FAIR Implementation Profile

Barbara Magagna

20 January 2025

IR0000032 – ITINERIS, Italian Integrated Environmental Research Infrastructures System
(D.D. n. 130/2022 - CUP B53C22002150006) Funded by EU - Next Generation EU PNRR-Mission 4 “Education and Research” - Component 2: “From research to business” - Investment 3.1: “Fund for the realisation of an integrated system of research and innovation infrastructures”

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani

Agenda

- 13:00-13.30 (30 minutes) Welcome and recap (Erik)
- 13:30-14:00 (30 minutes) From FAIR Principles to resource types (breakout)
- 14:00-14:10 (10 minutes) Breakout presentation
- 14:10-14:45 (45 minutes) FER & FSR Typology (Barbara)
- 14:45-15:00 (15 minutes) Break
- 15:00-15:40 (40 minutes) FER types I (Barbara)
- 15:40-16:00 (20 minutes) FER types II (Sander)
- 16:00-16:10 (15 minutes) Break
- 16:10-16:50 (40 minutes) Explore and create FERs (breakout)
- 16:50-17:00 (10 minutes) Q&A
- 17:00 Adjourn

OSF Project Space
Common notes

From FAIR Principles to resource types

13:30-14:00
(breakout)

FAIR Principles

Box 2 | The FAIR Guiding Principles

To be Findable:
F1. (meta)data are assigned a globally unique and persistent identifier
F2. data are described with rich metadata (defined by R1 below)
F3. metadata clearly and explicitly include the identifier of the data it describes
F4. (meta)data are registered or indexed in a searchable resource

To be Accessible:
A1. (meta)data are retrievable by their identifier using a standardized communications protocol
A1.1 the protocol is open, free, and universally implementable
A1.2 the protocol allows for an authentication and authorization procedure, where necessary
A2. metadata are accessible, even when the data are no longer available

To be Interoperable:
I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
I2. (meta)data use vocabularies that follow FAIR principles
I3. (meta)data include qualified references to other (meta)data

To be Reusable:
R1. meta(data) are richly described with a plurality of accurate and relevant attributes
R1.1. (meta)data are released with a clear and accessible data usage license
R1.2. (meta)data are associated with detailed provenance
R1.3. (meta)data meet domain-relevant community standards

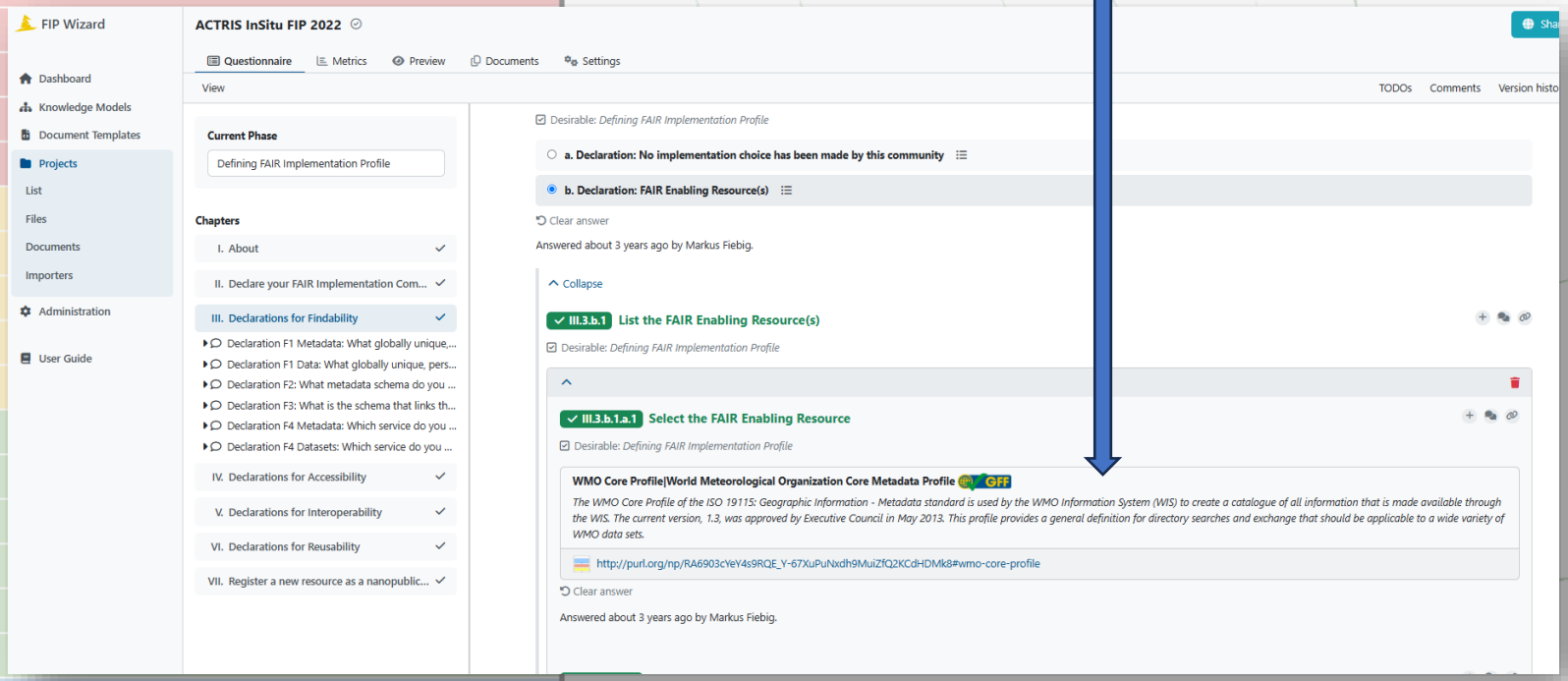
FIP questions

FAIR principle	FIP Question
F1	What globally unique, persistent, resolvable identifiers do you use for metadata records?
F1	What globally unique, persistent, resolvable identifiers do you use for datasets?
F2	Which metadata schemas do you use for findability?
F3	What is the technology that links the persistent identifiers of your data to the metadata description?
F4	In which search engines are your metadata records indexed?
F4	In which search engines are your datasets indexed?
A1.1	Which standardized communication protocol do you use for metadata records?
A1.1	Which standardized communication protocol do you use for datasets?
A1.2	Which authentication & authorisation technique do you use for metadata records?
A1.2	Which authentication & authorisation technique do you use for datasets?
A2	Which metadata longevity plan do you use?
I1	Which knowledge representation languages (allowing machine interoperation) do you use for metadata records?
I1	Which knowledge representation languages (allowing machine interoperation) do you use for datasets?
I2	Which structured vocabularies do you use to annotate your metadata records?
I2	Which structured vocabularies do you use to encode your datasets?
I3	Which models, schema(s) do you use for your metadata records?
I3	Which models, schema(s) do you use for your datasets?
R1.1	Which usage license do you use for your metadata records?
R1.1	Which usage license do you use for your datasets?
R1.2	Which metadata schemas do you use for describing the provenance of your metadata records?
R1.2	Which metadata schemas do you use for describing the provenance of your datasets?
R1.3	Who is the community, and what are their domain-relevant community standards?

FIP questions

FAIR principle	Resource Instances
F1	
F1	
F2	
F3	
F4	
F4	
A1.1	
A1.1	
A1.2	
A1.2	
A2	
I1	
I1	
I2	
I2	
I3	
I3	
R1.1	
R1.1	
R1.2	
R1.2	
R1.3	

Find instances im FIP Wizard



Derive types from instances

FAIR principle	Resource Instances
F1	
F2	
F3	
F4	
A1.1	
A1.2	
A2	
I1	
I2	
I3	
R1.1	
R1.2	
R1.3	

FAIR principle	Resource Types
F1	
F2	
F3	
F4	
A1.1	
A1.2	
A2	
I1	
I2	
I3	
R1.1	
R1.2	
R1.3	

FER and FSR typology

13:30-14:00
(breakout)

Definitions

All types are defined in the FIP Ontology and more details can be found in the FIP wiki.

FAIR Enabling Resources (FER)

A service, a specification or a data policy that is **essential to the operationalization of the FAIR Principles**, i.e., puts FAIR into action. A FAIR-Enabling Resource (FER) provides a function needed to achieve some aspect of FAIR behavior and is explicitly linked to one or more FAIR principles.

FAIR Supporting Resources (FSR)

Any resource that **supports** FAIR Data Stewardship. FSRs are represented as FAIR Digital Objects (using the nanopublication framework) with Globally Unique, Persistent, Resolvable Identifiers (GUPRI) that resolve to machine-readable metadata about the resource.

A FER is a subclass of a FSR!

FER types

FAIR principle	FER Type	Definition
F1	*Identifier Service*	A **service** that provides for any digital object (1) algorithms guaranteeing global uniqueness, (2) policy document that guarantees persistent and (3) resolution of the identifier to machine-actionable metadata describing the object and its location.
F2	*Metadata Schema*	A **specification** that specifies the structured representation of metadata describing attributes of data or other digital objects in terms of semantics, syntax and optionality.
F3	*Metadata-Data Linking Schema*	A **specification** that provides a unique, persistent, (ideally) bi-directional, machine-actionable link between metadata and the data they describe.
F4	*Registry*	A **service** that indexes metadata and data and provides search over that index.
A1.1	*Communication Protocol*	A **specification** of how messages are structured and exchanged.
A1.2	*Authentication & Authorisation Service*	A **service** that mediates access to digital objects according to specifed conditions.
A2	*Metadata Preservation Policy*	A data **policy** that describes the conditions under which metadata should be provided in the future.
I1	*Knowledge Representation Language*	A language **specification** that enables knowledge to be processed by machines.
I2	*Structured Vocabularies*	A **specification** for a controlled list of uniquely identified and unambigous concepts with their definitions represented using web standards.
I3	*Semantic Model*	A **specification** that defines qualified relations between entities describing data or other digital objects according to the Linked Data principles. This can include semantic data models and ontologies.
R1.1	*Data Usage License*	A data **policy** that specifies legal restrictions on the reuse of the data.
R1.2	*Provenance Model*	A **specification** that specifies metadata describing the origin and lineage of data or other digital objects.
R1.3	*The FAIR Implementation Profile*	A **FAIR Implementation Profile (FIP)** is a list of declared technology choices intended to implement each of the FAIR Guiding Principles, made as a collective decision by the members of a particular community of practice.

FER types

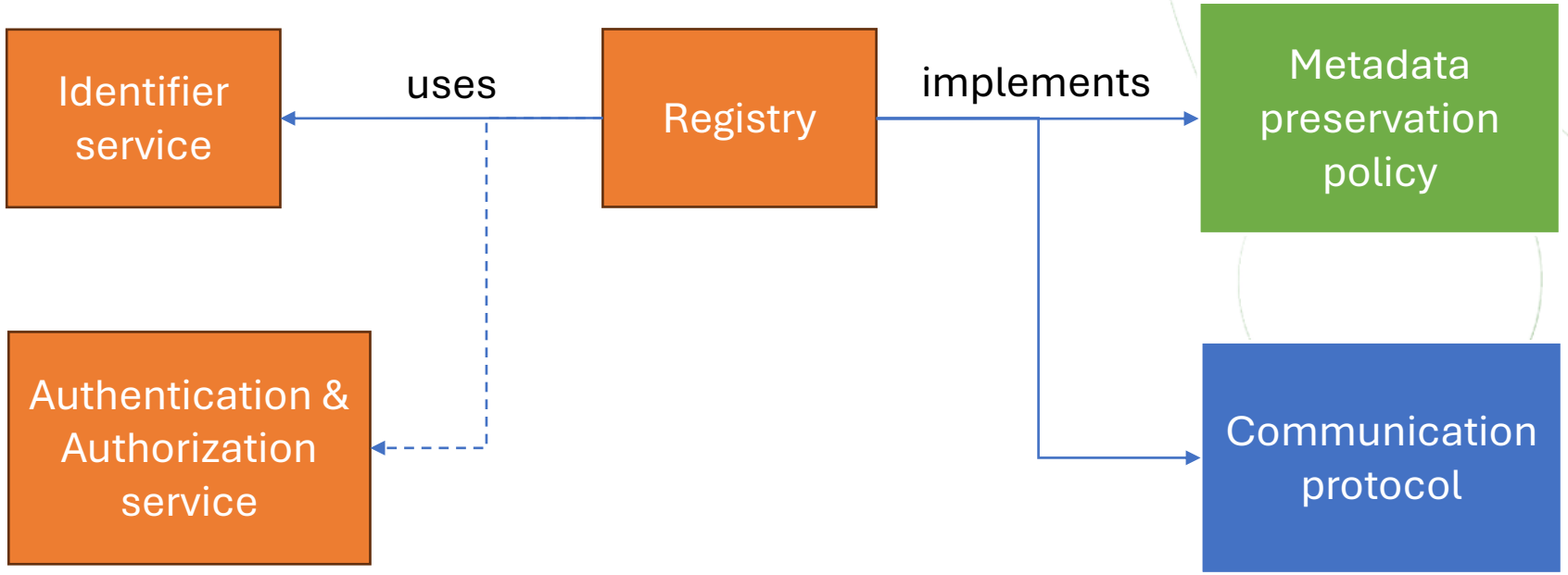
FAIR principle	FER Type	Definition
F1	*Identifier Service*	A **service** that provides for any digital object (1) algorithms guaranteeing global uniqueness, (2) policy document that guarantees persistent and (3) resolution of the identifier to machine-actionable metadata describing the object and its location.
F2	*Metadata Schema*	A **specification** that specifies the structured representation of metadata describing attributes of data or other digital objects in terms of semantics, syntax and optionality.
F3	*Metadata-Data Linking Schema*	A **specification** that provides a unique, persistent, (ideally) bi-directio…etadata and the data they describe.
F4	*Registry*	A **service** that indexes metadata and data and provides search over th…
A1.1	*Communication Protocol*	A **specification** of how messages are structured and exchanged.
A1.2	*Authentication & Authorisation Service*	A **service** that mediates access to digital objects according to specifec…
A2	*Metadata Preservation Policy*	A data **policy** that describes the conditions under which metadata sh…
I1	*Knowledge Representation Language*	A language **specification** that enables knowledge to be processed by machines.
I2	*Structured Vocabularies*	A **specification** for a controlled list of uniquely identified and unambigous concepts with their definitions represented using web standards.
I3	*Semantic Model*	A **specification** that defines qualified relations between entities describing data or other digital objects according to the Linked Data principles. This can include semantic data models and ontologies.
R1.1	*Data Usage License*	A data **policy** that specifies legal restrictions on the reuse of the data.
R1.2	*Provenance Model*	A **specification** that specifies metadata describing the origin and lineage of data or other digital objects.
R1.3	*The FAIR Implementation Profile*	A **FAIR Implementation Profile (FIP)** is a list of declared technology choices intended to implement each of the FAIR Guiding Principles, made as a collective decision by the members of a particular community of practice.

→ Specifications

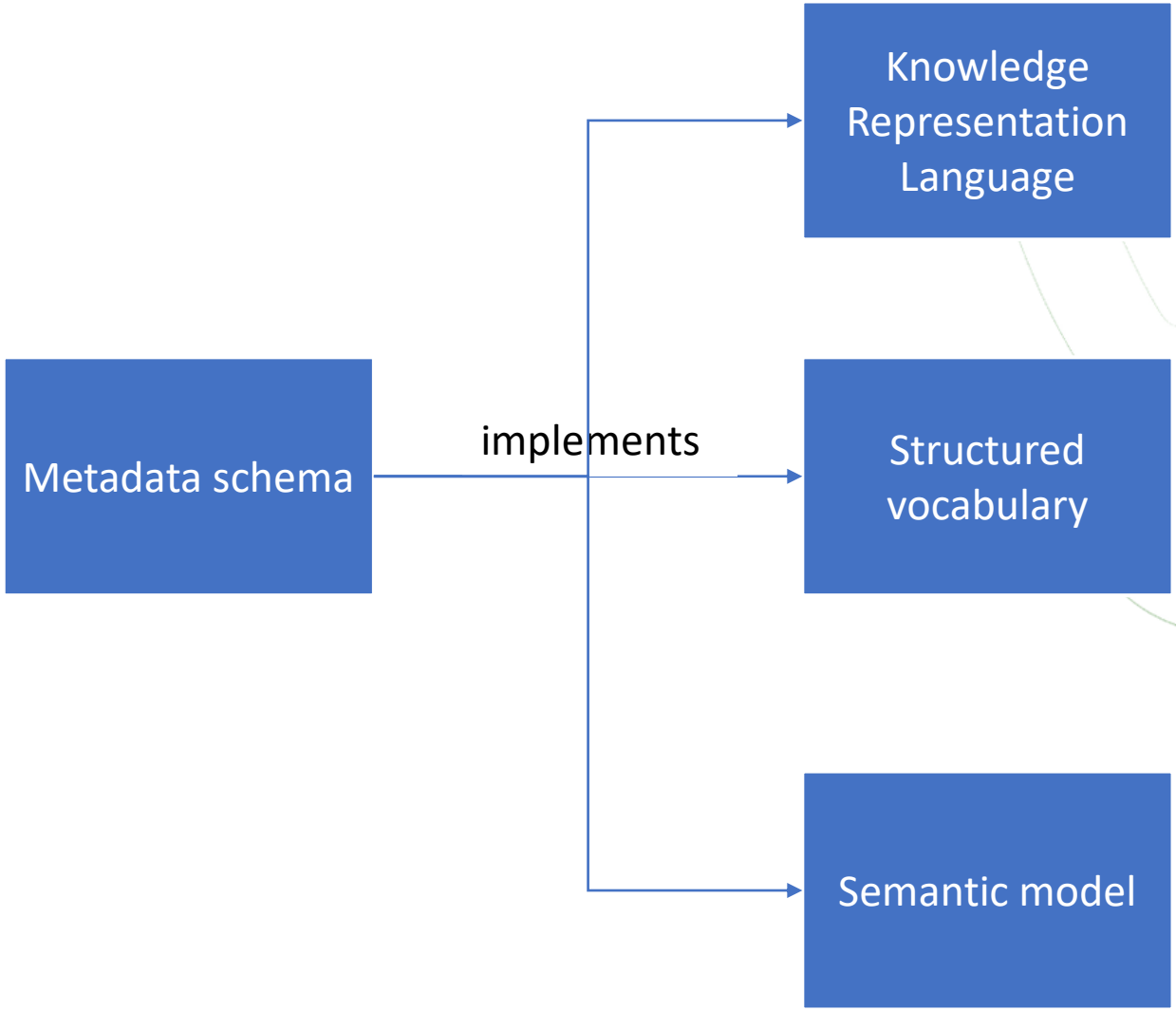
→ Services

→ Data Policies

Dependencies between FER types



Dependencies between FER types



A FER instance

Required feature:

- A minimal degree of FAIRness
- Enables FAIRification in some aspects

FERs can be:

- Available or
- To be developed (in development)

Available FERs can be:

- currently used by the community
- planned to be used in future
- planned to be replaced

FERs in development can be:

- planned to be used in future

A FER instance

Some **instances** of the type **specification** can be of **multiple subtypes**:

E.g. a metadata schema can also be a structured vocabulary, a semantic model and a provenance when it includes concepts with definitions organized in an ontology which includes elements that describe the provenance of data.

FERs are spread over the web and normally not well described and identifyable.

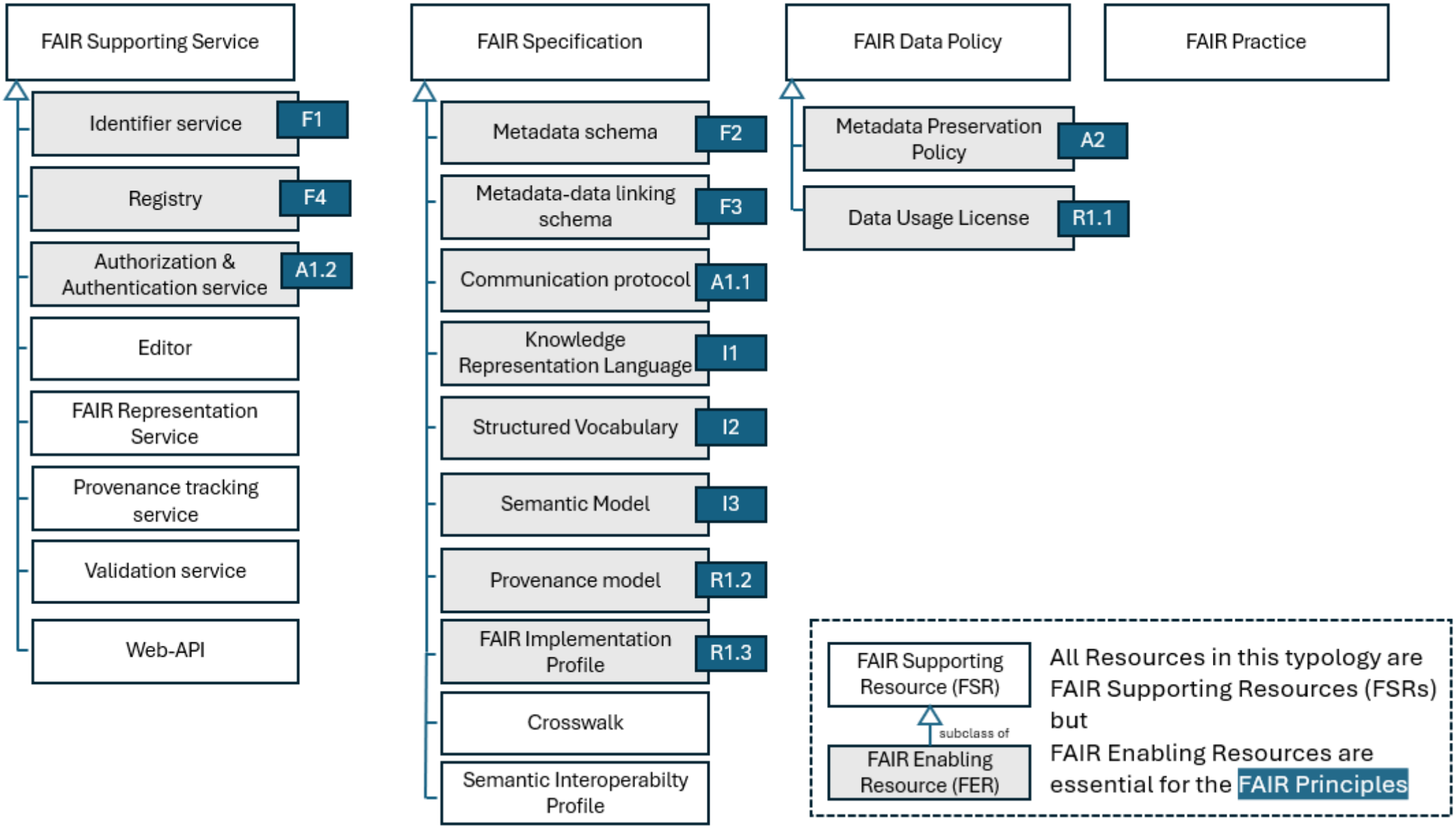
FAIRsharing provides metadata for many FER instance but not for all types (as it focuses on registries and standards). We borrow metadata from FAIRsharing where they exist.

Within the **FIP environment** we provide:

- Persistent identifiers to reference them
- Minimal metadata to describe them
- Machine actionable metadata using nanopublications

Metadata for FERs can be created in the FIP Wizard or in nanodash and are.

FAIR Supporting Resource Types



Creating FERs in the FIP Wizard

VII. Register a new resource as a nanopublication

Why nanopublications?

If your resource is not already referenceable as a nanopublication you can create it here in the Wizard environment. The new nanopublication you cr
(in the form of a Persistent URL or PURL) and a minimal metadata description. The nanopublication will allow your resource to be retrievable by the FI
question. As a nanopublication, your resource will also be considered by the GO FAIR Foundation for qualification assessment with respect to it
Nevertheless you can already use your newly minted nanopublication in the FIP Wizard or elsewhere.

Create new nanopublications

ncing your **FAIR Implementation Community (FIC)** please use this template

ncing your **FAIR-Enabling Resource (FER)** with metadata from a FAIRsharing record please use this template

ncing your **FAIR-Enabling Resource (FER)** please use this template

Create Project

Name

From project template | From knowledge model

Project templates are prepared projects with knowledge models, question tags, and document templates setup, so you don't have to start from scratch.

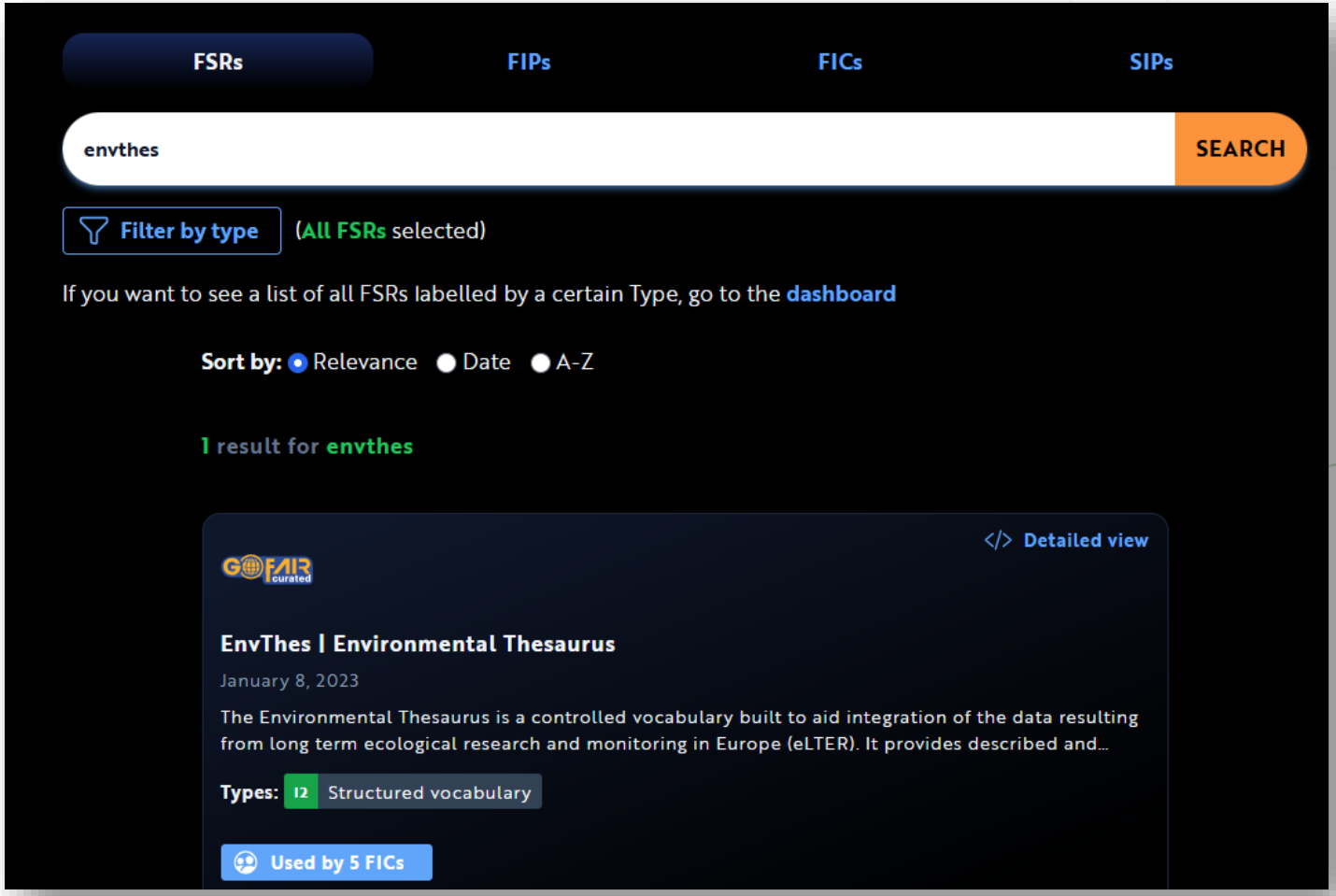
Project Template

F **FAIR Enabling Resource**
Define a new FAIR Enabling Resource ×

Cancel | Create

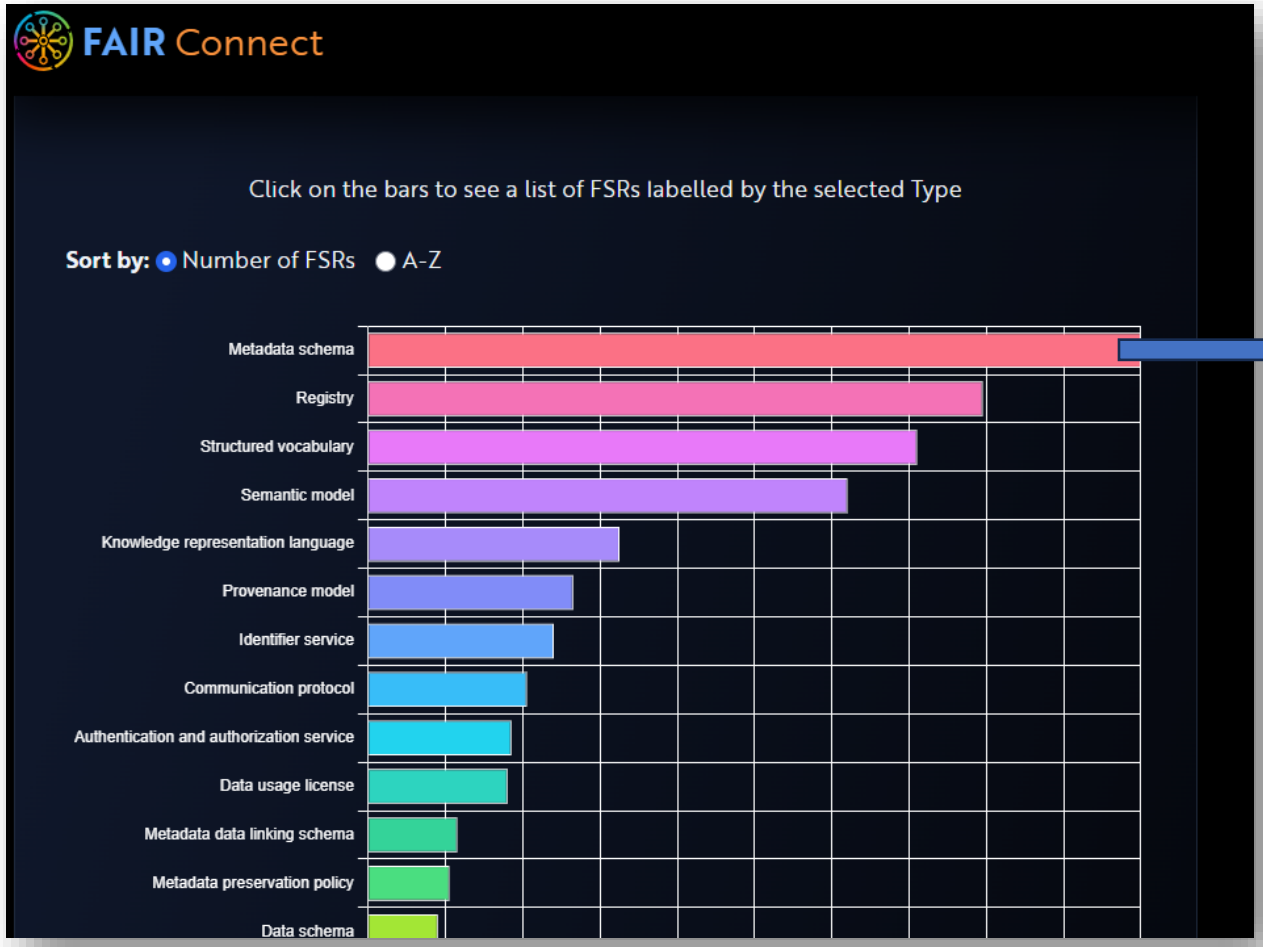
Finding FERs via FAIR Connect search

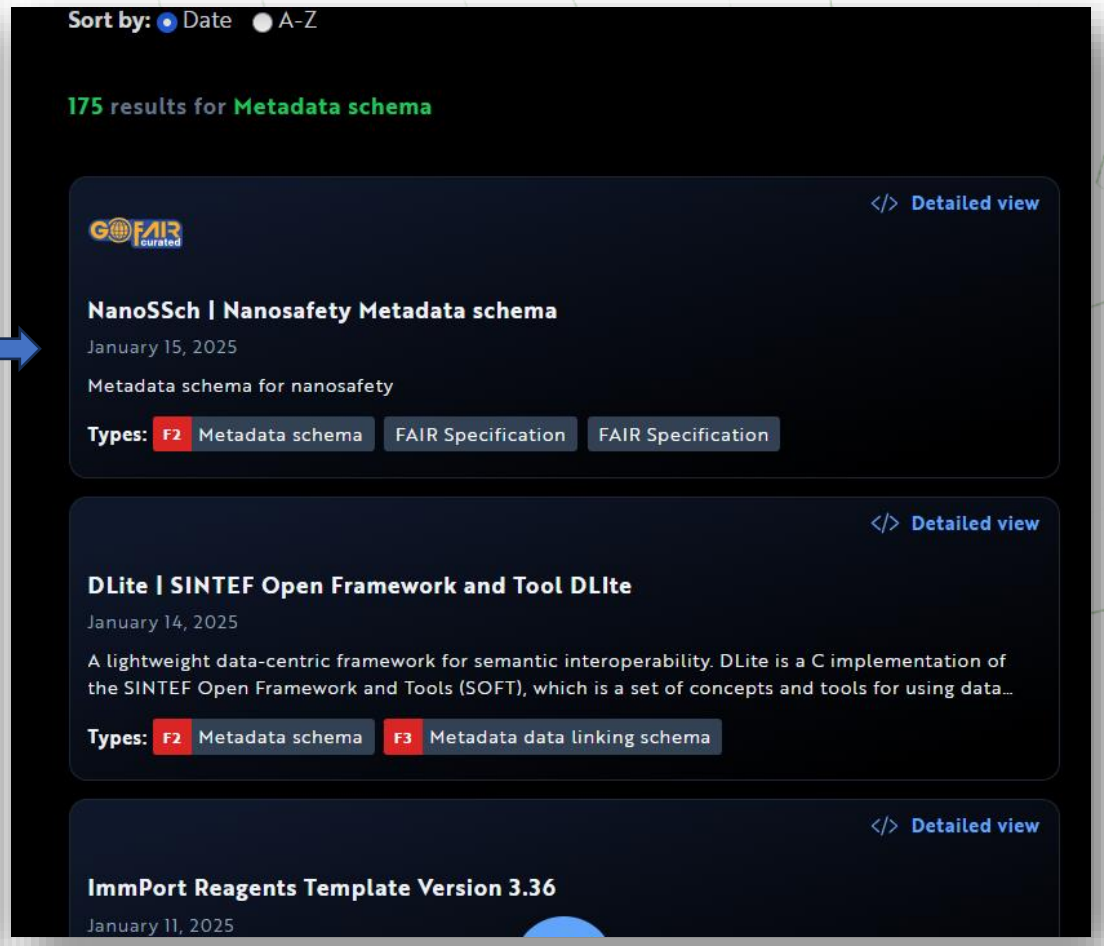
https://fairconnect.pro/search-fair-nanopublications/



Finding FERs via FAIR Connect dashboard

https://fairconnect.pro/dashboard/





FER Types I: Specifications & Policies

15:00-15:40
(Barbara)

Descriptions for FSR include following basic metadata

Required metadata:

- typeOf: one of the 12 FER types
- typeOf: avaialble FER or FER to be developed
- label: a preferred name
- comment: a description

Recommended metadata, when available:

- seeAlso: a link to a website with further information
- url: an URL where you can access the resource

Optional metadata:

- implements: a specification that this resource implements
- exactMatch (or alternatively, close or related match): an identifier of the resource to wich it is mappable.

Some types require additional metadata.

FAIR Specification

FAIR Enabling Resources

- Metadata schema
- Metadata-data linking schema
- Communication protocol
- Knowledge representation language
- Structured vocabulary
- Provenance model
- FAIR Implementation Profile

FAIR Supporting Resources

- Crosswalk
- Semantic Interoperability Profile

FAIR Specification metadata

The template used for all FAIR specifications includes all basic metadata and optionally:

- created by: FAIR Implementation Community or any other creator referenced via a ROR identifier or an ORCID
- registered on: Registry where the resource is indexed.

Metadata schema

An metadata schema is a FAIR Enabling Resource that enables FAIR principle F2.
IRI: https://w3id.org/fair/fip/terms/Metadata-schema

Definition: A specification that specifies the structured representation of metadata describing attributes of data or other digital objects in terms of semantics, syntax and optionality.

A metadata schema must not be a FAIR metadata schema to be a FAIR Enabling Resource. A resource that is at the same time a metadata schema that is based on a semantic model and uses a FAIR structured vocabulary is a FAIR metadata schema.

Metadata-Data Linking schema

A Metadata-Data Linking Schema is a FAIR Enabling Resource that enables **FAIR principle F3**.

IRI: https://w3id.org/fair/fip/terms/Metadata-data-linking-schema

Definition: A **specification** that provides a unique, persistent, (ideally) bi-directional, machine-actionable link between metadata and the data they describe.

Both data and metadata need to be referenced by their persistent identifiers and the specification should provide a relation between these two identifiers.

Communication Protocol

A communication protocol is a FAIR Enabling Resource that enables **FAIR principle A1.1**.

IRI: https://w3id.org/fair/fip/terms/Communication-protocol

Definition: A **specification** of how messages are structured and exchanged.

Knowledge Representation Language

A knowledge representation language is a FAIR Enabling Resource that enables FAIR principle I1.

IRI: https://w3id.org/fair/fip/terms/Knowledge-representation-language

Definition: A language specification that enables knowledge to be processed by machines.

It is important to note, that a resource needs to have some degree of FAIRness to be accepted as a FER. For knowledge representation languages, this means that the language should be understood by machines, and thus, not all formats can be approved.

Structured vocabulary

A Structured vocabulary is a FAIR Enabling Resource that enables FAIR principle **I2**.

IRI: https://w3id.org/fair/fip/terms/Structured-vocabulary

Definition: A specification for a controlled list of uniquely identified and unambigous concepts with their definitions represented using web standards.

A synonym for structured vocabulary is a controlled vocabulary but the principle I2 asks for the use of FAIR vocabularies. The knowledge representation language implemented is mostly RDF, more specifically SKOS. Vocabularies include thesauri and taxonomies.

GFF has decided to accept for this type a resource with some degree of FAIRness. For structured vocabularies, this means that a controlled list that is available as a PDF will not be accepted unless it is classified as an FER to be developed. If your community wants to FAIRify it, you can define it as a FER to be developed. Otherwise you can add this information in the considerations box and answer with NO when asked to add a FER.

Semantic Model

A Semantic model is a FAIR Enabling Resource that enables **FAIR principle I3**.

IRI: https://w3id.org/fair/fip/terms/Semantic-model

Definition: A **specification** that defines qualified relations between entities describing data or other digital objects according to the Linked Data principles.

Semantic models include semantic data models and ontologies.

Semantic artefact is used as an umbrella term for vocabularies, thesauri, data models and ontologies.

Provenance Model

A Provenance Model is a FAIR Enabling Resource that enables **FAIR principle R1.2**.

IRI: https://w3id.org/fair/fip/terms/Provenance-model

Definition: A **specification** that specifies metadata describing the origin and lineage of data or other digital objects.

FAIR Implementation Profile

Definition: A FAIR Implementation Profile (FIP) is a list of declared technology choices intended to implement each of the FAIR Guiding Principles, made as a collective decision by the members of a particular community of practice.

A FIP enables FAIR Subprinciple R1.3

A FIP can have different purposes:

- AsIsFIP: to express the status quo of the FAIR Enabling Resources used a FAIR Implementation Community;
- ToBeFIP or also reference FIP: to express the FAIR Enabling Resources to be used by a FAIR Implementation Community;
- FIP: to express the status quo and the planned FAIR Enabling Resources to be used by a FAIR Implementation Community.

A FIP can be defined at different levels:

- for all resources used by a FAIR Implementation Community
- for all resources used in a repository
- for all resources used for a specific digital object type
- for all resources used in a specific case study or project

Semantic Interoperability Profile

IRI: https://w3id.org/fair/fip/terms/Semantic-Interoperability-Profile

Definition: A Semantic Interoperability Profil (SIP) is a list of FAIR Supporting Resources (FSRs) chosen by a community to support semantic interoperability of (meta)data. It includes the semantic artefacts and their supporting services chosen by a community for a specific case study and type of data.

A SIP does not include all FER types, only those related to interoperability, on the other hand it includes resources that support those interoperability FERs:

- FAIR-Practice

FAIR Specification:

- Metadata schema
- Knowledge representation language
- Structured vocabulary
- Semantic model
- Provenance Model
- Crosswalk

FAIR Supporting Service:

- Registry
- Editor
- Web-API
- Validation service
- Provenance tracking service
- FAIR representation service

Crosswalk

IRI: https://w3id.org/fair/fip/terms/Crosswalk

Definition: A specification consisting of a set of rules that define how (meta)data elements or attributes from one schema can be aligned and mapped to (meta)data elements or attributes in another schema that share the same constraints and thus share the same semantic role.

Additional metadata:

has purpose:

- data model
- metadata schema
- semantic model
- strucutred vocabulary

has input: URL of input resource

has output: URL of output resource

FAIR Data Policy

- Metadata Preservation Policy
- Data Usage License

Metadata Preservation Policy

A Metadata Preservation Policy is a FAIR Enabling Resource that enables FAIR principle **A2**.

IRI: https://w3id.org/fair/fip/terms/Metadata-preservation-policy

Definition: A data policy that describes the conditions under which metadata should be provided in the future.

Additional metadata:

- metadata preservation in years: number of years
- has data policy document: link to data policy document

Data Usage License

A Data usage license is a FAIR Enabling Resource that enables FAIR principle **R1.1**.

IRI: https://w3id.org/fair/fip/terms/Data-usage-license

Definition: A data policy that specifies legal restrictions on the reuse of the data.