Competency Questions in Practice: Metadata Schema for Terahertz Research

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

Yaser Alashloo*, Sophia Leimer, Sonja Hendriks, & Stephanie Rehwald

NFDI4Ing Conference 2022

October 26, 2022

Outline

- **Introduction**
  - SFB MARIE
  - FAIR data
- **Methodology**
  - Competency questions
- **Results**
- **Conclusion**

SFB MARIE

- Involving almost 100 people namely professors, post-docs and PhD candidates
- Collaboration between 5 universities and 2 institutes

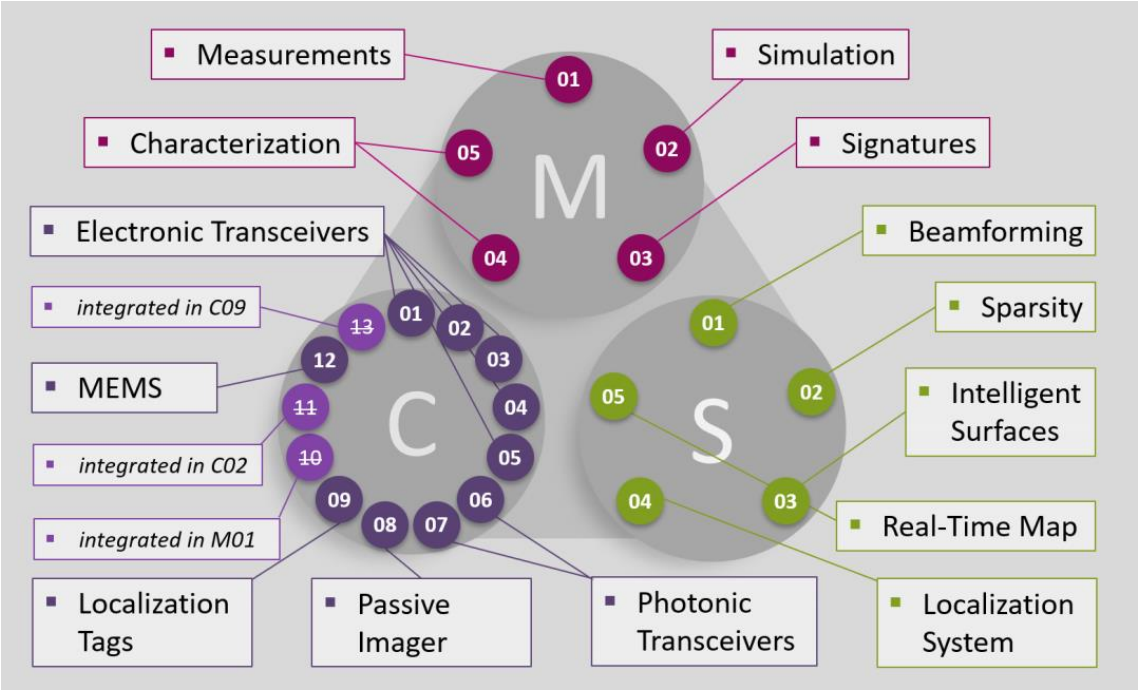
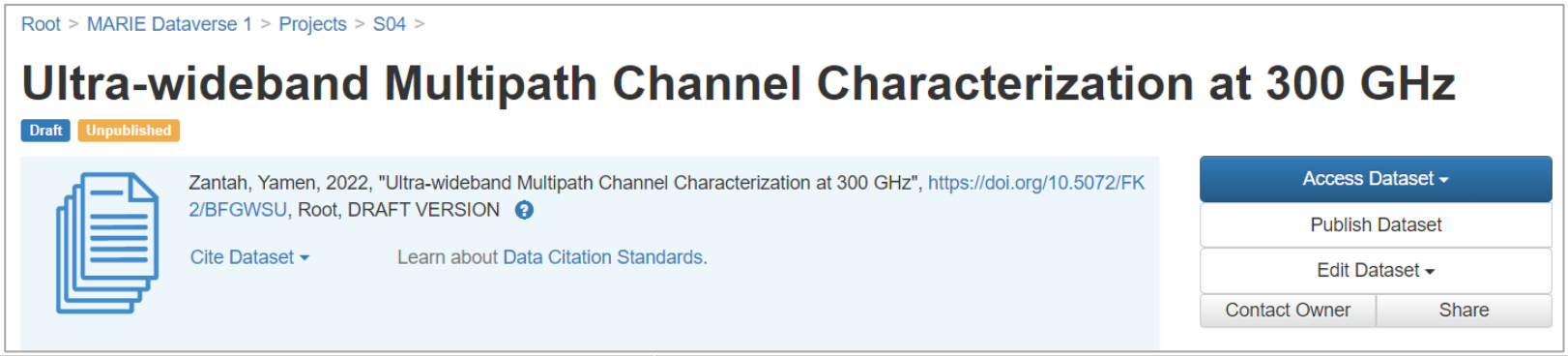


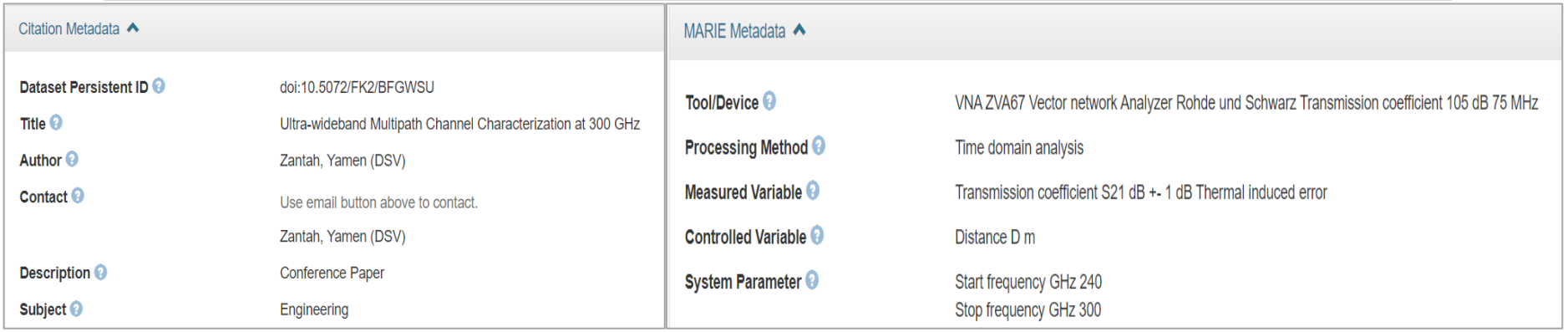
Table 2: MARIE Data types

#	Data type	Number of subprojects for which the respective data type is generated
1	Software/algorithms	15
2	Simulation data	14
3	Measurement data	13
4	Design data/circuit diagrams	10

Metadata Schema

- Dataverse features
  - Support for FAIR Data Principles
  - Versioning
  - APIs for interoperability
  - Preview and analysis of tabular files





NFDI4Ing

- Aims to develop, disseminate, standardize and provide methods and services to make engineering research data FAIR.

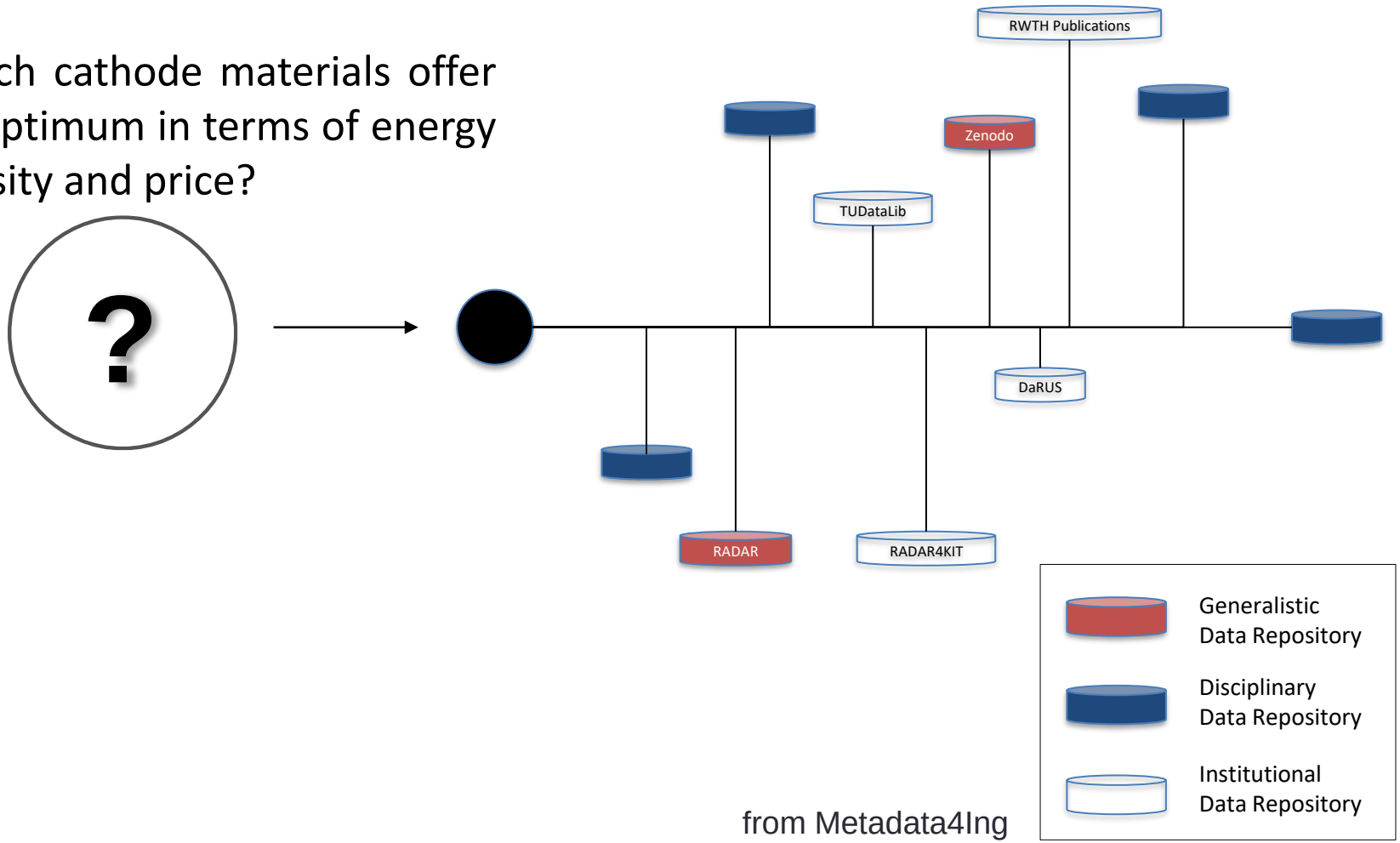
- ✓ **Metadata4Ing (in SIG metadata & ontologies)**
  - develops a generic ontological metadata model for the engineering community.
  - uses a modular and hierarchical modelling approach that offers a high level of specificity and flexibility while maximizing reusability and interoperability.

- ✓ **CRC MARIE:**
  - Participant of Metadata4Ing
  - Bring models to practice

Metadata Schema

- Vision for the future

✓ Which cathode materials offer an optimum in terms of energy density and price?

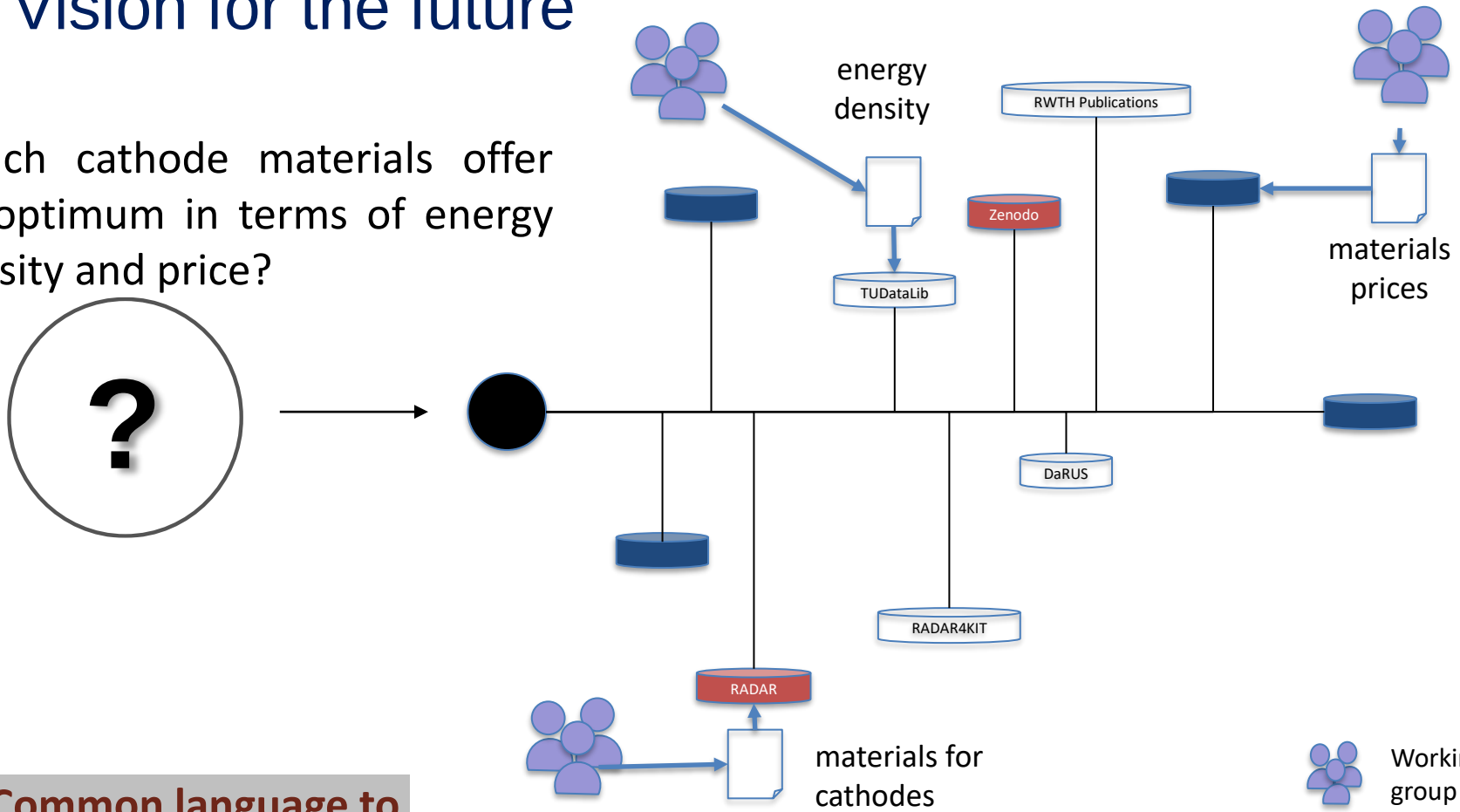


from Metadata4Ing

Metadata Schema

- Vision for the future

Which cathode materials offer an optimum in terms of energy density and price?



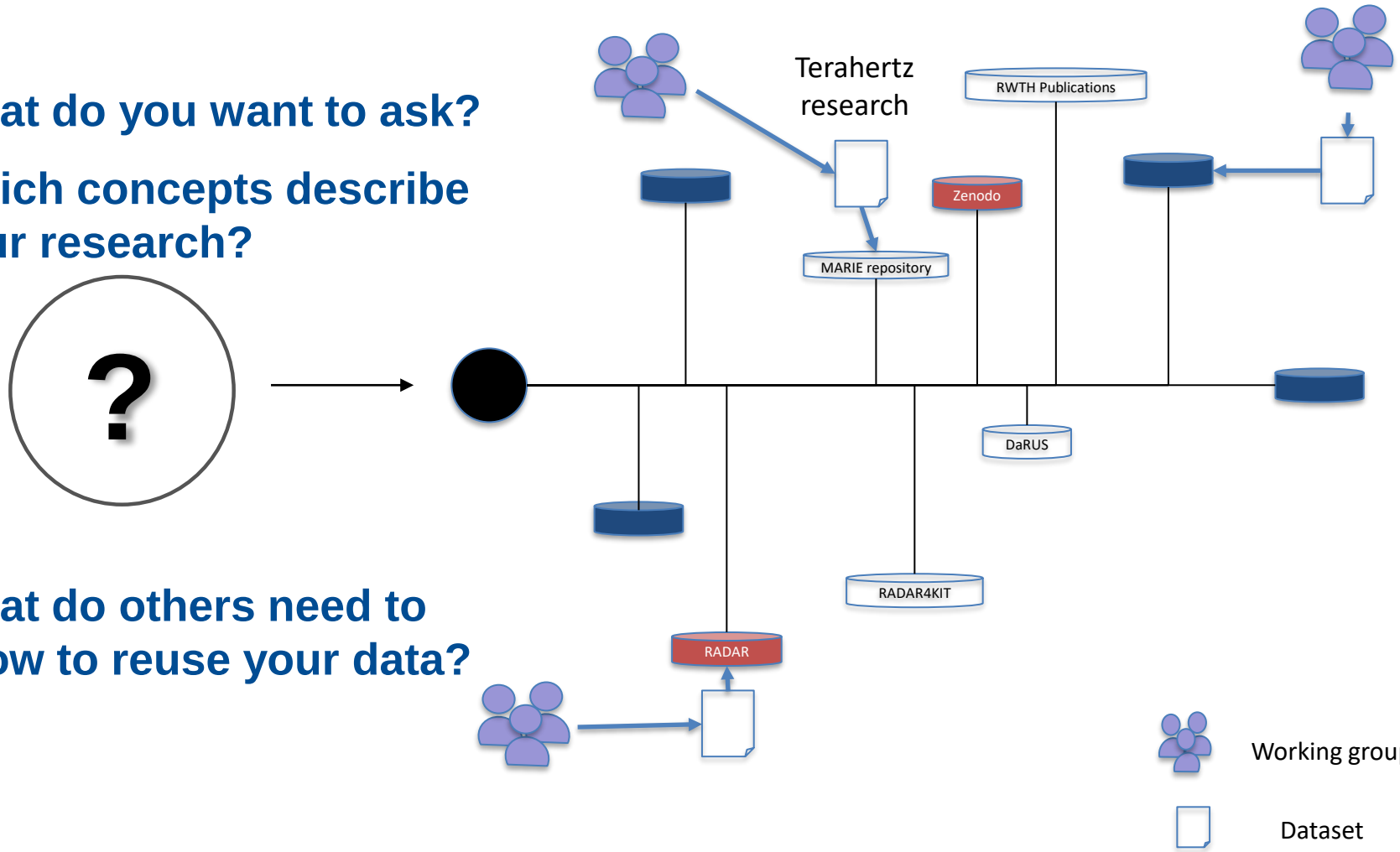
✓ **Common language to describe engineering data!**

from Metadata4Ing


Methodology:
Competency Questions

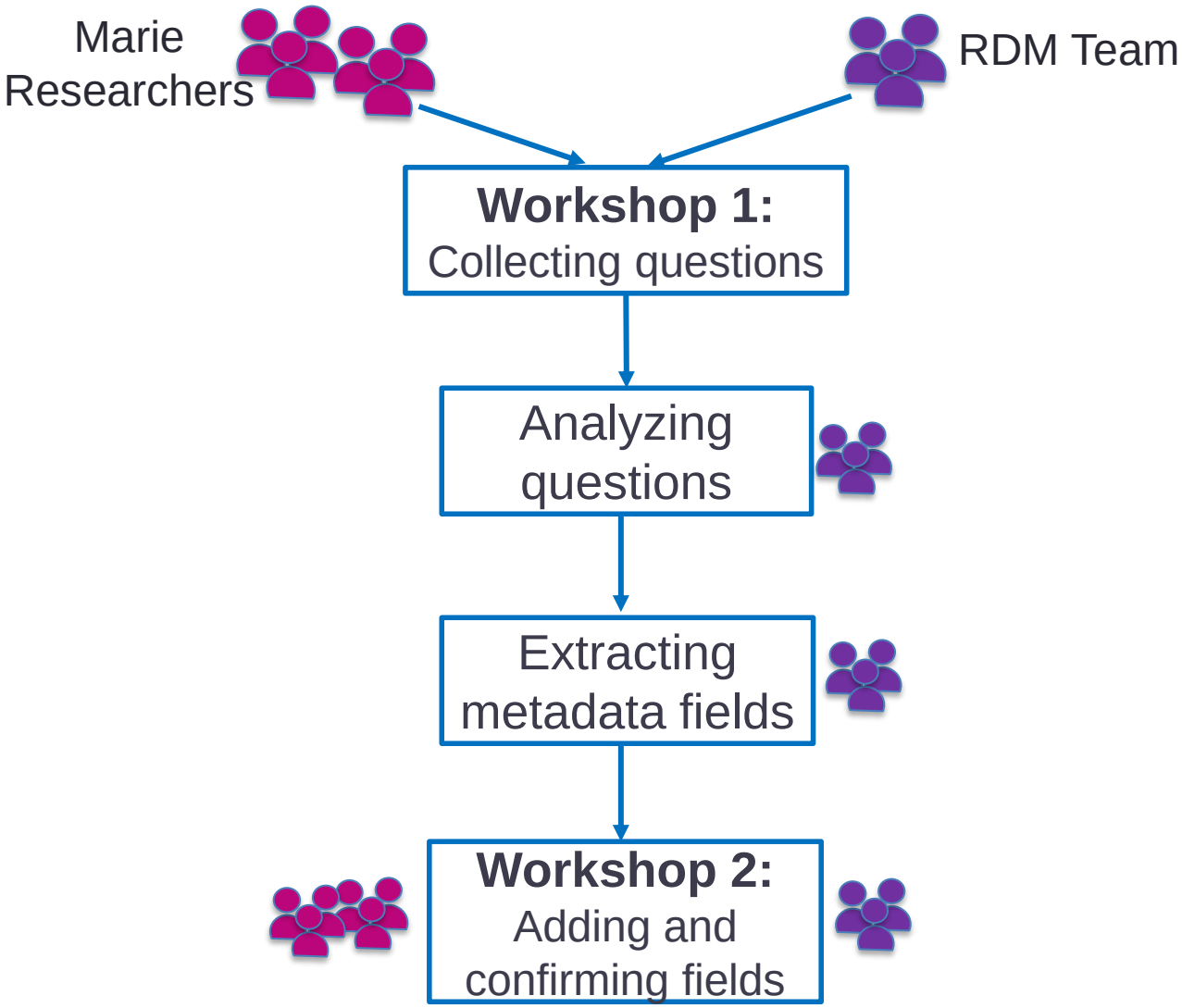
Concept of Competency Questions:

- ✓ What do you want to ask?
- ✓ Which concepts describe your research?
- ✓ What do others need to know to reuse your data?



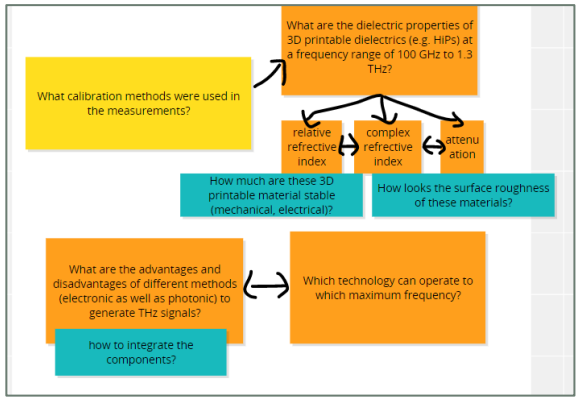
from Metadata4Ing

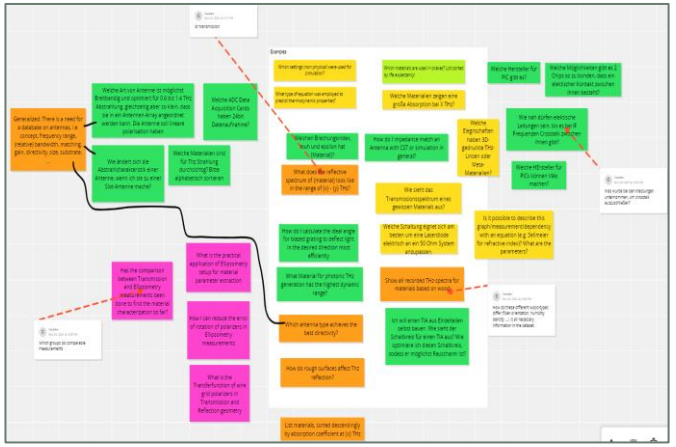
Competency questions workflow:



Workshop 1: Collecting questions

- Concept was presented
- An online whiteboard utilized
- 4 breakout sessions
- 100 questions were collected





Analyzing questions:

- Questions were sorted:
  - ✓ 84 acceptable, 16 out of scope

Categories for sorting	Values
Complexity	simple/complicated
Administrative	administrative
Out of scope	out of scope
Type of data	measurement/simulation/software/design/multiple
Sub-classes	method, instrument, software, variable, material
Instruments/tools	type, name, description, version
Material	type, name, description (e.g. composition), e.g. Optical material properties
Method	Type (generation, processing, analysis, others), name, description
Software	name, version, description, open-source, programming language, Operating system?
Variable	name, constant (value) if not: measured minimum value, measured maximum value

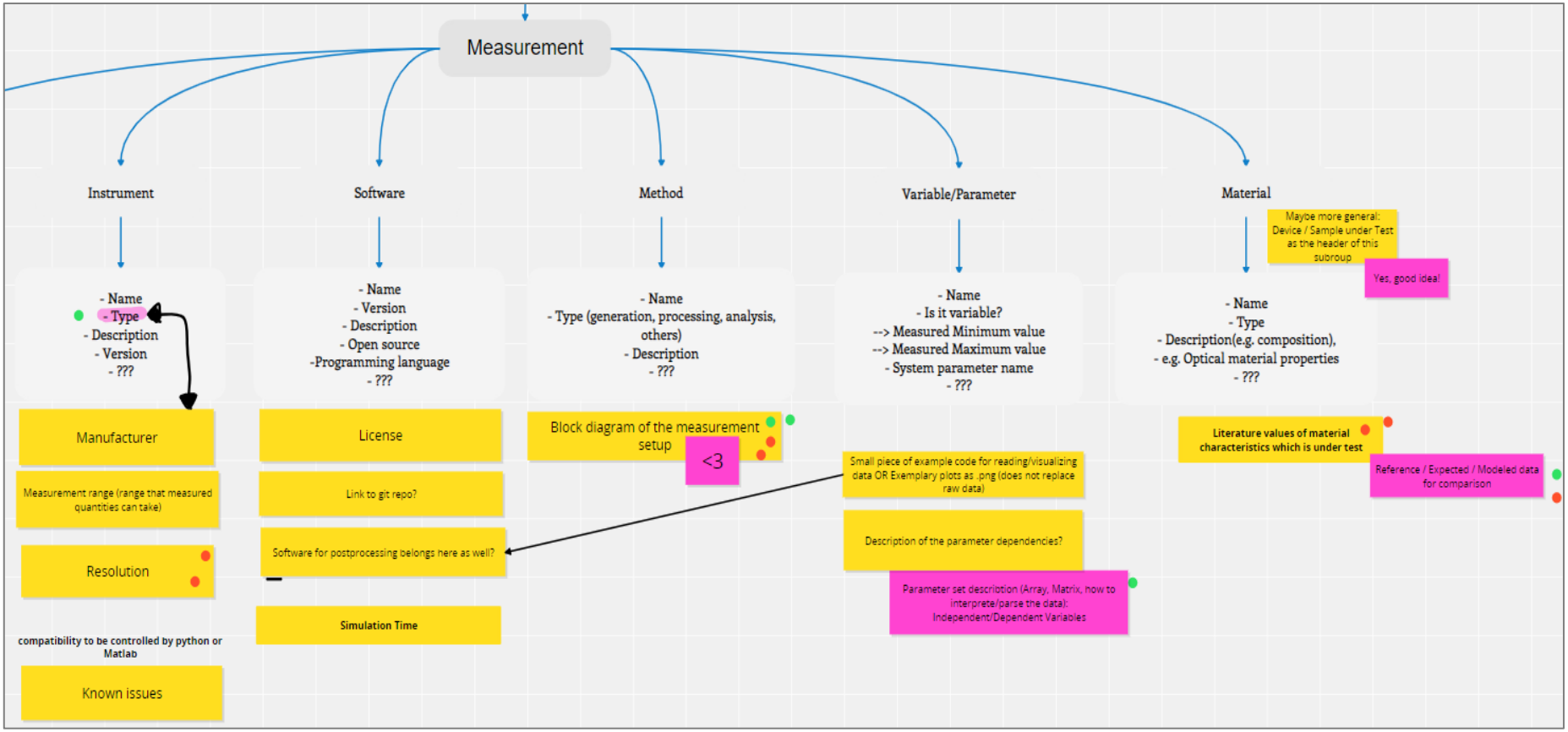
Analyzing questions:

- Questions and sorting: examples

	Questions	Complexity	administrative	Type of data	Sub-classes
1.	What **algorithm** has been used for post processing, what were the **parameters**?	simple		**software**	Software, parameters
2.	How does this measurement compare to [insert golden standard here]?	complicated		**measurement**	method
3.	Give me all details about the **sample** (composition, geometry, reference measurements, **other groups measurement**, photos)	simple	administrative	**multiple**	material, parameter
4.	Circuit data: Does the simulation results include layout effects (using EM simulation, parasitic extraction, ...)	simple		**simulation**	method, tool
5.	Is there a simulation for this experiment? Is there an experiment for this simulation?	simple		**multiple**	method
6.	How easy are the measurements to reproduce / replicate i.e., how difficult were they to perform	complicated		**measurement**	method
7.	What are the **dielectric properties** of **3D printable dielectrics** (e.g. HiPs) at a frequency range of 100 GHz to 1.3 THz? (Comment / Additions: relative refrective index / complex refrective index / attenuation.	simple		**design**	parameter
8.	Which samples show similar results? What is similar?	complicated		**Measurement**	parameter, material
9.	Which calibration method has been used?	simple		**multiple**	method

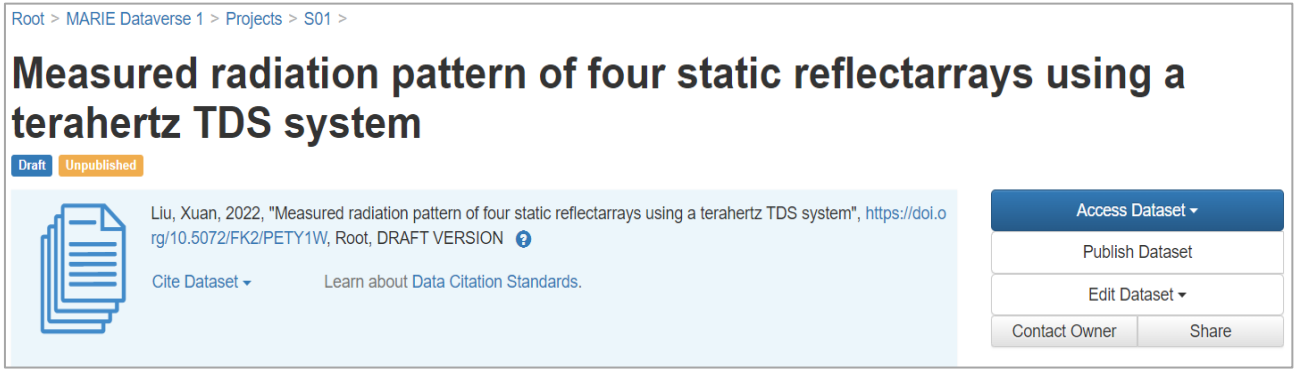
Workshop 2: Enhancing metadata

- Asking researchers to check and enhance the fields
- Breakout sessions with people in similar fields
- Receiving feedbacks and suggestions



Metadata Schema Evaluation:

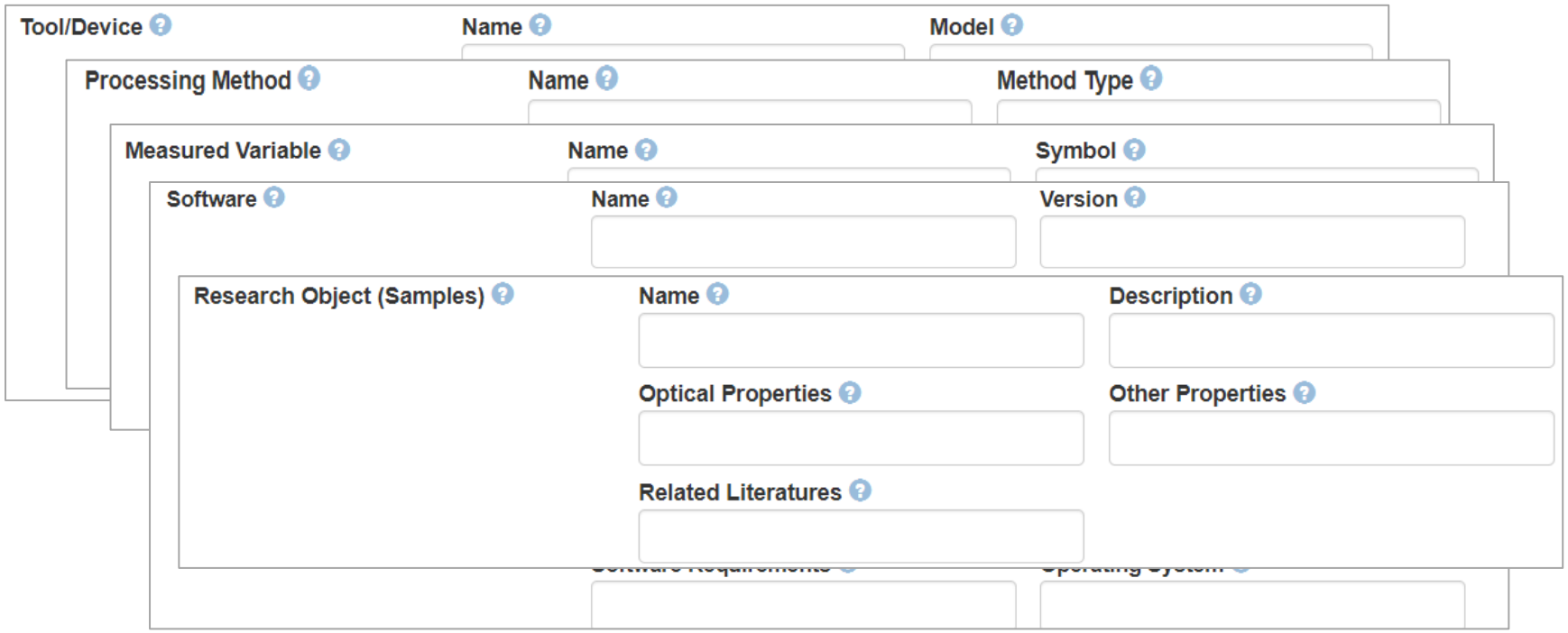
- Implemented in Dataverse



MARIE Metadata

Tool/Device	All fiber-coupled terahertz spectrometer TERA K15 MenloSystems terahertz bandwidth 5.5THz spectral resolution < 0.7GHz
Processing Method	zero padding time-domain windowing
Controlled Variable	angle φ degree 35 65
System Parameter	scan wondow of the delay line τ ps 200
Research Object (Samples)	static reflectarray

Developed Metadata Schema



Conclusion

- ✓ The collection of competency questions has proven to be a good tool to develop a metadata schema.
- ✓ The Terahertz metadata schema effectively covers the researchers needs.

❖ We thank **Metadata4Ing** for providing the slides and the workshop concept!

Thank you for your attention!