The Materials Commons consists of:

– 2 full time professional staff
– >25 PRISMS faculty and grad students as users.
– A 390 TB Isilon storage cluster data repository
– A website for uploading and downloading data, adding provenance, sharing and searching for data
– An application installed on your computer for uploading and downloading data
– A REST based API to access and extend the capabilities of the repository

• Interacting with NIST, Globus etal, (and you?) to share best practices, schemas, etc.
• Scheduled for PRISMS Community use in Summer 2015.
Example Scientific Workflow
(Just one 'box' in the integrated framework)
• Primary objects in the system are:
  - *Files*
  - *Samples*
  - *Processes*

• Organized into *Projects*
  - With permission, all objects can be shared between multiple *Projects*
  - Have an associated directory tree containing *Files*
• Primary objects in the system are:
  – *Processes*
  – *Files*
  – *Samples*

• Organized into *Projects*
  – With permission, all objects can be shared between multiple *Projects*
  – Have an associated directory tree containing *Files*
Materials Commons Organization

- Primary objects in the system are:
  - Processes
  - Files
  - Samples
- Organized into Projects
  - With permission, all objects can be shared between multiple Projects
  - Have an associated directory tree containing Files
Primary objects in the system are:
- Processes
- Files
- Samples

Organized into Projects
- With permission, all objects can be shared between multiple Projects
- Have an associated directory tree containing Files

Diagram:
- File
- Process
- Sample
- Sample Attributes
- Project
- Statistical Mechanics
- Dislocation Dynamics
- Phase Field
- Crystal Plasticity
- DFT
Materials Commons Workflow

1) Upload files or add measurements to a Sample
   – Use the Data Set Loader to copy files directly
   – Via the website: materialscommons.org

2) Add provenance
   – Describe the process that created the files
   – Process templates describe what information is expected
   – Soon: Automatic upload from computational jobs
   – Soon: Automatic parser of select experimental instrument output

3) Notify collaborators
Materials Commons: Provenance Entry

Create Provenance

- Show overview

APT Data Collection
- Sample
- Settings
- Files
- Files
- Done

Name:

Tags:

Add Tag

Description:

Add Run

Runs:
Materials Commons: Provenance Entry

• *Process* templates specify what provenance information must / should / can be given.
  – Examples types: APT, SEM, Fatigue Test, Monte Carlo Calculation, Continuum Plasticity Calculation, etc.
  – Instrument settings, process conditions, simulation or analysis input files, output data files, etc.

• We develop templates with users to limit their number
• Save 'drafts' to avoid repeated data entry
• Provides dictionary of search terms
Search and Use Data

Provenance:

The attribute A, of sample S, is value V, as determined by process P.
Provenance:

The attribute $A$, of sample $S$, is value $V$, as determined by process $P$.

In PRISMS technical areas:
- Composition
- Length/width
- Crystal structure
- Grain/phase/defect size, orientation, shape, distribution
- Interface structure & chemistry
- Strain
- Dislocation density & configuration
- Young's modulus
- Yield Strength
- Strain to fracture
- Cycles to failure
- etc.
Search and Use Data

Provenance:

The attribute A, of sample S, is value V, as determined by process P.

Samples may be experimental or computational.

Samples define a set of related attributes that may be linked with a constitutive model.

Samples may be transformed by a process, such as a heat treatment or tensile test. Then certain attributes might change.

New samples may be created by a process, such as sectioning.
Search and Use Data

Provenance:

The attribute $A$, of sample $S$, is value $V$, as determined by process $P$.

May be:
- Scalar
- Tensor
- Function
- Description
- Other...

Formatted in JSON

Track units!
Search and Use Data

Provenance:

The attribute A, of sample S, is value V, as determined by process P.

May be multiple measurements obtained by alternate or repeated processes. Must record which is must trusted and why.

In PRISMS technical areas, experimental processes:
• Characterization: Optical, SEM, TEM, APT, XRD, EPMA, SAXS/SANS
• Monotonic stress-strain
• Fatigue: Cyclic stress-strain, LCF, HCF, FCG
• Hardness
• Resistivity – aging

Computational Processes:
• Density functional theory
• Monte Carlo, KMC
• Phase Field
• Dislocation dynamics
• Crystal & Continuum plasticity
Scientific Knowledge Informs Provenance

- File
- Process
- Sample
  - Text
- Sample Attributes
  - Supplier
  - Composition
  - Geometry
  - Treatment
  - Location
  - Orientation

- Sectioning

- Optical
  - Images
  - Measurements
    - Grain size
    - Morphology
- SEM
  - Images
  - Measurements
    - Ppt. size
    - Grain orientation
- TEM
  - Images
  - Measurements
    - Ppt. chemistry

- Treatment

- Machining

- Fatigue
  - Images
  - Measurements
    - $N_f$ vs. $\Delta \varepsilon$

- Analysis

- SEM
  - Images
  - Measurements
    - Crack location

Samples with Attribute: Composition C1
Scientific Knowledge Informs Provenance

Sample Attributes
- Supplier
- Composition
- Geometry
- Treatment

Sectioning
- Location
- Orientation

Surfaces
- Optical
- SEM
- TEM

Grain Size
- GS1
- GS2

Fatigue
- $N_f$ vs. $\Delta \varepsilon$

Analysis
- Images
- Measurements

Machining
- Location
- Orientation

SEM
- Crack location

Prisms
Scientific Knowledge Informs Provenance

Sample
- Supplier
- Composition
- Geometry
- Treatment

Samples with Attribute:
- Orientation O1
- Orientation O2
- Orientation O3

Sample
- Location
- Orientation

Sectioning

Sample
- Optical
- SEM
- TEM

Sample Attributes
- Text
- Grain size
- Morphology
- Ppt. size
- Grain orientation
- Ppt. chemistry

Treatment
- Machining
- Fatigue
- Analysis

Fatigue
- $N_f$ vs. $\Delta \varepsilon$

Analysis
- Images
- Measurements

SEM
- Crack location

PRISMS
Provenance:

The attribute A, of sample S, is value V, as determined by process P.

Search:

• General:
  • Find all samples that have attributes X, Y, Z and underwent process of type T.
  • Find all samples that have attribute X, as measured by process of type T.

Define materials using a pre-defined filter:

• AZ91:
  • Select all samples that have composition (wt%):
    • Mg=90.8±0.1, Al=8.25±0.01, Zn=90.8±0.1, Si=0.035±0.005, Mg=0.22±0.01
• AZ91-T6
  • Select all samples that match AZ91, with solution treatment and aging
• Phase: HCP-Mg
  • Select all samples that are mostly Mg, and P6₃/mmc
Use Data

- Flexibly visualize and analyze
- Fit constitutive models
- Fit process models
- Provenance graph -> workflow -> optimization:
  - Multi-scale integration
  - Material design
  - Model reduction
  - Optimal design of experiments and computations
PRISMS-IntegrationTools

• Pass arbitrary functions in human readable form
  – Regular or piecewise functions
  – Specify basis sets for series functions
  – Automatic code generation using GiNaC: http://www.ginac.de

• Pass fields in .vtk format

• Evaluate functions, fields, 1\(^{st}\) and 2\(^{nd}\) derivatives

• C++ software library with C and Python interfaces
  – https://github.com/prisms-center/IntegrationTools
Development Plans

• Continuously improve the user interface and workflow
• Short term plans:
  – Provenance linked sample model
  – Enhanced search capability
  – Python API, automated computations provenance upload
  – “Automated” data collection from selected instruments
• Prepare for growing user base
  – First limited release anticipated mid-2015
• Longer term plans (Note: opportunities for leveraging and collaborative development)
  – Crowd curation
  – Comparison plotting – experiment vs simulation, etc.
  – Image libraries (e.g. Lightroom API)
  – Data analysis, image analysis
  – Federation
PRISMS Center – Faculty and Staff

Faculty
• John Allison, Center Director
• Samantha Daly
• Krishna Garikipati
• Vikram Gavini
• Margaret Hedstrom
• H. V. Jagadish
• J. Wayne Jones
• Emmanuelle Marquis
• Veera Sundararaghavan
• Katsuyo Thornton
• Anton Van der Ven - UCSB

Staff
• Larry Aagesen
• Brian Puchala
• Shiva Rudraraju
• Stravya Tamma
• Glenn Tarcea
New institute led by UM, Ohio State U & Edison Welding Institute (Awarded March 2014)
- 60 member consortia of industry, academia and national labs
- Focus is manufacturing innovation for Al, Fe, Ti
- ICME will be a key technology focus
- Materials Commons proposed as primary collaboration platform
Materials Commons Technologies

- API/REST Services are all JSON based
- Data Sync is moving to MsgPak, Curve, and ZeroMQ
- Website uses Socket.IO for real time updates, REST/JSON for service access
  - Just started integrating this in
- JSON Document store on backend (RethinkDB)
- Website written using AngularJS
- Backend is a mix of Python and Go
- Some Erlang mixed in
- RabbitMQ used for pipeline processing
Architecture Block Diagram

HTTPS

NGINX

SSL

Reverse Proxy

API Services

RabbitMQ

HTTP

Website

Web Service User

Scripts

Data Upload

JSON

GOB (MsgPak, Curve)

Data Services

Shared Service Access

elasticsearch Cluster

Database Cluster

Data Loader Services