

(Meta) Data Standards for agricultural research data management and approaches towards evaluation

An overview

17.06.2024, FAIRagro Community Summit

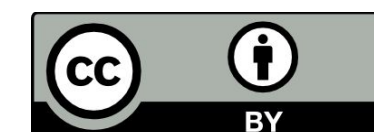
Jascha Jung (Kuratorium für Technik und Bauwesen in der Landwirtschaft, KTBL)

Nils Reinosch (Kuratorium für Technik und Bauwesen in der Landwirtschaft, KTBL)

Daniel Martini (Kuratorium für Technik und Bauwesen in der Landwirtschaft, KTBL)

Gabriel Schneider (ZB MED - Informationszentrum Lebenswissenschaften)

on behalf of the FAIRagro consortium



All contents of this presentation are, unless otherwise stated, licensed under the [Creative Commons BY 4.0 International](https://creativecommons.org/licenses/by/4.0/)

Metadata in FAIRagro



- Mainly organized in Measures 3.1 (Standards for Digital Resources) and 3.2 (Standards for Data Management, FAIRness and Discoverability)
- Focus on FAIRification, Reusability and Reuse



Inventory of Standards

Compile an inventory of all meaningful data standards for representing (meta)data in agrosystems research

- Generic Ontologies and Vocabularies:
 - SSN/SOSA
 - PROV
 - ODRL
 - DQV
 - Domain Specific Ontologies and Vocabularies:
 - Crop Ontology
 - AGROVOC
 - Geospatial Data and Metadata:
 - INSPIRE
 - ISO19115
 - GML
 - Plant Phenotyping Data and Metadata
 - MIAPPE
 - ISA-Tab
 - Modeling Data:
 - ICASA
 - AgMIP
 - Protocol Standards:
 - OGC WMS/WFS
 - BrAPI
 - OAI-PMH
 - Informal Terminologies/Codesystems:
 - EPPO
 - Pesticide Registration Database
 - ...
- different domain specificity
- different interoperability level: syntax, semantics, protocol standards
- different level of formalization: simple term list vs. full-fledged ontologies

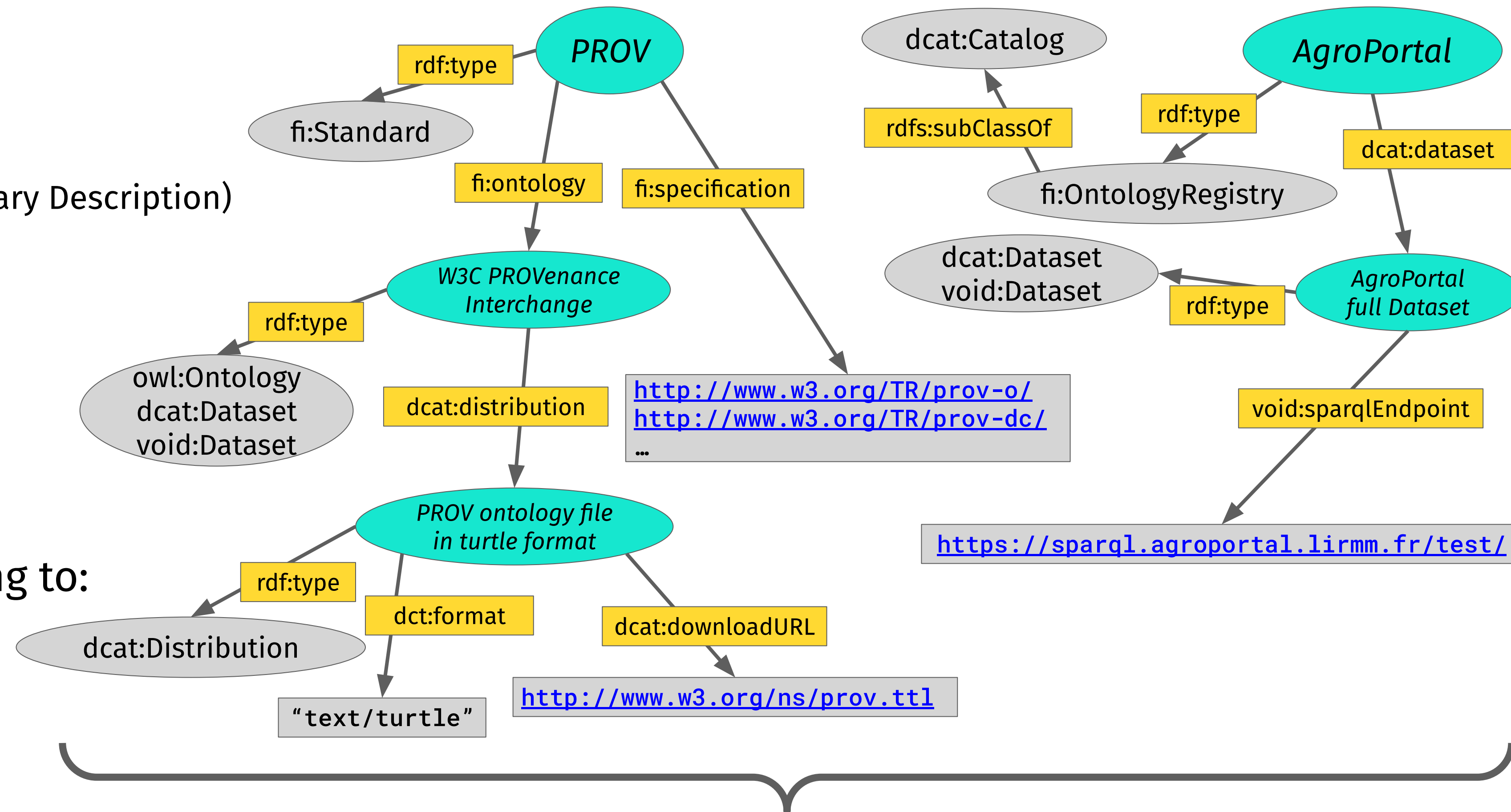
Standards for Describing Standards

The inventory is itself based on standards:

- DCAT (Data Catalog Vocabulary)
<https://www.w3.org/TR/vocab-dcat-3/>
- VOID (Vocabulary of Interlinked Datasets)
<https://www.w3.org/TR/void/>
- VANN (Vocabulary for Annotating Vocabulary Description)
<https://vocab.org/vann/>
- OWL (Web Ontology Language)
<https://www.w3.org/TR/owl2-overview/>
- ...
- + a minimal “glue” vocabulary

Minimum Information Linked Data:
A Knowledge Graph of links pointing to:

- Specifications
- Tutorials
- Ontology namespaces
- Ontology Serializations
- Terminological Datasets
- Repositories
- Web Services (like SPARQL)



User specific views by query and metadata retrieval, e.g.:

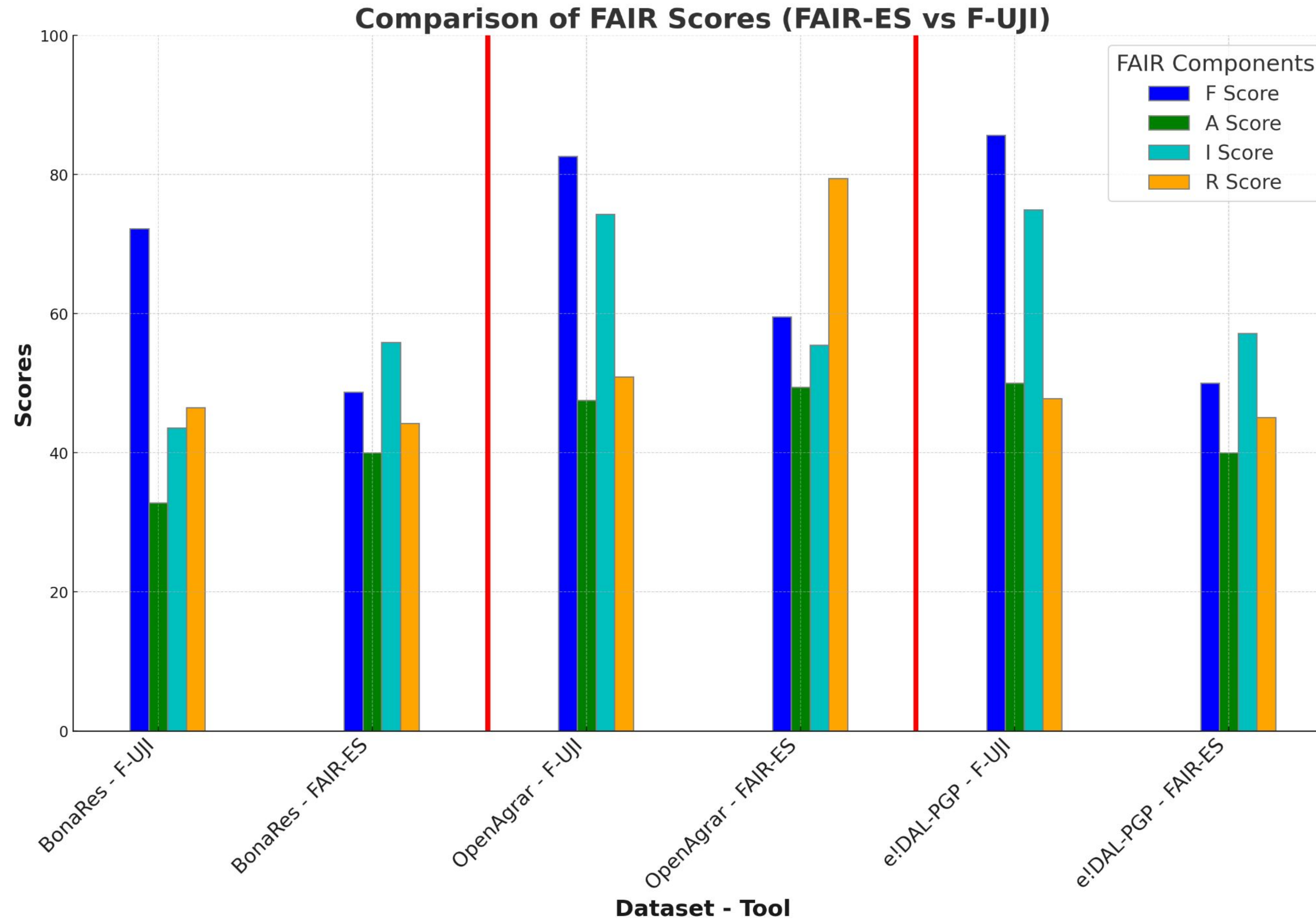
- List of all standards, that are distributed as formal ontologies already
- List of all terms available in a certain vocabulary
- ...

FAIR assessment



- FAIR: **F**indable, **A**ccessible, **I**nteroperable, **R**eusable
- Motivation:
 - Evaluate and improve data **FAIR**ness
 - Assess and automate assessment services
 - Identify important metrics
 - Establish standards and guidelines
- Assessment Tools:
 - [FAIR Evaluation Service](#): 22 metrics
 - [F-UJI](#): [16 out of 17 core FAIR object assessment metrics](#)

FAIR assessment



The visualization illustrates the FAIR assessment by two services. The datasets are exemplary and do not necessarily represent the entirety of each dataset.

Extension of Schema.org/Bioschemas for the agrosystem domain

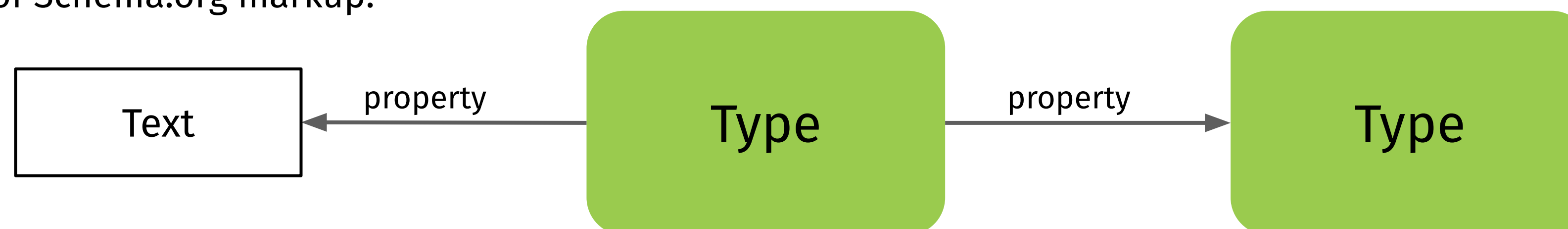
- Schema.org is a lightweight vocabulary for describing resources on the web → research datasets
- Consumed by search engines → increased Findability
- Due to its generic approach, Schema.org is not fully sufficient for the requirements of our Use-Cases → use of extension mechanism

○ Joining Bioschemas



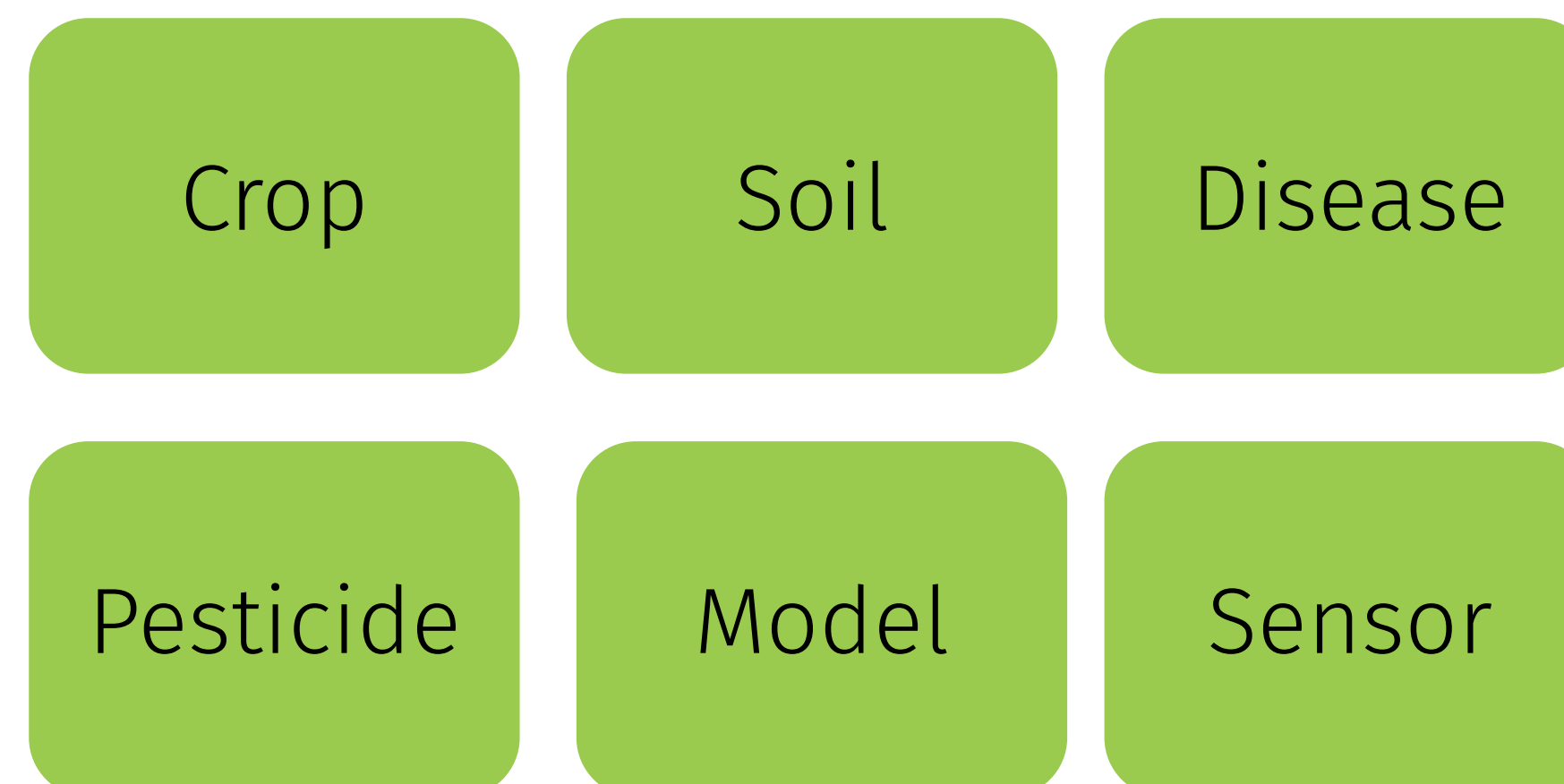
Bioschemas Logo. Available under the [Creative Commons Attribution-ShareAlike License](https://creativecommons.org/licenses/by-sa/4.0/). <https://bioschemas.org/about/logos>.

Basic structure of Schema.org markup:



Extension of Schema.org/Bioschemas for the agrosystem domain

- Collection of use case requirements on finding datasets to identify central entities

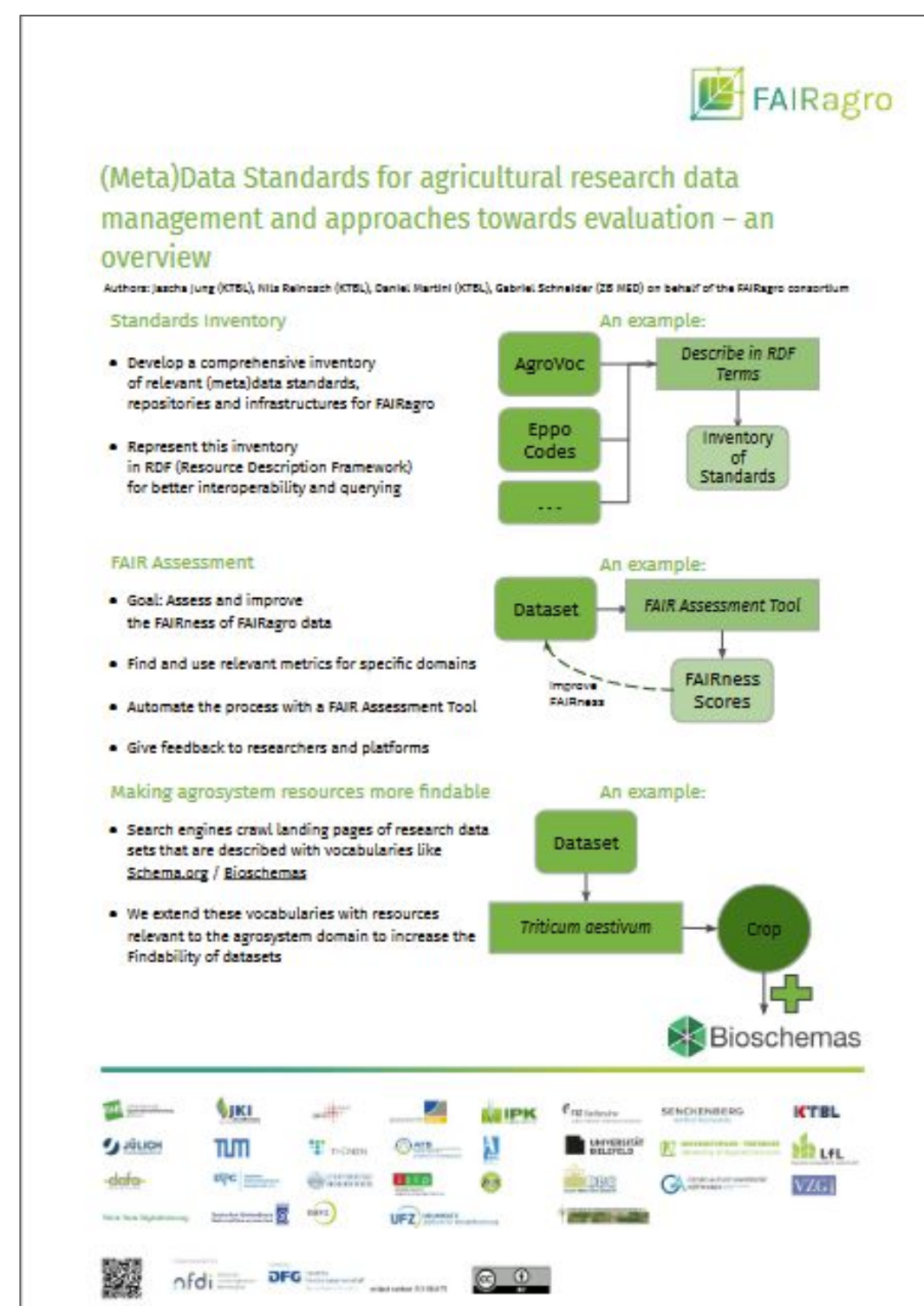


- Modeling of example datasets with existing Schema.org/Bioschemas resources to identify gaps
- Organization in a Bioschemas Working Group → Join us if you are interested

Meet us at the poster



- If you have questions, want to learn more and discuss talk to us at our poster



Contacts

Jascha Jung, KTBL: J.Jung@ktbl.de

Nils Reinosch, KTBL: N.Reinosch@ktbl.de

Daniel Martini, KTBL: D.Martini@ktbl.de

Gabriel Schneider, ZB MED: schneiderg@zbmed.de

Thank you!