Integrative Materials Design (IMaD) - Leverage, Innovate, and Disseminate

B. Blaiszik[1,2], I. Foster[1,2], J. Allison[3], D. Morgan[4], D. Trinkle[5], P. Voorhees[6]

[1] University of Chicago [2] Argonne National Laboratory [3] University of Michigan [4] University of Wisconsin-Madison
[5] University of Illinois at Urbana-Champaign, [6] Northwestern University

OBJECTIVES AND DESIGN GOALS

Leverage and integrate the expertise, combined knowledge base, facilities, and services available through a consortium of academic, industrial, and government collaborators

Innovate via MDF's scalable data infrastructure and cloud services to promote materials data publication, sharing, and discovery to expedite and advance materials designs for the aeronautics and automotive sectors

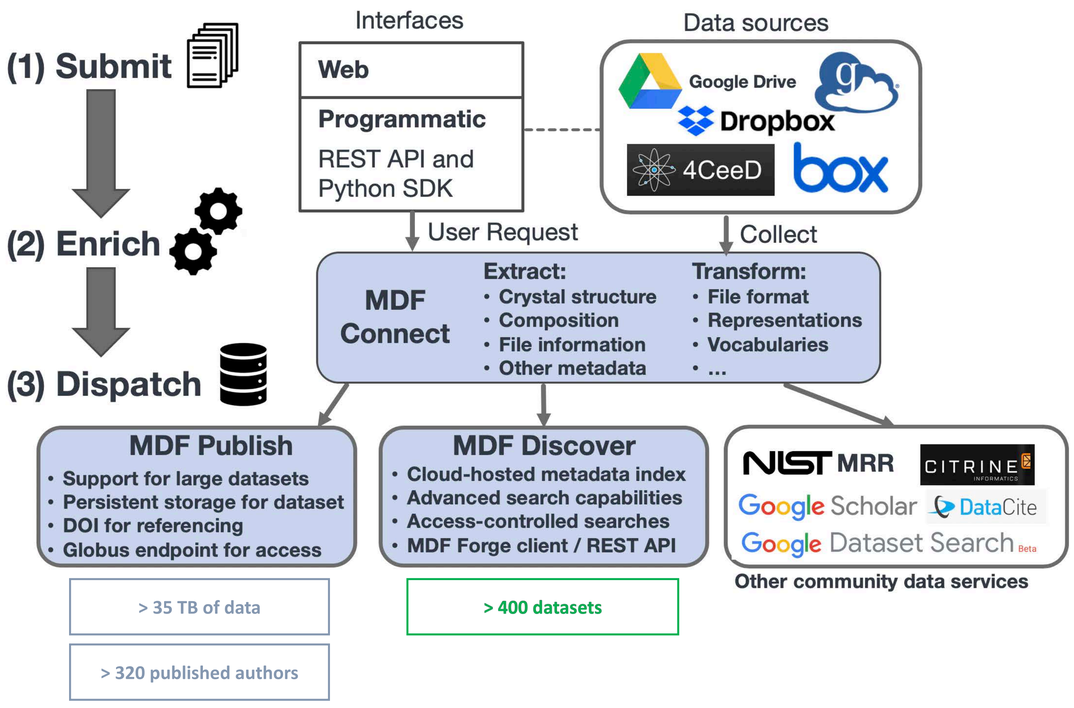
Disseminate distributed and distinct Midwest materials data resources to increase access and reuse of valuable materials data assets

Partners

Building an Open Materials Data Ecosystem

- Emerging ecosystem includes many separate data services to solve specific problems
- MDF Connect acts as the glue to bind these together into a cohesive system.
- One interface to deposit into many services
- Strictly opt-in for cross-posting datasets
- Data may be deposited through Web UI or API

CHiMaD NIST



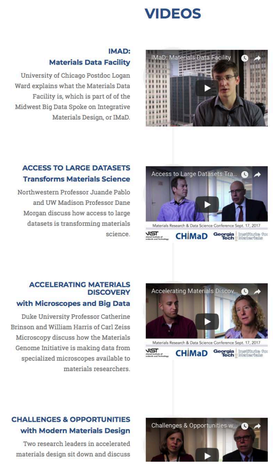
Outreach

- Visited each partner site and created a video showcasing people, facilities, and data services
- Inform the community and public about the materials informatics work being done in the Midwest through videos news articles, webinars, tutorials, workshops, etc.

Science Communication Program

Elena Bruess is currently producing a podcast called "Materials Data Matter"

Led by Laura Bartolo



Industry Engagement

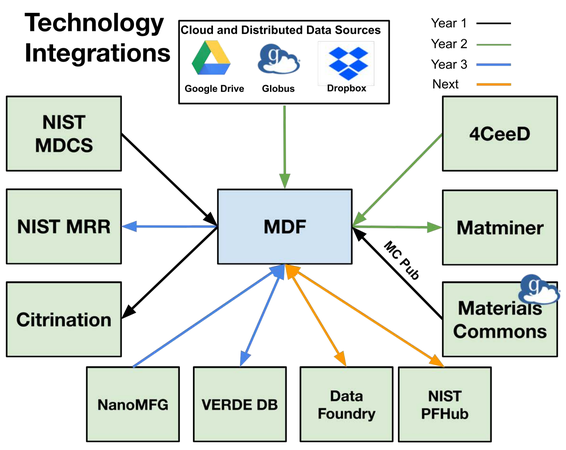
- Held on site training
- Developed plan to use IMaD data ecosystem for Questek data
- **Awarded joint DOE SBIR Phase I building a data service to speed thermoelectric material design**

- PI Foster won a DOE HPC for Manufacturing (HPC4MFG) award to work with 3M to optimize fiber spinning process with machine learning and advanced data capabilities
- **New funded effort with industrial partner (partial match)**

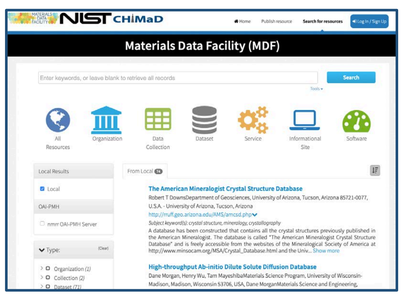
- Bimonthly meetings to discuss service integrations and machine learning
- Resulted in 2 joint papers
- **New joint industry-funded project**

- The IMaD data ecosystem (via MDF Connect) is being used to the Air Force Research Laboratory Additive Manufacturing Modeling Challenge
- The IMaD data ecosystem (via MDF Connect) is being used to host the Sandia Fracture Challenge experimental results

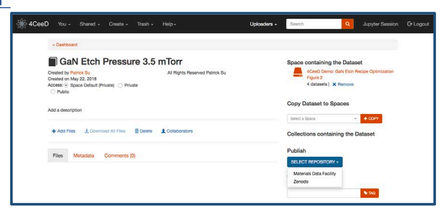
Data Integrations



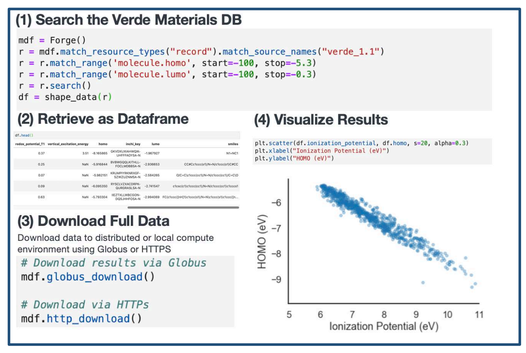
NIST/CHiMaD Materials Resource Registry Instance



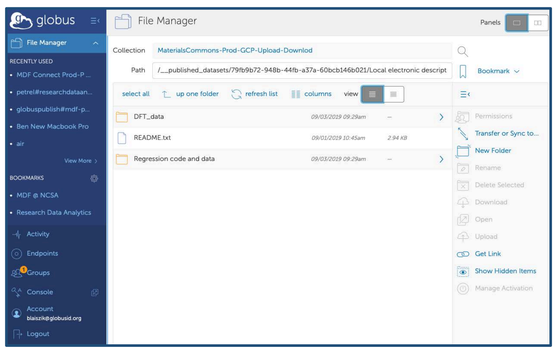
4CeeD to MDF



VERDE Materials DB Interface



Materials Commons data via Globus



Future Directions

- Continue building data integrations between community services and tools
  - MAST-ML
  - Materials Commons
- Release new tutorial videos to speed user onboarding
- Finish recording and release podcast
- Hold workshop to promote usage of the available data ecosystem

Papers and References

- A Data Ecosystem to Support Machine Learning in Materials Science DOI: 10.1557/mrc.2019.118
- Matminer: An open source toolkit for materials data mining DOI: 10.1016/j.commatsci.2018.05.018
- Machine Learning Prediction of Accurate Atomization Energies of Organic Molecules from Low-Fidelity Quantum Chemical Calculations DOI: 10.1557/mrc.2019.107
- Strategies for accelerating the adoption of materials informatics DOI: 10.1557/mrs.2018.204
- Automated Data Curation for Electron Microscopy using the Materials Data Facility

This work was also supported by the National Science Foundation as part of the Midwest Big Data Hub under NSF Award Number: 1636950 "BD Spokes: SPOKE: MIDWEST: Collaborative: Integrative Materials Design (IMaD): Leverage, Innovate, and Disseminate".

Northwestern University · Wisconsin–Madison · University of Michigan · University of Illinois at Urbana-Champaign · The University of Chicago · Argonne National Laboratory