



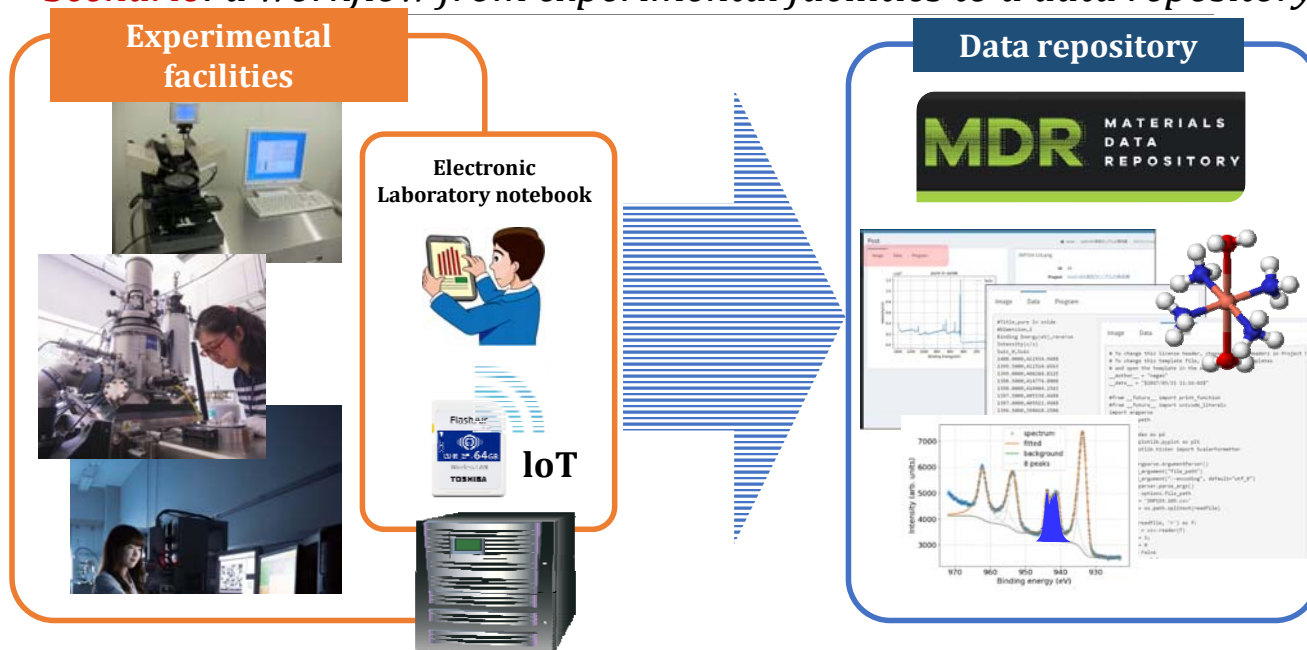
Mikiko Tanifuji
 Managing Director, Materials Data Platform Center
 National Institute for Materials Science, Tsukuba (NIMS), Japan

COAR Annual Meeting & General Assembly, May 21-23, 2019 @ Lyon, France



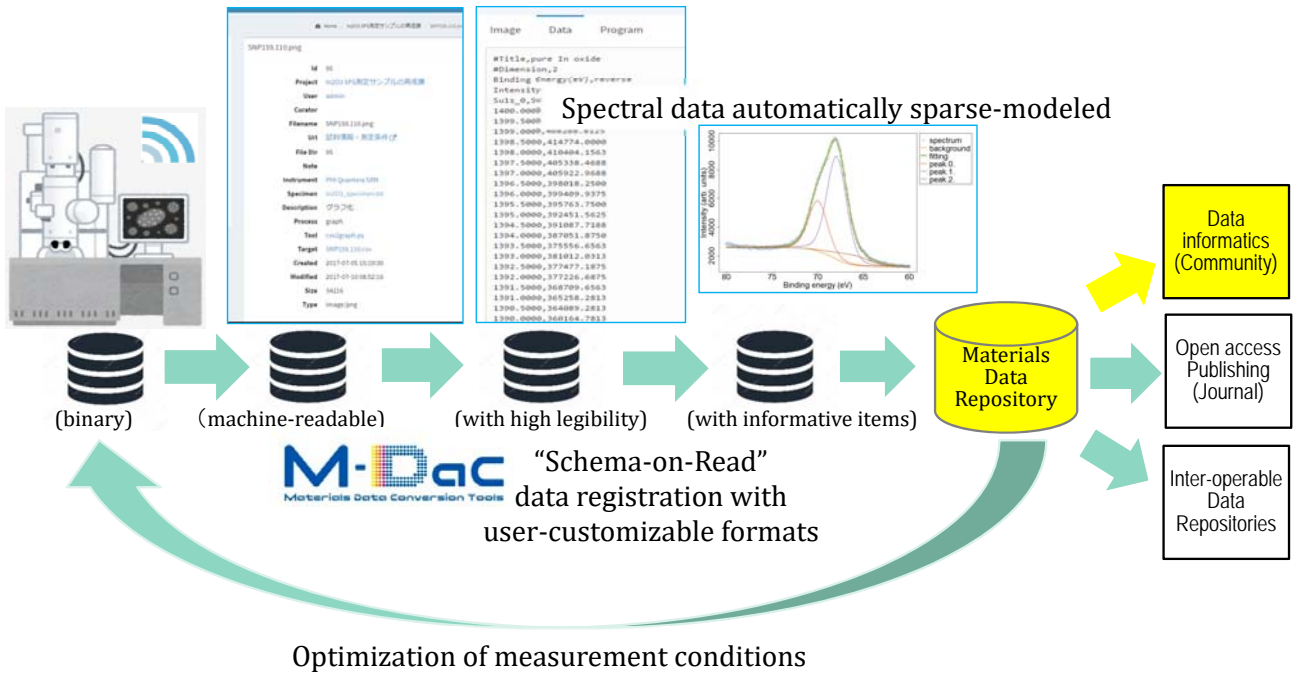
Materials Data Repository

- *Scenario: a workflow from experimental facilities to a data repository*



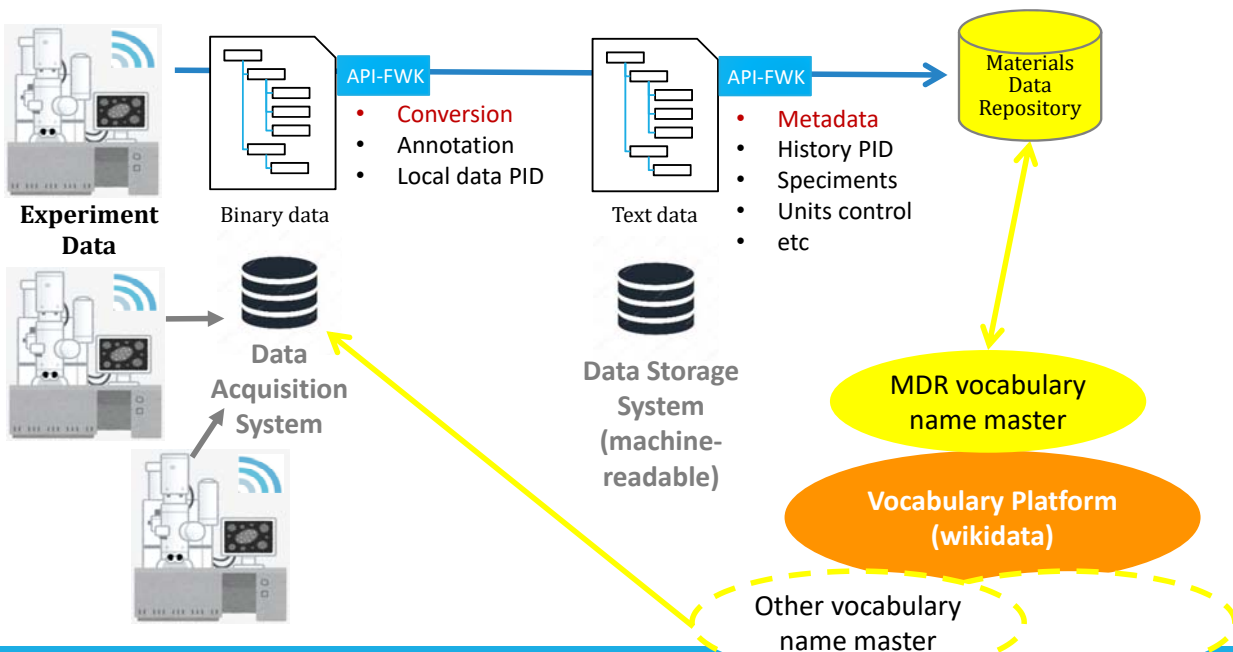
Materials Data Repository

- Scenario: a workflow from experimental facilities to a data repository

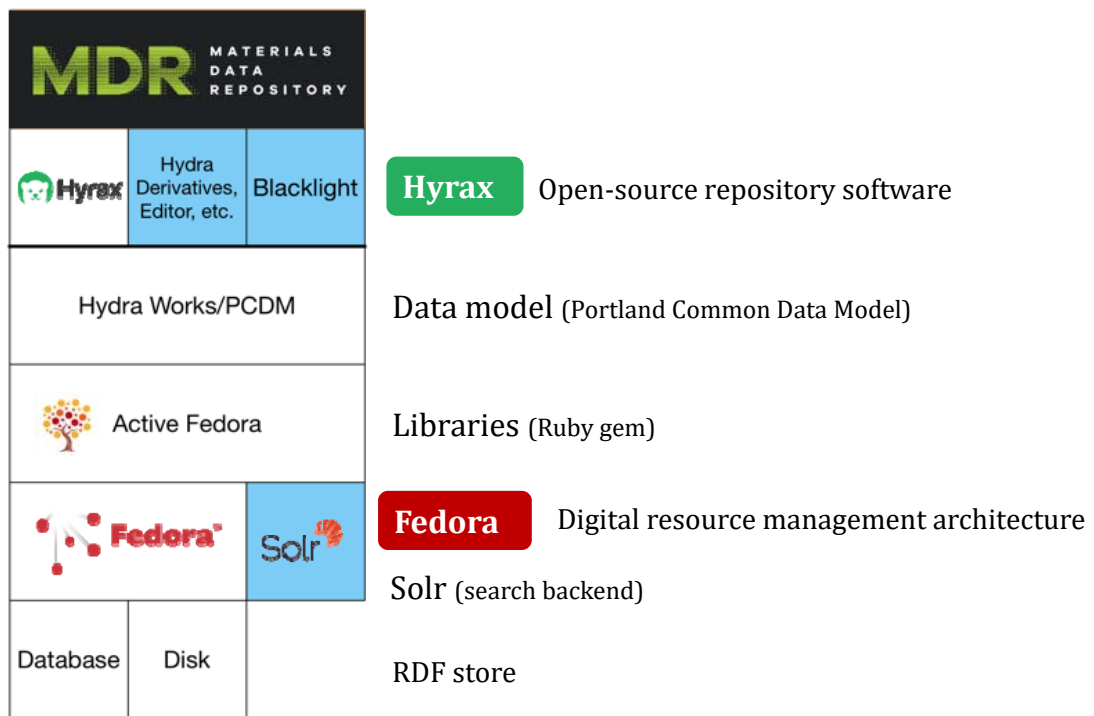


Technical challenges 1:

Metadata as a data recipe: how to assist researchers?



Data-centric-repository system



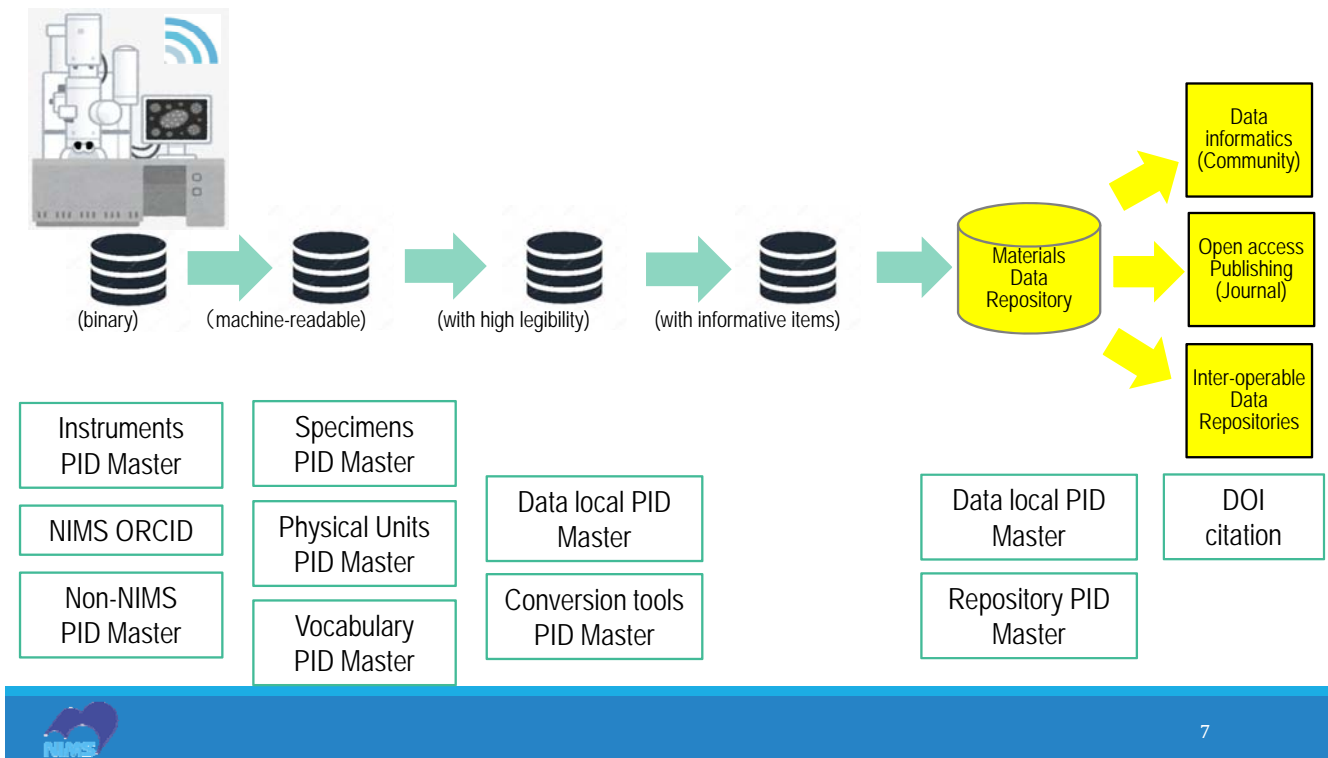
Technical challenges 2:

ResourceSync

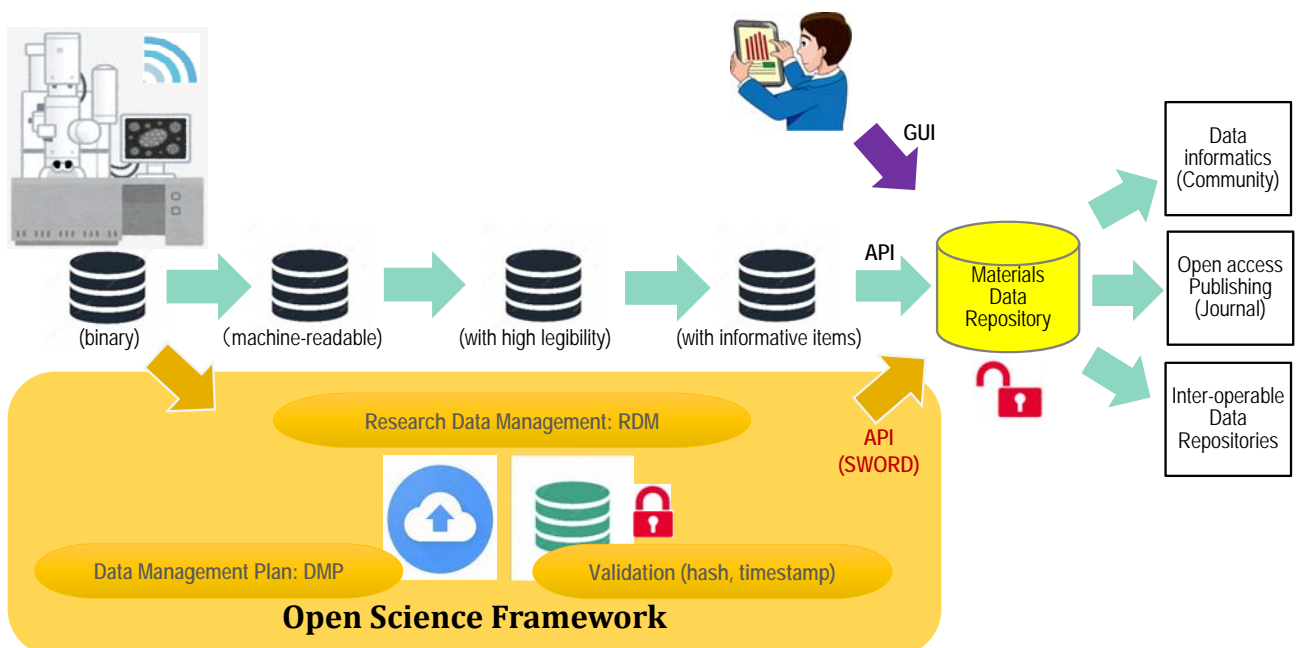
1. ResourceSync is recommended by the COAR Next Generation Repository report as a successor to OAI-PMH.
2. ResourceSync is implemented in the MDR (also OAI-PMH)
3. It will allow both the metadata and (in some cases) the content (research data, publications) to be harvested by other services on the network.
4. NIMS will be testing this with the Open University's Core aggregator system in the next few weeks.



Technical challenges 3: Using Persistent Identifiers



Technical challenges 4: SWORD



5 Challenges of Materials Data Platform for Open Science

1. Quality

- Identify who/what/when/how
- Integrity of the data

2. Accessibility with Open Data

- Open data for data publishing
- Linked data
- Data search for machine learning

3. Usability

- Machine-readability
- Metadata as data recipes for informatics
- Data licensing (CC, CC-BY-NC, MIT, etc.)

4. Security and Preservation

- Open data policy
- User identifications
- Data preservation
- Cyber security

5. Research-aids on the platform

- Vocabulary assistance (for data curation, collection, conversion for AI)
- Data analysis software

Summary: How the MDR is following the NGR recommendations

1. Exposing Identifiers
2. Declaring Licenses at the Resource Level
3. Discovery Through Navigation
4. Interacting with Resources (Annotation, Commentary, and Review)
5. Resource Transfer
6. Batch Discovery
7. Collecting and Exposing Activities
8. Identification of Users
9. Authentication of Users
10. Exposing Standardized Usage Metrics
11. Preserving Resources