

Metador: A metadata-centric framework for enabling FAIR research (meta)data handling

Anton Pirogov¹ // Mustafa Soylu¹ // Fiona D'Mello¹ // Volker Hofmann¹ // Stefan Sandfeld¹
{a.pirogov, m.soylu, f.dmello, v.hofmann, s.sandfeld}@fz-juelich.de

¹ Materials Data Science and Informatics (IAS-9), Forschungszentrum Jülich GmbH

Motivation

- development of high **quality** standards and software **is difficult** and **slow**
- bottom-up efforts often solve similar RDM problems with incompatible solutions

Science needs better software tools



Tools require (meta)data standards

Project Goals

Metador will support scientists by:

- helping to **organize** research **(meta)data**
- facilitating** bottom-up, pragmatic and **incremental harmonization** of schemas
- accelerating development** of useful metadata-driven tools and services

Findable

search for (meta)data of certain types, making datasets findable

Accessible

Metadata-driven automatic **dashboards** for data introspection

Interoperable

JSON Schema and JSON-LD export, aligned with **schema.org** / **ROcrate**

Reusable

an ecosystem of **reusable** and **generic** software components

Current Status

- operational container API and plugin system
- prototype of schema and widget system
- Metador dashboard integrated into InvenioRDM <https://inveniosoftware.org/products/rdm>

Next Steps

- Implementation of scientific pilot use case
- Development of improved user-oriented tooling

Interested? Get in touch!

<https://materials-data-science-and-informatics.github.io/metador-core>



Architecture of the METADOR Framework

Metador = Plugin System + (Meta)data Packaging Format

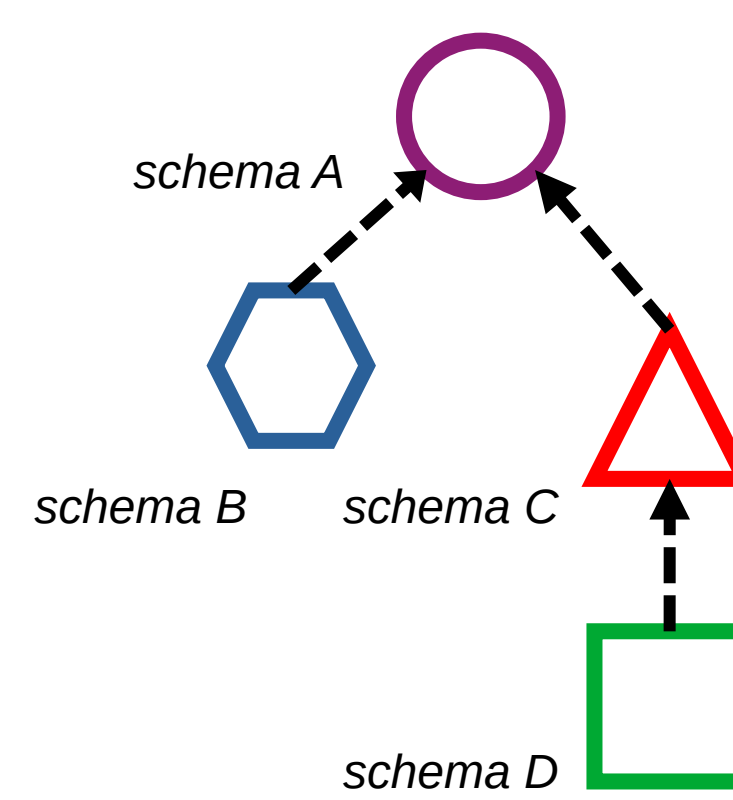
(Python ≥ 3.8)

(entrypoint - based)

(lightweight, domain-agnostic, extensible)

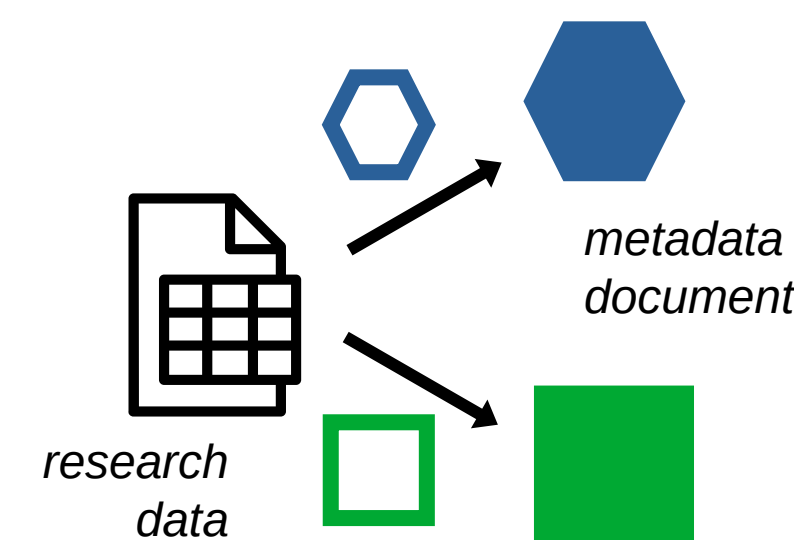
Metadata Schemas

- based on **pydantic**
- ensure strong validation
- support field inheritance
- JSON Schema export



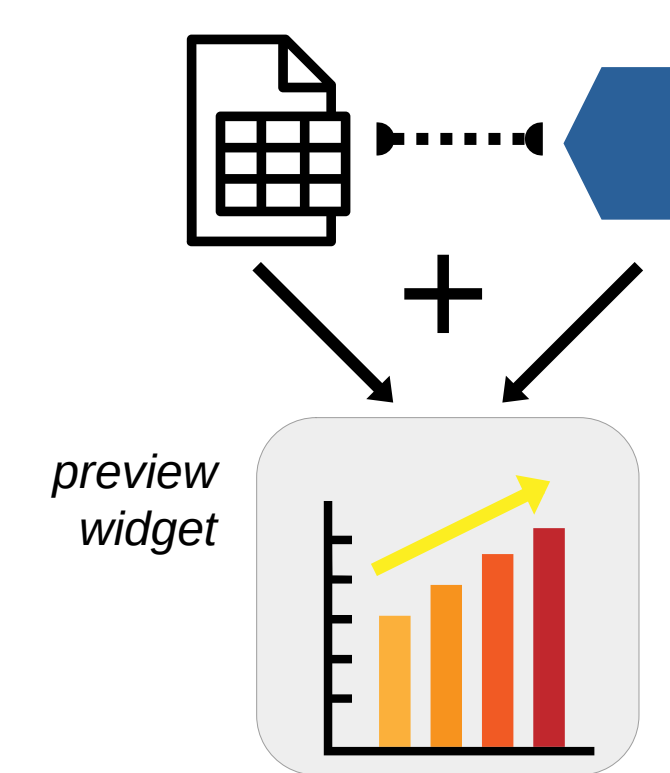
Metadata Harvesters

- can be pipelined
- extract metadata into schema-compliant objects
- building block to support packaging automation



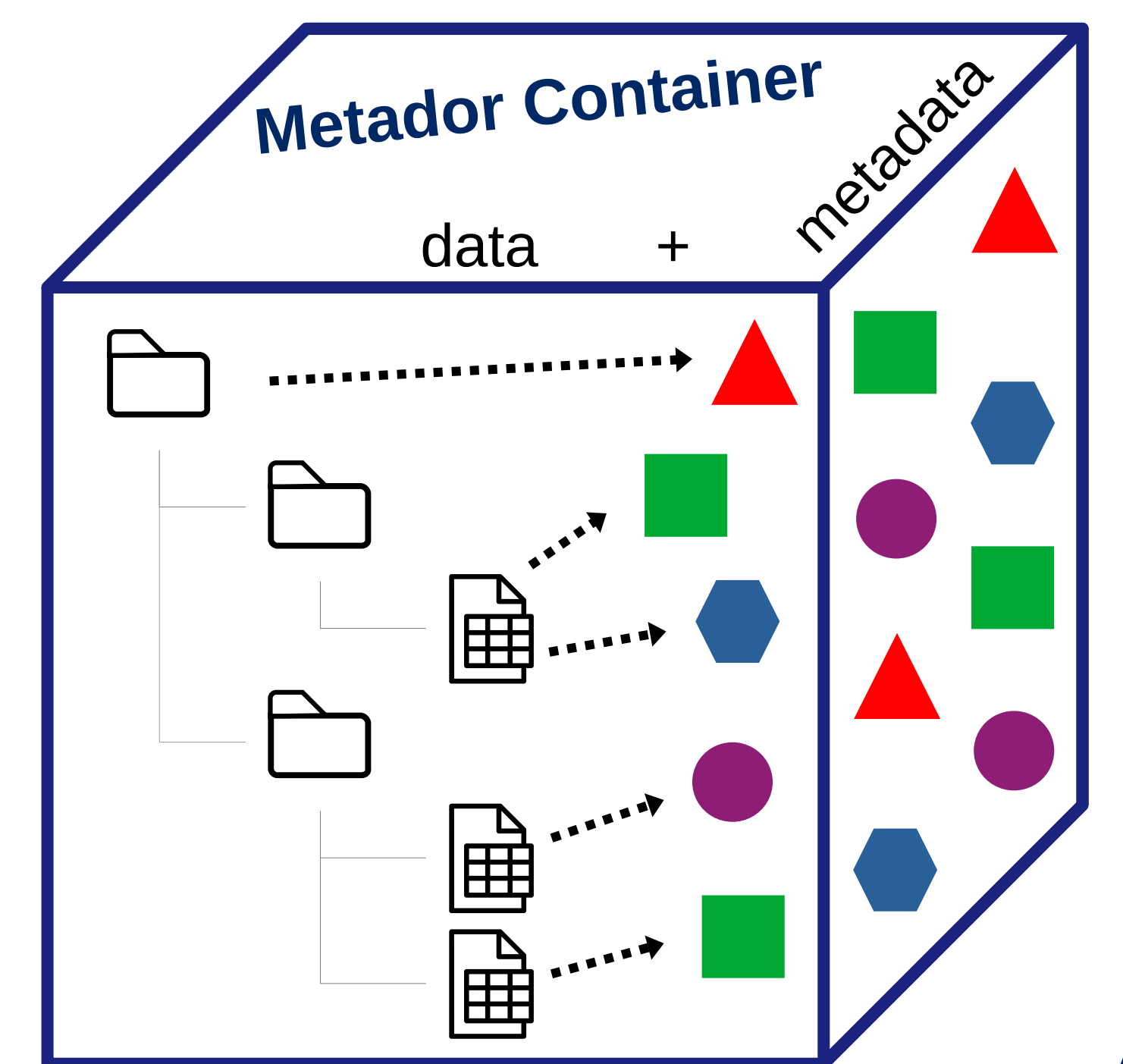
(Interactive) Widgets

- based on **bokeh** + **panel**
- support Jupyter Notebooks
- embeddable in a website
- used for (semi-)automatic container dashboard



Other Domain-Agnostic Tools

- technical:** e.g. **flask** blueprint supporting dashboard integration in websites
- user-oriented:** tools to help with metadata(-schema) preparation, creation and management of Metador-enabled research data containers



Sketch of Planned Solution for the Pilot Use-Case

