MaRDA MDI Provider Integration and Interoperability WG Collaborative Notes

Thursday, May 6, 2021

Quick Links

- This Document: [MaRDA-MDI3-2021-05-06](#)
- MaRDA: https://www.marda-alliance.org/
- MaRDA Slack: https://mardaworkspace.slack.com
- Discussion Board: https://matsci.org/c/marda/45
- WG GitHub: https://github.com/marda-alliance/MDI-Provider-Integration-Interoperability
- 2021 Annual Meeting Documents: https://bit.ly/marda2021documents
- **Slido Poll: https://app.sli.do/event/vmnlkvyk**
- **Slido ID: #MDI-May**

Agenda

- Welcome
- Introduction to MaRDA and this Working Group
- Slido Poll
- Open discussion

Attendance

Please sign in:

Name	Email	ORCID
Zachary Trautt	zachary.trautt@nist.gov	0000-0001-5929-0354
Suhas Somnath	somnaths@ornl.gov	0000-0002-5398-3050
Joe Eklund	joe@llnl.gov	
Michael Tynes	mtynes@lanl.gov	0000-0002-5007-1056
Branden Kappes	bkappes@kmmd.io	0000-0001-9478-3496
Patrick Huck	phuck@lbl.gov	0000-0001-8237-2227
Marcus Hanwell	mhanwell@bnl.gov	0000-0002-5851-5272
Matthew Jacobsen	matthew.jacobsen.1@us.af.mil	

Sam Chance	sam@cambridgesemantics.com	
Joshua Latimer	joshua.latimer.ctr@us.af.mil	
James Fourman	james.fourman.ctr@afresearchlab.com	0000-0002-6783-2800
Tristan Uribe	tristan.uribe.ext@afresearchlab.com	0000-0003-4477-7703
Grant Phillips	grant.phillips.3.ctr@afresearchlab.com	
Jason Thiese	jason.thiese.1.ctr@afresearchlab.com	
Kareem Aggour	aggour@ge.com	
Timur Bazhirov	timur@exabyte.io	0000-0003-3719-522X
Santanu Chaudhuri	schaudhuri@anl.gov	
Alexander Zech	alexander.zech@berkeley.edu	0000-0002-4332-8546

Comments and Notes

- Interoperability was the dominant entry in the word cloud, along with vocabularies, ontologies, and schemas.
- Discussion on hackathon format. Large preference for in-person. Hybrid format can utilize in-person "satellites", where geographic subsets assemble to work on specific areas of research, and then report back to the larger group.
- Lightning talks - overwhelmingly favorable, early topics could include Research Data Framework (to be covered at the June meeting), and the FAIR implementation framework.
- Communicate activities - increase awareness of technologies and methods, avoid duplication of efforts
- Need a page of upcoming events, the projects we are all working on
- Group's preference on a communication channel - slack vs. matsci website - ambivalent / no response
- Industry perspective - GE has successfully used the hackathon model to share tools and data between business units
- Domain experts at hackathons, what is the tool that you don't have to use your data
  - Not necessarily very effective since domain scientists often give thumbs up regardless of what was delivered.
  - Keeping domain scientists engaged might be a challenge but it would be great to try to do.

- - Interoperability will be challenging without having the domain experts and MDI developers in the same place.
- May need breakout groups to ensure that conversations continue to take place and are productive. People should return from the breakout rooms and debrief before the meeting ends
  - We will need topics for each group to work on / discuss.
  - Structure will coalesce over the next couple of meetings

Slido Polls

What problems can this group address?

ontology (schema) definition
Unified Data Standards
Interactivity
Adoption of schemas
JSON
Sustainability
Ontologies
Vocabulary
Security
Rapid
Reusability
REST
Interoperability
Codeless
IoT
common vocabulary
learn about technologies
knowledge graphs
Schemas
devops
Reduce FAIR principles to practice
classification of currently available schemas

What are your thoughts on virtual, hybrid, and in-person hackathons?

- Honestly don't work well, tend to avoid them and hope for better days!
- Maybe a hybrid approach with some in-person interaction followed by virtual coding

- Local hackathons made in-person but clusters grouped virtually - it needs little bit more in-person driving
- I think in-person > virtual > hybrid in terms of productivity and experience. Discovering and forming groups / teams might be a challenge in anything that is not in-person. However, we can always use polls or similar to solve this problem. Hybrids may cause the remote participants to lose interest / be forgotten about in the live conversations
- In-person would be preferable. Using virtual tools at the planning phase of the hackathon would be nice so that when it is time for the in person portion of the event, we can get right into the work.
- We have done several virtual hackathons with varied success at LLNL. In person definitely more effective.
- A lot of value in just pointing at things, sketching on whiteboards, and looking over people's shoulders when in-person.
- Virtual hackathon needs better planning - the spontaneity is lost
- In-Person hackathons are a much better solutions due to communication and technology issues with a hybrid model or virtual.
- In-person. Virtual lacks the interactivity and dynamism of in-person interaction.
- In-person are ideal. Having virtual options can work, assuming the right

Should we have a lightning talk each month?

Support

88%

Neutral

13%

Oppose

0%

Introductory Slides

MaRDA Working Group:
Materials Data Infrastructure
Provider Integration and Interoperability

Chairpersons

- Santanu Chaudhuri (Argonne National Laboratory)
- Marcus D. Hanwell (Brookhaven National Laboratory)
- Matthew D. Jacobsen (Air Force Research Laboratory)
- Klara Nahrstedt (University of Illinois at Urbana–Champaign)
- Suhas Somnath (Oak Ridge National Laboratory)
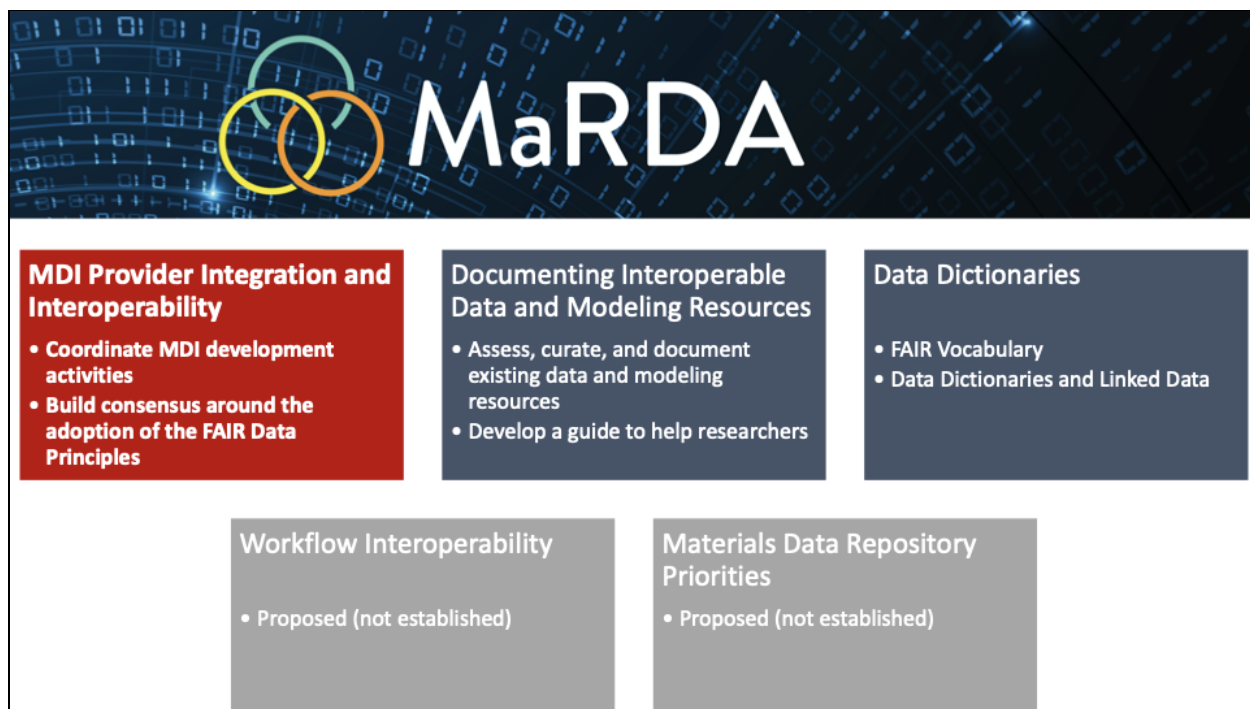- Zachary T. Trautt (National Institute of Standards and Technology)

Who are we?

MaRDA is a community network focused on developing the open, accessible, and interoperable materials data that fuels the Materials Genome Initiative (MGI).

MaRDA is a convergence of people and ideas working together to **connect materials data infrastructure** to **accelerate discovery**, enable new insights into **materials mechanisms**, and lay a foundation for both human-centered and artificial intelligence-assisted approaches to **materials design**.

https://www.marda-alliance.org/

MaRDA

MDI Provider Integration and Interoperability
- Coordinate MDI development activities
- Build consensus around the adoption of the FAIR Data Principles

Documenting Interoperable Data and Modeling Resources
- Assess, curate, and document existing data and modeling resources
- Develop a guide to help researchers

Data Dictionaries
- FAIR Vocabulary
- Data Dictionaries and Linked Data

Workflow Interoperability
- Proposed (not established)

Materials Data Repository Priorities
- Proposed (not established)

MDI Provider Integration and Interoperability WG

Goals
- Coordinate MDI development activities
- Build awareness and consensus within the MDI provider community
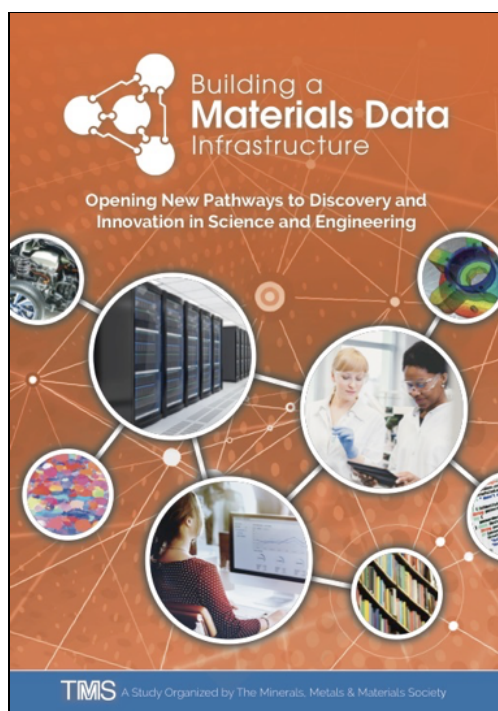
Work Plan
- **Host Monthly virtual meetings**
- **Host specific, task-oriented, events** (e.g., **hackathons**, workshops, etc.) as identified and planned during regular business during monthly meetings
- Coordinate our activities with other groups (e.g., MaRDA, RDA, CODATA, GO FAIR, etc.)

Federated MDI Development Loop

Survey

Monthly Online Meetings

Operate

Plan

Independent DevOps Loop

Hackathons

Develop

Prototype

Incubate New Working Groups
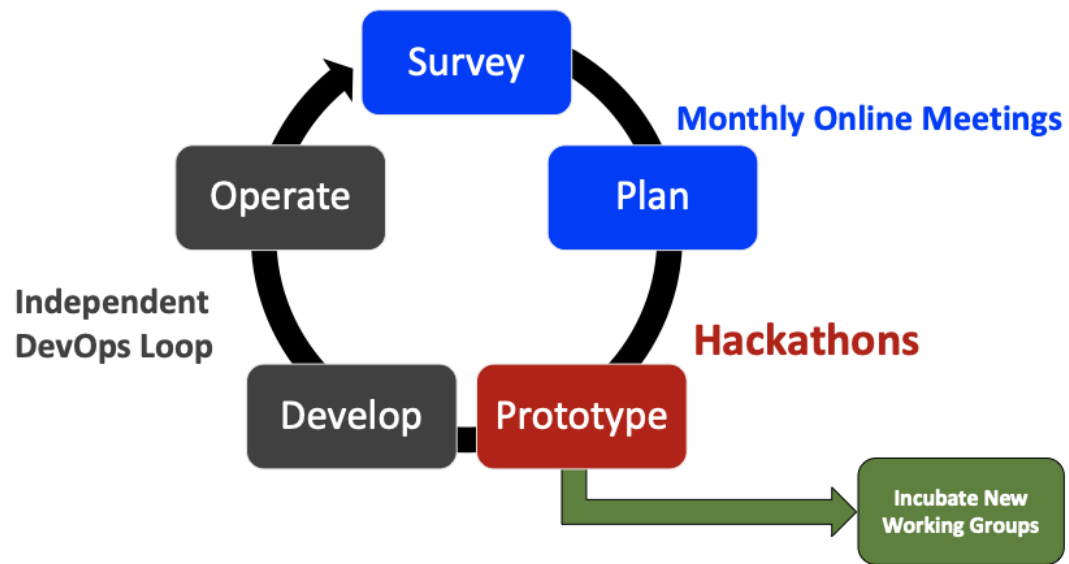
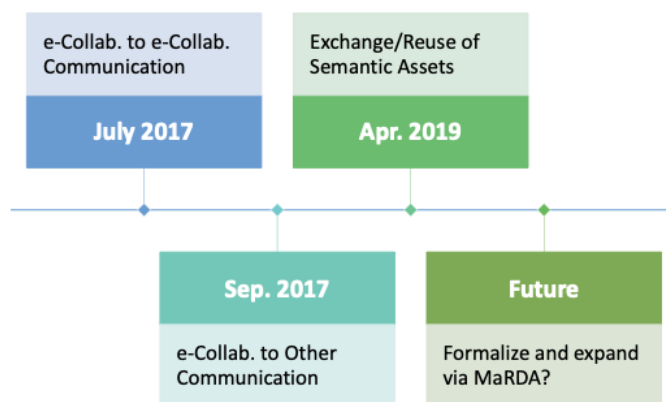# Past Hackathons

NIST-AFRL Hosted Events

Recommendation 4: Develop demonstration projects and cross-disciplinary community efforts that enhance and accelerate adoption of the MDI
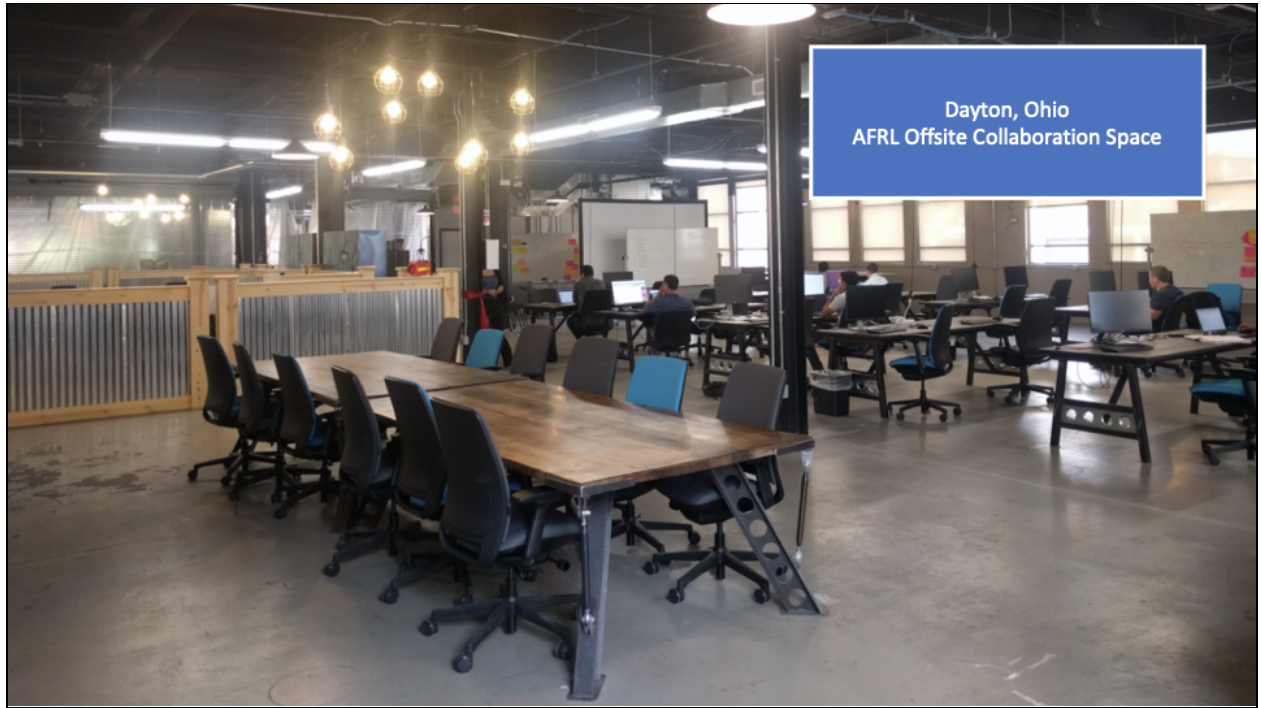
Tactic #3: *Launch targeted community efforts to help achieve MDI critical mass in the MSE community*

"Critical mass" in the present context refers to the number of implementers and users of the MDI needed to make it self-propagating, or organic, in its widespread adoption, growth, and sustainability. Although some specific types of community efforts are discussed in this tactic for reaching critical mass of users and implementers, many of the other recommendations and tactics in this report could also contribute to obtaining such critical mass, which will in turn accelerate the widespread adoption of the MDI.

Targeted efforts toward reaching a critical mass could encompass four different elements: (1) the service provider-level of MDI integration, (2) individual, user-level engagement of the MDI, (3) institutional-level MDI-focused working groups or committees, and (4) a broader community wide-level engagement (such as data-digitization projects).

https://tms.org/mdistudy

Hackathon Events

| e-Collab. to e-Collab. Communication |
| July 2017 |

| Sep. 2017 |
| e-Collab. to Other Communication |

| Exchange/Reuse of Semantic Assets |
| Apr. 2019 |

| Future |
| Formalize and expand via MaRDA? |

Dayton, Ohio
AFRL Offsite Collaboration Space

Rockville, MD
NIST, NCCoE

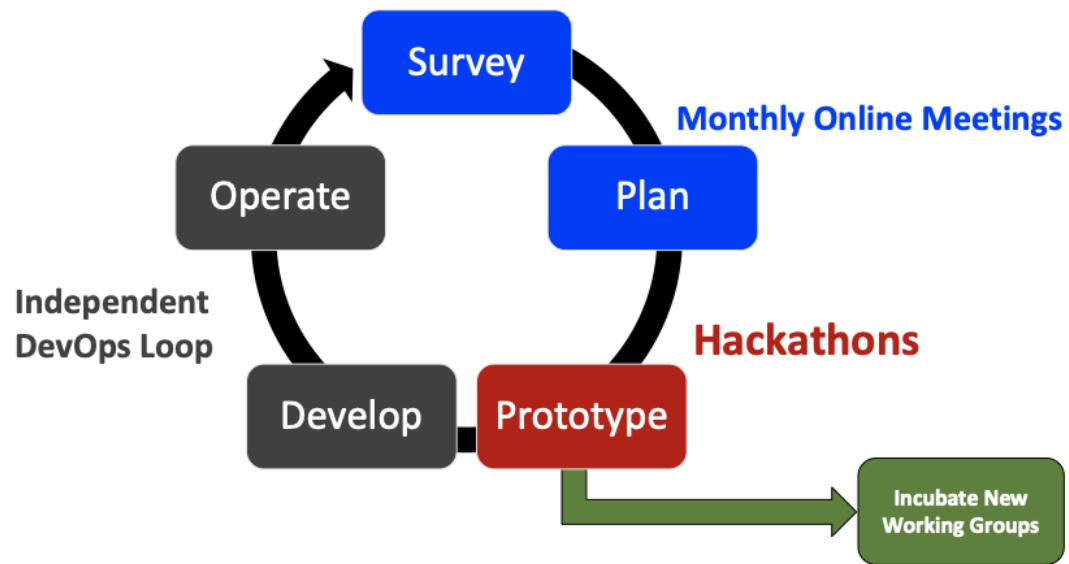|	Monday	Tuesday	Wednesday	Thursday	Friday
8:30 - 9:00	Pass Security	Pass Security	Pass Security	Pass Security	Pass Security
9:00 - 9:30	Talks/Demos	Hacking	Hacking	Hacking	Present Results
9:30 - 10:00	Talks/Demos	Hacking	Hacking	Hacking	Present Results
10:00 - 10:30	Talks/Demos	Hacking	Hacking	Hacking	Present Results
10:30 - 11:00	Hack Planning	Hacking	Hacking	Hacking	Discussion
11:00 - 11:30	Hack Planning	Hacking	Hacking	Hacking	Discussion
11:30 - 12:00	Hack Planning	Hacking	Hacking	Hacking	Closing Remarks
12:00 - 12:30	Lunch	Lunch	Lunch	Lunch	Lunch
12:30 - 13:00	Lunch	Lunch	Lunch	Lunch	Lunch
13:00 - 13:30	Networking	Networking	Networking	Networking	
13:30 - 14:00	Networking	Networking	Networking	Networking	
14:00 - 14:30	Hacking	Hacking	Hacking	Hacking	
14:30 - 15:00	Hacking	Hacking	Hacking	Hacking	
15:00 - 15:30	Hacking	Hacking	Hacking	Hacking	
15:30 - 16:00	Hacking	Hacking	Status Updates	Status Updates	
16:00 - 16:30	Hacking	Status Updates	BOF Discussions	BOF Discussions	
16:30 - 17:00	Status Updates	Lightning Talks	BOF Discussions	BOF Discussions	

Legend

	Downtime
	Presentations
	Group Discussion
	Perform Work

Federated MDI Development Loop

Monthly Online Meetings

Hackathons

Independent DevOps Loop

Survey · Plan · Prototype · Develop · Operate

Incubate New Working Groups

Questions

- *What are some open challenges?*
- *How do we plan more and better hackathons?*
- *How do we support virtual/hybrid hackathons?*