



ESCAPE

European Science Cluster of Astronomy &
Particle physics ESFRI research Infrastructures

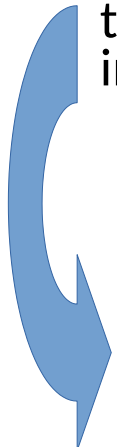
EOSC portal onboarding & IVOA Registry resource publishing overlaps and lessons learned

Marco Molinaro – INAF

ESCAPE EEB Extended discussions day

28 September 2021

- Task 4.1 “Integration of astronomy VO data and services into the EOSC”
- among the activities
 - Interfacing the VO framework with the EOSC (This comprises the inclusion of the VO registry)
 - Build an Astronomy portfolio of VO services to be contributed to the EOSC Marketplace and/or analyse the possibility to operate a customized marketplace for Astronomical VO services and, if necessary, implement it



Assess the methods for contributing VO services as an Astronomy Portfolio to the EOSC Marketplace.

Analyse the possibilities for operating a customized marketplace for astronomical VO services.

Establish a test plan for contributing selected VO services to the EOSC Marketplace (*dependent on the results of the above steps*).



The screenshot shows the VO Registry interface. At the top, there are navigation links: GUIDELINES, COMMUNITIES, FACETED SEARCH, SEARCH GUIDE, AB, B2FIND, EUDAT, GUIDELINES, COMMUNITIES, FACETED SEARCH, SEARCH GUIDE, ABOUT. Below the navigation, there are logos for B2FIND and EUDAT. The main content area shows a search for 'IVOA' resulting in 24,341 datasets. A map on the left allows filtering by location. The 'GAVO Data Center Registry' is highlighted, with a description: 'The publishing registry for the GAVO Heidelberg Data Center. On request, we also host other registry records. Use the contact address for more information.' Below the description are filter buttons for 'interdisciplinary a...', 'observational astro...', and 'virtual observatories'. A table provides details for the GAVO registry:

Identifier	
Source	http://dc.zah.uni-heidelberg.de/_system_/services/registry/info
Metadata Access	http://dc.g-vo.org/rr/q/pmh/pubreg.xml?verb=GetRecord&metadataPrefix=oai_datacite&identifier=ivo://org.gavo.dc/_system_/services/registry

On the right side of the screenshot, there is a social media section with buttons for Twitter and Facebook. Below that, there is a section for 'The VO @ ASTRON TAP service' with a description and a table of metadata:

Identifier	
Source	https://vo.astron.nl/_system_/tap/run/info
Metadata Access	http://dc.g-vo.org/rr/q/pmh/pubreg.xml?verb=GetRecord&metadataPrefix=oai_datacite&identifier=ivo://astron.nl/tap

Below the metadata table, there is a 'Provenance' section with the following details:

Provenance	
Creator	ASTRON
Publisher	ASTRON
Contributor	International Virtual Observatory Alliance (IVOA)
Publication Year	2009
OpenAccess	true
Contact	Michael Sipior <vo(at)astron.nl>

At the bottom of the screenshot, there is a 'Representation' section with the following details:

Representation	
Resource Type	AstroData: Service
Discipline	Observational Astronomy; Interdisciplinary Astronomy

- Included in EUDAT B2-FIND

- (GAVO) IVOA Full Registry RegTAP interface
- extended with DataCite metadata
- harvested by B2-FIND

- Preserves resources description

- linking back to VO Registry

- Lacks

- some EUDAT / DataCite based metadata (not originally available in VOResource)
- actionable data linking from collection to service (integral part of the VOResource model)
- coverage metadata content and technology mapping from VO to EOSC



Astronomy portfolio directly in EOSC

- Steps for the assesment & contribution guidelines
 - setup a Provider within EOSC
 - test the onboarding procedure
 - from non VO Registered resources
 - test the onboarding procedure
 - starting from VO Registered resources
 - report guidelines minimising duplicated efforts



Registry integration & Portal onboarding

IVOA

- authority
- Model
 - Flexible, complex, extensible
- distribute & harvest
- metadata granularity
- linked collection to interface
- matured over time
- community “weather” forecast

EOSC

- provider, AARP
- Model
 - single resource, general domain
- form based
 - programmatic API available
- external repositories inclusion:
 - domain, national, cluster (?)
- high details on policies



Portal onboarding experience

DONE

- Setup of a Provider for INAF in EOSC
 - included political hurdles
 - ARP identification
 - resource metadata setup
 - wait (manual & tech) approval
- Selected a couple of services to onboard
 - H2020 NEANIAS collaboration
 - preparation of policy, plans, helpdesk documents and services
 - resource metadata population
 - automated validation and approval

unclear/discussed “service” definition

TO DO

- VO registered resource onboarding
 - starting from scratch
 - long due resources to be registered available
- identify metadata overlap
 - estimate level of effort duplication
- check “linked” data-to-service availability
- try programmatic API solution



Open questions / Discussion topics

- Registry inclusion/harvesting vs. direct resource onboarding
 - effort duplication, repo redundance, visibility, use-ability
- Domain, national, cluster, project repository inclusion
 - how-to, long term sustainability, effort duplication, model governance
- Resource granularity and types
 - resource relationships, actionable link from discovery to access

