

## Automatic Metadata Registration using regimo

Enhancing data interpretation and interfacing in energy systems analysis are key concepts to make research in the energy domain more FAIR and, thus, more efficient.

Data analysts are faced with heterogeneity of data and an incompatibility of definitions. Referring to ontologies is not only a remedy to those problems, but it is also a precondition for model coupling and semantic analysis of data.

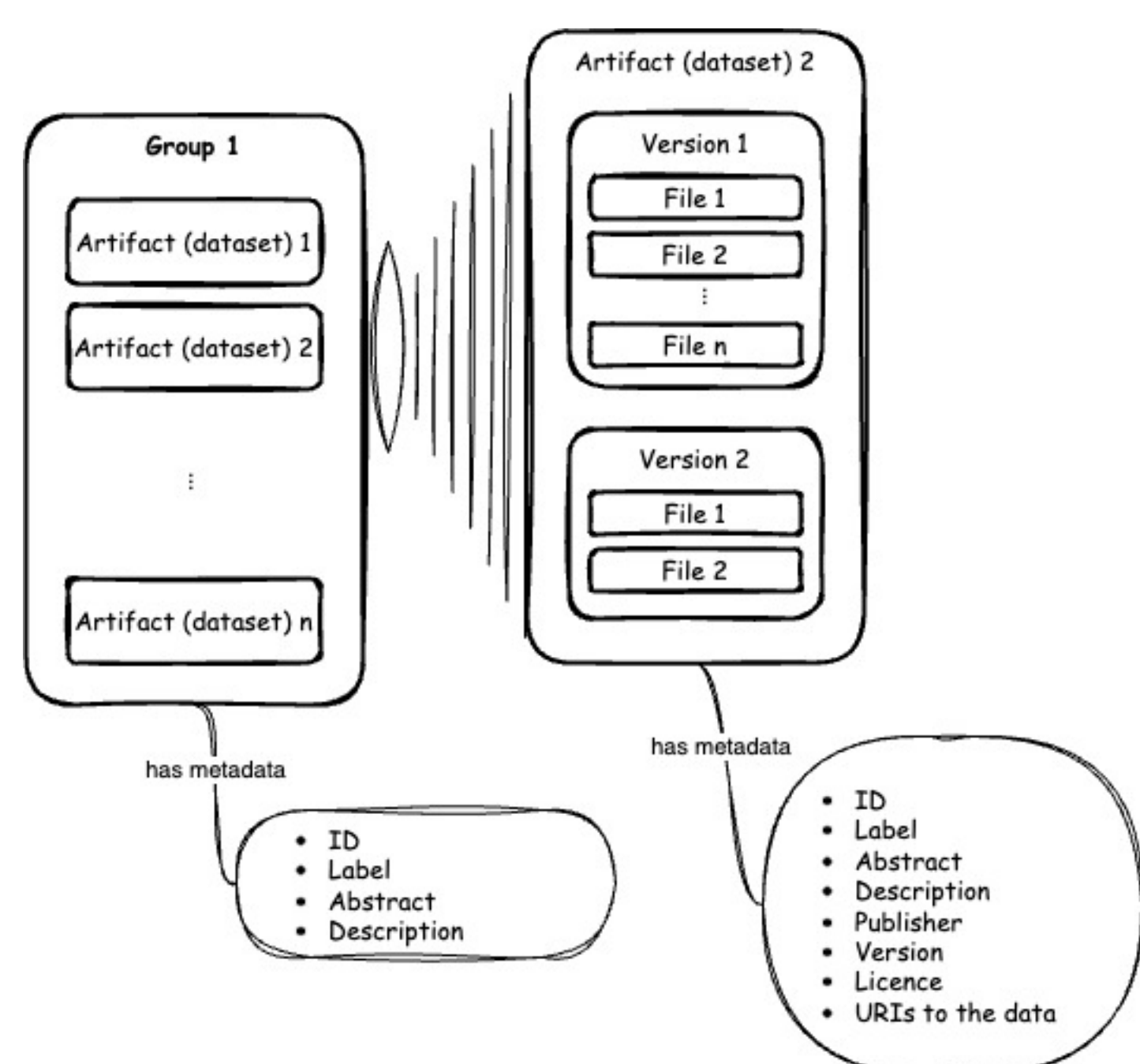
When producing data, their enrichment with metadata, describing them in a standardized way, is a challenge for every researcher. The publication of these metadata on a distributed data infrastructure (e.g. the databus within the open energy family) enhances their findability by other fellows, without alternating the ownership.

A tool to automatize the publication of metadata on a databus, i.e. registration of data, is proposed here.

### DBpedia Databus Service

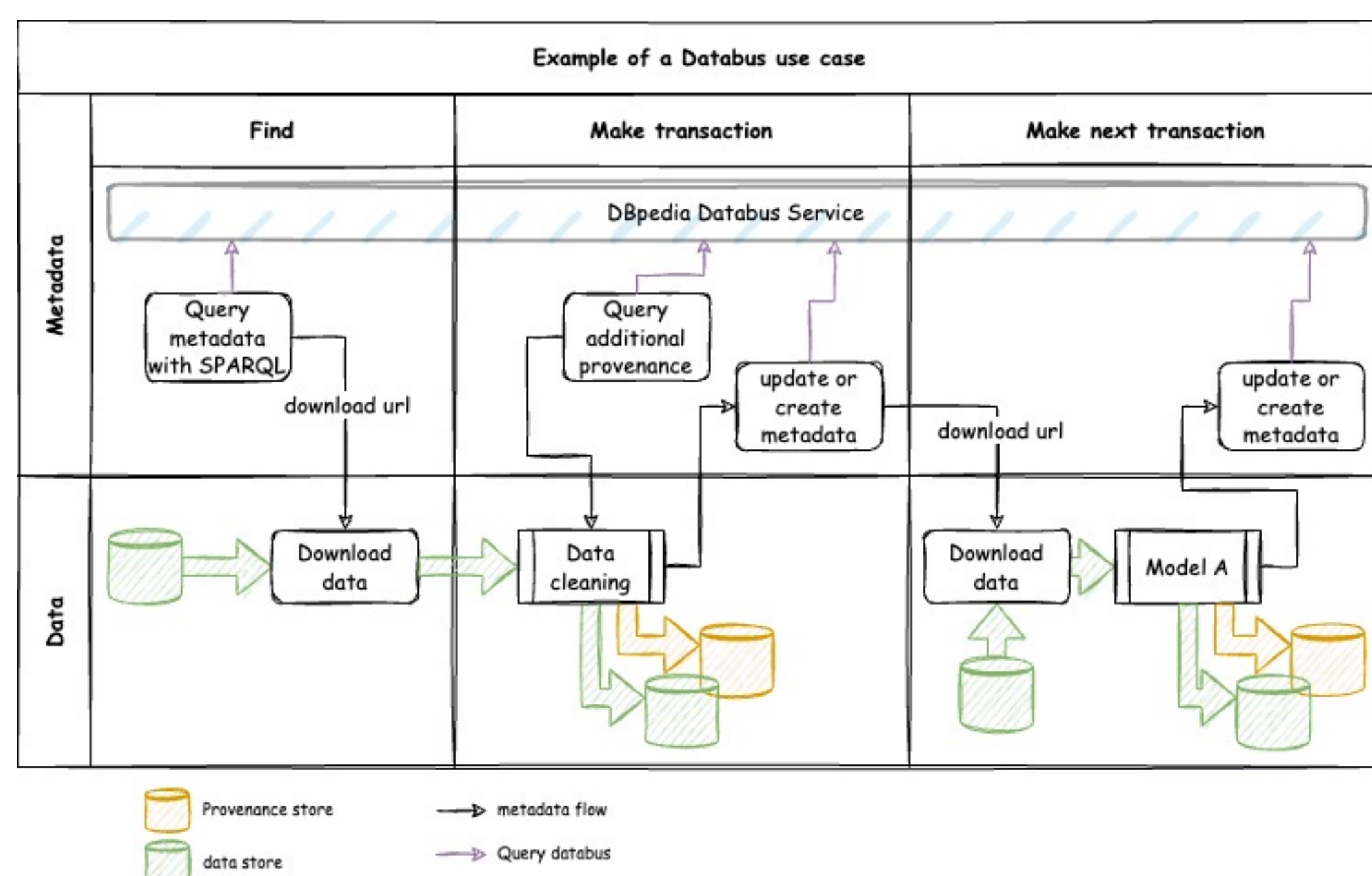
The open energy platform maintains an instance of a DBpedia databus.

The data model on the databus is as represented here:



- Metadata attributes list is extensible when needed
- Attributes can be entered by using a web-wizard or a dedicated API

In the following we tried to depict a typical use case of the databus.



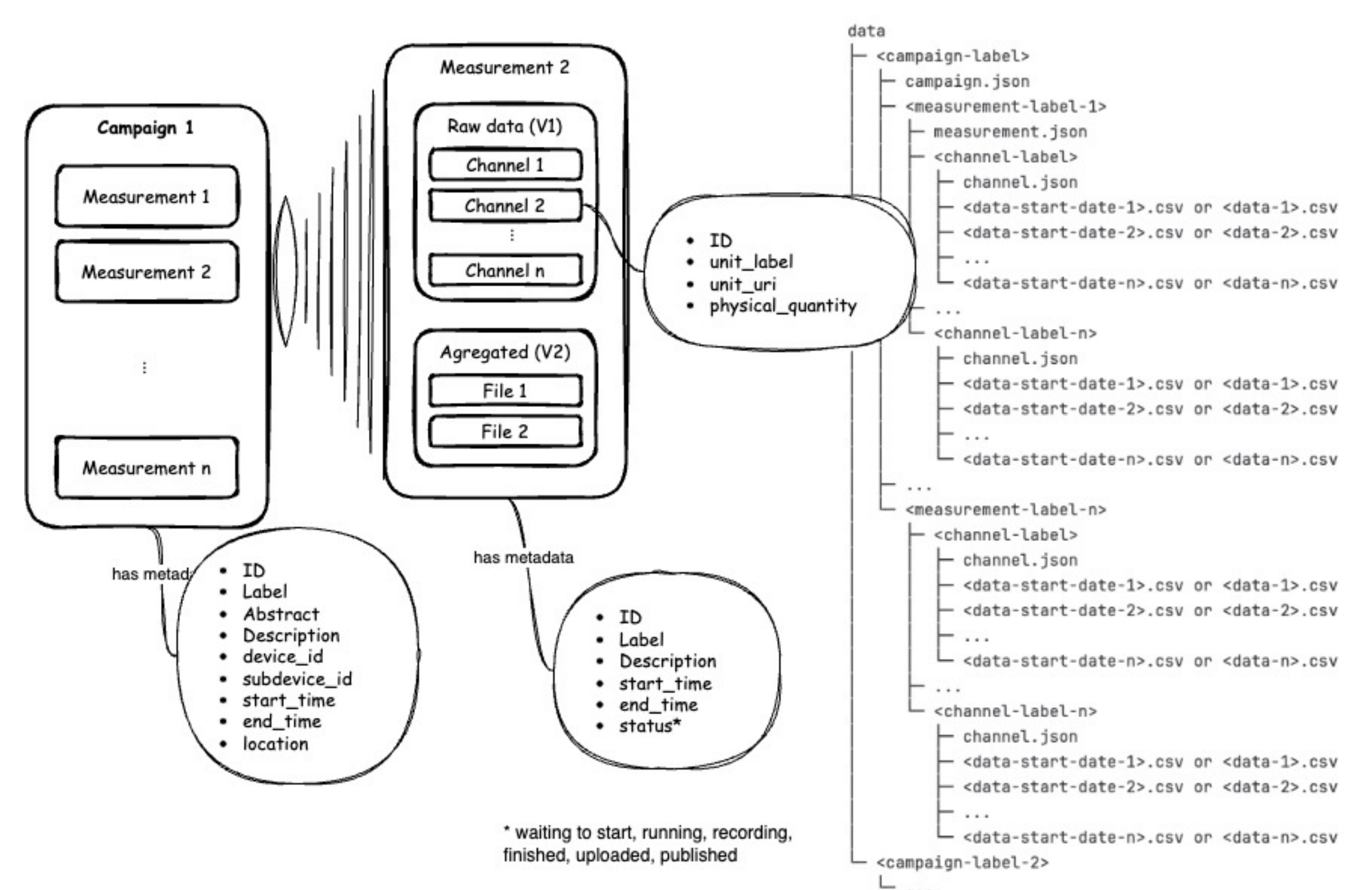
- Provenance is represented in the W3C provenance standard, so that one can use ProvStore to manage and visualize them according to different points of view (flow of information, conducted processes or assigned responsibilities)
- As member of the Open Energy Family, the Databus service accepts terminologies from the Open Energy Ontology to be used, in order to assign well defined concepts to Artifacts.

### The Databus offers

- a service to manage and search registered metadata
- persistent identifiers for tracing data processing and citing data (PID's as pointers to digital objects)
- in conjunction with the Open Energy Ontology, semantic searches for data in the domain of energy systems analysis
- an improvement of open data exchange, model coupling, tracing of workflows and collaboration

### Generated data

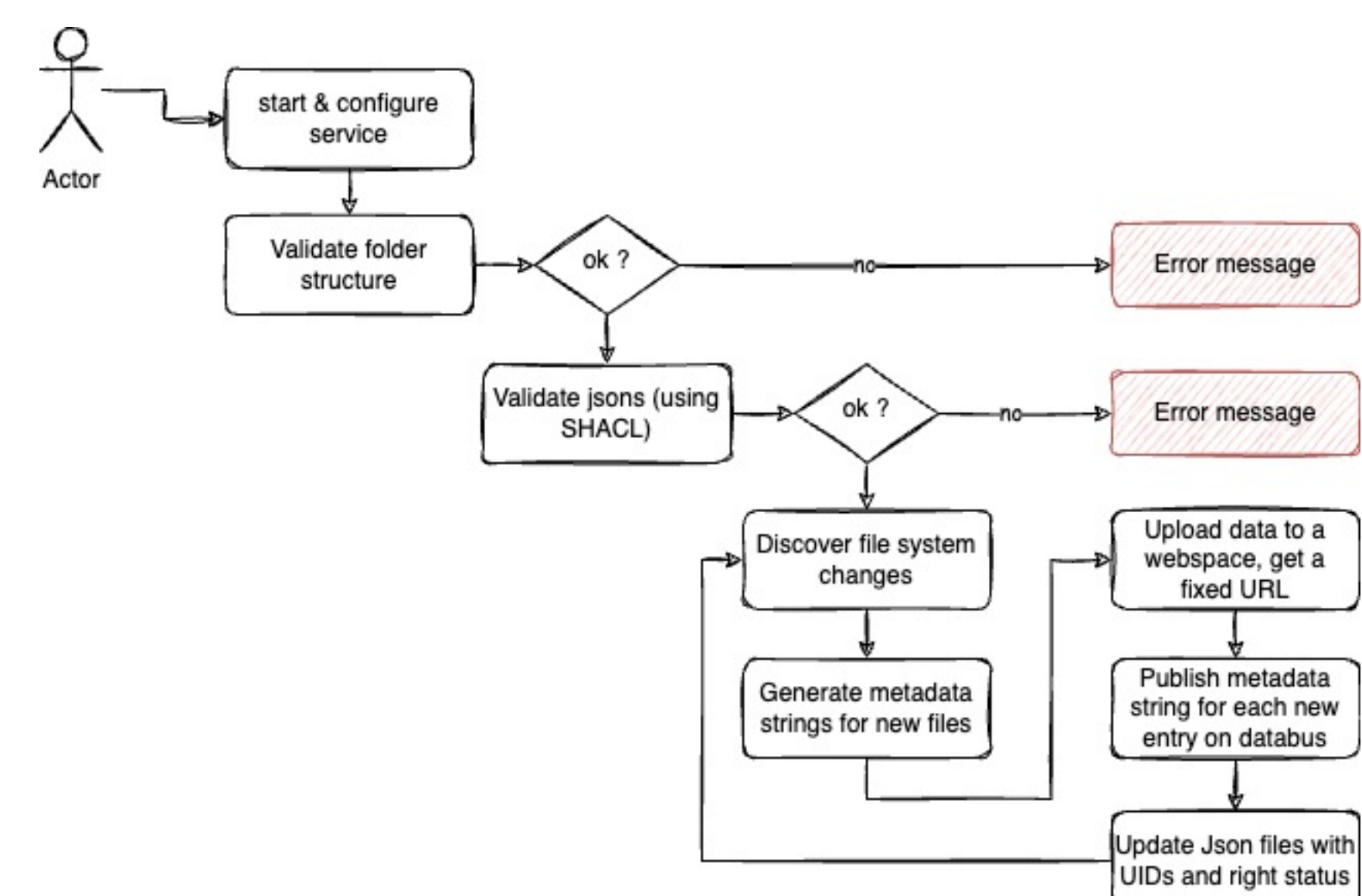
The generated data are gathered during campaigns and are typically organized in a folder structure similar to the following.



- The json files contain the actual metadata to the campaign, measurements and channels.
- The folder structure can be established automatically in a planning step.

### Workflow of regimo

Regimo is a service which can be installed on the measurement device. It observes the configured folder and registers metadata automatically on the Open Energy Databus as soon as a measurement is done.



In the future, the planning phase of the campaigns shall facilitate the application of further automation steps throughout the entire process. The use of Sensor Management Systems and Enterprise Resource Planning tools, as well as the unique identification of devices to capture their calibration states during the measurement, is an ongoing work.

### References

[1] C. Hoyer-Klick et al. Demonstration and best practices, July 2023.  
[2] J. Frey et al. Distributed data infrastructure. Technical report, DLR, July 2023.  
[3] M. Booshehri et al. Introducing the open energy ontology: Enhancing data interpretation and interfacing in energy systems analysis. Energy and AI, 5:100074, 2021.