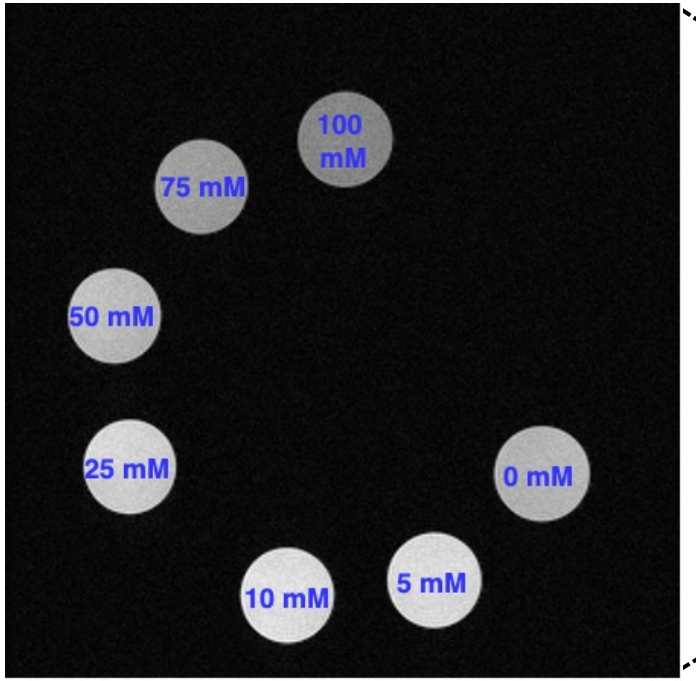
DaMaLOS 2023 @ ESWC

# FAIR Data Management Workflow for MRI Data

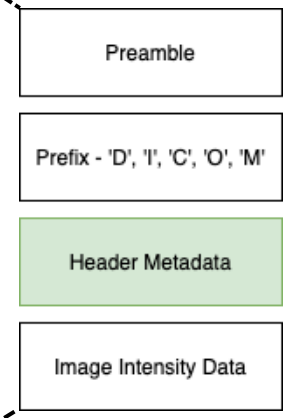
Nicolas Blumenröhr[1], Neil MacKinnon[2], Rossella Aversa[1]
[1] SCC, Karlsruhe Institute of Technology
[2] IMT, Karlsruhe Institute of Technology

# Motivation

Magnetic Resonance Imaging (MRI) data in the material sciences and medical imaging
→ different samples, same storage file format, i.e. DICOM




MRI image of $CuSO_4$ samples

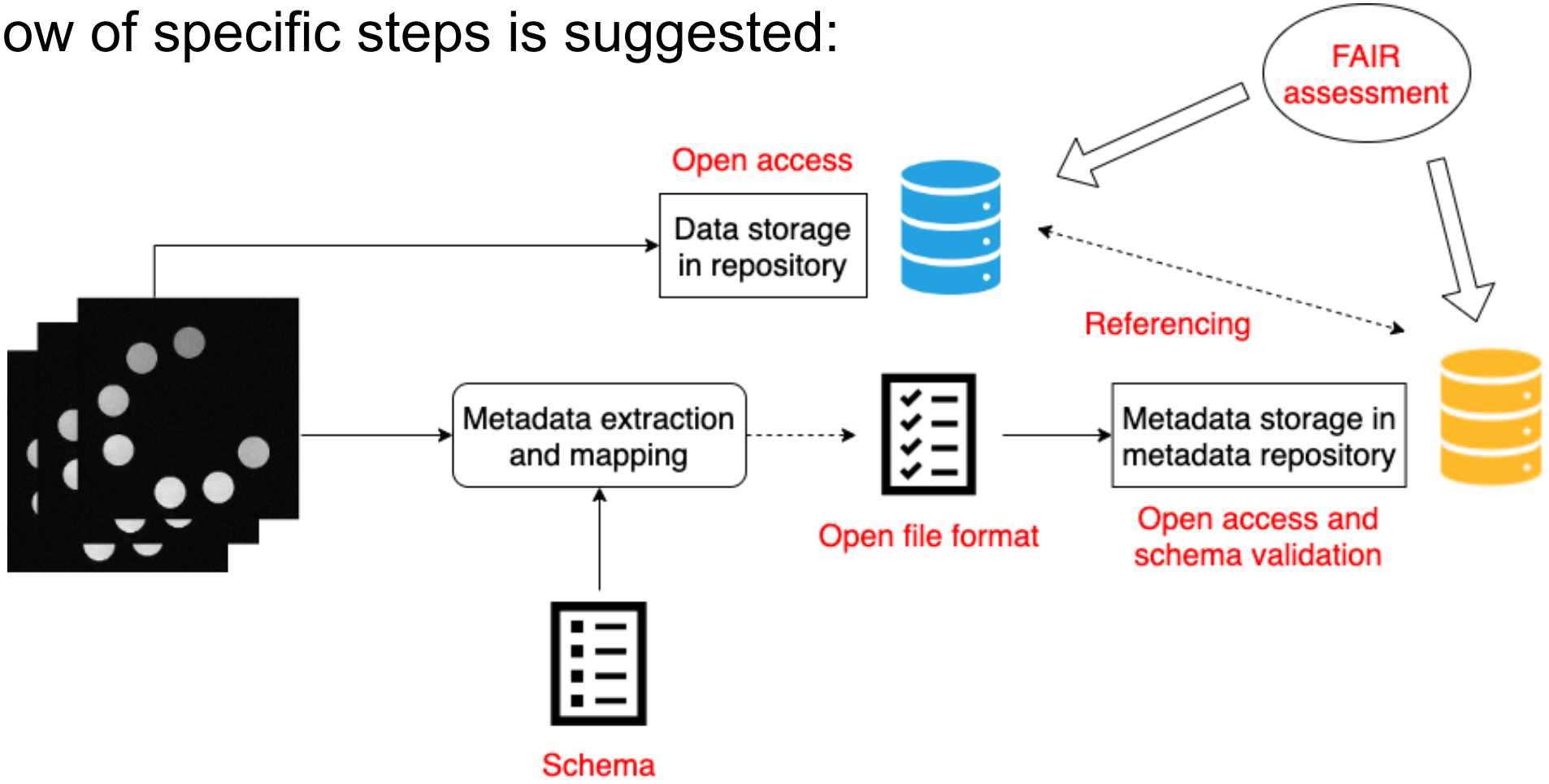



| Tag | Name | Value |
|---|---|---|
| (0008, 0020) | Study Date | 20211015 |
| (0010, 0010) | Patient's Name | Phantom |

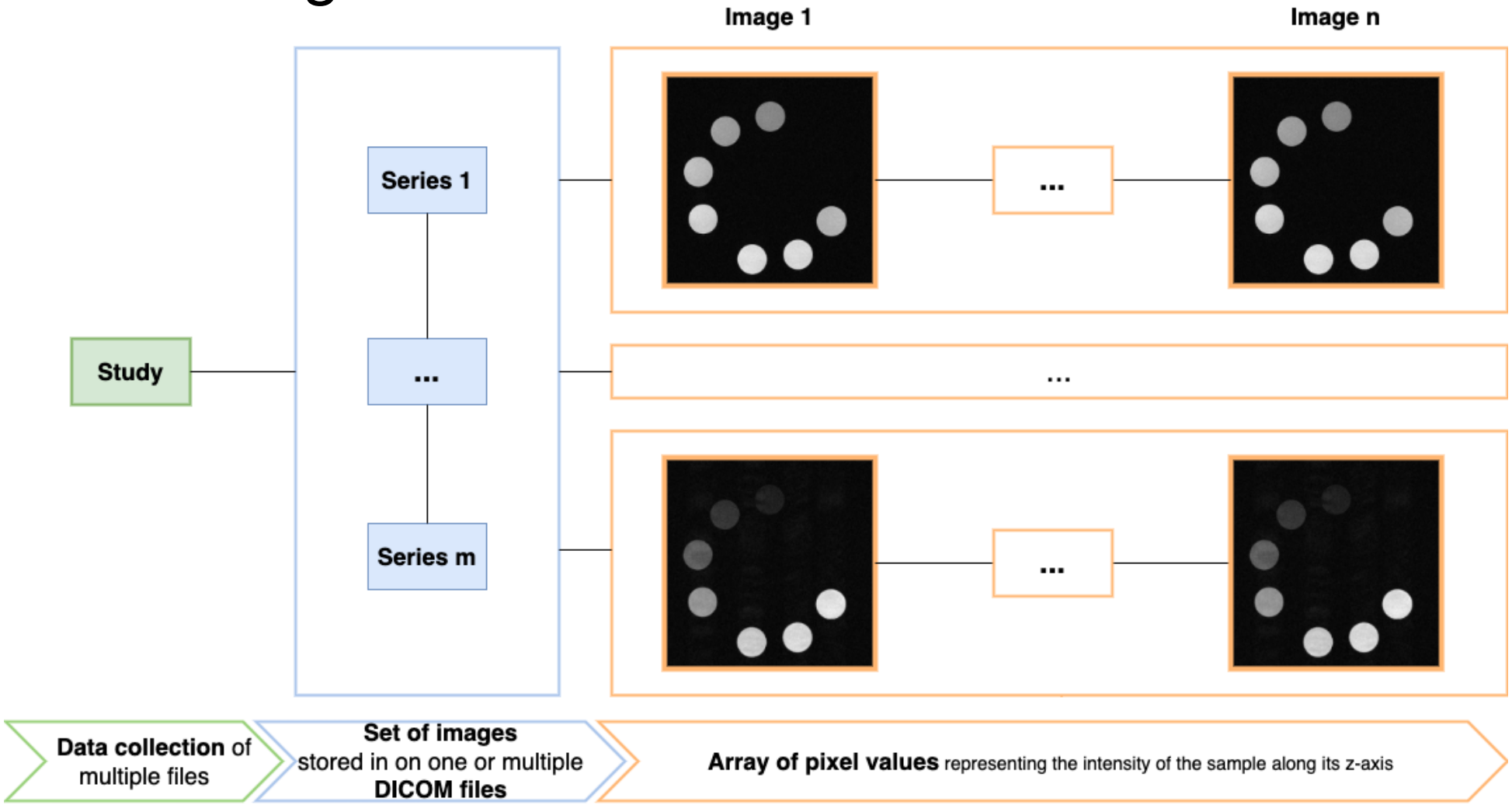
- Adaption of metadata vocabulary to the material sciences
- Metadata should be machine-actionable
- Deposition in open-access storage systems required

# How Can We Make the Data More FAIR Compliant?

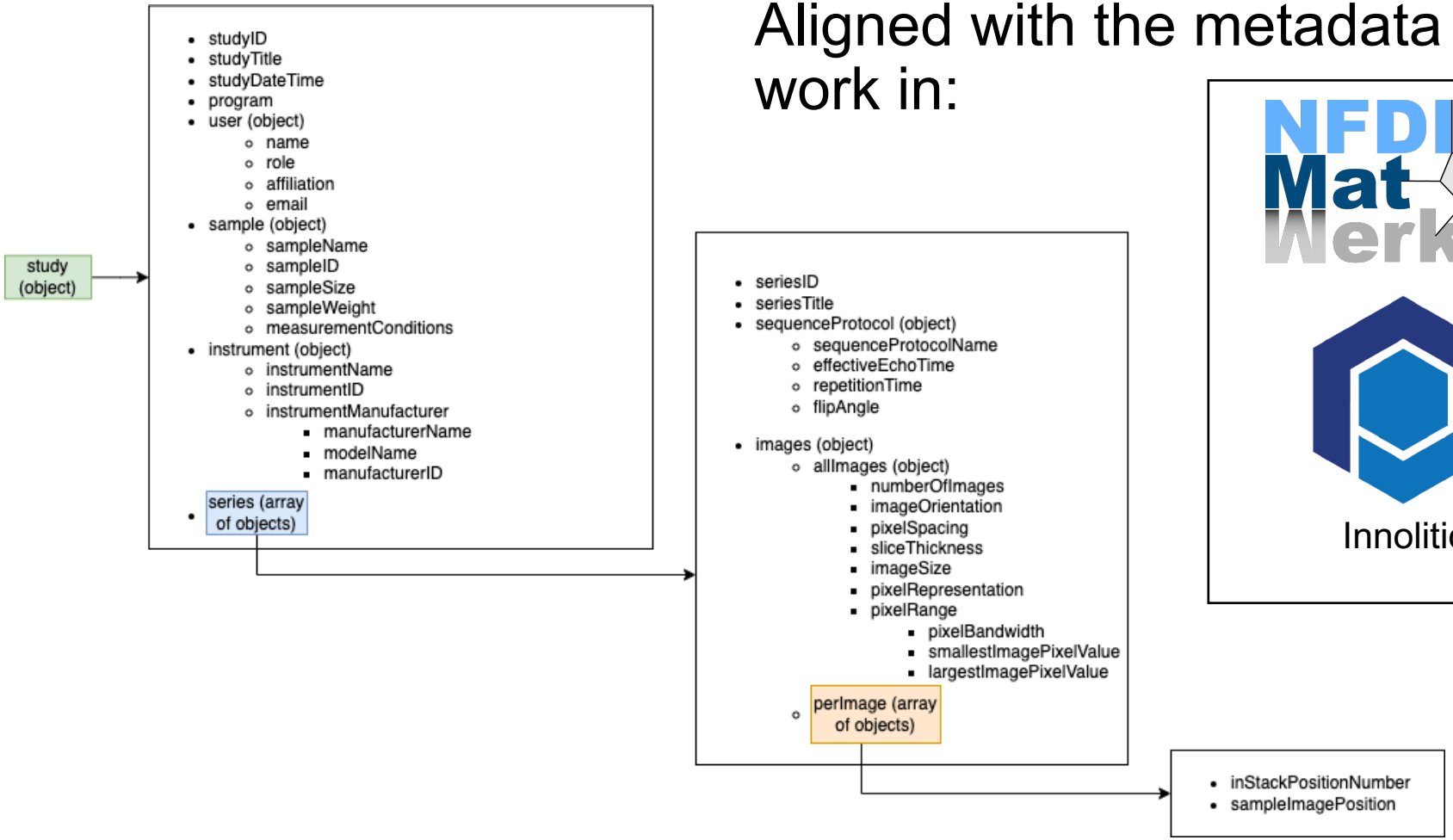
A workflow of specific steps is suggested:



# MRI Data Organization



# MRI Metadata Schema




Aligned with the metadata vocabularies from the work in:

NFDI MatWerk

nffa.eu | PILOT

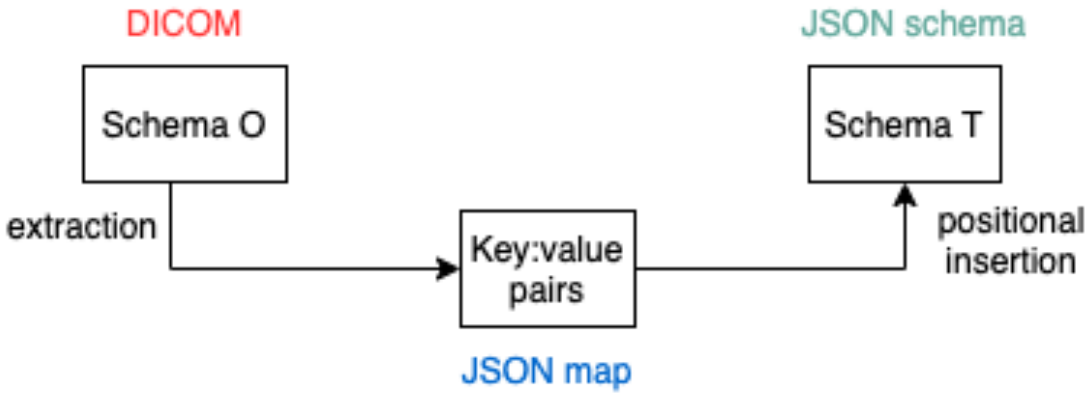
Innolitics

HELMHOLTZ MDMC

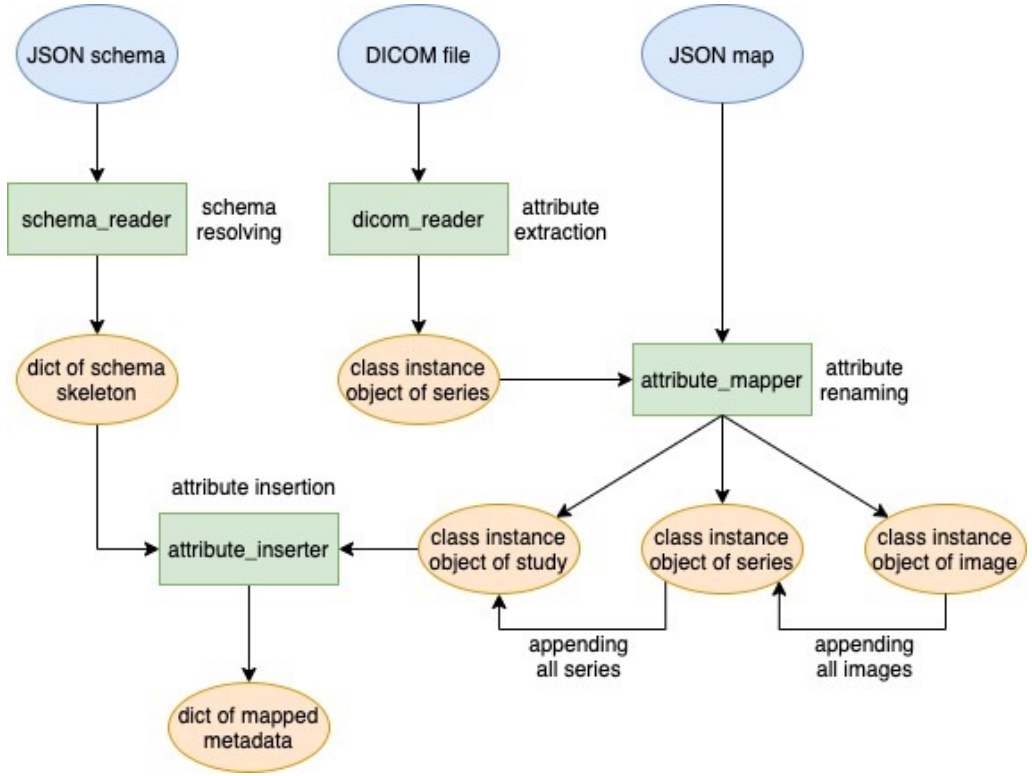
# Metadata Mapping

## JaMMaTo: JSON Metadata Mapping Tool for attribute extraction and transfer into schema format

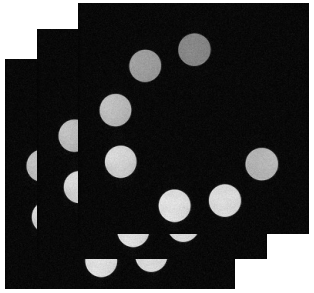
### Mapping Process




### Software Architecture



# MRI (Meta)data Storage



zenodo

- User friendly interface
- Set of administrative metadata
- Referencing to external metadata


```
{ "study":
        { "studyID": "1.2.840.10008.5.1.4.1.1.4.1",
          "studyTitle": "7 samples\nCuSO4 0 - 100 mM",
          "studyDateTime": "2021-10-15T13:48:31Z",
          "program": "['Acquisition PV-360.3.1', 'ParaVision 360.3.1']",
          "user": { "name": "nmrsu",
                    "affiliation": { "institutionName": "Karlsruher
Institut fuer Technologie" }
                  },
          ...
```


nffa.eu

**MetaStore Frontend for NFFA EU Pilot**

Schema and Metadata Management

- User friendly interface
- Metadata schema registration and validation
- Referencing to data
- Metadata document search

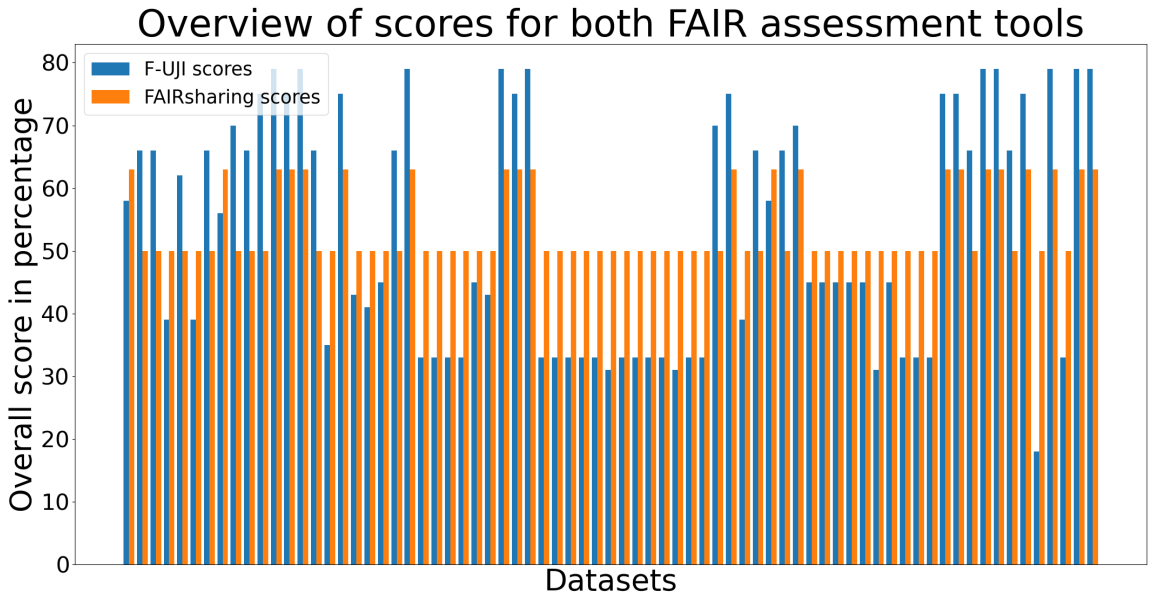
# FAIR Assessment of MRI (Meta)data

- Quantitative comparison using F-UJI and FAIRsharing


- Total datasets: 73
- Average score: 52.29 %
- Lowest score: 18 %
- Highest score: 79 % (7 reference datasets and our two MRI datasets)

# FAIR Assessment of MRI (Meta)data

- Quantitative comparison using F-UJI and FAIRsharing




- Total datasets: 73
- Average score: 52.29 % 53.47 %
  Lowest score: 18 %, 50 %
- Highest score: 79 % (7 reference datasets + our 2), 63 % (19 reference datasets + our 2)

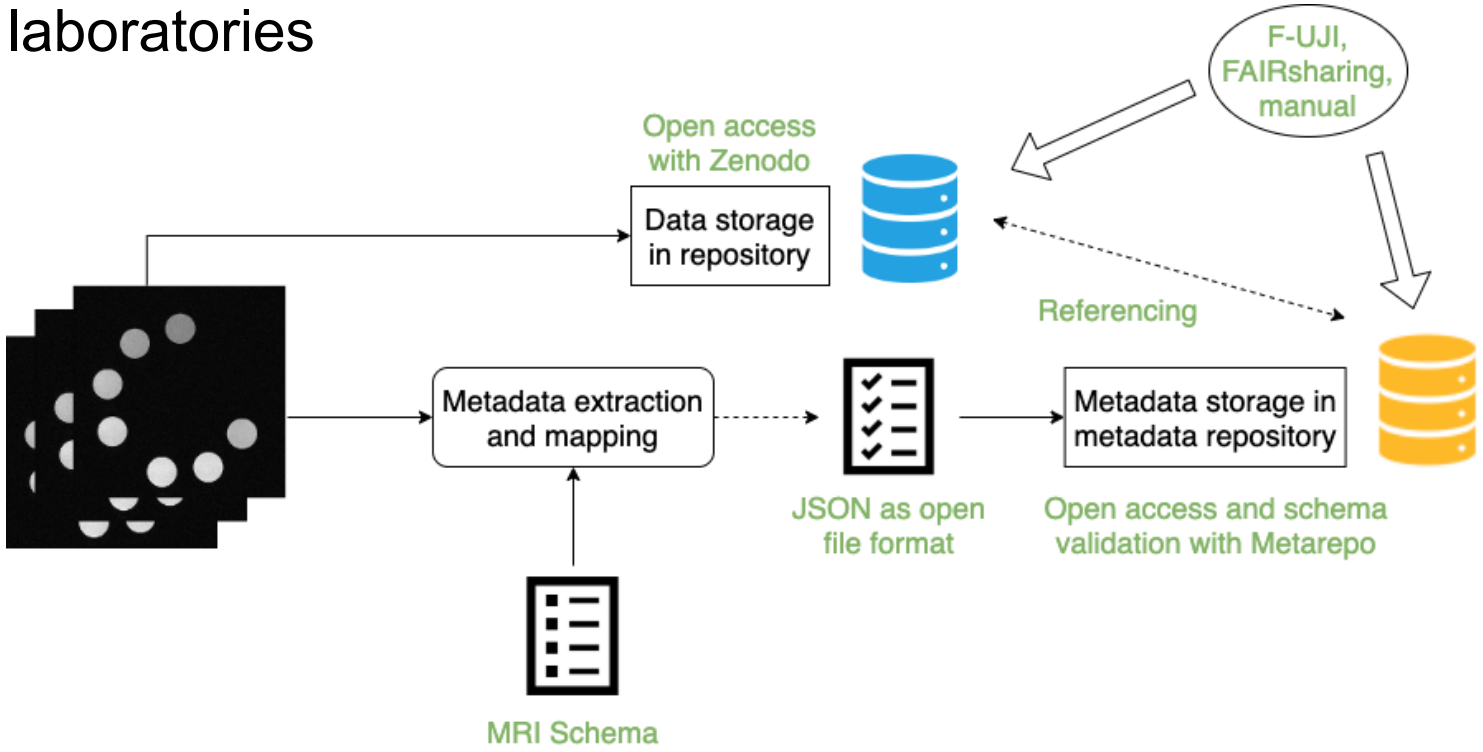
- Some commonalities and some variations between the tests
- A similar ratio of scores for datasets within each tool

# Manual Assessment Considering Missing Aspects

- I2 (usage of vocabularies that follow the FAIR principles)
  - Terms in MRI schema defined in public glossaries, accepted by the community
  - Need to be included in a vocabulary to be resolvable by PIDs
- R1 (plurality of accurate and relevant attributes)
  - MRI schema contains rich metadata with definitions, confirmed by community experts
  - Can be mapped to PROV-O for a more general context
- R1.2 (detailed provenance)
  - MRI schema contains all details needed for reproducing the output and performing the imaging experiment

# Conclusions

- Improved FAIRness of MRI data, validated by F-UJI, FAIRsharing, and a manual assessment
- Workflow and components are reusable by the MRI community or other scientific laboratories




MRI Schema

MRI data – $CuSO_4$

MRI data - Glucose

JaMMaTo

# References

- MRI schema: https://metarepo.nffa.eu/api/v1/schemas/mri_schema?version=8
- MRI data:
  - https://doi.org/10.5281/zenodo.6107721
  - https://zenodo.org/record/7778338#.ZGNCGOxBz0o
- JaMMaTo: https://github.com/kit-data-manager/JaMMaTo
- F-UJI: https://www.f-uji.net/
- FAIRsharing: https://fairsharing.github.io/FAIR-Evaluator-FrontEnd/ - !/