

Why Mappings Matter and How to Make them FAIR?

Organised in collaboration with the
SSSOM community

Short outline

- Welcome and Housekeeping rules
- Why this workshop?
- What about a bit of context? Short intro to FAIR Impact
- Why FAIR Mappings are a good idea?



Welcome and Introduction

Yann Le Franc, PhD
CEO and Scientific Director, e-Science
Data Factory, France
ORCID:0000-0003-4631-418X

Agenda

	Topic	Presenter
14:00-14:10	Welcome and Introduction	Yann Le Franc
14:10-14:30	A brief state of the art briefing and why FAIR mappings would be a good idea	Yann Le Franc
14:30- 15:30	Perspective from different communities	Henriette Harmse Keith Jeffery Luca De Santis Claus Weiland
15:30-15:45	What about some coffee?	
15:45-16:00	Perspective from the EOSC Semantic Interoperability Task Force	Milan Ojstersek
16:00-16:15	SSSOM: a machine actionable model for simple mappings	Nicolas Matentzoglu
16:15- 16:30	FAIRCORE4EOSC: toward a mapping registry	Tommi Suominen
16:30- 17:30	BREAK OUT Discussions	
17:30- 17:45	Results from the breakout	
17:45-18:00	Mapping with ChatGPT	Kristian Garza

Housekeeping rules and documents

- Collaborative Note taking:
<https://docs.google.com/document/d/1c1huuflwK0IW4hP09Yr4SYtdEF5rEM79bRkK50P9Rq0/edit?usp=sharing>
- Miro Board: https://miro.com/app/board/uXjVMTALLhA=/?share_link_id=977694451072

Breakout topics

- Requirements for FAIR Mappings
- What are the different types of mappings?
- What are the existing mapping tools and what would be the requirements for such mapping tools?
- What are the existing methodologies for creating mappings?

A framework for FAIR Mappings

- Technical requirements for FAIR mappings
- FAIR Mappings implementation Cookbook
- Mapping Cookbook: how to do mappings in practices with examples?
- Guidelines for publishing mappings (target: mapping and semantic artefact repositories)
- Extending and testing SSSOM to be used as mapping exchange model (use-cases)

Why this workshop?

- Starting discussing and understanding community practices
- Collecting information and views from practitioners
- Establishing a community consensus to ease the process of working with mappings
- Collaborate on the definition of FAIR requirements and FAIR Mapping framework content



About FAIR Impact

Yann Le Franc, PhD
CEO and Scientific Director, e-Science
Data Factory, France
ORCID:0000-0003-4631-418X

Expanding FAIR Solutions across Europe



Call
**HORIZON-INFRA-2021-
EOSC-01-05**
*Enabling discovery and
interoperability of federated
research objects across scientific
communities*

**Expanding FAIR
solutions in Europe**

**Partly following up on
FAIRsFAIR**



EU funded project

**Coordination and
Support Action**

10 million euro

**36 months, starting
1 June 2022**



**28 partners and
affiliate entities**

**From 10 EU
member states:
NL, FI, FR, DK, IT,
DE, ES, NO, BE,
RO**

and the UK



FAIR-IMPACT overall objective



WHAT:

to realise a FAIR EOSC by **supporting the implementation** of FAIR-enabling practices across scientific communities and research outputs at a European, national, and institutional level;

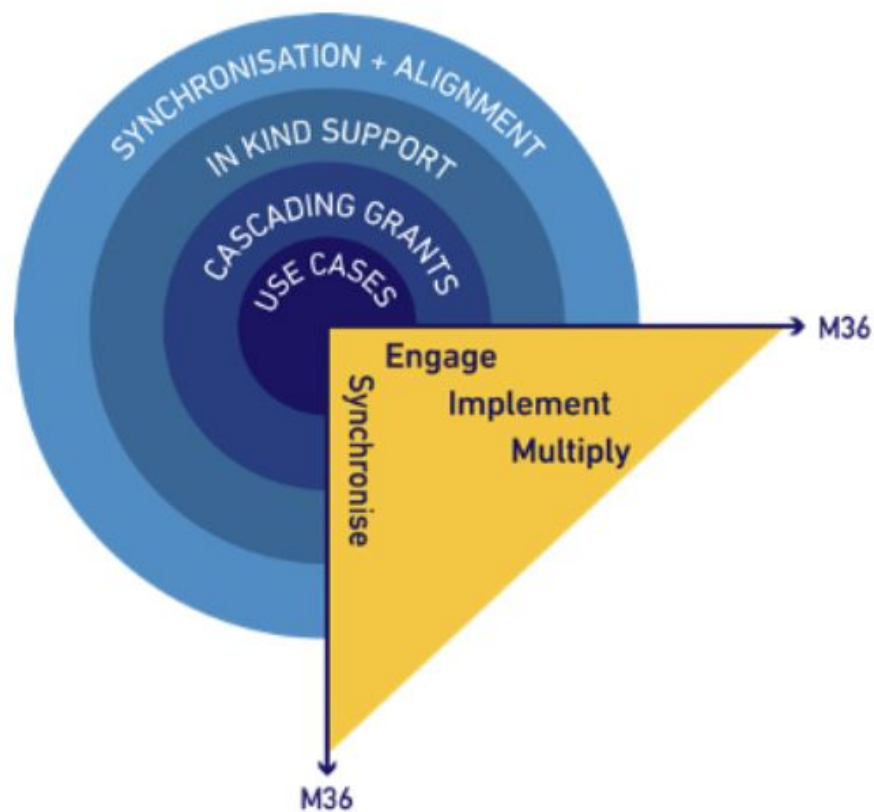
HOW:

- **identifying** current and emerging components for enabling FAIR (practices, policies, tools & technical specifications);
- **translating** viable solutions, guidelines and frameworks that have been developed for one domain or research output and **supporting** their application in others;
- taking the next step in implementation by **defining** the support, governance, and coordination mechanisms required to ensure the continuous function of FAIR-enabling practices in the EOSC.

FAIR-IMPACT Work Packages



Expected outcomes



Improved FAIRness of data and other research outputs by **coordinating the alignment of FAIR data practices and the implementation of frameworks**

Tier 1: integrated use cases

- Act as a **bridge** to the four scientific domains
- **Provide expertise**
- **Translate** what works for them



Social Sciences
and Humanities



Photon and
neutron science

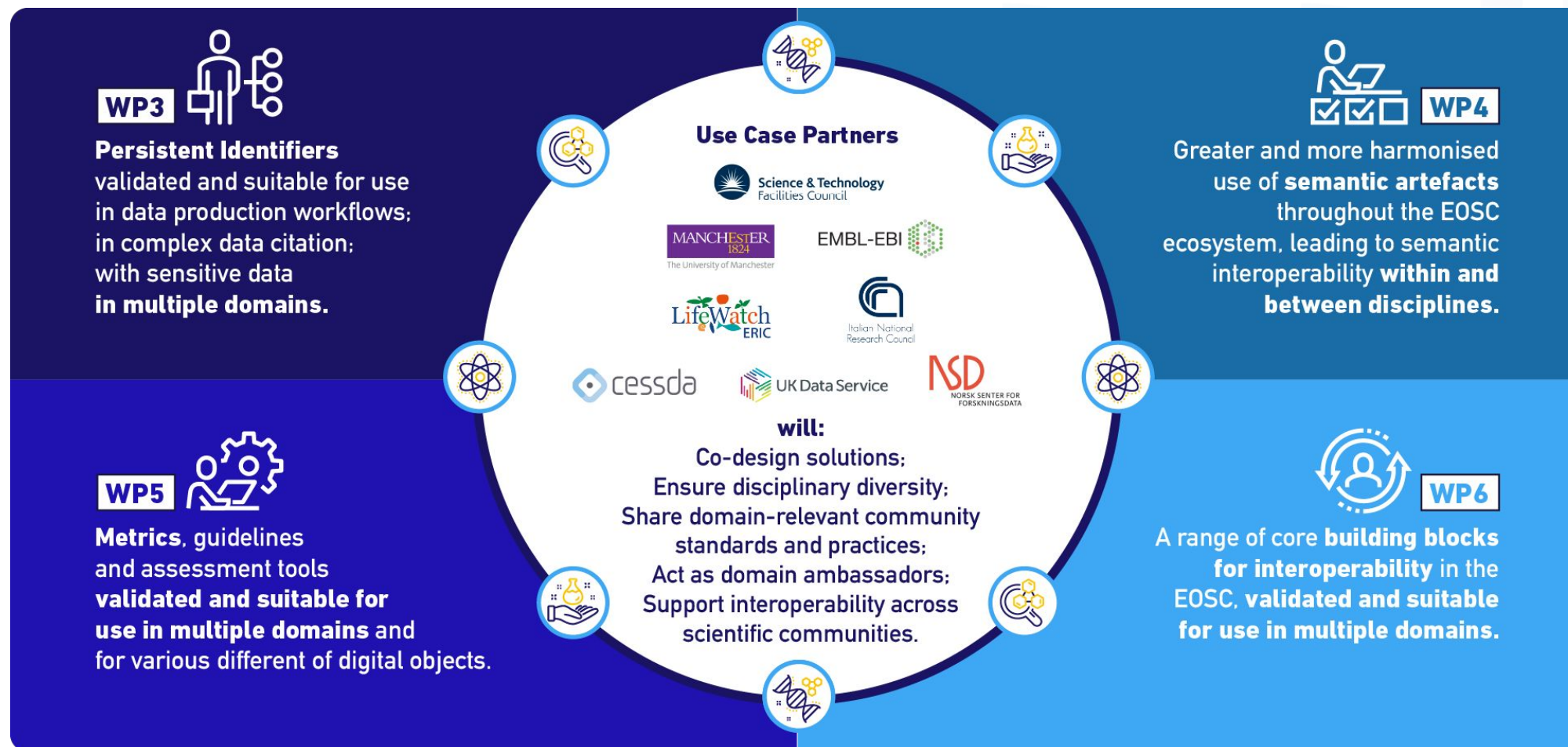


Life science

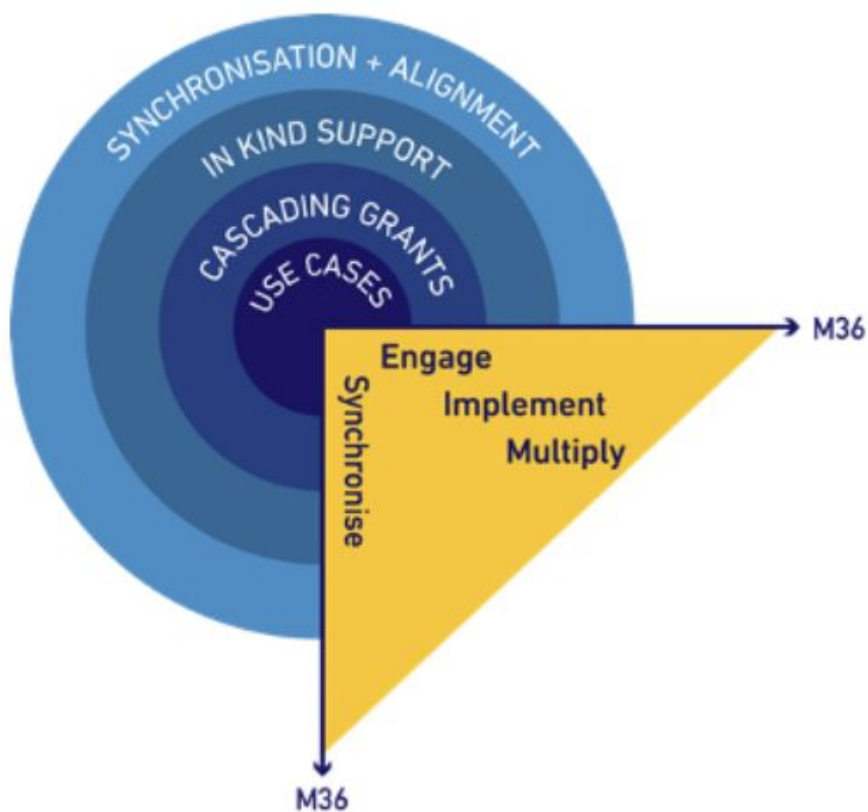


Agri-food
Environmental

Tier 1: integrated use cases



Expected outcomes



Wide uptake of and compliance with FAIR data principles and practices by national and European research data and metadata providers and repositories

Tier 2 and 3: cascading grants and in-kind support

Cascading grants

- open calls to support
- €500,000
- ca 50 organisations

In-kind support

Provision of in-kind support by the partners through open calls

The first Open Call is open!

Two support actions to enhance the FAIRness of data, semantic artefacts and data-related services.

[Learn more and apply!](#)

Home /

FAIR-IMPACT Open calls for support

In recent years, a number of tools, solutions and approaches have been developed to help different stakeholders implement the FAIR principles. Many of these have been developed through European Commission supported activities such as the [regional](#) and [thematic](#) INFRAEOSC projects and international initiatives such as the Research Data Alliance.

It can be challenging to know which tools and/or approaches to use for different purposes and how they should be implemented most effectively. To help take some of the guesswork out of implementing FAIR, the FAIR-IMPACT project will provide a set of clearly defined support actions that will help successful applicants learn how to use specific tools/methods/approaches which will help them start (or continue) their journey to becoming more FAIR-enabling and active contributors to the European Open Science Cloud.

Visit

<https://fair-impact.eu/fair-impact-open-calls-support> for more details/webinars/FAQs and to submit your application before June 1st 2023!

31st March 2023: the two first support actions launched

Applicants will be able to apply to join the first two support actions, which are designed to enhance the FAIRness of data, semantic artefacts and data-related services from March 31st, 2023, through our [grants platform](#). For a brief summary of the open call and details on the first two support offers, please read this [blog post](#). Detailed information about each of the support offers is provided below. On March 27, FAIR-IMPACT hosted a webinar to introduce potential applicants to the open call and to provide details about the two support actions on offer.

The support actions

#1: FAIRness assessment challenge

More harmonised use of semantic artefacts such as ontologies, terminologies, taxonomies, thesauri, vocabularies, metadata schemas and standards is a key element to achieving a high level of FAIRness. However, it can often be difficult to find and use semantic artefacts as they themselves are not always FAIR.

This targeted support action will help a cohort of dataset providers or semantic artefact developers to self-assess and work towards maximising the level of FAIRness of their resources. Successful applicants will be supported to employ a variety of assessment tools and methods, including [F-UJI](#), [O'FAIRe](#) and [FOOPS](#). FAIR-IMPACT mentors will provide guidance and advice on how to interpret their scores and to improve the overall FAIRness of their resources. The duration of effort required, including attendance at mandatory virtual workshops and completing the assigned 'homework', will be approximately five days.

[Learn more & apply before 1 June 2023](#)

#2: Enabling FAIR Signposting and RO-Crate for content/metadata discovery and consumption

The findability of both data and metadata is central to the FAIR principles. In this support action, FAIR-IMPACT will provide a method to increase the discoverability of the metadata and content. It will do so using a combination of two approaches; RO-Crate and Signposting. These two approaches are being used in combination as a pragmatic approach to making digital scholarly/research objects more FAIR. The duration of effort required to take part in this one will be five days of effort over a one-month period, again most likely in early autumn 2023.

[Learn more & apply before 1 June 2023](#)

Expected key results to look forward to:

1. A range of **governance, coordination and collaboration mechanisms** necessary for a functioning FAIR research ecosystem;
2. **Adoption of FAIR-enabling components and mechanisms** at multiple levels among a diverse group of stakeholders using a wide range of tools and Support Tiers
3. A **persistent identifier support programme** and guidance for the provision of PID services in EOSC, incl. mechanisms for adoption and implementation;
4. A **semantic framework** for the governance, creation, mapping, sharing, reuse, FAIRness assessment and interoperability of semantic artefacts for EOSC;
5. **Guidelines and prototypes** to support both enabling FAIR and assessing the FAIRness and trustworthiness of different types of digital objects in multiple domains.
6. Components and services for increased **legal, organisation, semantic and technical interoperability** within and across disciplines.



WP4 - Metadata and Ontologies

Semantic artefacts* are a key elements to achieving FAIR and these artefacts and their catalogues have to be FAIR too

*ontologies, terminologies, taxonomies, thesauri, vocabularies, metadata schemas and standards...

WP4 will develop and foster the uptake of a semantic framework for the governance, creation, mapping, sharing, reuse, FAIRness assessment and interoperability of semantic artefacts for EOSC.



WP4

Greater and more harmonised use of **semantic artefacts** throughout the EOSC ecosystem, leading to semantic interoperability **within and between disciplines.**

*...implementation
of FAIR-enabling
practices across
communities and
research outputs*



WP4's use case include

- Agri-food (INRAE with AgroPortal, EMPHASIS, ANAEE)
- Ecology/biodiversity (LifeWatch with EcoPortal)
- Earth sciences (CNRS with DataTerra EarthPortal)
- Astronomy/materials science/physics (UKRI-STFC)

*...projecting
the FAIR
principles to
other types of
research
objects*



WP4's research objects

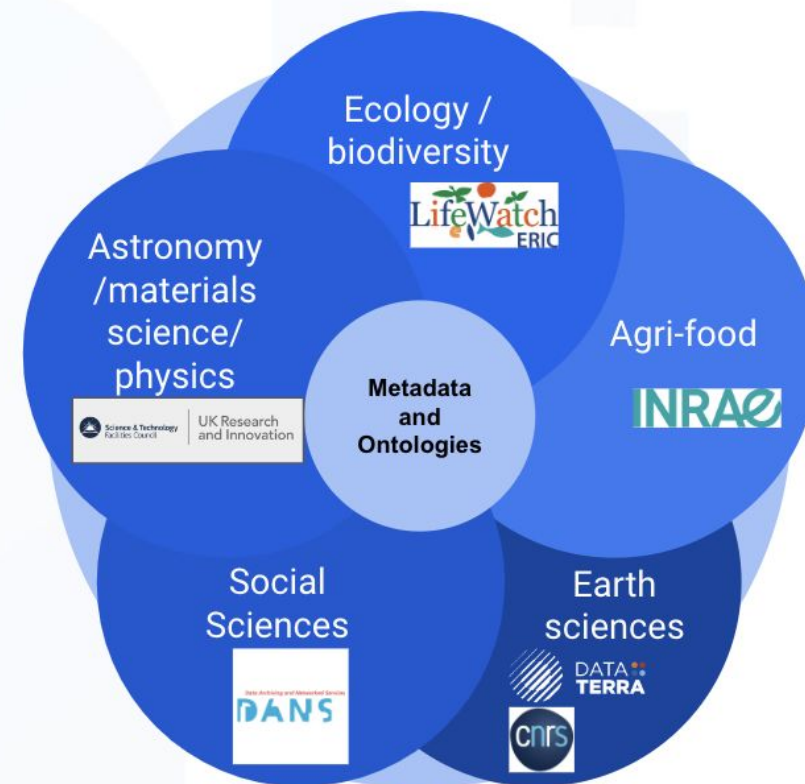
- *semantic artefacts, research software,
crosswalks & mappings*

WP4 use case objectives

- Broader and more harmonized use of semantic artefacts in EOSC.
- Guidelines to collect and curate research software metadata.
- A framework for metadata crosswalks and mappings between semantic artefacts.
- Assert and validate the use of semantic artefacts within data repositories for better data search and indexing.

Taking into account points of view from:

- SA consumers (data repositories & end users)
- SA providers
- SA catalogues –based on OntoPortal (AgroPortal, EcoPortal, EarthPortal) or not (LOV, CESSDA Vocab Service)



WP4's tasks



T4.1 (governance)

Nicola Fiore



T4.2 (lifecycle and catalogues)

Daniel Garijo +
Clement Jonquet
+ Alejandra
Gonzalez-Beltran



T4.3 (for research software)

Morane
Gruenpeter



T4.4 (crosswalks and mappings)

Yann Le Franc



T4.5 (in-use in data repos)

Sophie Aubin

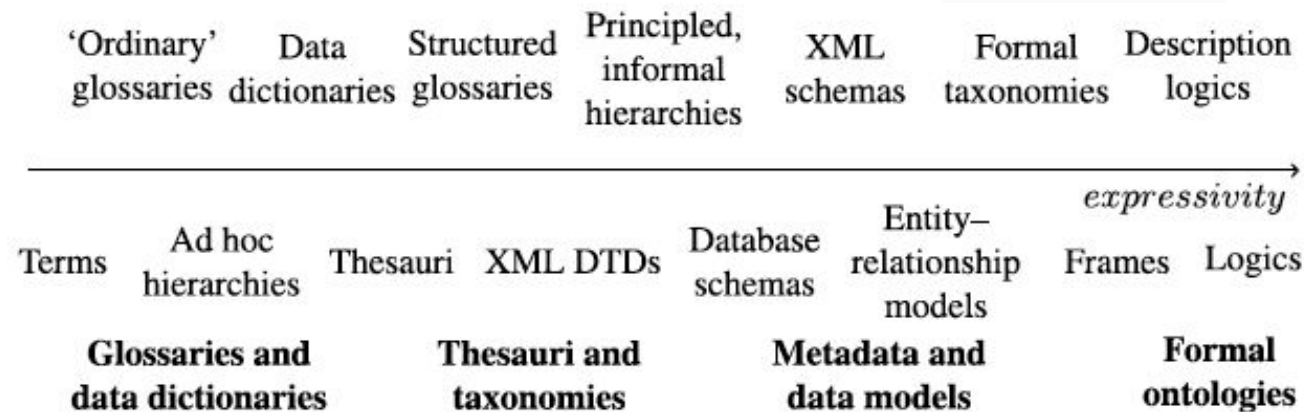




Why do we need mappings/crosswalks and why FAIR mappings/crosswalks would be a good idea?

Semantic artefacts?

A **semantic artefact** is defined within our work as a machine-actionable and -readable formalisation of a conceptualisation, enabling sharing and reuse by humans and machines. These artefacts may have a broad range of formalisation, from loose sets of terms, taxonomies, thesauri to higher-order logics. Moreover, semantic artefacts are serialised using a variety of digital representation formats, e.g., RDF Turtle, and OWL, using XML (RDF) and JSON-LD.



From « Ontology Matching », 2013, J. Euzenat, P. Shvaiko, Springer, 2nd ed. - DOI 10.1007/978-3-642-38721-0

What do we call mappings and crosswalks?

- No unique definition of mappings and crosswalks
- Related terms: ontology matching (process of finding relationships and correspondances), semantic alignment (set of correspondances between two or more ontologies), correspondance (relation holding between entities of different ontologies), mapping (directed relation between entities using mapping rules), mapping rule (correspondance)

From « Ontology Matching », 2013, J. Euzenat, P. Shvaiko, Springer, 2nd ed. - DOI 10.1007/978-3-642-38721-0

=> Need a common terminology with clear definition

When do we need mappings?



kitty
kitten



Cat



Felis
catus

When do we need mappings?



kitty
kitten



Cat

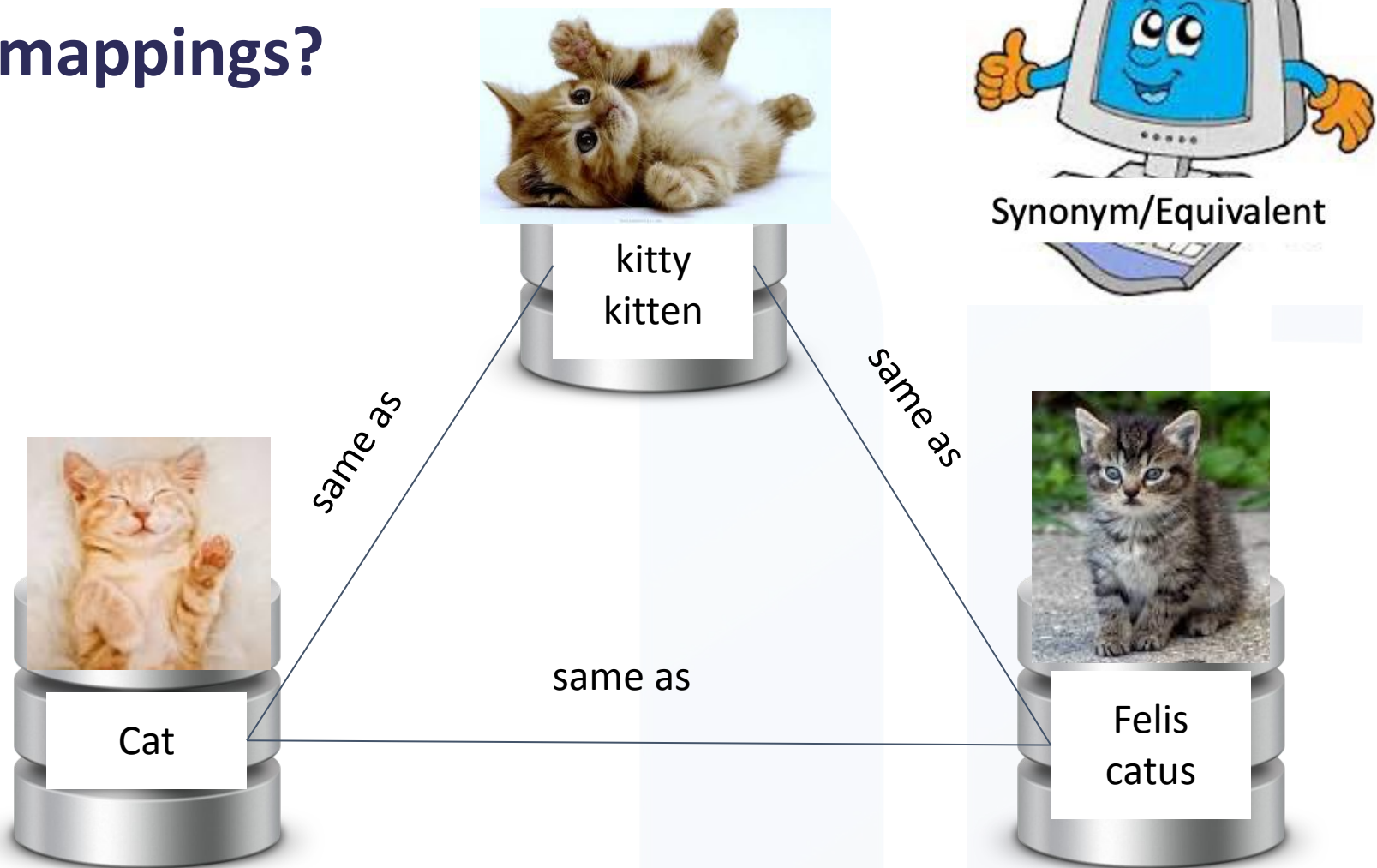


Synonym/Equivalent



Felis
catus

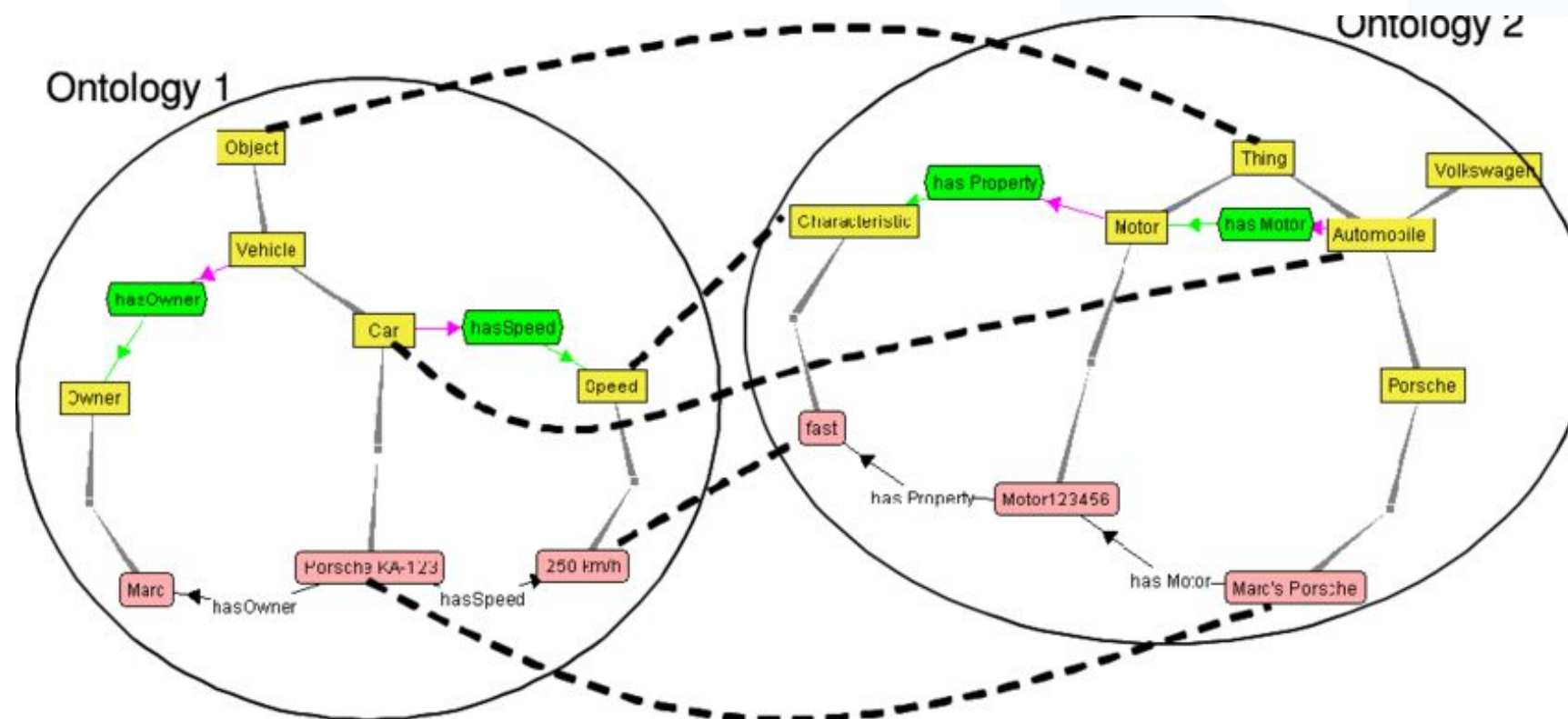
When do we need mappings?



When do we need mappings?

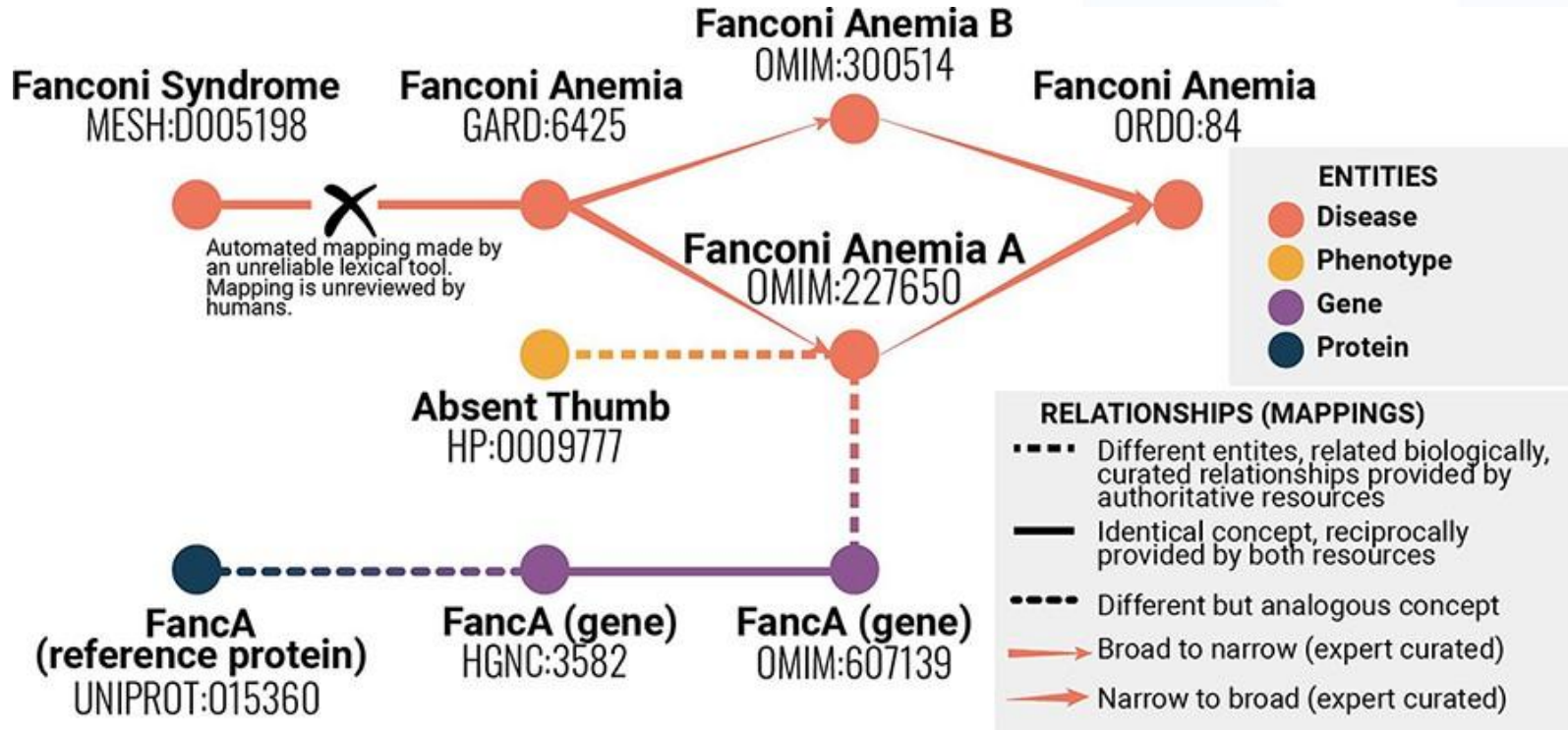
- Needed to integrate different information sources either as one information system or for enabling federated search
- Part of the ontology development process: reducing ambiguities in information systems

Mapping examples



From Ehrig, Marc & Staab, Steffen. (2004). QOM – Quick ontology mapping. 3298. 356-361. 10.1007/978-3-540-30475-3_47.

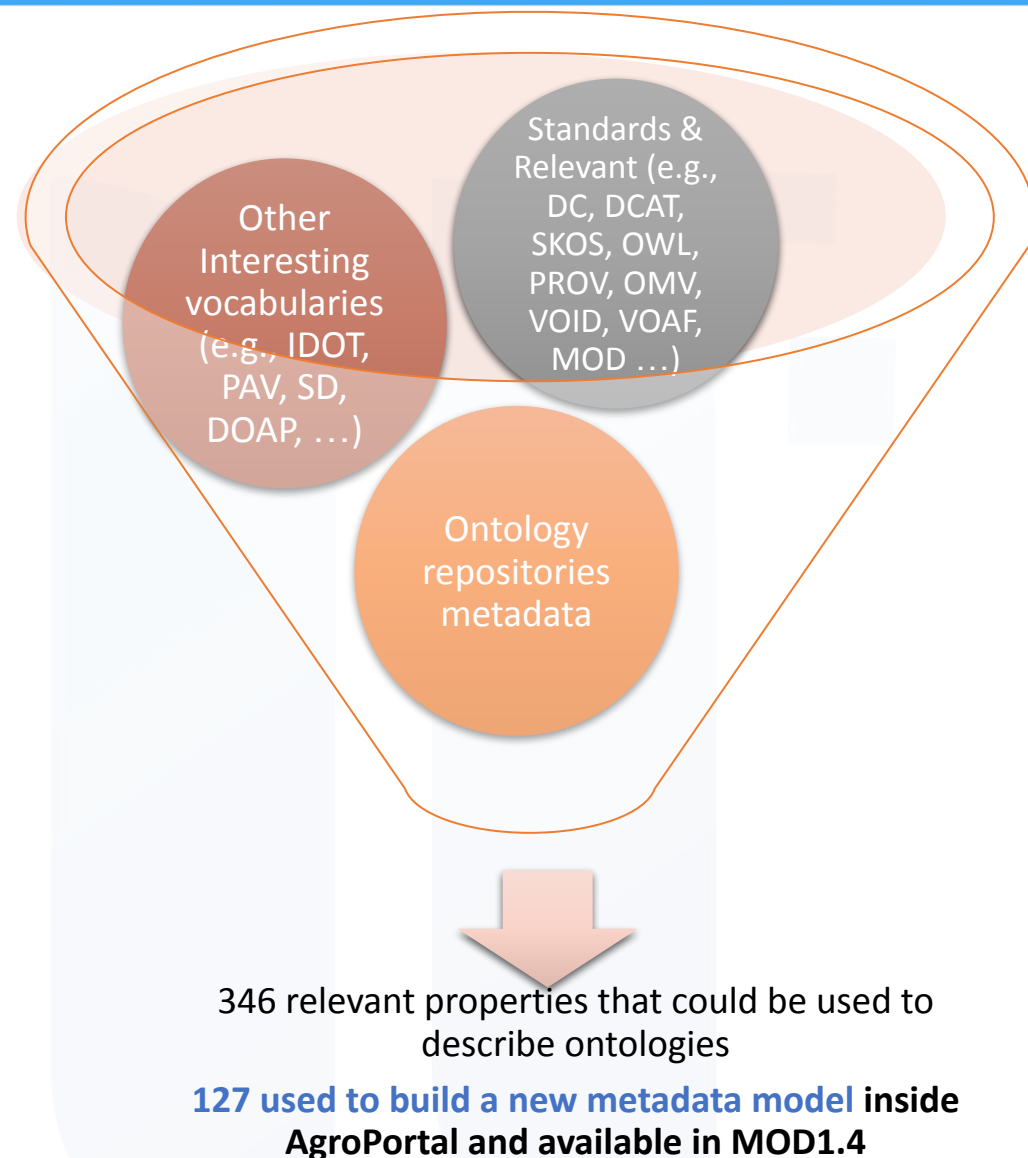
Mapping examples



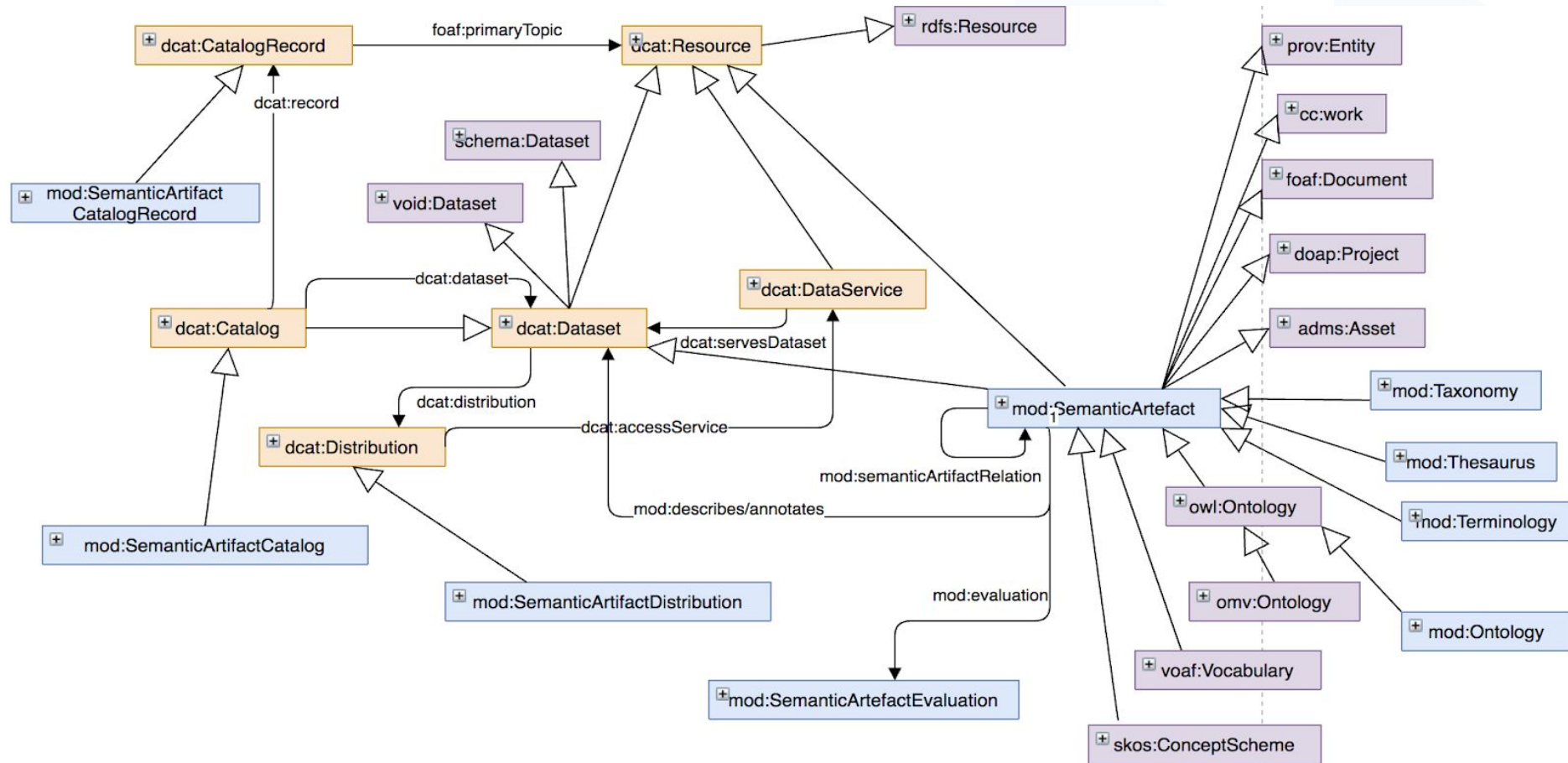
From N. Matentzoglou et al., "A Simple Standard for Sharing Ontological Mappings (SSSOM)", *Database*, Volume 2022, 2022, baac035, <https://doi.org/10.1093/database/baac035>

Review of ontology metadata practices to produce MOD

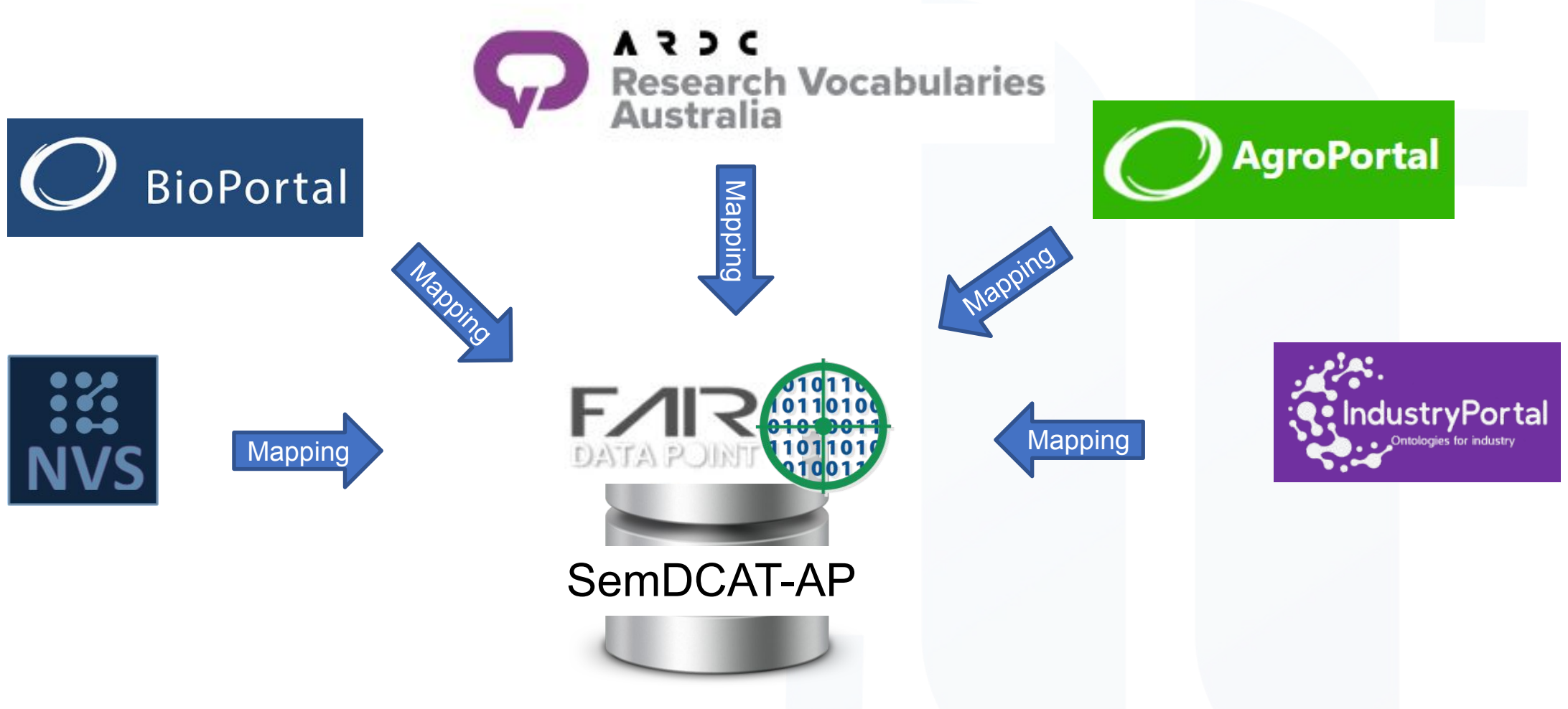
- Analysis of the **use of metadata vocabularies** in describing ontologies (by ontology developers)
 - 805 ontologies analyzed
- Analysis of the existing **metadata vocabularies**
 - 23 metadata vocabularies
- Analysis of the **uses of metadata vocabularies in various ontology libraries and repositories** (e.g., BioPortal, MMI, LOV, OBO Foundry)
 - 13 libraries



SemDCAT-AP: minimum metadata profile



Mapping ontology repositories to SemDCAT-AP



CodeMeta Project

The CodeMeta Project [HOME](#) [CROSSWALK](#) [CREATE](#) [TERMS](#)



 [Crosswalk for WikiData Properties](#)

 [Crosswalk for DataCite metadata](#)

 [Crosswalk for Debian packages](#)

 [Crosswalk for DOAP Ontology](#)

 [Crosswalk for GitHub API](#)

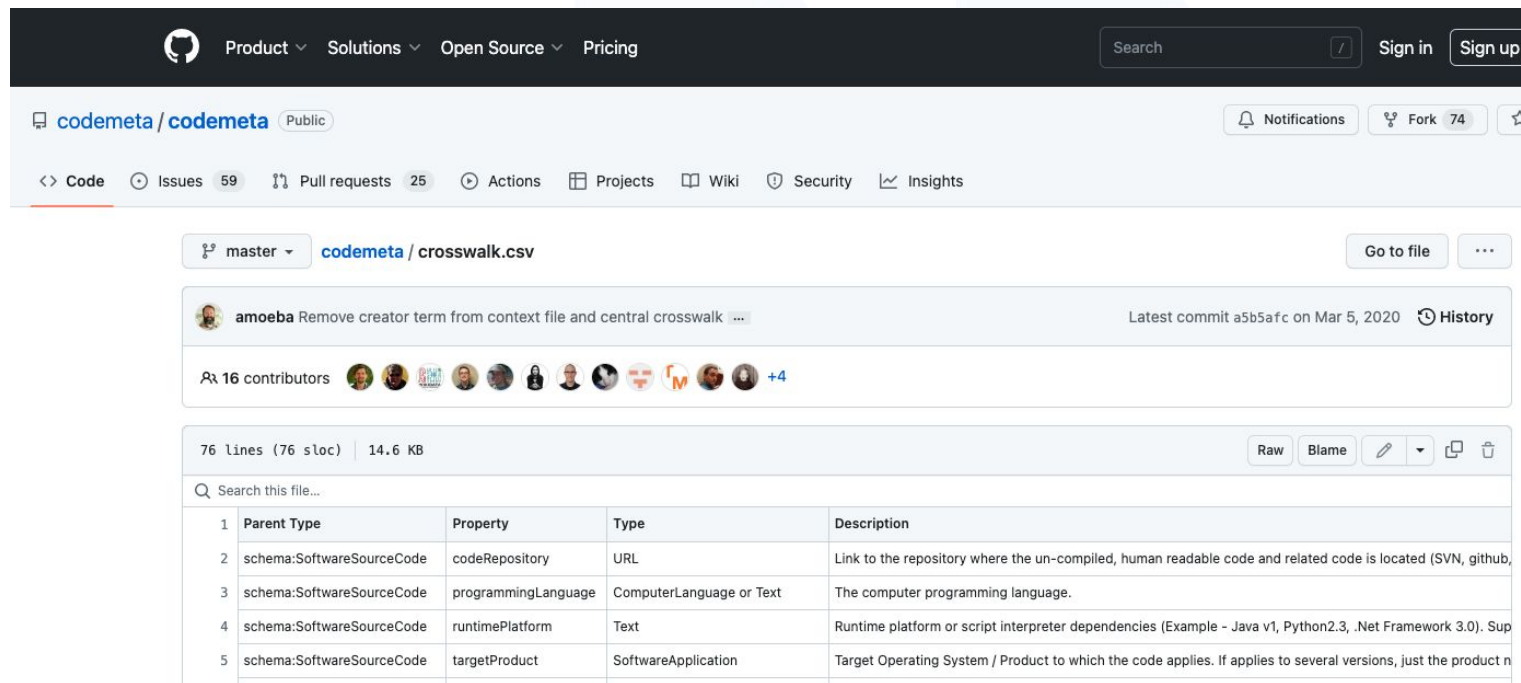
 [Crosswalk for Java's Maven metadata](#)

 [Crosswalk for NodeJS package.json](#)

 [Crosswalk for Python distutils](#)

 [Crosswalk for R Packages](#)

 [Crosswalk for Ruby gems](#)



Product Solutions Open Source Pricing Search Sign in Sign up

codemeta / codemeta Public Notifications Fork 74

Code Issues 59 Pull requests 25 Actions Projects Wiki Security Insights

master codemeta / crosswalk.csv Go to file

amoeba Remove creator term from context file and central crosswalk ... Latest commit a5b5af c on Mar 5, 2020 History

16 contributors +4

76 lines (76 sloc) 14.6 KB Raw Blame

	Parent Type	Property	Type	Description
2	schema:SoftwareSourceCode	codeRepository	URL	Link to the repository where the un-compiled, human readable code and related code is located (SVN, github,
3	schema:SoftwareSourceCode	programmingLanguage	ComputerLanguage or Text	The computer programming language.
4	schema:SoftwareSourceCode	runtimePlatform	Text	Runtime platform or script interpreter dependencies (Example - Java v1, Python2.3, .Net Framework 3.0). Sup
5	schema:SoftwareSourceCode	targetProduct	SoftwareApplication	Target Operating System / Product to which the code applies. If applies to several versions, just the product n



How are mapping stored?

- Within the ontology or as machine readable format
- In tabular format for simple correspondances shared on github.
- In research papers

Can I reuse mappings made by others?

- In few exemption, yes (e.g. Biomedical,...)
- Most of the time, no
 - Hard to find (need to know a guy who knows a guy who knows where to find them)
 - In some cases, not published
 - Not interoperable: relation is often not explicit and format varies.
 - Not reusable
 - No context explaining the reason for the mapping
 - No information about who created the mapping
 - Based on community standard practice and format requiring additional mappings (mappings of mappings)

Cost/Benefits of FAIR Mappings

COST	BENEFIT
<ul style="list-style-type: none">• Limited scope of the mapping• No documentation associated with the mappings• No means to assess the quality of mapping• No possibility to automate mapping process: manual burden.• Cannot reuse then redo	<ul style="list-style-type: none">• Add value to the effort of mappings• Support reusability instead of duplication• Mapping can be reused in other context (extend scope)• Reduction of mappings (deduplication,...) = reduction of the computing power necessary to use the mappings• Allow automation of information retrieval

Cost/Benefits of FAIR Mappings

Mapping now



Cost/Benefits of FAIR Mappings

Mapping now



FAIR
Principles



With FAIR Mapping?



Let's discuss and share about
mappings !!



@fairimpact_eu /company/fair-impact-eu-project



Funded by
the European Union