AGENDA

--

2015 UCSB Workshop on Collection and Analysis of Big Data in 3D Materials Science

University of California – Santa Barbara
May 12th-14th 2015

1. Objective & Rationale

One accepted definition of 'Big Data' is "high volume, high velocity, and/or high variety information assets that require new forms of processing to enable enhanced decision making, insight discovery and process optimization. The concurrent progression of new tools and methodologies for materials investigation occurring in tandem with innovative experimental techniques has become a characteristic descriptor of 3D Materials Science as a field and created the existence of Big Data in the materials community. As examples, femto-second laser machining, micro- and nano-tomography and multi-modal data collection have advanced by leaps and bounds within the last decade alone. As a result, the creation or adapting of established tools for collecting, processing, analyzing and representing this data are being forced to grow and adapt as well. The 2015 UCSB Workshop on Collection and Analysis of Big Data in 3D Materials Science will focus on recently developed and evolving tools for processing Big data in 3D materials science, exposing participants to research solutions derived from other fields which may be applicable to pressing materials challenges. The workshop will accomplish this two-fold goal by providing overview plenary talks as a survey of the field and hands-on tutorials for currently-available data analysis tools.

2. Workshop Agenda

Day One	**Scheduled Item**	**Description**
6:00 PM – 8:30 PM	Informal Welcome Reception & Dinner	Best Western South Coast Inn - patio
Day Two		**Elings Hall – 1601**
8:15 AM 8:30 AM – 9:30 AM	Van departures from the hotel ICMSE Data & Workflow Plenary – Dave Furrer (Pratt & Whitney)	Introduction to the data flow in the ICMSE construct. Will set table for following plenaries and tutorials
9:30 AM – 10:15 AM	Modeling Materials Processing Plenary – John Elmer (LLNL)	Overview of commercial & open-source software in the area of materials processing & necessary input data, resulting output & critical assumptions associated with the tools
10:15 AM – 10:30 AM	Coffee Break	
10:30 AM – 12:30 PM	Working Group Tutorial Session 1: Spparks – John Mitchell (Sandia) DREAM3D – Mike Groeber (AFRL)	Hands-on work with Spparks and DREAM3D.
12:30 PM – 1:30 PM	Lunch	
1:30 PM – 2:15 PM	Representing Materials Structure Plenary – Surya Kalidindi (Georgia Tech.)	Overview of commercial and open-source software in the area of structure representation and the effect of analysis and processing choices on the results of these tools
2:15 PM – 4:15 PM	Working Group Tutorial Session 2: FIJI/ImageJ/DREAM3D – Mike Uchic (AFRL) CTEMSoft – Marc DeGraef (Carnegie Mellon Univ.)	Hands-on work with FIJI/ImageJ/DREAM3D and CTEMSoft.
4:15 PM – 5:00 PM	Specific Issues/Challenges Round Table – All Participants	Panel discussion session to begin discussion of weak-links in the tools discussed and used during the day.

		Opportunity to seed brainstorming ideas for following day.
Day Three		**Elings Hall - 1601**
8:45 AM 9:00 AM – 9:45 AM	Van departures from the hotel Modeling Materials Properties Plenary – Dennis Dimiduk (BlueQuartz Software)	Overview of commercial and open-source software in the area of materials property prediction and the necessary input data, resulting output and critical assumptions associated with these tools
9:45 AM – 11:45 AM	Working Group Tutorial Session 3: FFT – Tony Rollett (Carnegie Mellon Univ.) DREAM3D – Mike Groeber (AFRL)	Hands-on work with LANL FFT and DREAM3D.
11:45 AM – 1:00 PM	Lunch	
1:00PM – 2:00PM	National Data Initiatives and MGI Plenary – Chuck Ward (AFRL) & Large Group Discussion: Brainstorming Path Forward – All Participants	Overview of initiatives at the national level for the management and standardization of data for materials engineering & Panel discussion to note 'missing pieces of puzzle' and discuss efforts to fill them.
2:00 PM – 3:00 PM	UCSB Lab Tours & Group Photo – Tresa Pollock (Univ. California – Santa Barbara)	Opportunity to see laboratories that working on materials processing (SX Casting), structure collection (Tri-Beam) and property measurement (TBC testing?)
3:00 PM – 5:00 PM	Working Group Tutorial Sessions Cont. – All Participants	Flex-time for more hands-on work with previously introduced tools. Encouragement for attendees to bring own data or problems to tutorial leaders to begin transitioning learned tools to actual practice

3. Plenary Abstracts

ICME Data & Workflow Plenary

This talk will focus on introducing the construct of ICME and more specifically the data flow through the construct. An overview of the key 'modules' will be presented with context towards the information/data that flows into and out of each step in the ICME process. The talk will spend time where necessary highlighting the current broken or missing links in the data workflow chain. Insights into the manner in which and the need to track how uncertainty and error propagate through the process outlined will be offered. The talk will serve to set the stage for the following plenaries to speak to the individual 'modules' in more detail, but should provide the larger context of the need to manage data flow.

Modeling Materials Processing Plenary

This talk will focus on introducing the field of materials processing simulation codes. An overview of the field, from large-strain forming to heat-treatment and casting/solidification will be presented at a high-level. Focus will be paid to the available tools in each of these spaces, both from a commercial and open-source/research-grade perspective. The deep inner-workings of the codes will not be presented, but rather a specific focus will be paid to the input and output information that is typical of software tools that sit within this space of the Integrated Computational Materials Engineering (ICME) paradigm. For example, the need to have explicit or statistical representations of (micro-)structure as input to processing simulations will be discussed. Further, insight about the barriers to including (micro-)structure in tools

that currently do not treat it will be offered. The talk will also comment on some typical or common post-processing steps applied to output information in preparation for its passing to subsequent tools in the ICME workflow. The goal of this talk is not to create experts in any one area of this vast subspace of ICME, but rather provide a context and understanding of some of the information that flows into and out of these tools and highlight some of the assumptions and limitations of these tools as the stand today. Specific examples of processing tools may include, but are not limited to: DEFORM, ProCast, PhaseField, grain growth codes (eg SPParks), multiphysics tools (eg Albany), etc.

Representing Materials Structure Plenary
This talk will focus on introducing the field of materials structure representation codes. An overview of the approaches to quantifying the internal structure, from feature distributions to n-point statistics to principle component analysis (PCA) will be presented at a high-level. Focus will be paid to the available tools in each of these spaces, both from a commercial and open-source/research-grade perspective. The deep inner-workings of the codes will not be presented, but rather a specific focus will be paid to the input and output information that is typical of software tools that sit within this space of the Integrated Computational Materials Engineering (ICME) paradigm. Concepts related to sampling statistics, representative volume elements (RVEs) and similarity metrics will be discussed as areas of growth in the objective description of structure. The talk will address how the requirements of the modeling domains of ICME establish requirements for structure representation and should drive techniques in this area. Some focus will also be paid to the synergy that must exist between the collection devices and the structure analysis tools to provide the most accurate and useful descriptions of structure. Specific examples of structure representation/processing tools may include, but are not limited to: DREAM3D, FIJI/ImageJ, Aviso, MIPAR, MTEX, etc.

Modeling Materials Properties Plenary
This talk will focus on introducing the field of materials property simulation codes. An overview of the field, from the Finite Element Method (FEM) to Fast Fourier Transform (FFT) to Dislocation Dynamics (DD) to empirical models will be presented at a high-level. Focus will be paid to the available tools in each of these spaces, both from a commercial and open-source/research-grade perspective. The deep inner-workings of the codes will not be presented, but rather a specific focus will be paid to the input and output information that is typical of software tools that sit within this space of the Integrated Computational Materials Engineering (ICME) paradigm. For example, the need to have explicit or statistical representations of (micro-)structure as input to processing simulations will be discussed. Further, insight about the barriers to including (micro-)structure in tools that currently do not treat it will be offered. The talk will also comment on some typical or common pre-processing and post-processing steps applied to input and output information in preparation for its use in these tools or its passing to subsequent tools in the ICME workflow. The goal of this talk is not to create experts in any one area of this vast subspace of ICME, but rather provide a context and understanding of some of the information that flows into and out of these tools and highlight some of the assumptions and limitations of these tools as the stand today. Specific examples of property simulation tools may include, but are not limited to: Abaqus, Ansys, LANL FFT, ParaDis, multiphysics tools (eg Albany), spreadsheet models etc.

National Data Initiatives and MGI Plenary
This talk will focus on overviewing the current national efforts to establish practices and standards for managing, storing and working with data in materials engineering. The intent of the talk is to educate the community of common efforts ongoing across the country and encourage collaboration in data sharing and archiving.

4. Intended Participants

The intended audience for the workshop shall be graduate students, academics and government laboratory personnel who use or anticipate using currently available commercial or non-commercial tools designed for treating and analyzing large datasets for materials science challenges in three dimensions. A target attendance of approximately 40 persons is sought with a subset of 20 – 30 persons of the aforementioned 40 being graduate and undergraduate students.