



2022

A Taxonomy of Tools and Approaches for FAIRification

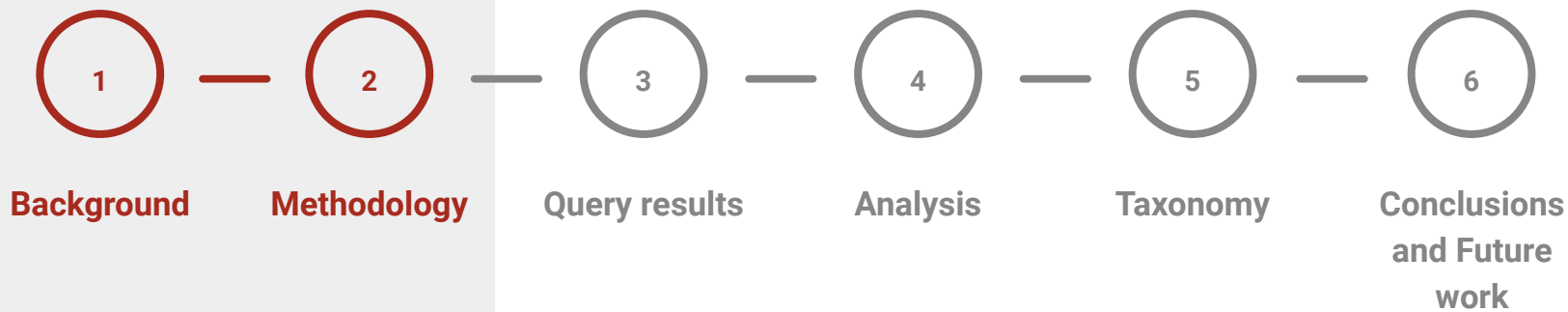
Dario **Mangione**

Leonardo **Candela**

Donatella **Castelli**

Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo" - Consiglio Nazionale delle Ricerche

Outline



Background

→ FAIR principles¹

- 4 foundational principles
 - Findability
 - Accessibility
 - Interoperability
 - Reusability
- 15 guiding principles to maximise the added-value of research objects
- nothing new to the digital libraries (just the presentation format)
- recently recognised at European level (REGULATION 2021/695 establishing Horizon Europe; open research data pilot 2017)
- open to interpretation²

→ FAIRification

- practical process
- different implementations

¹ Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1), Article 1. <https://doi.org/10.1038/sdata.2016.18>

² Jacobsen, A., de Miranda Azevedo, R., Juty, N., Batista, D., et al. (2020). FAIR Principles: Interpretations and Implementation Considerations. *Data Intelligence*, 2(1–2), 10–29. https://doi.org/10.1162/dint_r_00024

Methodology

There is FAIR (Findable, Accessible, Interoperable and Reusable) and

- FaIR (climate model)
- FAIR (Facility for Antiproton and Ion Research)
- ...

Need to improve precision and recall

- filters
- stopwords
- OpenAIRE subjects (559)

243,815 RESEARCH OUTCOMES, PAGE 1 OF 24,382

Publication · Research · Preprint · Article · Report · 2002
How fair is fair trade?

- × Publications
- × Software
- × Open Access
- × Report
- × Project deliverable
- × Project milestone
- × English

- trade ×
- value ×
- play ×
- treatment ×
- price ×
- indivisibility ×

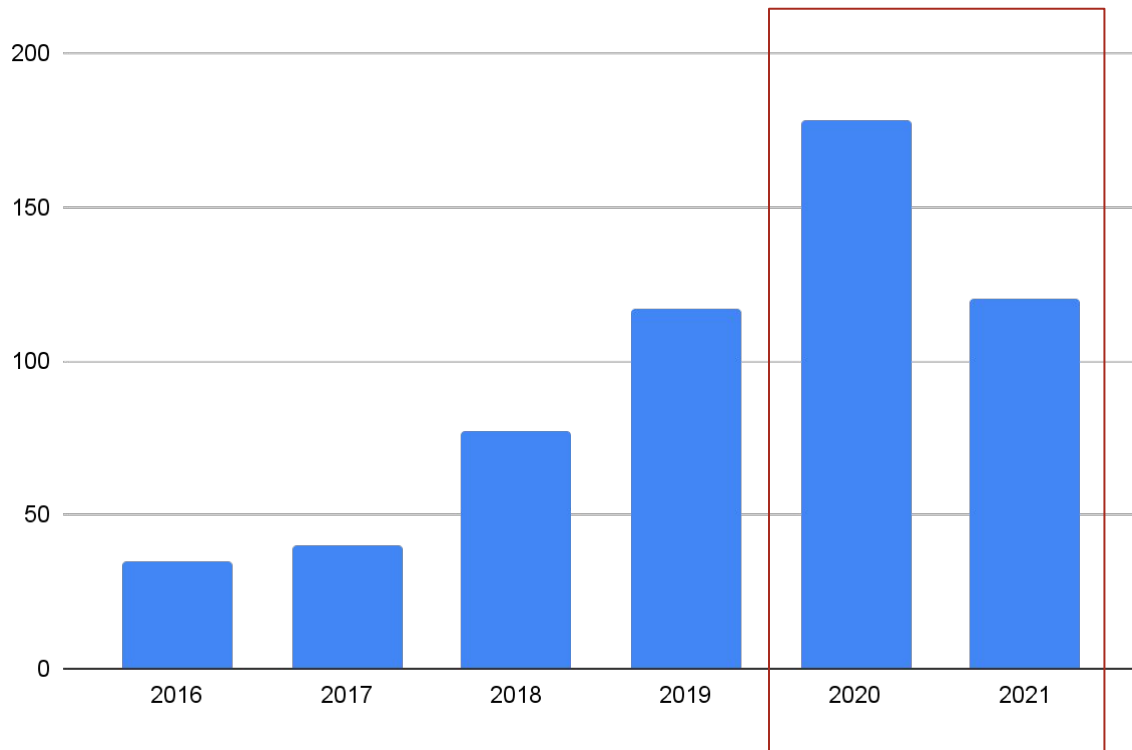
...

Query results

567 publications

- Project deliverables
- Project milestones
- Reports

389 software entries



Analysis

477 corpus-derived entries (277)

→ dataset published into 7 CSV³

1. OpenAIRE query results
2. tool/source (duplicates)
3. tools (unique)
4. tools/classification
5. tools/FAIR principles
6. tools/FAIR resource
7. tools/domains

³ Mangione, Dario, Candela, Leonardo, & Castelli, Donatella. (2022). A Taxonomy of Tools and Approaches for FAIRification (Version 1) [Data set]. Italian Research Conference on Digital Libraries (IRCDL), Padova, Italy. Zenodo. <https://doi.org/10.5281/zenodo.6037509>

Analysis

FAIR principles-derived categories

FAIR principle	tool/service category
F1 globally unique and persistent identifiers	GUPRI helper
F2 rich metadata	Metadata helper
F3 identifier of the data	Metadata helper
F4 indexed in a searchable resource	Indexing and discovery service
I1 language for knowledge representation	Metadata helper Converter
I3 qualified references to other (meta)data	Metadata helper
R1.1 data usage licence	Licence helper
R1.2 provenance	Metadata helper
R1.3 community standards	Metadata helper Converter

Analysis

FAIR principles-derived categories

→ Accessibility as a general reference

- “A1. (Meta)data are retrievable by their identifier using a standardised communications protocol”
- “A1.1 The protocol is open, free, and universally implementable”
- “A1.2 The protocol allows for an authentication and authorisation procedure, where necessary”
- “A2. Metadata are accessible, even when the data are no longer available”
- depends on F1
- depends on the protocols and the policies adopted

→ I2 does not establish any relationship

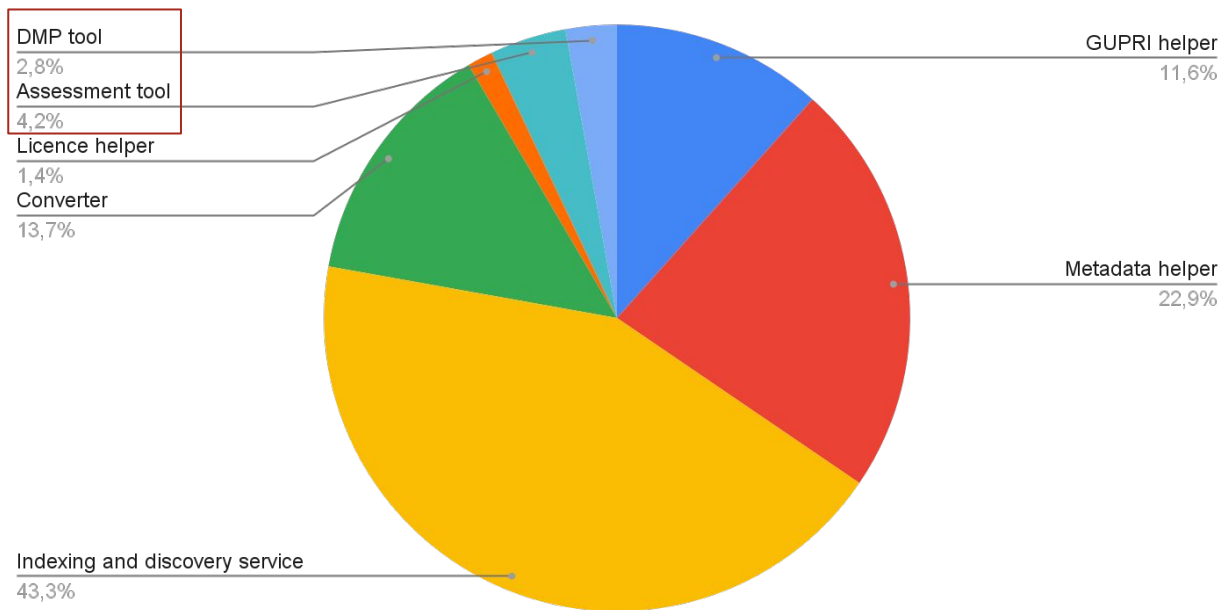
- “I2. (Meta)data use vocabularies that follow FAIR principles”
- recursivity

→ R1 does not require a new category per se

- “R1. (Meta)data are richly described with a plurality of accurate and relevant attributes”
- its subdivisions do

Analysis

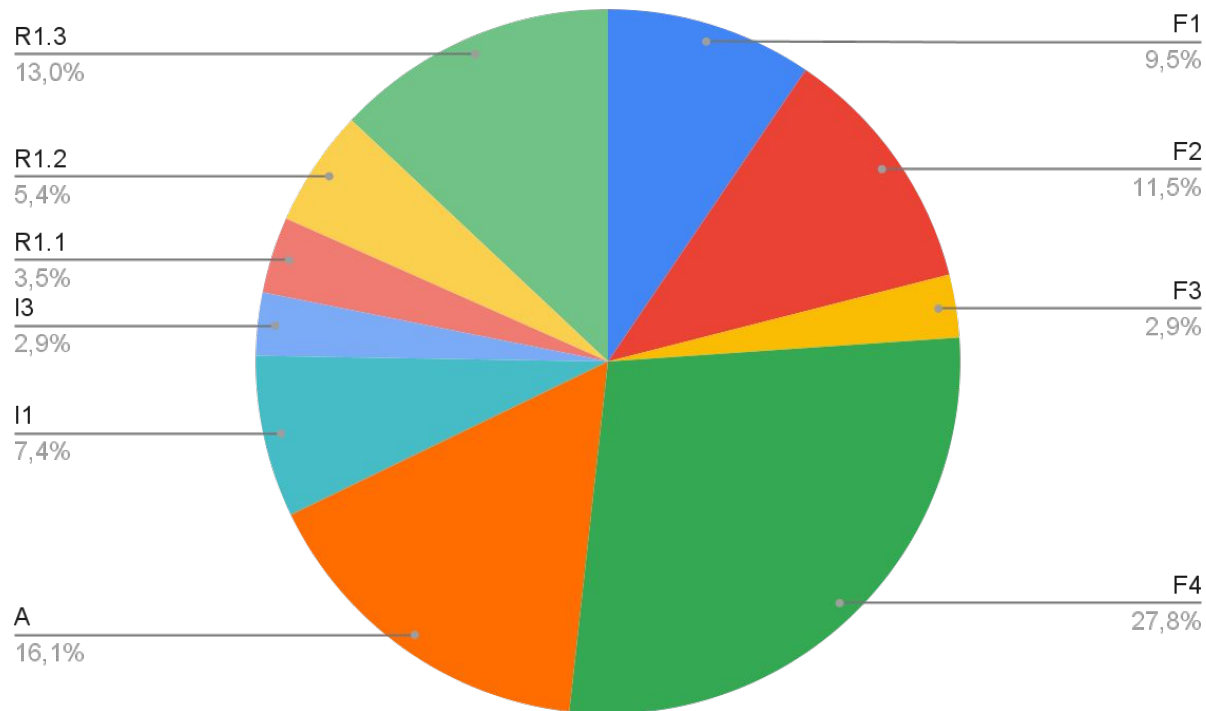
Entries and categories



Analysis

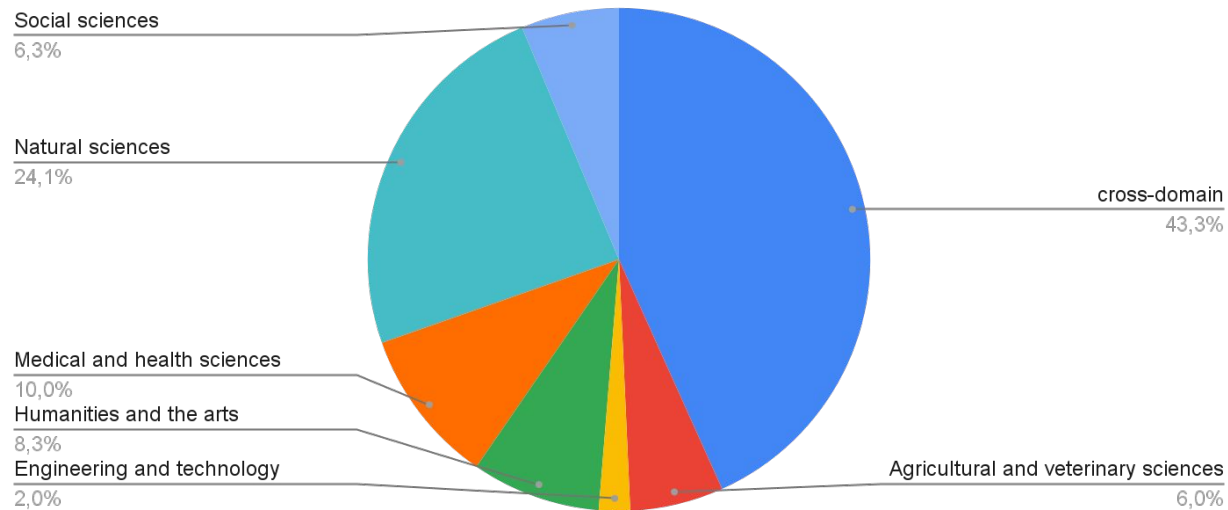
Entries and FAIR principles

- Findable 51,7%
- Accessible 16,1%
- Interoperable 10,3%
- Reusable 21,9%



Analysis

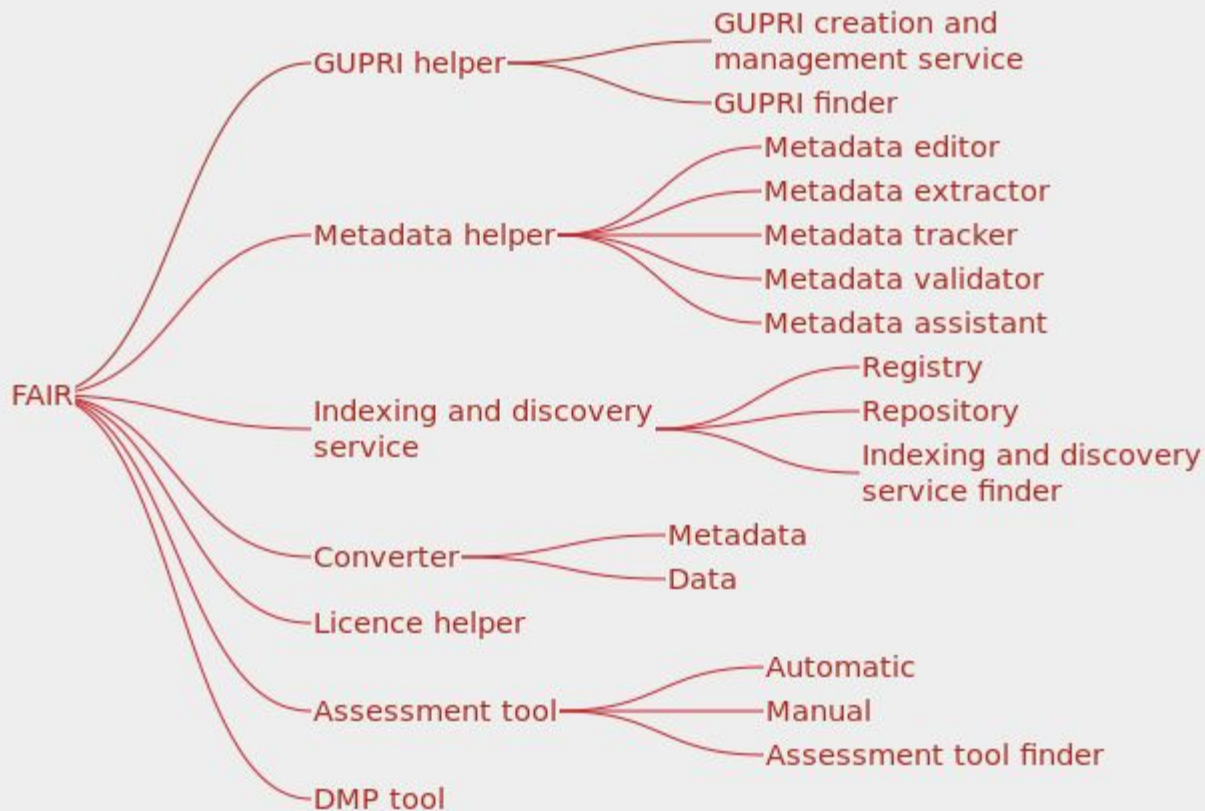
Entries and scientific domains



OECD. (2015). *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development*. 59. OECD. <https://doi.org/10.1787/9789264239012-en>

Taxonomy

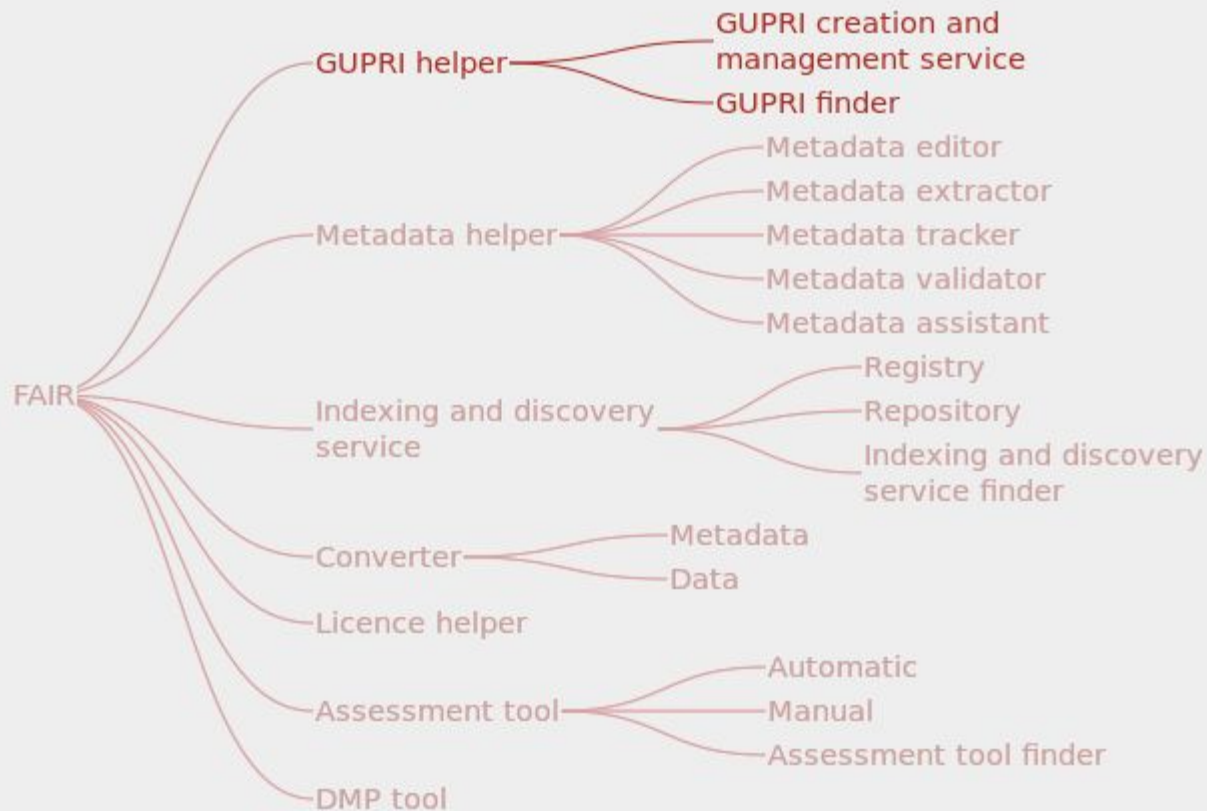
- 7 classes
- 15 subclasses



Taxonomy

Globally Unique, Persistent and (machine) Resolvable Identifier helper (33)

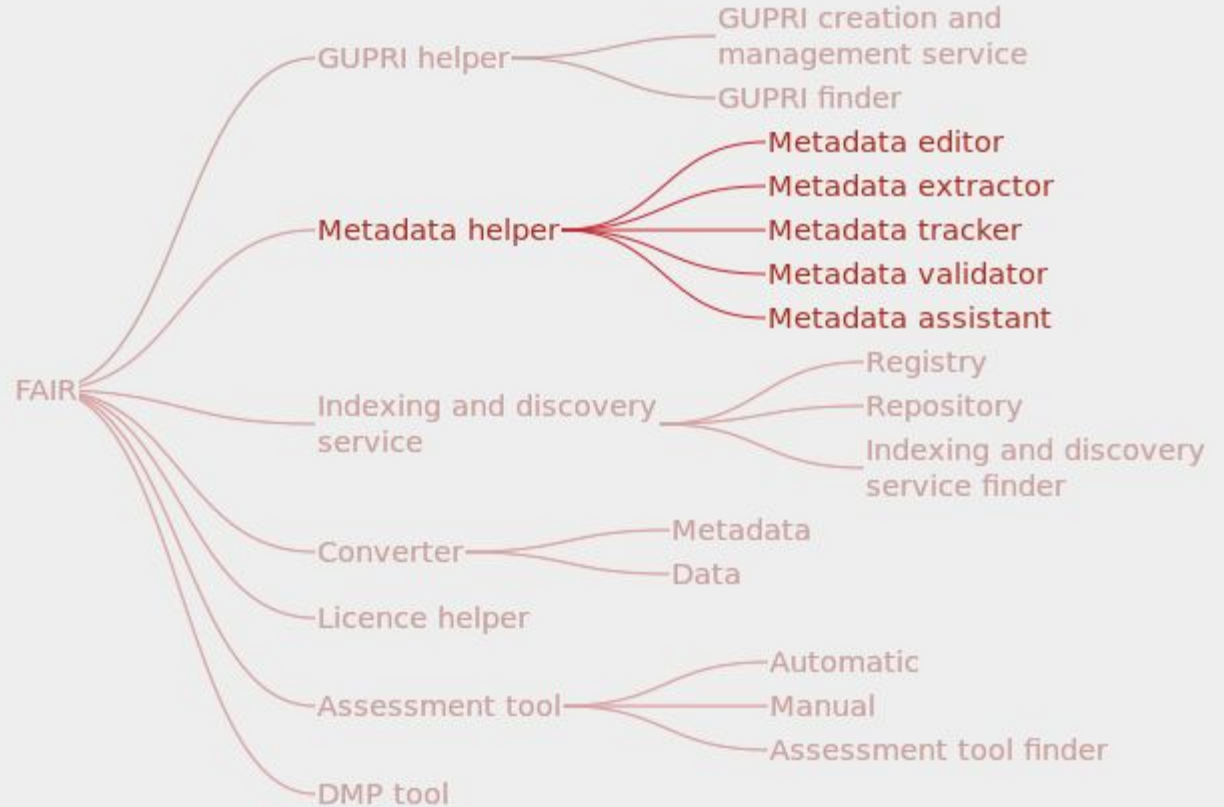
- create GUPRIs
- maintain GUPRIs
- find providers



Taxonomy

Metadata helper (65)

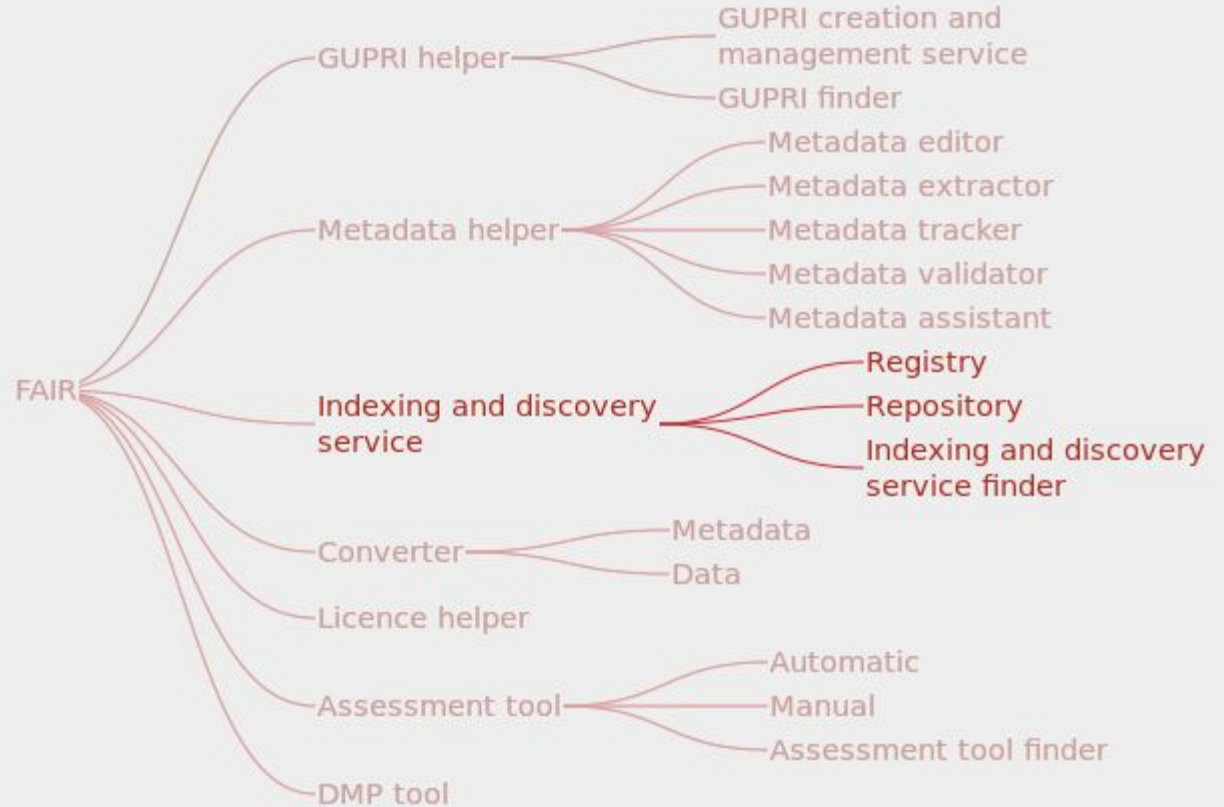
- edit metadata
- extract metadata
- automatically add metadata
- validate metadata
- suggest metadata



Taxonomy

Indexing and discovery service (123)

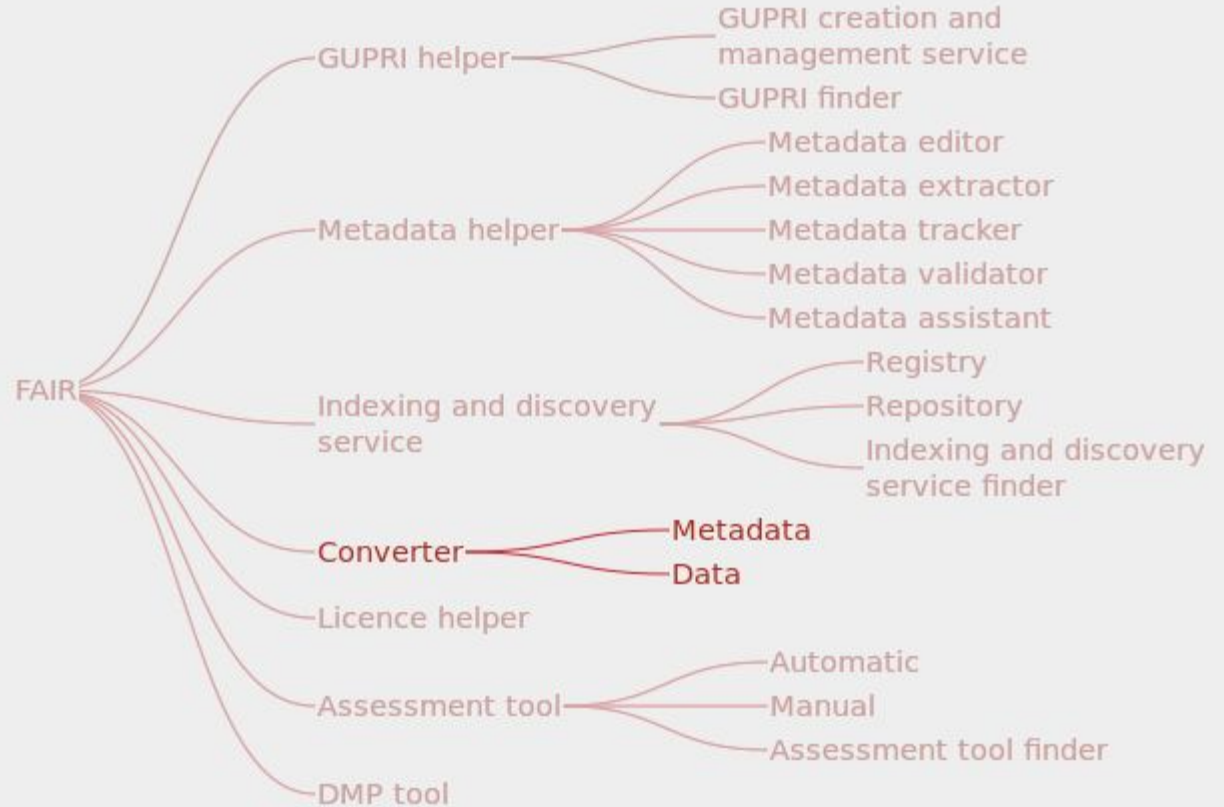
- index resources
- search resources
- access resources
- find services



Taxonomy

Converter (39)

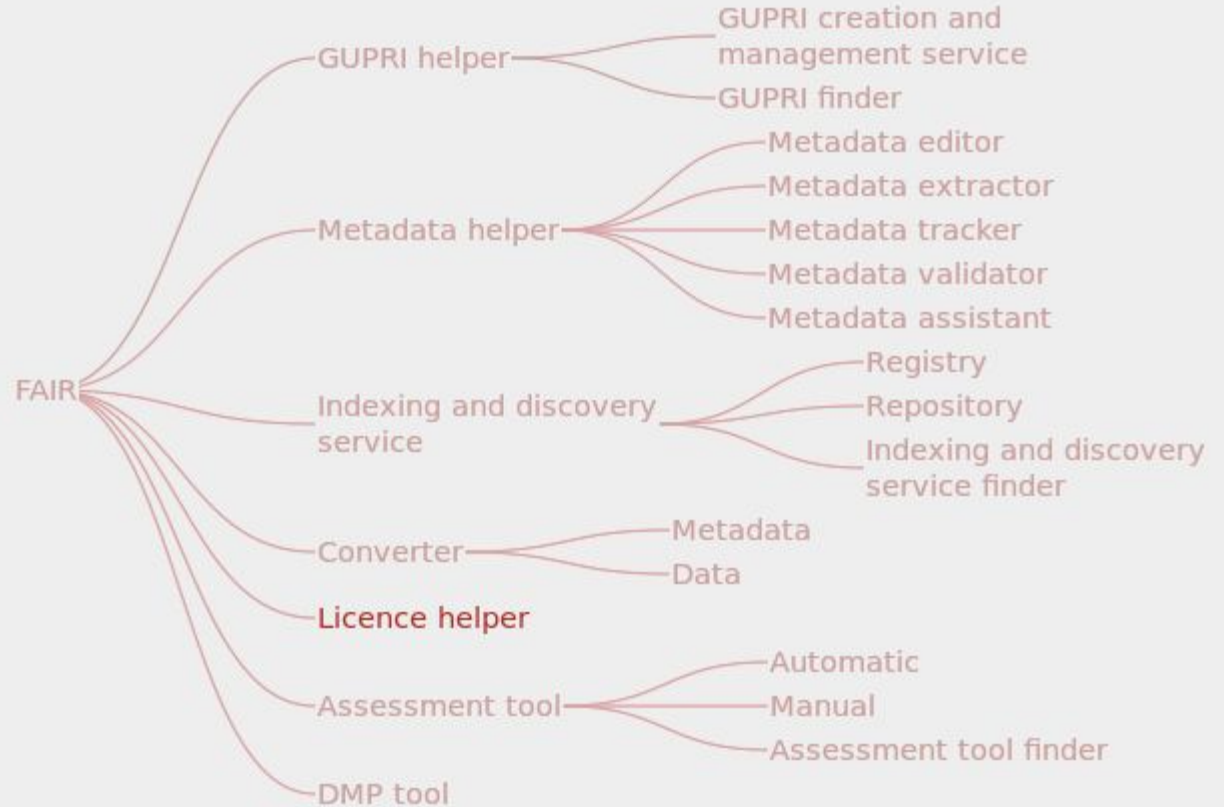
→ convert (meta)data



Taxonomy

Licence helper (4)

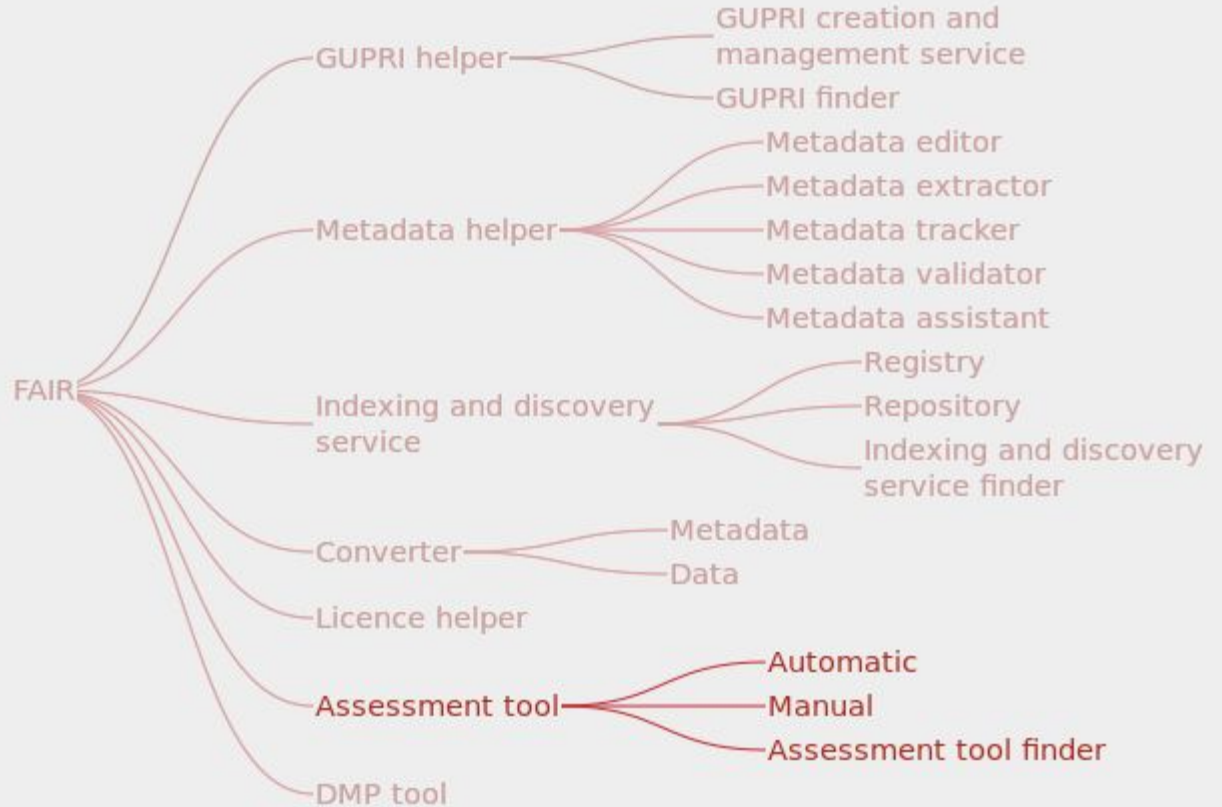
→ choose a licence



Taxonomy

Assessment tool (12)

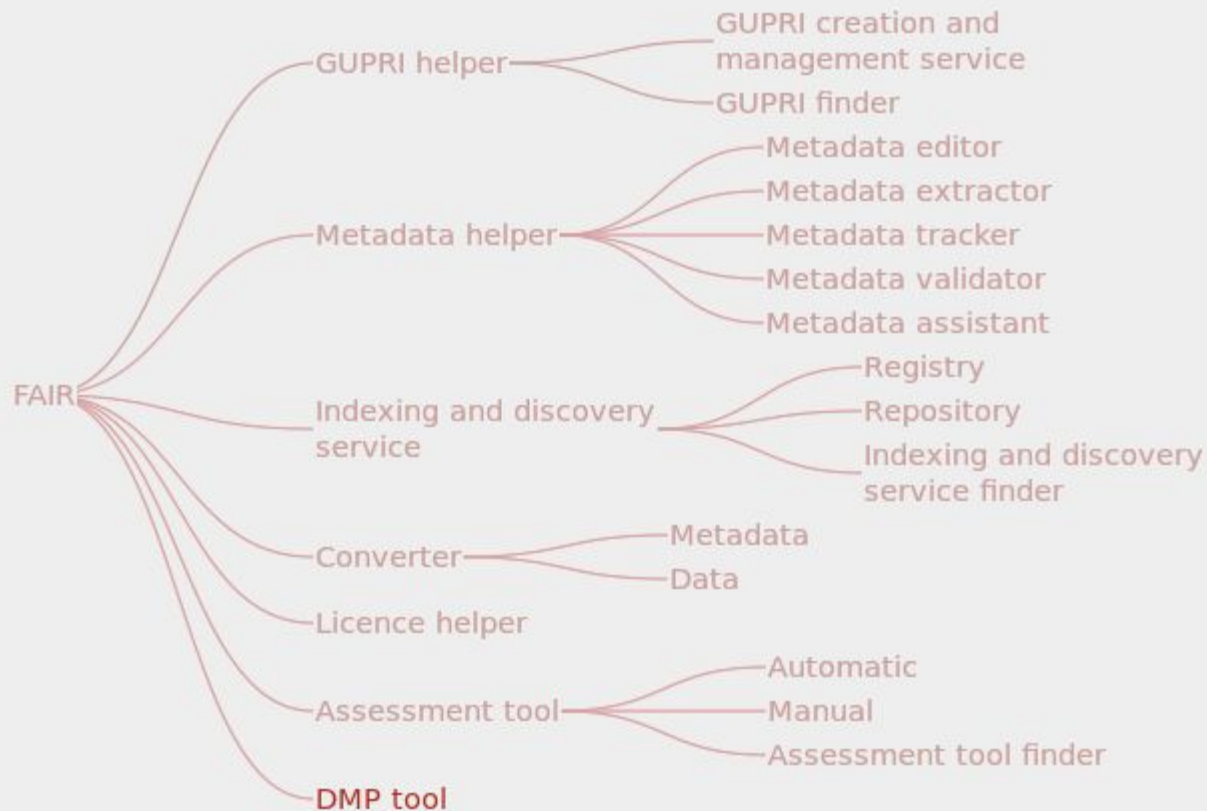
- evaluate FAIRness
- find assessment tools



Taxonomy

Data Management Plan tool (8)

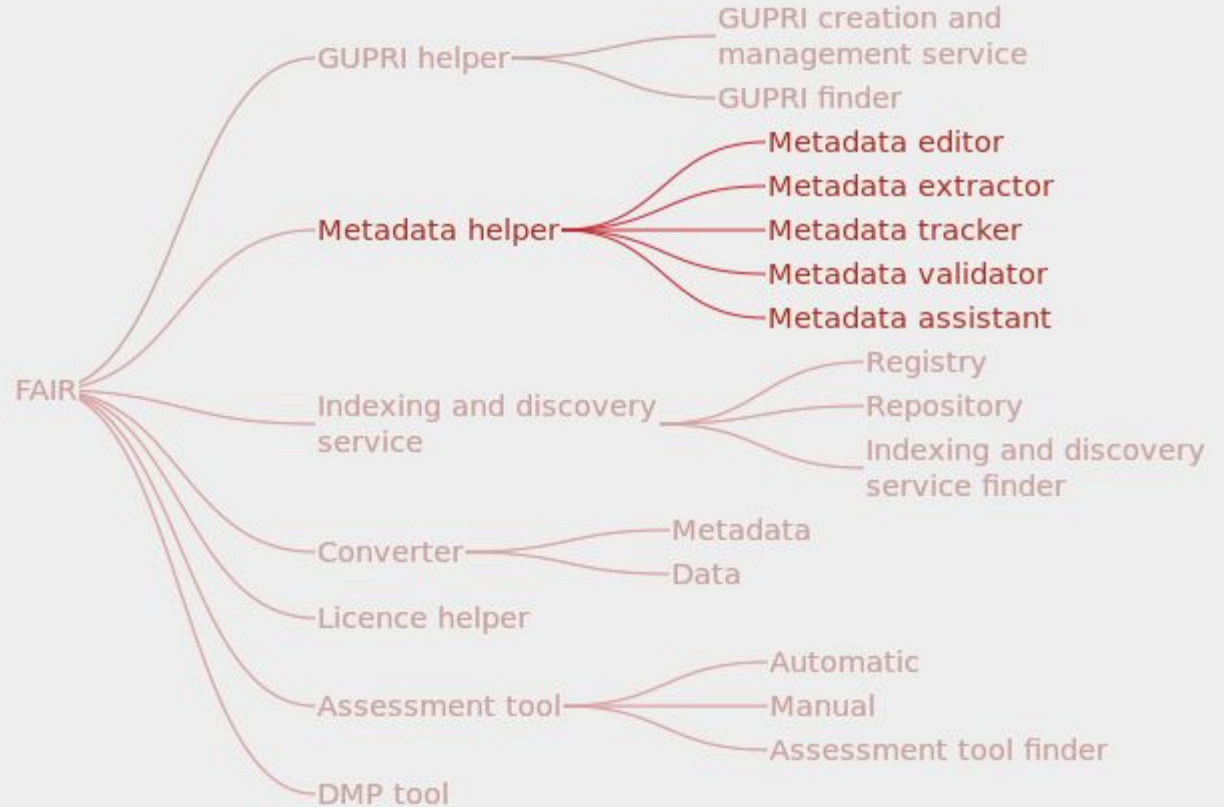
→ create data management plans



Taxonomy

Metadata helper (65)

- edit metadata
- extract metadata
- automatically add metadata
- validate metadata
- suggest metadata



Taxonomy

FAIR principles reference

GUPRI helper	GUPRI creation and management service GUPRI finder	F1
Metadata helper	Metadata editor Metadata extractor Metadata tracker Metadata validator Metadata assistant	F2 F3 I1 I3 R1.2 (R1.3) F2 F2 R1.2 I1 R1.3 R1.3
Indexing and discovery service	Registry Repository Indexing and discovery service finder	F4 F4 A
Converter	Metadata Data	I1 I3 R1.3 R1.3
Licence helper		R1.1
Assessment tool	Automatic Manual Assessment tool finder	F A I R F A I R
DMP tool		

Conclusions and future work

→ Findings

- 225 publications, 95 software entries
- 277 tools (the majority are cross-domain)
- taxonomy (7 classes and 15 subclasses; generally disjoint)
- (sub)classes/FAIR principles != 1 to 1
- no all-in-one FAIRification solutions

→ Value

- state of the art of European-related FAIRification activities
- FAIRification process support tool
- FAIRification software development support tool

→ Further assessment

- literature
- communities of practice
- FAIR implementation profiles⁴

⁴ Schultes, E., Magagna, B., Hettne, K. M., et al. (2020). Reusable FAIR Implementation Profiles as Accelerators of FAIR Convergence. In G. Grossmann & S. Ram (Eds.), *Advances in Conceptual Modeling* (pp. 138–147). Springer International Publishing. https://doi.org/10.1007/978-3-030-65847-2_13

Thanks!

Questions?

dario.mangione@isti.cnr.it

leonardo.candela@isti.cnr.it

donatella.castelli@isti.cnr.it