

Identifier in VO Paris Data Centre

Pierre Le Sidaner Jonathan Normand
Jerome Berthier
Observatoire de Paris

status

ID syntax

purpose

Conclusion

□ Identifier in the resources

- Clearly define in the document
 - IVOA Identifiers Version 1.12
 - But still need examples for construction.
 - Need to clearly understand the purpose/goal/use of this identifier

- How it appear
 - As a primary key for registry ?
 - As a visibility for the resource provider ?
 - As a id for Archive Storage (OAIS) ?
 - As a reference to retrieve data for a user ?

How they have been used

▣ Our experience

- ID where everything, stabilized in the version evolution of registry

ivo://obspm.fr/planets

ivo://obspm/PDR

ivo://obspm.fr/SSA_HIG

ivo://obspm.fr/cai/ESOR.

Lets take the example of exoplanet CS

Initialy ivo://obspm.fr/planets

It come from a research department calling LUT “Laboratoire
Univers et Théories”

Member of Observatoire de Paris

Participant of project VO Paris Data Centre

status

ID syntax

purpose

Conclusion

How we are trying to describe

□ It start with a naming authority

○ For us it's double

- It should be `ivo://vopdc` for description of the transversal project
- It should be `ivo://vopdc.obspm` for resources shared by vopdc in Observatoire de Paris or `ivo://vopdc.iap` in case of institut d'astrophysique de Paris.

○ Then we have define as an Organization

- The partners inside an institute like research department for example `ivo://vopdc.obspm/luth` as “Laboratoire Univers et Theories”

status

ID syntax

purpose

Conclusion

ID construction

Then we will come to the service

For simple example we add the service name

ivo://vopdc.obspm/luth/exoplanet and the service
have as publisher

status

ID syntax

```
<publisher ivo-id="ivo://vopdc.obspm/luth">VO-Paris  
Data Centre – LUTH</publisher>
```

purpose

And to define context

```
<relatedResource ivo-id="ivo://vopdc.obspm">VO-  
Paris Data Centre</relatedResource>
```

Conclusion

Case of multiple ressources

Gepi/CAI promote multiple surveys of southern emisphere

status

ID syntax

purpose

Conclusion

naming

ivo://vopdc.obspm/

Organisation

ivo://vopdc.obspm/gepi

Project

ivo://vopdc.obspm/gepi/vopsat

Survey

ivo://vopdc.obspm/gepi/vopsat/esor

ivo://vopdc.obspm/gepi/vopsat/srcj

ivo://vopdc.obspm/gepi/vopsat/pose

Capabilities could be SIA, CatalogService

Questions

□ Talking about ID

- **Why do we need ID ?**
- **Why the approach of IVOA could change from a classical archive. ?**
- **Why IVOA seem a good structure to propose long term ID system ?**

status

ID syntax

purpose

Conclusion

Why do we need ID

To give a unique access to a dataset/product/service

As it is describe in the documents :

Long time URI

<http://www.astro.gla.ac.uk/users/norman/ivoa/long-term-uris.html>

IVOA Identifiers Version 1.12

<http://www.ivoa.net/Documents/REC/Identifiers/Identifiers-20070302.html>

Give a formal name to object.

Give a URI to use as a machine-readable information.

Can also be used as an internal primary key.

status

ID syntax

purpose

Conclusion

Difference between classical Archive and VO approach

status

Unique identifier is also a key element of classical archive in the OAIS Open Archive Initiative System.

ID syntax

Archive is done to retrieve data in a format describe that can be native and protocol that can even be FTP.

purpose

In the VO, the access to data is already standardized and be usable by a machine. It give you direct access to the data in standardized format self described.

Conclusion

DAL provide access to data usually not by ID.

In Archive ID must be fixed for long time as privileged access system. His construction is specific (library ...)

In VO the ID could be representative of the data provider and the root is stored in the registry. That mean that for a data ID could change by time. But VO always proposed an easy solution to retrieve data.

It's still a good way to represent origin and credit to data

Challenge of long term preservation

status

If ID construction for a dataset is dependant on the research group that promote it. In can not be stable for long period.

ID syntax

ID is construct from registry root and the sustainability depend on registry lifetime.

purpose

Registries are harvested that avoid failure point and sustain by governments and spacial agencies.

Conclusion

Astronomical Observatories that host registries are one of the oldest scientific institute. Their lifetime is measured in 100 years.

Potential users for registries

Other VO project used a IVOA like protocols.

status

* VAMDC used TAP-XSAMP and K. Benson have developed a registry for their services

ID syntax

http://msslxv.mssl.ucl.ac.uk/vamdc_registry/main/index.jsp

And services used an ID like :

purpose

`ivo://vamdc/basecol_views/ceaApplication`

`ivo://vamdc/UDFA`

Conclusion

* Europlanet will adopt IPDA protocol calling PDAP where ID is mandatory calling `dataset_id` or `product_id`.

Due to similarity if IVOA DAL and PDAP, IVOA registry could be a good candidate to register services. Should ID be `ivo://epn/` or `ivo://ipda/`
Or `enp://` `ipda://`

What impact in the registry software ?

status

ID syntax

purpose

Conclusion

- ❑ **ID is an important parameter in IVOA system**
 - Construction should carefully be understood by service providers
 - IVOA document could be completed by many examples to facilitate implementation.
 - IVOA ID should be open to other structure. It's a proof of good standard quality and long time used