



# Vocabulary for the Materials Data Platform at the National Institute for Materials Science

Asahiko Matsuda and Kadohira Takuya

Materials Data Platform Center,  
National Institute for Materials Science



Asahiko Matsuda

- Engineer, Materials Data Platform Center
- National Institute for Materials Science (NIMS)
- Tsukuba, Japan

## **1. Name of our vocabulary**

- Still under development, no formal name
- Let's call it Materials Data Platform (MDPF) Vocabulary for now

## **2. Application domain**

- Materials science



- **3. Intended purpose**

- NIMS's Materials Data Platform is a collection of about ten different systems
  - Data Collection System gathers research data created by materials researchers, and adds metadata.
  - Materials Data Repository is the place for publishing the materials datasets, with metadata.
- All systems share a common metadata model, which can take values from a vocabulary service.
- Our vocabulary service uses Wikibase, the software behind Wikidata.

- **4. World representation**

- Up to the researchers! (They can submit their own new terms)



## 5. Upper level concepts

- No upper level concepts yet, taking bottom-up approach.
- Our current highest concepts follow the NIST vocabulary ([doi.org/10.18434/T4/1435037](https://doi.org/10.18434/T4/1435037)):
  - Data origin
  - Material types
  - Structural features
  - Properties addressed
  - Characterization methods
  - Synthesis and processing



## 6. Use cases

- A user browsing a materials data repository can look for materials datasets using a search engine that's aware of synonyms, related concepts, and concept classifications.
- A text and data mining researcher can annotate or mine the literature using systems that are connected to the vocabulary engine, allowing for easier annotation or more semantic-aware text and data mining.

## 7. Overlaps with other taxonomies

- Our vocabulary overlaps heavily with NIST's Materials Vocabulary as we are building on top of it.