

Observation Facilities in the VO

[Data Discovery Use Case]

B. Cecconi (1), L. Debisschop (1)
M. Louys, (2), E. Perret (2),

(1) Observatoire de Paris, Meudon, France;

(2) CDS, Strasbourg, France



Why

- **ObsTAP / EPN-TAP** have “instrument_host_name” / “facility” keywords.

Standard nomenclature needed for efficient data mining

- No official nomenclature/standard for “observation facilities” names.
- Several lists identified (some have >1000 items).
- Merging is complex (by hand or programmatically)

Goal for this study

- Use case: **Data discovery**
 - **step 1:** define what should be stored (observatory/telescope/space mission/spacecraft...), and if relations are needed (e.g., telescope to observatory)
 - **step 2:** match lists and catalogues, build a lookup table with alternate names
 - **step 3:** define maintenance procedure
 - **step 4:** build a name resolver for data discovery clients, or to help provider to select a name

Connecting with WikiData

- Wikidata = free and open knowledge base
 - => structured data (export in RDF, JSON)
 - (each item has properties + values)
 - => connected with wikipedia, Wiktionary
- Wikidata can be **queried** and **edited** manually, or with API, SPARQL and other tools (e.g., QuickStatements, OpenRefine...)
 - <https://meta.wikimedia.org/wiki/QuickStatements>
 - <https://www.wikidata.org/wiki/Wikidata:Tools/OpenRefine>
- For *Observation Facilities*:
 - many identifiers are already connected
 - model/properties are fuzzy but adequate
 - curation / extension is feasible
 - semantics quality is variable



Wikidata record examples

Item [Discussion](#) [Read](#) [View history](#)

Hubble Space Telescope (Q2513)

unmanned space telescope launched into outer space by NASA and ESA in April 1990 [edit](#)

HST | Hubble

[In more languages](#)
[Configure](#)

Language	Label	Description	Also known as
English	Hubble Space Telescope	unmanned space telescope launched into outer space by NASA and ESA in April 1990	HST Hubble
French	télescope spatial Hubble	télescope spatial	HST Hubble Space Telescope Hubble télescope Hubble Telescope spatial Hubble Large Space Telescope Télescope spatial Hubble
Spanish	telescopio espacial Hubble	telescopio en órbita alrededor de la Tierra lanzado en 1990	HST Telescopio Hubble Telescopio Espacial Hub... Telescopio espacial Hubble TEH Hubble (telescopio)
German	Hubble-Weltraumteleskop	Weltraumteleskop für sichtbares Licht, Ultraviolett- und Infrarotstrahlung	HST Hubble Space Telescope Hubble-Space-Telescope Hubbleteleskop Hubble-Teleskop Hubble

[All entered languages](#)

Statements


instance of	space observatory edit
	0 references
	+ add reference

Alias

Property

Qualifier

Wikidata record examples



Item [Discussion](#) Read [View history](#)

Mars Orbiter Mission (Q2156739)

Indian Mars orbiter, launched in 2013 [edit](#)
MOM | Mangalyaan

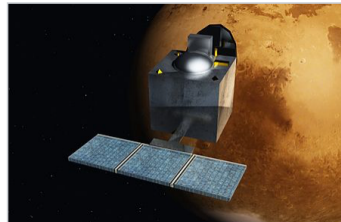
[In more languages](#)
Configure

Language	Label	Description	Also known as
English	Mars Orbiter Mission	Indian Mars orbiter, launched in 2013	MOM Mangalyaan
French	Mangalyaan	No description defined	
Spanish	Mars Orbiter Mission	sonda espacial cuyo lanzamiento fue el 5 de noviembre de 2013	
German	Mars Orbiter Mission	indische Marsmission	

[All entered languages](#)

Statements

instance of [space probe](#) [1 reference](#)

image 

Identifiers

COSPAR ID	2013-060A edit
	2 references
	+ add value
Encyclopædia Britannica Online ID	topic/Mars-Orbiter-Mission edit
	0 references
	+ add reference
	+ add value
Freebase ID	/m/0knt6hh edit
	1 reference
	+ add value
NAIF ID	-3 edit
	1 reference
	+ add value

Wikidata record examples



WIKIDATA

- Main page
- Community portal
- Project chat
- Create a new Item
- Recent changes
- Random Item
- Query Service
- Nearby
- Help
- Donate

Lexicographical data
Create a new Lexeme

English Not logged in Talk Contributions Create account Log in


Item **Discussion** Read **View history** Search Wikidata

Mount Wilson Observatory (Q466863)


astronomical observatory in Los Angeles County, California [edit](#)

[In more languages](#)

Statements

instance of	 astronomical observatory edit
	0 references
	+ add reference
	+ add value

Identifiers

Minor Planet Center observatory code	 672 edit
	0 references
	+ add reference
	+ add value

SparQL query example



Wikidata Query Service

Exemples

Assistant de requêtes

Aide

Davantage d'outils

français



```
1 PREFIX schema: <http://schema.org/>
2 PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
3 PREFIX wikibase: <http://wikiba.se/ontology#>
4 PREFIX bd: <http://www.bigdata.com/rdf#>
5 SELECT
6   ?item
7   ?itemLabel
8   (GROUP_CONCAT(DISTINCT ?alias; SEPARATOR="|") AS ?aliases)
9 WHERE
10 {
11   ?item p:P31 ?stat .
12   #item instance of
13   {?stat ps:P31 wd:Q148578 .} # space observatory
14   OPTIONAL {?item skos:altLabel ?alias .}
15   SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE],en". }
16 }
17 GROUP BY ?item ?itemLabel
```



147 résultats en 1109ms

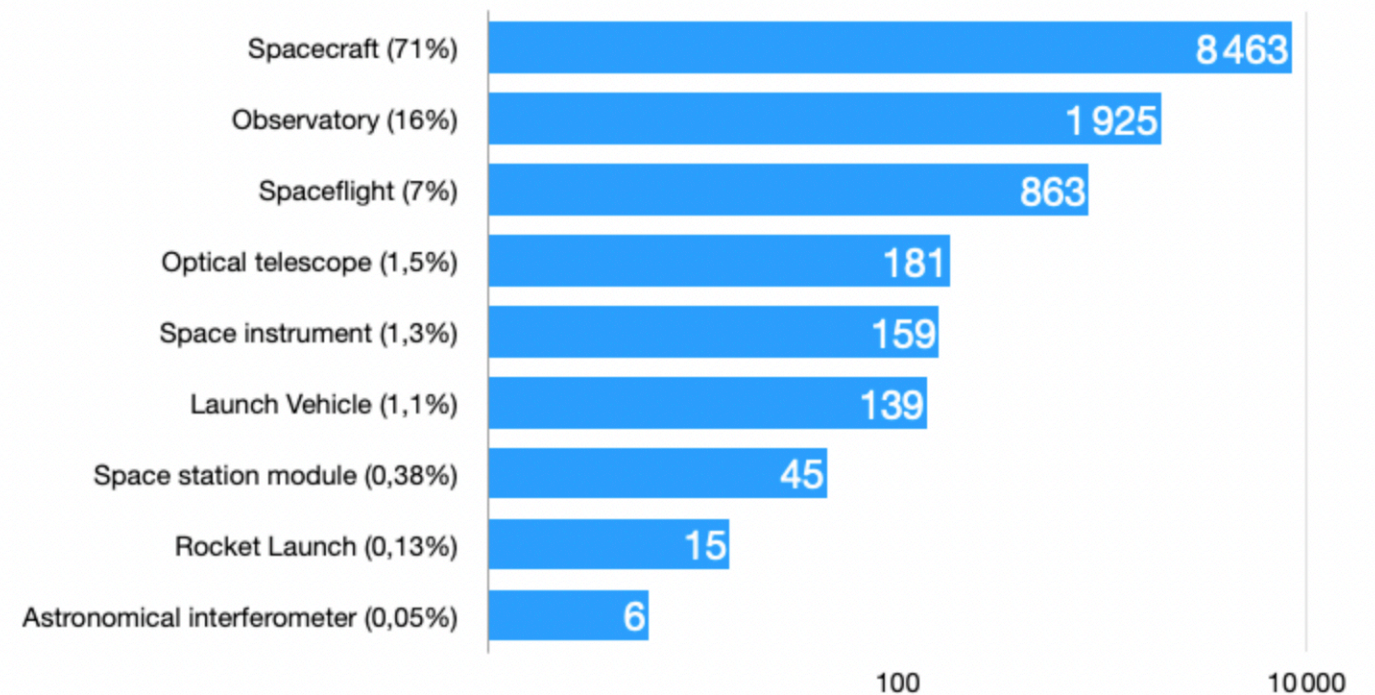
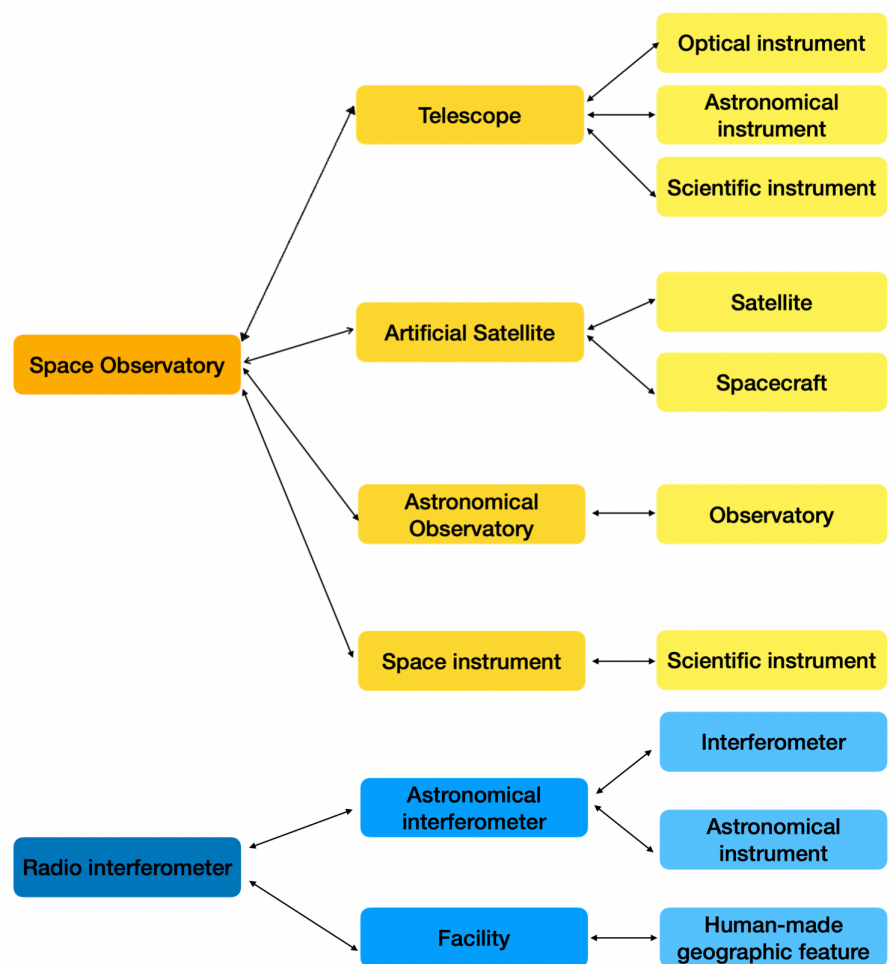
Code

Télécharger

Lien

item	itemLabel	aliases
wd:Q19764	Astron	
wd:Q48633	Corot	CoRoT Convection Rotation and Planetary Transits CONvection ROTation and planetary Transits Convection, Rotation and planetary Transits CoRoT コロト衛星 COROT space telescope CONvection ROTation and planetary Transits
wd:Q49694	EXOSAT	Sat European X-ray Observatory Satellite

Selecting records in Wikidata



Possible to simplify queries using subclasses:

```
?item wdt:P31/wdt:P279* wd:Q40218 . #spacecraft
```

Name resolver prototype

- Implemented on Elasticsearch + Kibana.
- Export from WikiData:
 - identifier (Qxxxxxx), label, aliases, known external identifiers (IAU, NSSDC, UAT...)
- Import into « obsfacility_index »
- Use Kibana for exploring dataset
- Currently: 14200 items ingested
Curation still needed.

Examples

free-text search

The screenshot shows the Elastic search interface. At the top, there is a search bar with the text 'meudon'. Below the search bar, there are several filters and a table of results. The table has columns for 'label', 'aliases', 'iau-mpc', 'uat', and 'uri'. The results are as follows:

label	aliases	iau-mpc	uat	uri
> meudon-observatory	Observatoire de Paris, Observatoire de Paris-Meudon, Observatoire de Paris, Observatoire de Paris-Meudon, Observatoire de Paris-Meudon	005	(empty)	http://www.wikidata.org/entity/Q13104410
> solar-observatory-tower-meudon	(empty)	(empty)	(empty)	http://www.wikidata.org/entity/Q488850
> paris-observatory-psl-university	Observatoire de Paris, Observatoire de Paris, Observatoire de Paris, Pařížská hvězdárna, Observatoire de Paris, Pariser Sternwarte, Sternwarte Paris, Observatorio de Paris-Meudon, Observatoire de Paris - PSL, Observatoire de Paris, Observatoire de Paris-Meudon, Osservatorio astronomico di Parigi, パリ王立天文台, ムードン天文台, Paräiser Observatorium, Observatoire de Paris, Pariserobservatoriet, Observatoire de Paris, Королевская обсерватория, Медонская обсерватория, Observatoire de Paris, Observatoire de Paris-Meudon, Observatorij v Parizu, Pariški observatorij, Observatoire de Paris	007	(empty)	http://www.wikidata.org/entity/Q461340

Examples

known id search

The screenshot shows the Elastic search interface. The search bar contains the query 'naif: -82'. The left sidebar shows a filter for 'obsfacility_index' with '1 hit'. The main results area displays a table with one row for 'cassini'.

label	aliases	iau-mpc	uat	uri
> cassini	Cassini probe, Cassini spacecraft, Cassini probe, Cassini spacecraft, Cassini probe, Cassini spacecraft	(empty)	(empty)	http://www.wikidata.org/entity/Q2941291

The screenshot shows the Elastic search interface. The search bar contains the query 'iau-mpc: 000'. The left sidebar shows a filter for 'obsfacility_index' with '1 hit'. The main results area displays a table with one row for 'royal-observatory'.

label	aliases	iau-mpc	uat	uri
> royal-observatory	Royal Observatory, Greenwich, Royal Observatory, Observatori de Greenwich, Observatori de Greenwich, , Reial Observatori de Greenwich, Royal Greenwich Observatory, Observatori Royal, Hvězdárna Greenwich, Královská greenwichská observatoř, Yr Arsyllfa Frenhinol, Flamsted House, Flamsted House, Greenwich-Observatorium, Royal Observatory, Αστεροσκοπίο του Γκρίνουιτς, Greenwich Observatory, Royal Greenwich Observatory, Obserwejo Grenviča, Reġa Observatorio de Greenwich, Observatorio de Greenwich, Royal Greenwich Observatory, Royal Observatory, رصدخانه سلطنتی گرینویچ, Observatoire de Greenwich, Royal Greenwich Observatory, מצפה גריניץ', Observatorium Greenwich, Royal Greenwich Observatory, Observatorio Reale di Greenwich, Royal Greenwich Observatory, グリニッジ王立天文台, 旧グリニッジ天文台, 그리니치천문대, 왕립 그리니치 천문대, Кралска опсерваторија во Гринич, Greenwich observatoriet, Roval Greenwich Observatorv, Roval Observatorv Greenwich. K	000	(empty)	http://www.wikidata.org/entity/Q192988

Operation & Maintenance

- Very preliminary... *doc in preparation, currently in French*
- Scripts to export from WikiData: SPARQL => json
- Comparison with known lists (fuzzy matching)
=> tagged results:
 - 'not_found': check for wikidata records, add record if necessary.
 - 'very_good_match': Ingest into Elastic Search database
- TBD:
 - SPARQL query retrieving only updated or new records.
 - procedure to ingest results with existing content in Elastic Search
 - translate doc in english...

Wikidata

as a knowledge base

- **Proposal:** use Wikidata as our curated knowledge base
- **Data Discovery** use case: *looks good.*
 - => regular extraction into a name resolver database (using SPARQL queries)
 - => curation of metadata directly in Wikidata
 - => name-resolver (e.g., with elastic-search API)
- Two documents in preparation (IVOA Notes):
 - <https://github.com/BaptisteCecconi/ObsFacilityWikidata>
 - update and maintenance document (not online yet)