CHiMaD Data Schema Working Group

Report
Topics of Today's Discussion

• Thermomechanical Processing History
• Hardness Testing
• Nanomine
• CALPHAD Protodata
• Workflow tool integration
• NoMaD Repository
• ASM Taxonomy
Goals of 3 Month Project

• Identify dataset(s) for curation in MDCS
• Discuss and agree on shared terminology for
  – metadata (e.g., strain rate)
  – data (e.g., load vs. displacement)
  – derived data (e.g., modulus of elasticity)
• Create reusable XML schema types
• Create scripts for data transformation
• Curate dataset(s) in MDCS
Materials Data

- Discoverable
- Accessible
- Interoperable

- Discoverable
- Accessible
- Maybe Not Interoperable

Organized big data

Long-tail data

Literature limit

Unpublished and dark data

Nat Neurosci, 17(11), 1442-1447.
doi:10.1038/nn.383
A Configurable Data Curation System

Configurable Interface

Data Management & Search Engine

Structured Data

Large/Binary Files

Your Data Repository or Your Resource Registry
Long Tail

Nat Neurosci, 17(11), 1442-1447.
doi:10.1038/nn.383
NIST MGI APPROACH TO LONG TAIL DATA

- **Discoverable**
  (via the Registry)
  https://mgi.nist.gov/Zkp

- **Accessible**
  (via the Curator)
  https://mgi.nist.gov/ZkS

- **Interoperable**
  (via Community Data Standards)
  https://mgi.nist.gov/ZkG
Group Activities/Discussion
Draft Thermomechanical Processing History Schema

- **step**
  - **number**
    - Type: `xsd:int`
  - **type**
    - Type: `Restriction of 'xsd:string'`
  - **description**
    - Type: `xsd:string`
  - **start-date-time**
    - Type: `xsd:dateTime`
  - **thermal-parameters**
  - **mechanical-parameters**
  - **environment**
    - Type: `Restriction of 'xsd:string'`
  - **parameter**
    - Type: `parameter-type`
  - **comment**
    - Type: `xsd:string`
Draft Hardness Testing Schema

- measurement-method
  - Type: Restriction of 'xsd:string'

- applied-load
  - Type: physical-quantity-ML-TT-type

- dewell-time
  - Type: physical-quantity-T--type

- indentor-size
  - Type: physical-quantity-L--type

- temperature
  - Type: physical-quantity-0--type

- hardness
  - Type: physical-quantity----type

- Vickers
  - brinell
  - rockwell b
  - rockwell a
  - rockwell c
  - rockwell d
  - rockwell e
  - rockwell f
  - rockwell g
  - rockwell k
  - rockwell l
  - rockwell m
  - rockwell n
  - rockwell p
  - rockwell q
  - rockwell r
  - rockwell s
  - rockwell v
  - rockwell w
  - rockwell x
Nanomine

Statistical Learning and Analysis Module Tools

Statistical learning and analysis modules include web and downloadable packages that can be used to pre-process and analyze structure and material property data. Each of the modules will specify required format of input and output data, and provide a brief introduction of mechanism of the algorithm.

NIBLACK BINARIZATION

Descriptor Characterization is a modular tool that takes input from a micrograph image of a microstructure of material and generate statistical descriptors that can characterize the structure information. More details to follow.

DESCRIPTOR CHARACTERIZATION

Descriptor Characterization is a modular tool that takes input from a micrograph image of a microstructure of material and generate statistical descriptors that can
Nanomine
Curating Diffusion Data

Sample Information
- Sample Id,
- Owner
- Date of Experiment

End Member Material Information
- Phase name
- Crystal structure
- Phase Fraction
- Composition
- Processing

Experimental Procedures

Diffusion Annealing Conditions

Collected Data
- Spreadsheet
- Micrograph
CALPHAD Protodata

Self Diffusion Resource

http://www.ctcms.nist.gov/~gkl/selfdiffusion.html
## CALPHAD Protodata

<table>
<thead>
<tr>
<th>Method</th>
<th>Structure</th>
<th>Diffusion Direction</th>
<th>Frequency Factor D0 (m2/s)</th>
<th>Activation Energy Q (kJ/mole)</th>
<th>Temperature (K)</th>
<th>Details</th>
<th>Reference</th>
<th>Raw data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experim.</td>
<td>fcc</td>
<td>isotropic</td>
<td>1.710e-04</td>
<td>142.4</td>
<td>729 - 916</td>
<td></td>
<td>T.S. Lundy, J.F. Mur..</td>
<td>MDCS</td>
</tr>
<tr>
<td>Experim.</td>
<td>fcc</td>
<td>isotropic</td>
<td>2.200e-04</td>
<td>144.4</td>
<td>673 - 883</td>
<td></td>
<td>M. Beveler, Y. Adda..</td>
<td>MDCS</td>
</tr>
<tr>
<td>Experim.</td>
<td>fcc</td>
<td>isotropic</td>
<td>1.760e-05</td>
<td>126.5</td>
<td>358 - 482</td>
<td></td>
<td>T.E. Velin, R.W. Ball..</td>
<td>MDCS</td>
</tr>
<tr>
<td>Experim.</td>
<td>fcc</td>
<td>isotropic</td>
<td>1.370e-05</td>
<td>123.6</td>
<td>515 - 770</td>
<td></td>
<td>R. Messer et al.: Pr..</td>
<td>MDCS</td>
</tr>
<tr>
<td>First-prin.</td>
<td>fcc</td>
<td>isotropic</td>
<td>8.510e-04</td>
<td>131.2</td>
<td>654 - 934</td>
<td></td>
<td>S.-L. Shang et al.: ...</td>
<td>not available</td>
</tr>
<tr>
<td>Experim.</td>
<td>fcc</td>
<td>isotropic</td>
<td>1.000e-05; 9.00e-04*</td>
<td>121.7; 172.8*</td>
<td>515 - 916</td>
<td></td>
<td>S. Dais, R. Messer..</td>
<td>not available</td>
</tr>
<tr>
<td>CALPHAD</td>
<td>fcc</td>
<td>isotropic</td>
<td>1.080e-05</td>
<td>126.7</td>
<td>300 - 900</td>
<td></td>
<td>Y. W. Cui et al.; J. P..</td>
<td>not available</td>
</tr>
<tr>
<td>CALPHAD</td>
<td>fcc</td>
<td>isotropic</td>
<td>1.710e-04</td>
<td>142.0</td>
<td>300 - 900</td>
<td></td>
<td>A. Engstrom, J. Agre..</td>
<td>not available</td>
</tr>
<tr>
<td>CALPHAD</td>
<td>fcc</td>
<td>isotropic</td>
<td>8.233e-05</td>
<td>123.1</td>
<td>300 - 900</td>
<td></td>
<td>L. Zhang, Y. Du; NI..</td>
<td>not available</td>
</tr>
<tr>
<td>CALPHAD</td>
<td>hcp</td>
<td>⊥ c axis</td>
<td>2.380e-05</td>
<td>79.79</td>
<td>300 - 900</td>
<td></td>
<td>Y. W. Cui et al.; J. P..</td>
<td>not available</td>
</tr>
</tbody>
</table>

### Method: Experimental  
**Structure:** FCC  
**Direction:** Isotropic  
**Author:** Lundy  
**Year:** 1962

\[
D = D_0 \exp\left(-\frac{Q}{RT}\right)
\]

\[
* D = D_0^1 \exp\left(-\frac{Q^1}{RT}\right) + D_0^2 \exp\left(-\frac{Q^2}{RT}\right)
\]

\[
** D = D_0 \exp\left(-\frac{Q}{RT}\right) \exp\left(\frac{(T_M)^2}{T^2}\right)
\]
Workflow Tool Integration

Experimental Analysis Groups

NAIVE
Proprietary Software e.g: TA SW v2.0

INTERMEDIATE
Excel formulas and charts

EXPERT
Python Scripting and Matlab

Difficulties in sharing
Workforce Tool Integration

Reproducibility: MS Galaxy workflow
Workflow Tool Integration

Capabilities for MS-Galaxy users

How different levels of user can take advantage of galaxy

- **NAIVE**
  - Store data and perform analysis

- **INTERMEDIATE**
  - Design workflows

- **EXPERT**
  - Create custom tools for MS galaxy
NoMaD Repository
NoMaD Repository
ASM Taxonomy
Closing Remarks

- New Schemas
- Integration