IVOA - June 2026

Observation Facilities & Observation Instruments Vocabularies: First Release, open to review

Liza FRETEL[1], Baptiste CECCONI[1], Laura DEBISSCHOP[1]

[1] LIRA, Observatoire de Paris, Université PSL, Sorbonne Université, Université Paris Cité, CY Cergy Paris Université, CNRS, 92190 Meudon, France

This work is supported by the OPAL cascading grant funded by the OSCARS project (Grant Agreement: 101129751). The authors also acknowledge support from CNRS and Observatoire de Paris.

Furthermore, the authors would like to thank the CDS (Centre de Données de Strasbourg) for their constant implication and their participation in the quality review of the data.

Summary

1. Presentation of the Observation Facilities & Instruments Vocabularies
- 2. How did we build it?
- 3. Current shortcomings and discussions around those limitations
4. Towards an Unified Vocabulary: how can the community contribute?

LIRA Observatoire de Paris | PSL

1. Presentation of the Observation Facilities & Instruments Vocabularies

OBSF : Observation Facilities

- Can host instruments
- Can be broader to other facilities (observatory broader of telescope)

Sphinx Observatory (Switzerland)

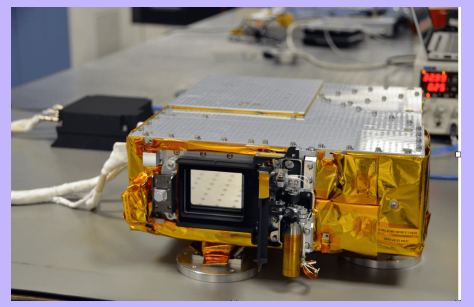
Green Bank Telescope (USA)

Hubble Space Telescope, HST

Hosted by

OBSI : Observation Instruments

- Take measures, like cameras
- Sometimes no clear distinction with facilities
- Not in this release but planned for 2nd release and further (SPASE, PDS)



MIRS infrared imaging spectrometer for Japanese MMX space mission

High Resolution Camera on Chandra

LIRA Observatoire de Paris | PSL

IVOA Vocabulary: Observations Facilities

This is the description of the vocabulary `http://www.ivoa.net/rdf/obsfacility` as of 2026-06-09.

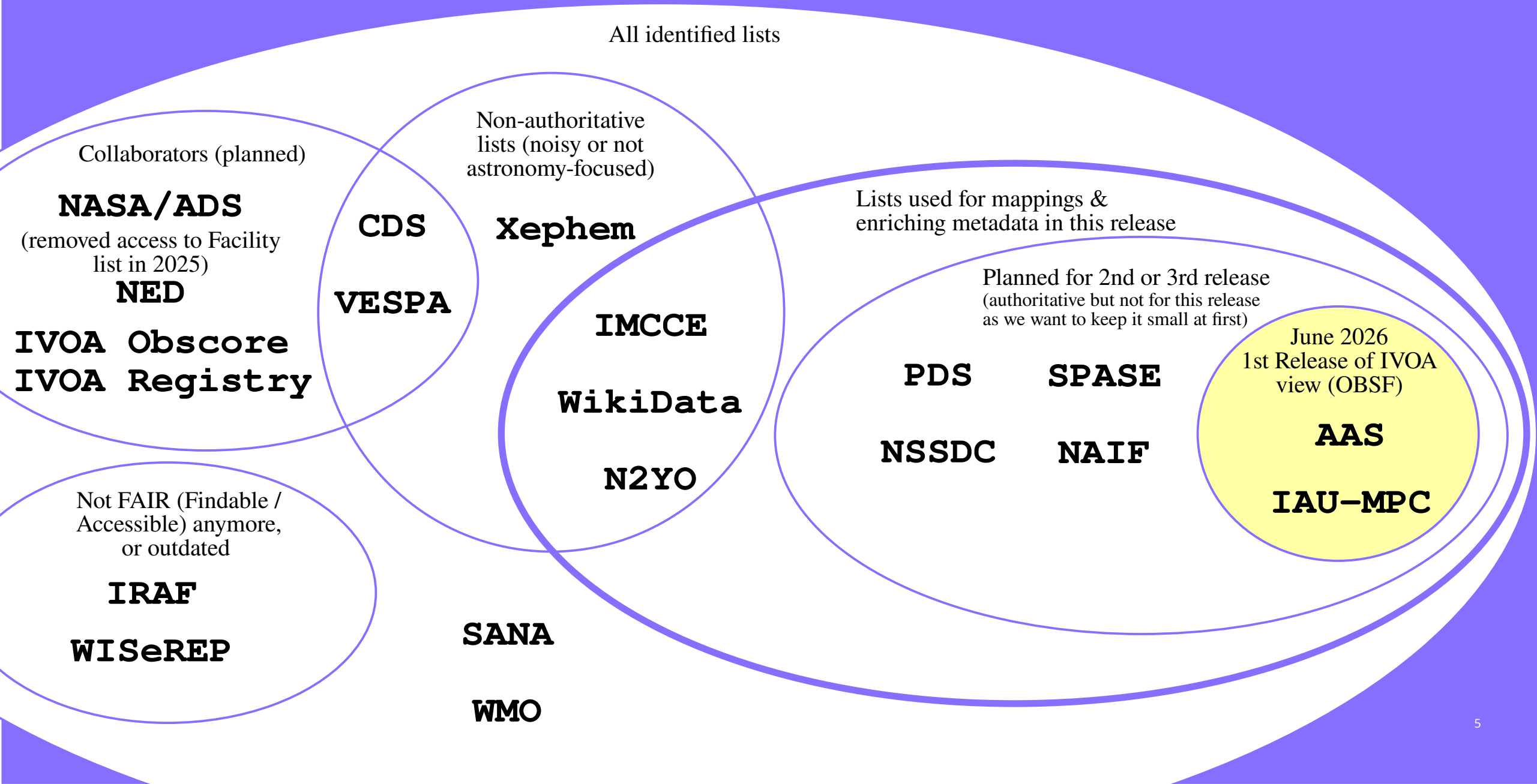
> This vocabulary is not yet approved by the IVOA. This means that terms can still disappear without prior notice.

A vocabulary giving identifiers, standard labels, and aliases for observatories, space probes, telescopes and alike.

ivoa.net/rdf/obsfacility

Term	Label	Description	Parent	More
0-15m-infrared-telescope-in-space-part-of-space-flyer-unit (Preliminary)	0.15m Infrared Telescope in Space part of Space Flyer Unit	The 0.15m Infrared Telescope in Space, also known as IRTS, was an instrument aboard the Space Flyer Unit operated by the Institute of Space and Astronautical Science (ISAS). It had an aperture of 0.15 meters.		
0-15m-robotic-bochum-twin-telescopes-at-cerro-armazones-observatory (Preliminary)	0.15m Robotic Bochum Twin Telescopes at Cerro Armazones Observatory	The 0.15m Robotic Bochum Twin Telescopes are located at the Cerro Armazones Observatory in South America and are operated by Ruhr University Bochum and Catholic University of the North (UCN).	#0-25m-berlin-exoplanet-search-telescope-ii-at-cerro-armazones-observatory	
0-15m-solar-telescope-at-catania-astrophysical-observatory (Preliminary)	0.15m Solar Telescope at Catania Astrophysical Observatory	The 0.15m Solar Telescope at Catania Astrophysical Observatory is a telescope located in Europe, specifically designed for observing the Sun with an aperture of 0.15 meters.		
0-25m-berlin-exoplanet-search-telescope-ii-at-cerro-armazones-observatory (Preliminary)	0.25m Berlin Exoplanet Search Telescope II at Cerro Armazones Observatory	The 0.25m Berlin Exoplanet Search Telescope II is part of the Cerro Armazones Observatory in Chile, which includes several telescopes managed by various institutions for astronomical research.		Narrower
0-25m-big-data-for-small-bodies-telescope-at-bucharest-astronomical-institute (Preliminary)	0.25m Big Data for Small Bodies Telescope at Bucharest Astronomical Institute	A 0.25-meter telescope at the Bucharest Astronomical Institute, designed for studying small celestial bodies with big data techniques.	#the-big-data-for-small-bodies-0-25m-telescope-at-bucharest-astronomical-institute	
0-25m-rapid-action-telescope-for-transient-objects-at-la-silla-observatory-chile (Preliminary)	0.25m Rapid Action Telescope for Transient Objects at La Silla Observatory, Chile	The 0.25m Rapid Action Telescope for Transient Objects (TAROT) at La Silla Observatory, Chile, is designed to quickly observe transient celestial events with an aperture of 0.25 meters.	#la-silla-observatory-chile	
0-25m-telescopea-action-rapide-pour-les-objets-transitoires-rapid-action-telescope-for-transient-objects-at-calern-observatory (Preliminary)	0.25m Telescopea Action Rapide pour les Objets Transitoires, Rapid Action Telescope for transient objects at Calern Observatory	A rapid action telescope for observing transient objects, the 0.25m Telescopea Action Rapide pour les Objets Transitoires (TAROT) is part of Calern Observatory in Europe and has an aperture of 0.25 meters.	#rapid-action-telescope-for-transient-objects-at-calern-observatory	
0-27m-convection-rotation-and-planetary-transits-satellite-mission (Preliminary)	0.27m Convection, Rotation, and planetary Transits Satellite Mission	The CoRoT satellite mission, operated by the French Space Agency (CNES) and European Space Agency (ESA), is designed to study stellar convection, rotation, and planetary transits		

1. Presentation of the Observation Facilities & Instruments Vocabularies



2. How did we build it?

1. Gather resources, harmonize format & metadata, clean identified errors in data

2. Merge identical entries with NLP tools & LLM as validators

3. Generate views

- IVOA views: CSV
- OntoPortal views: TTL

LIRA Observatoire de Paris | PSL

2. How did we build it?

1. Gather resources, harmonize format & metadata, clean identified errors in data

Identified more than 20 resources *more or less FAIR* that list observation facilities or instruments. Right now, only 9 are exploited.

+NED & CDS's data will be included in the near future

2. Merge identical entries with NLP tools & LLM as validators

3. Generate views

- IVOA views: CSV
- OntoPortal views: TTL

LIRA Observatoire de Paris | PSL

2. How did we build it?

Need a human review of the generated mappings using annotation tools (the LLM makes mistakes).

Views are generated from this output ontology, every error should be fixed at this level (on OntoPortal-Astro).

1. Gather resources, harmonize format & metadata, clean identified errors in data

2. Merge identical entries with NLP tools & LLM as validators

3. Generate views

- IVOA views: CSV
- OntoPortal views: TTL

LIRA Observatoire de Paris | PSL

2. How did we build it?

1. Gather resources, harmonize format & metadata, clean identified errors in data

2. Merge identical entries with NLP tools & LLM as validators

3. Generate views

- IVOA views: CSV
- OntoPortal views: TTL

Views are generated from entities having at least one authoritative lists as their source.

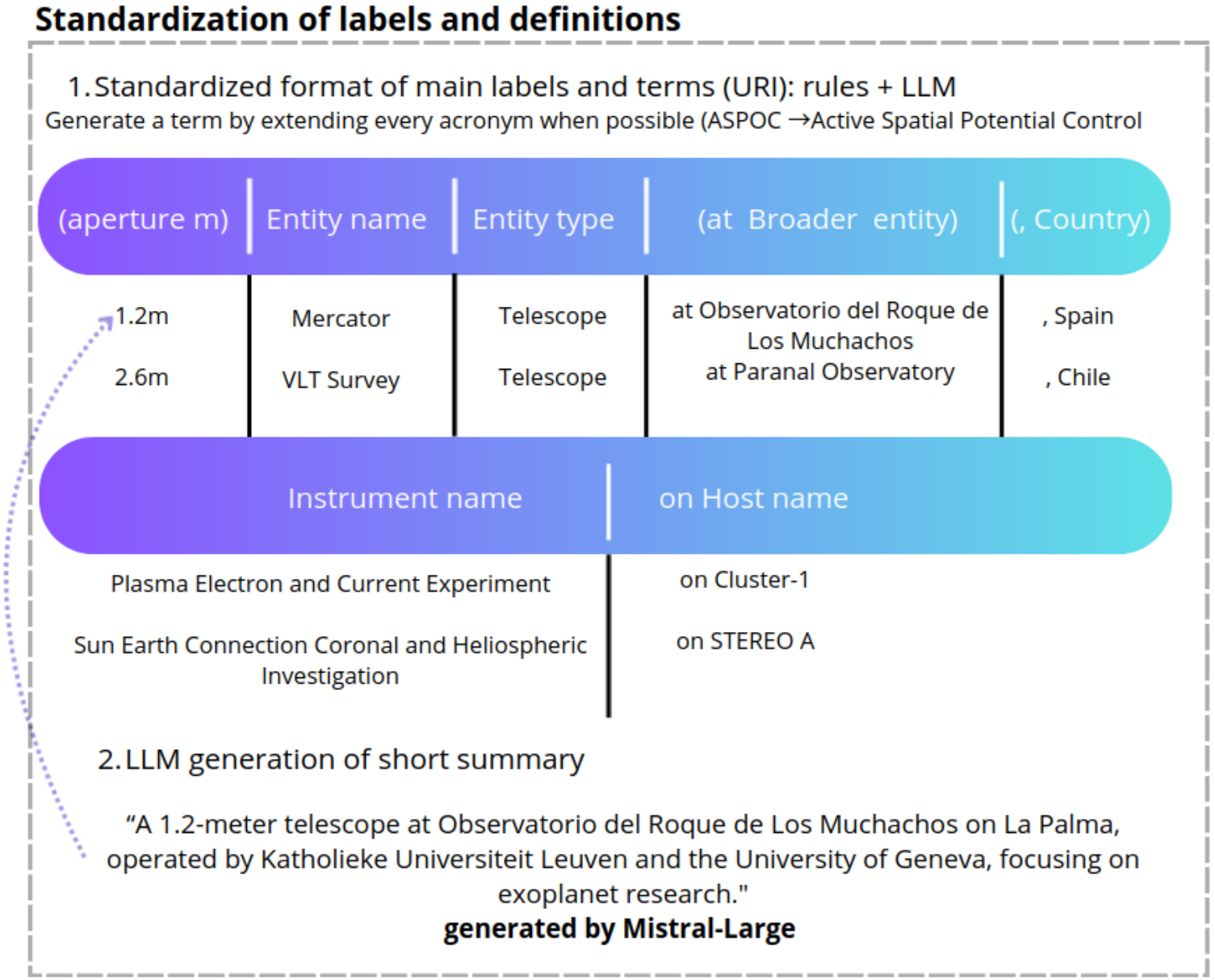
This beta version does not include every source, only AAS, IAU-MPC, NSSDC.

In the future, we will add every authoritative list to the view generation.

LIRA Observatoire de Paris | PSL

2. How did we build it?

The post processing step uses an LLM to generate **labels** and **definitions** from the entities' merged metadata.



3. Current shortcomings and discussions around those limitations

1) Standardizing means taking decisions

Country in the label?

Taiwan / China ? Crimea → Russia / Ukraine ?
Israel / Palestine ?

Different country names
Russia → Federation of Russia
Currently we **use a tool (geopy)** that sets a country to entities according to their coordinates, and **ask the LLM to include the country in the label.**

Acronyms or full text?

MIT → do not extend
NASA → do not extend

How about entities' acronyms (HST for Hubble Space Telescope, SSO for Siding Spring Observatory etc?)
The LLM does not know where to stop AND may lack information to extend labels.

3. Current shortcomings and discussions around those limitations

2) IVOA labels to identifiers breaks the OBSF-OBSI vocabularies

1. Original labels are duplicate

Example: Cassini (mission) VS Cassini (spacecraft)

In order to be distinct, labels are **extended** with their broader entity and country to make them distinct.

As a result, the **identifiers become too long**.

2. Harder maintenance of a large scale vocabulary

Labels will evolve → entities in the IVOA will become **deprecated** with useInstead relations.

Harder to maintain with ~100 yearly label updates

Conclusion: Not the same scale as other IVOA vocabularies (**10k+** entities VS 10 to 100 for most of the IVOA vocabularies), harder to control each entity's identifier and label pairs.

Remove the constraint of having identical label-URI?

3. Current shortcomings and discussions around those limitations

3) Lack of controlability over the LLM output

Non standard formats

Example 1:
Label: "**300 foot** Radio Telescope at National Radio Astronomy Observatory, United States"
Definition: "The 300-foot Radio Telescope was a former radio telescope located in Green Bank, West Virginia, United States, with an aperture of approximately **91 meters**."

Example 2:
Label: "iTelescope Observatory"
Definition: "The iTelescope Observatory is a ground-based astronomical facility located in Beryl Junction, Iron County, Utah, **United States**.";

Possible causes

LLM-related

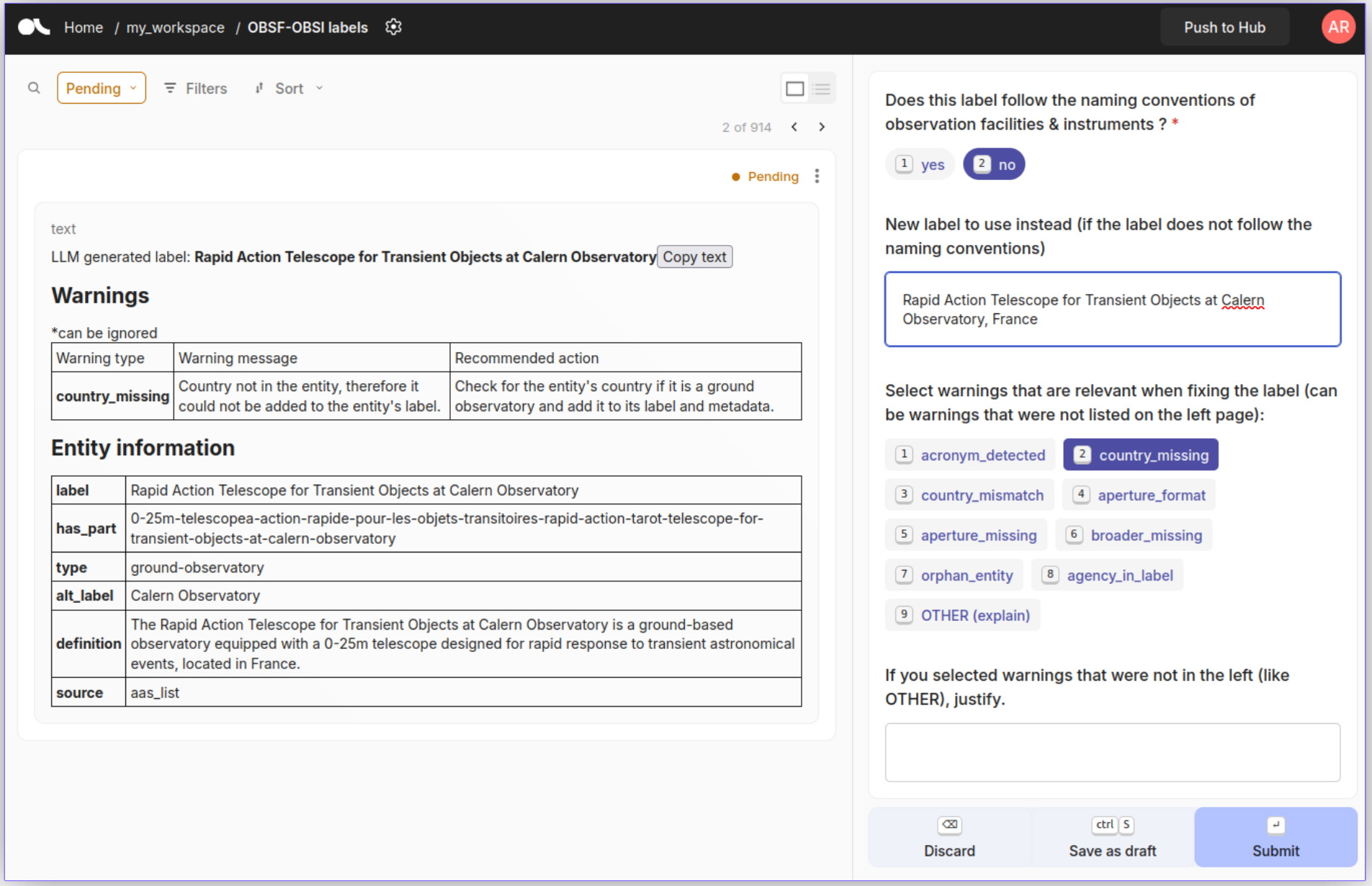
- LLM did not respect the prompt
- Information missing in output
- LLM hallucinates

Data related

- Lack of information in the entity

4. Towards an unified vocabulary: how can the community contribute?

Argilla – annotation tool for final product fixes

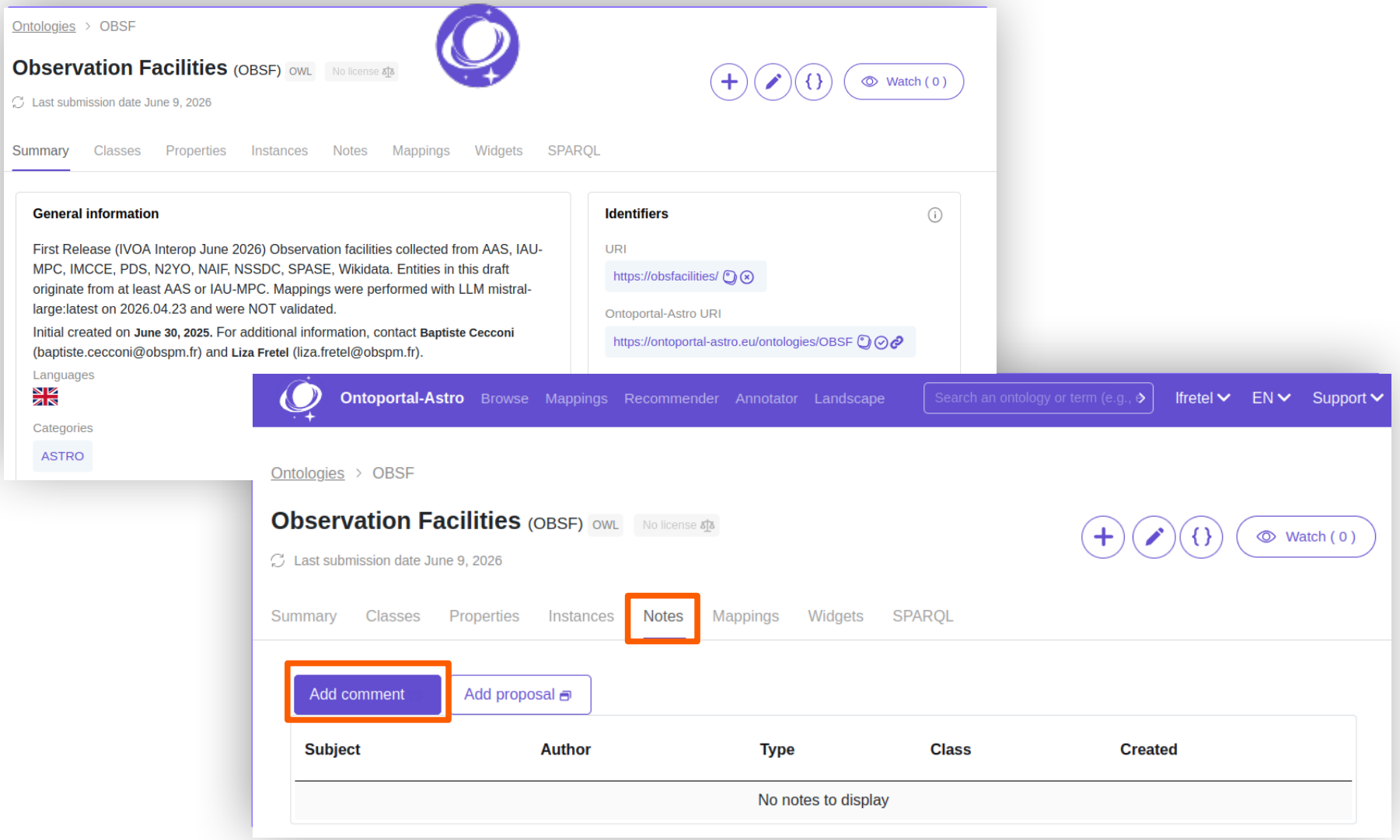


Argilla https://docs.argilla.io/latest/

- Deployed at Paris Observatory (not publicly available yet)

4. Towards an unified vocabulary: how can the community contribute?

OntoPortal-Astro – base ontology that served the generation of the IVOA vocabulary view



https://ontoportal-astro.eu/ontologies/OBSF

- Full merged ontology of Facilities is now available on OntoPortal-Astro!

LIRA Observatoire de Paris | PSL

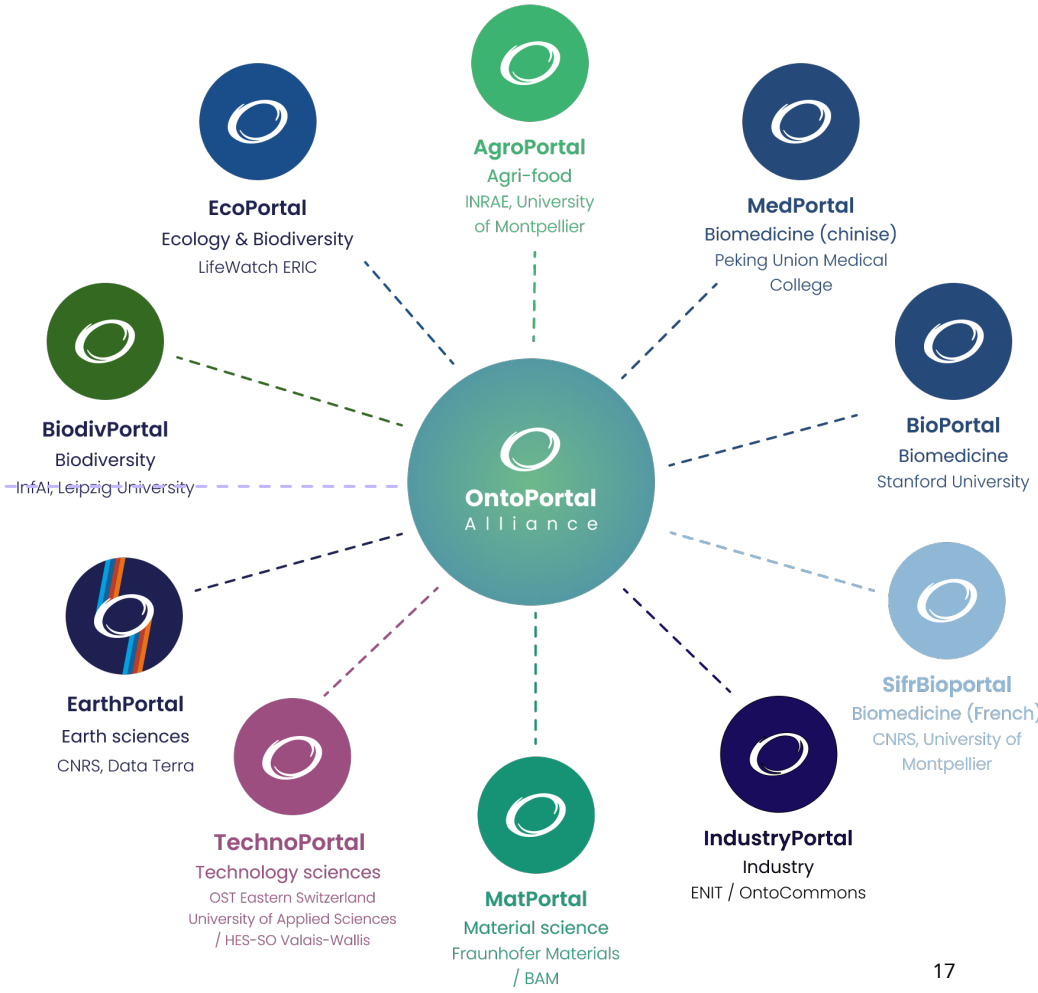
APPENDIX

Context
Objectives

- Create and issue a cross-community vocabulary for the **IVOA** (International Virtual Observatory Alliance)

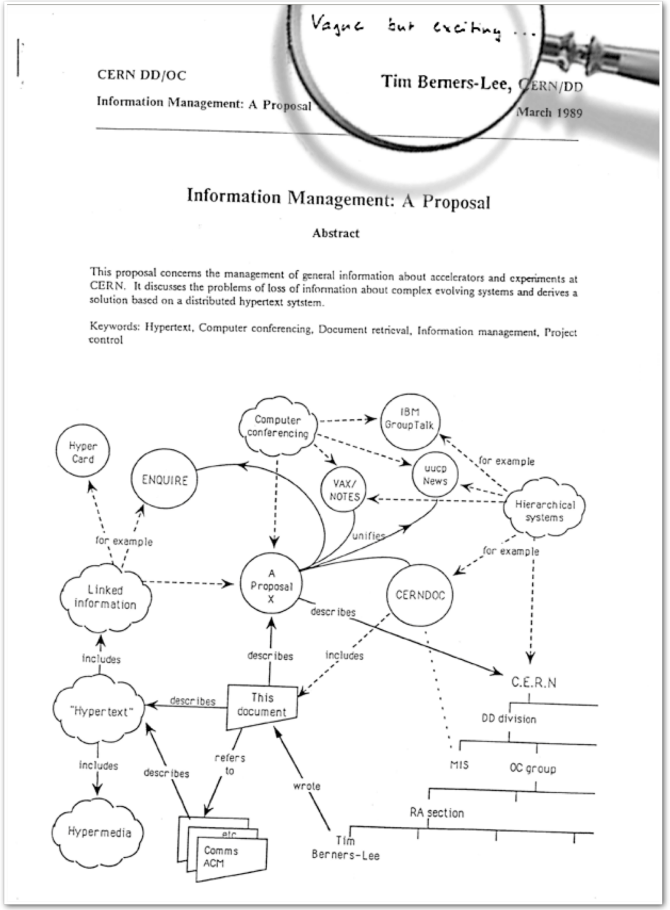
- Merge data and their metadata from semantic lists issued by different communities into one ontology and make it available on **OntoPortal-Astro**
  http://ontoportal-astro.eu/ontologies/OBSF *temporary private link

- Set a recommended label per observation facility to help researchers lookup and use the conventional labels (**Name Resolver**)



Case 1 – Facility List

Context – What are observation facilities in Semantics Web ?



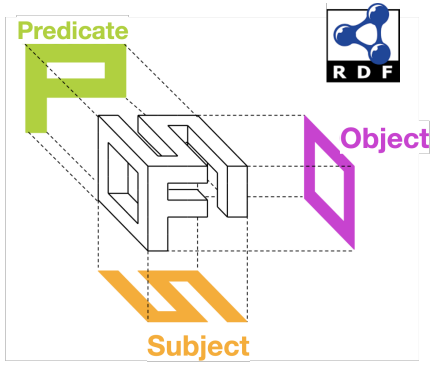
Semantics web:

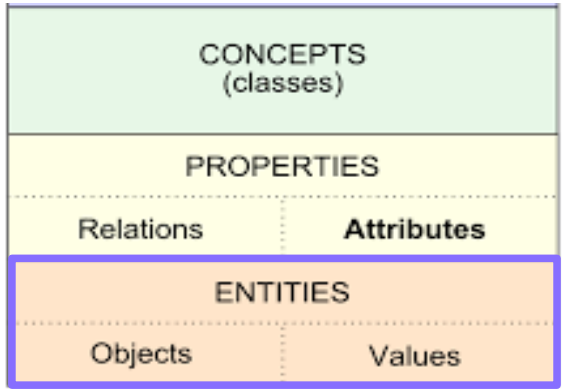
Knowledge graphs (or ontologies) are adapted for **sparse and connected** data representation.

Example: a telescope that is part of an observatory.

Built on RDF triple banks : **subject-predicate-object**

subjects & objects = data, values
predicates = relation between data, attributes



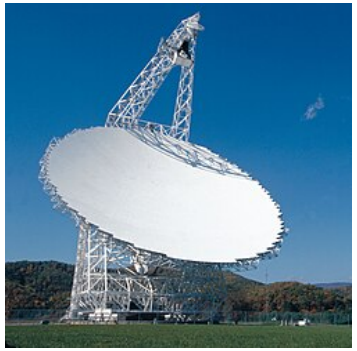


Ontology layers

Astronomical observation facilities: entity level

Research infrastructure designed, constructed, and operated for the **purpose of acquiring astronomical data** that enables the **observation of celestial objects and phenomena** under controlled scientific conditions.

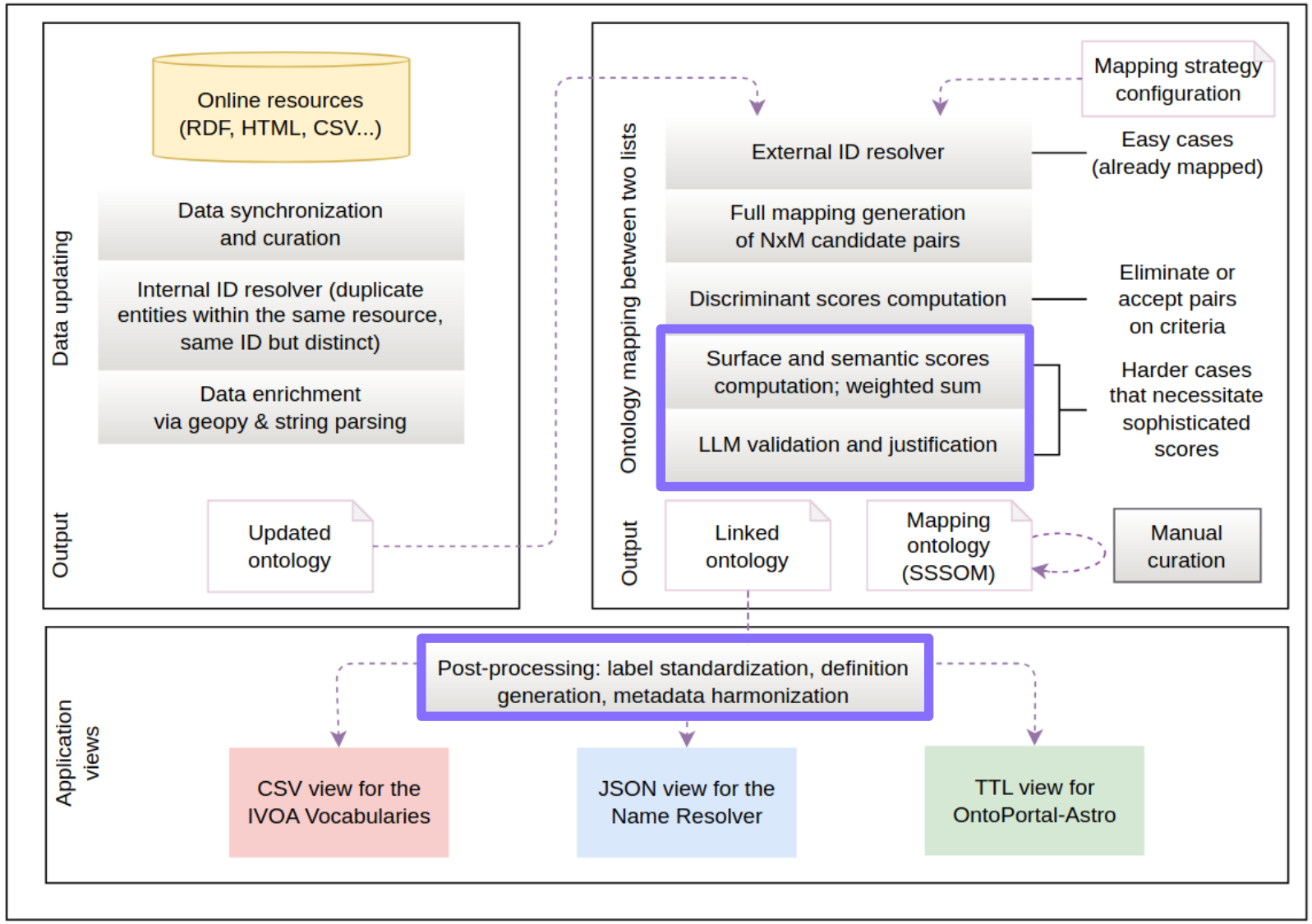
Sphinx Observatory (Switzerland)



Green Bank Telescope (USA)

Hubble Space Telescope, HST

LIRA Observatoire de Paris | PSL

Case 1 – Facility List

Pipeline Overview