

# Semantic Artefacts Alignment for Improving Interoperability in Astronomy

## Overview

The astronomy community is structured in several sub-communities, with matured but siloed semantic artefact ecosystems. **This use case brings all semantic artefacts in the same catalogue.** The goal is to improve the semantic interoperability between the astronomy communities, and in turn the semantic artefact FAIRness.

The astronomy community is composed of three main semantics sub-communities:

- **celestial astronomy** (objects are referenced to with their sky coordinates, e.g., stars, galaxies, etc);
- **planetary sciences** (the study of the Solar System objects, e.g., planets, comets, asteroids, etc);
- **heliophysics** (the study of the Sun, the plasma environments throughout the Solar System).

Each of these sub-communities have developed interoperability and semantic ecosystems, which are rather siloed up to now.

## Context and objectives

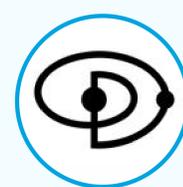
The sky astronomy community is organized around the IVOA, **which maintains an interoperability operating framework used by data repositories and scientific application platforms around the world.** In this community, semantic artifacts (SA) are composed of terms (vocabularies) and schemas (data models). The IVOA Semantics Working Group **manages the vocabularies used in the IVOA standards**, which are located on a dedicated web page (<https://ivoa.net/rdf>) and can be accessed with IVOA or RDF tools. The planetary science community is less organized than the astronomical community. Two main frameworks with different purposes co-exist: **the IPDA** (<https://planetarydata.org>), which proposes an advanced information model of data storage for planetary exploration datasets, and **the OGC** (<https://www.ogc.org/>), used by teams studying planetary surfaces.

The heliophysics community is organized around the **IHDEA** (<https://ihdea.net>), which proposes a set of tools and standards for searching and accessing datasets in this domain. The semantic artifacts of this community have historically been of two types:

- the SPASE ontology (<https://spase-group.org>) includes a list of terms, properties and classes to define various objects (people, observers, tools, datasets, archives, etc.);
- the SOLARNET set of keywords (dedicated to solar observations).

**The IVOA, IPDA and IHDEA alliances are all worldwide, consensus-driven, bottom-up working groups based on "best-effort" contributions.** Interdisciplinary linkages between these communities have been developed through the Europlanet/VESPA project (<http://www.europlanet-vespa.eu/>), focused on discoverability and plugin implementation to extend the capabilities of existing tools. Semantic interoperability between sub-communities has recently begun with the ongoing development of two common semantic artifacts: **one vocabulary for "observation facility"** and **another for "coordinate systems."** The goal is to enable semantic interoperability among the sub-communities of astronomy at an early stage and to explore it with neighboring fields such as earth and environmental sciences or particle physics.

## Contributors



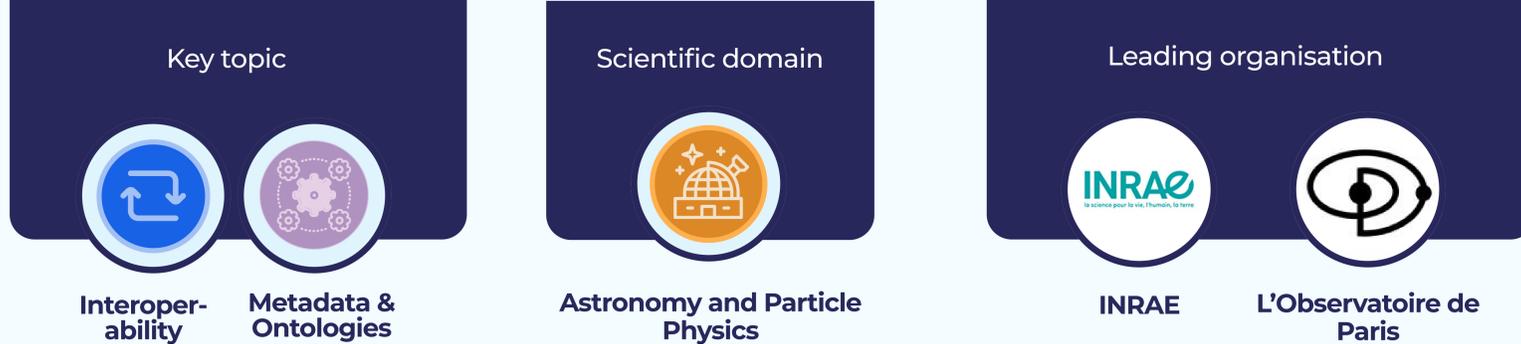
**Baptiste Cecconi**  
Observatoire de Paris



**Laura Debisschop**  
Observatoire de Paris



**Sophie Aubin**  
INRAE



## Challenges and solutions implemented

As part of FAIR-IMPACT task 4.2, an instance of OntoPortal was created with the goal of **collecting the semantic artifacts of the various astronomy sub-communities in the same place**. The astronomy ontology portal is now available and several relevant semantic artifacts (39 SA, at the time of writing) covering sky astronomy and heliophysics have been included in it.

The main challenge of this use case is to produce semantic artifacts in form (RDF, OWL, or SKOS) for inclusion in the OntoPortal instance. Most current semantic artifacts (except those in sky astronomy) are in different forms, from lists of terms in XML schemas to unformatted lists of metadata in specification documents.

This lifting will be done by the semantic working groups or authorities of the communities concerned with the support of the FAIR-IMPACT group of ObsParis.

In the IVOA context, all semantic artifacts were available on a web page and an RDF version of each SA was available on the respective target pages. IVOA vocabularies are managed according to an IVOA recommendation that defines rules and conventions, specifically how the RDF version of IVOA SAs should be designed, with a limited subset of SKOS and OWL properties. This design choice aims to limit external dependencies and ensure the sustainability of the IVOA infrastructure.

Semantic artifact management in IVOA is based on Vocabulary Enhancement Proposal (VEP), which is a process for proposing, updating, and deprecating a term. The VEP process involves a consensus-based decision after a community discussion.

In the IHDEA context, a general review of semantic artifacts has been initiated since the 2023 IHDEA meeting. The previous state was based on the SPASE information model, serialized solely by an XML schema. Lists of allowed values are embedded in the SPASE schema and require the frequent release of new versions (e.g., for each new entry in a list of allowed values). It became clear that many term lists needed to be updated and that convergence with IVOA semantic artifacts was desirable. The first work on a common semantic artifact concerns a vocabulary for “solar system reference frames,” which will be merged with the IVOA RefFrame vocabulary.

**The goal of the FAIR-IMPACT use case with the IVOA and IHDEA communities is to improve the quality of the semantic artifact, especially with regard to interoperability.** Part of the expected results is to update semantic artifact management practices (e.g., a new version of the Vocabulary in the VO recommendation).

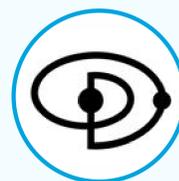
## Expected/Measured Impacts

From the perspective of interoperability alliances, the creation of the OntoPortal instance and the evaluation/preparation of semantic artifacts have been a quantitative and qualitative improvement.

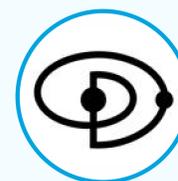
- **IVOA context:** For example, a revision of the document “Vocabularies in the VO” (<https://ivoa.net/documents/Vocabularies>) is being prepared to include a set of terms needed for the creation of SKOS-based semantic artifact catalogs in the IVOA. This sparked a discussion about reusing external semantic artifacts versus keeping things “simple” (but disconnected) by reducing external semantic dependencies.
- **IHDEA context:** Several teams have begun work on producing linked data and metadata using RDF tools. The OntoPortal instance for astronomy was an important incentive for the adoption of this framework.

From the user's perspective, the increased FAIRness of semantic artifacts will improve the FAIRness of published datasets. Semantic artifacts can be embedded in intelligent DMP tools (URIs for terms, rather than free text) or search interfaces, in turn enabling more refined queries and selections on data discovery interfaces. In this regard, **the development of the OSTRails project (<https://ostrails.eu>) for the astronomy thematic pilot will rely on the semantic artifact catalog tools developed through FAIR-IMPACT.**

## Contributors



**Baptiste Cecconi**  
Observatoire de  
Paris



**Laura Debisschop**  
Observatoire de  
Paris



**Sophie Aubine**  
INRAE