We are developing a materials data platform for researchers.

### Metadata for materials data

1. **Metadata as a custom schema**
   - "One JSON schema to rule them all"?
   - **Common metadata**
     - IDs
     - Authors
     - Data origin
     - Date ...
   - **Characterization metadata**
     - Synthesis metadata
     - Calculation metadata
   - Difficulties in implementing a complex structure.
   - Massive edit forms → Users were overwhelmed.

2. **Metadata as directories**
   - Researchers already manage their data using folders.
   - Pre-defined structure:
     - OSC-100 (Instrument ID)
     - u01234 (User Name/ID)
     - ProjABC (Project Name)
       - data.csv
   - Implemented as part of our IoT-assisted data collection system.
   - Mostly positive researcher feedback.

3. **Metadata in a data package**
   - Making full use of package metadata files.
   - Data packaging:
     - Making data self-documenting and self-contained. RO-Crate, BagIt, etc.
   - Assists wide distribution beyond our platform.
   - Parts can have different metadata from the whole.
   - Limited vocabulary for domain-specific metadata.
   - Potential topic for community discussions.
     - (See Bioschemas, an extension of Schema.org)

---

Thoughts

- **Bibliographic metadata** are well standardized, while domain-specific scientific metadata require discussions in respective communities.
- Dealing with heterogeneous data requires a common metadata schema, but it’s not always a good idea to expose the whole to the users.
- There’s more than one way to deal with metadata. Look for a practical combination of methods.