

Metadata in the Research Workflow: Tools for Enrichment and Validation of Structured Metadata

Anton Pirogov¹, Fiona D'Mello¹, Volker Hofmann¹, Stefan Sandfeld¹

¹Institute for advanced simulation - Materials Data Science and Informatics (IAS-9), Forschungszentrum Jülich, Jülich, DE

{a.pirogov, f.dmello, v.hofmann, s.sandfeld}@fz-juelich.de



solves the following problem:

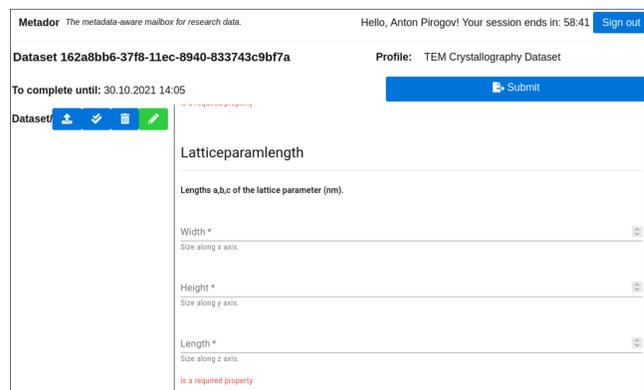
*Help enriching research data from an unstructured environment
With structured, domain-specific metadata to make it more FAIR*



Click for Video!

Key features:

- ORCID-based authentication
- Resumable uploads of large files
- Forms generated from **JSON Schemas**
- Combines data and relevant metadata
- Easy deployment and integration



Outlook:

Currently, we work on extending Metador into a **modular, extensible** and **generic** metadata-focused platform for “medium-hot” research data, fostering:

Interoperability (structural + semantic):

- JSON Schema, JSON-LD
- ROcrate** support built-in
- Custom ontologies can be used
- Planned integrations with InvenioRDM + JupyterHub



Data (Re-)Use:

- Granular **metadata queries** to quickly find relevant datasets
- Automatic visual **dashboards** for quick overview and insights

Name	Category	By	Email	On	Actions
565a36e-2876-11e4-a555-d8-9437f9027	TEM	Author Name	author@email.com	Month/Date/Year	[A] [D]
679a3d8-2876-11e4-a22a-58-756a5851	TEM	Author Name	author@email.com	Month/Date/Year	[A] [D]



Interested? Get in touch! <https://github.com/Materials-Data-Science-and-Informatics/metador>



<https://github.com/Materials-Data-Science-and-Informatics/dirschema>

Interested? Get in touch!

solves the following problem:

Formalize and enforce conventions to be used and metadata to be provided for files in a dataset (e.g. for publishing on Zenodo or in preparation for data processing)



Key features:

- YAML-based specification with JSON Schema based keywords
- Allows to require presence or absence requirements for files and directories
- Allows to require the existence of a „companion“ metadata JSON file that must be valid according to an external or embedded JSON Schema
- Versatile – can be used on plain directory structures as well as e.g. HDF5 files, Can be easily extended to support other hierarchical archive formats

