Project Title	European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures
Project Acronym	ESCAPE
Grant Agreement No	824064
Instrument	Research and Innovation Action (RIA)
Topic	Connecting ESFRI infrastructures through Cluster projects (INFRA-EOSC-4-2018)
Start Date of Project	01.02.2019
Duration of Project	48 Months
Project Website	www.projectescape.eu

D4.4 - INTERMEDIATE ANALYSIS REPORT ON INTEGRATION OF VO DATA AND SERVICES INTO EOSC

Work Package	WP4, Connecting ESFRI projects to EOSC through VO
Lead Author (Org)	Marco Molinaro (INAF)
Contributing Author(s) (Org)	M. Demleitner (UHEI), D. Morris (UEDIN), M. Allen (CNRS), F. Genova (CNRS), A. Schaaff (CNRS)
Due Date	31.07.2020, M18 (delayed to 30.11.2020, M22 due to COVID)
Date	30.11.2020
Version	1.0

Dissemination Level

X	PU: Public
	PP: Restricted to other programme participants (including the Commission)
	RE: Restricted to a group specified by the consortium (including the Commission)
	CO: Confidential, only for members of the consortium (including the Commission)

Versioning and contribution history

Version	Date	Authors	Notes
0.1	14.10.2020	M. Molinaro (INAF)	First draft skeleton circulated to co-authors.
0.5	09.11.2020	M. Molinaro (INAF), M. Allen (CNRS), F. Genova (CNRS), D. Morris (UEDIN)	Second draft updated with main partners co-authoring contributions
0.6	10.11.2020	M. Molinaro	Editorial: acronyms, refs, biblio
0.7	17.11.2020	M. Molinaro	Added a few sentences in §4
0.9	23.11.2020	M. Molinaro, M. Demleitner (UHEI), F. Genova	F. Genova revision, updates from M. Molinaro and integration of M. Demleitner comments and corrections
0.95	23.11.2020	M. Molinaro, M. Demleitner (UHEI), F. Genova, M. Allen (CNRS)	M. Allen review and text editing.
1.0	24.11.2020	All	Final reading and editing before submission to Project Manager.

Disclaimer

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

Table of Contents

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

Acronym list

AAI: Authentication and Authorization Infrastructure

A&A: Authentication & Authorization

ADASS: Astronomical Data Analysis Software and Systems

CC-0: Creative Commons 0 ("no rights reserved") license

CERN: European Organization for Nuclear Research

CEVO: Connecting ESFRI projects to EOSC through the Virtual Observatory framework (ESCAPE Work Package 4)

CNRS: Centre National de la Recherche Scientifique

CNRS-ObAS: CNRS - Observatoire Astronomique de Strasbourg

CTA: Cherenkov Telescope Array

DIOS: (ESCAPE WP2) Data Infrastructure for Open Science

DOI: Digital Object Identifier

ELT: Extremely Large Telescope (was E-ELT)

EOSC: European Open Science Cloud

EOSC-Hub: Integrating and managing services for the European Open Science Cloud

ESAP: (ESCAPE WP5) ESFRI Science Analysis Platform

ESCAPE: European Science Cluster of Astronomy & Particle physics ESFRI research infrastructures

ESFRI: European Strategy Forum on Research Infrastructures

ESO: European Southern Observatory

EST: European Solar Telescope

FAIR: Findable, Accessible, Interoperable, Reusable *or* Facility for Antiproton and Ion Research

HEALPix: Hierarchical Equal Area iso Latitude Pixelation (of the sphere)

HiPS: Hierarchical Progressive Survey

HL-LCH: Civil engineering for the High-Luminosity

INAF: Istituto Nazionale di Astrofisica

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

IVOA: International Virtual Observatory Alliance

IVOID: IVOA Identifier

JIVE: Joint Institute for VLBI ERIC

KM3NeT: A multi-km3 sized Neutrino Telescope

MOC: HEALPix Multi-Order Coverage map

NEANIAS: Novel EOSC Services for Emerging Atmosphere, Underwater & Space Challenges

OAI: Open Archives Inititive

OAI-PMH: OAI - Protocol for Metadata Harvesting

ORCID: Open Researcher and Contributor ID

OSSR: (ESCAPE WP3) Open-source Software and Service Repository

RDA: Research Data Alliance

RegTAP: IVOA Relational Registry Schema

RI: Research Infrastructure

SKA: Square Kilometre Array

UEDIN: University of Edinburgh

UHEI: Ruprecht-Karls-Universität Heidelberg

URI: Uniform Resource Identifier

VESPA: Virtual European Solar and Planetary Access

VO: Virtual Observatory

W3C: World Wide Web Consortium

WP: Work Package

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

Project Summary

ESCAPE (European Science Cluster of Astronomy & Particle physics ESFRI research infrastructures) addresses the Open Science challenges shared by ESFRI facilities (SKA, CTA, KM3Net, EST, ELT, HL-LHC, FAIR) as well as other pan-European research infrastructures (CERN, ESO, JIVE) in astronomy and particle physics. ESCAPE actions are focused on developing solutions for the large data sets handled by the ESFRI facilities. These solutions shall: i) connect ESFRI projects to EOSC ensuring integration of data and tools; ii) foster common approaches to implement open-data stewardship; iii) establish interoperability within EOSC as an integrated multi-messenger facility for fundamental science. To accomplish these objectives ESCAPE aims to unite astrophysics and particle physics communities with proven expertise in computing and data management by setting up a data infrastructure beyond the current state-of-the-art in support of the FAIR principles. These joint efforts are expected to result in a data-lake infrastructure as a cloud open-science analysis facility linked with the EOSC. ESCAPE supports existing infrastructures such as the astronomy VO to connect with the EOSC. With the commitment from various ESFRI projects in the cluster, ESCAPE will develop and integrate the EOSC catalogue with a dedicated catalogue of open source analysis software. This catalogue will provide researchers across the disciplines with new software tools and services developed by the astronomy and particle physics communities. Through this catalogue ESCAPE will strive to provide researchers with consistent access to an integrated Open Science platform for data-analysis workflows. As a result, a large community “foundation” approach for cross-fertilization and continuous development will be strengthened. ESCAPE has the ambition to be a flagship for scientific and societal impact that the EOSC can deliver.

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union’s Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

Executive Summary

Work Package 4 of ESCAPE, CEVO, is focused on seamlessly connecting ESFRI and other astronomy research infrastructures to the EOSC through the VO framework. The CEVO Task 4.1, "Integration of astronomy VO data and services into the EOSC" is dedicated to identifying the relevant domain specific data access, discovery and manipulation standards and integrating them with the EOSC platform. This document provides an intermediate analysis report of the status of the integration, with success stories, ongoing efforts and description of the current and foreseen challenges. The integration activities have been identified in the CEVO Project Plan, ESCAPE D4.1 [1]; the progress on the activities is reported here with respect to the insights they provide on the VO-to-EOSC integration development. The analysis provided by this document, built on top of those activities, summarises the connection points, requirements and challenges that the VO community of data providers and consumers faces when trying to connect to the current EOSC infrastructure. Cross-WP activities relevant to the integration to EOSC are described, as well as connections to other EOSC related projects and participation in meetings and symposia in the EOSC landscape. This document also looks ahead to the remaining time span of the project and the remaining integration goals.

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

1. Introduction

ESCAPE Work Package 4, "Connecting ESFRI projects to EOSC through VO framework" (CEVO), plans to make the seamless connection of ESFRI and other astronomy and astroparticle research infrastructures to the EOSC through the VO framework. Task 4.1 partners within CEVO have been investigating, and then applying, solutions to integrate data and service resources, based upon the IVOA architecture, into the EOSC one. Progress so far has worked out more easily on the technical side, where the technologies involved have proved to be compatible, with the exemplary case of integrating the IVOA Resource Registry. The developments on other aspects were slower due to the challenges of interfacing with the evolving structure of the EOSC platform, for instance for service provision, or due to not-yet-provided features of the architecture, for instance for the integration of semantics and vocabularies.

A lot of effort has been put into finding the proper interface and contacts to guide the process of aligning the features and characteristics of the mature, community-driven architecture of the IVOA with the more recent architectural components of the EOSC. Participation in relevant EOSC project events (see Table in Appendix) has helped a lot in this activity.

In connection with Task 4.2, efforts were also put into describing the Open and FAIR approach, that comes natively with the IVOA vision, to other research domains and to projects whose purpose is to define building blocks for FAIR data sharing. This helped with the clarification of concepts with respect to data policies, licensing and user identification and permission granting. These concepts are a major consideration when providing resources but should not impede the easy access to public data holdings that form the predominant case in astrophysics.

Further activities involved connections to other ESCAPE Work Packages (Section 0), to make available a portfolio of software and services through EOSC, also through containerisation (WP3 - OSSR), to contribute to the EOSC Hybrid Cloud with technical standard solutions for the ESCAPE Science Platform (WP5 - ESAP), also considering the need for this latter service to have a standardised access to the Data Lake solution (WP2 - DIOS).

The outcome of the investigations done during the first half of the project is analysed in Section 4 and the open questions, challenges and vision for the remaining of the project in Section 0. The Task 4.1 activities were also published in a contribution to the ADASS XXIX Conference (Molinaro & al., [2]).

The main partners in the task described in the present document are INAF, CNRS-ObAS, UHEI and UEDIN, while all other project partners contribute to the task individually by testing the identified integration pathways.

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

2. Progress of integration activities

The activities described in this section have taken advantage of the networking activities, reporting and knowledge acquired by participating in relevant EOSC project events. The listing of the events attended is provided in the Appendix.

2.1. VO to EOSC interfacing

This section describes several aspects of the interfacing of the VO with EOSC: the registry of resources, resource on-boarding, and semantics vocabularies.

Registry of resources

The first activity towards integrating the IVOA architecture within EOSC has been to test the compatibility of and then include the IVOA Registry of Resources in the EOSC service catalogue. This activity started immediately at the beginning of the ESCAPE project, with the identification of contacts to check the technical details and the compatibility of metadata models. Contact was established with B2FIND, and the the resources of the IVOA Registry were then actually included into the B2FIND catalogue[1], taking advantatge of the fact that the IVOA Registry of resources is OAI-PMH compliant. This activity is described in Molinaro & al. [2]. Reports on the relevant activity during the IVOA Registry Working Group sessions in the IVOA meetings are captured in the Milestones documents M20 and M21 of ESCAPE CEVO (Progress and Priorities at IVOA 1 & 2, ref. [3] & [4]). The IVOA Relational Registry Schema standard interface (RegTAP [5]) proved to be an efficient tool for this task. The technical bases for B2FIND-IVOA interaction are the OAI-PMH transport protocol (already in use in both communities) and the DataCite metadata schema, which is an extension to the IVOA registry protocol. This extension is available on the GAVO (UHEI managed) full registry, which enabled the records to be harvested by B2FIND.

[1] This inclusion is visible at http://b2find.eudat.eu/group?q=ivoa (or http://b2find.eudat.eu/group/ivoa), and illustrated by Figures 1, 2, 3 & 4.

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

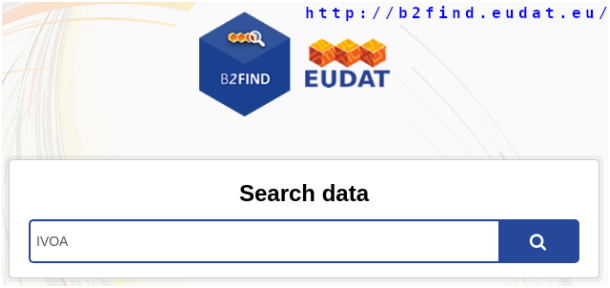

Figure 1: EUDAT B2FIND search box, initialised with a search using the term "IVOA".

Figure 1 shows how to start a quicklook of VO resources within EUDAT. A simple search driven by the *IVOA* word yeilds the results page shown in Figure 2.

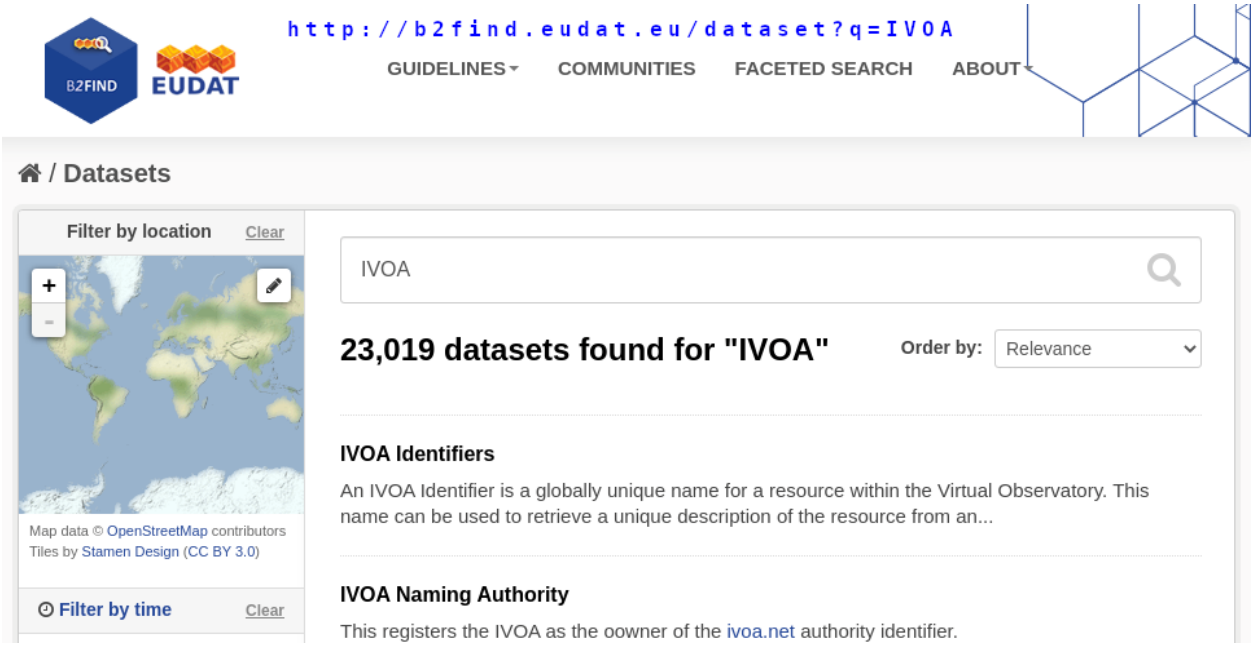

Figure 2: B2FIND "IVOA" results. The dataset count displays the actual content of the IVOA Registry at the moment the inclusion took place and the snapshot was taken.

The plain result listing of the IVOA Registry content in B2FIND shown in Figure 2 can then be filtered by features, that are taken from the metadata elements of the resources description. As shown in Figure 3, the mapping of the metadata content between VOResource metadata and B2FIND metadata works well on the high-level description (*Publisher* or *Keywords* in Figure 3) but starts to be less informative the more one enters into the domain specific details (e.g., *ResourceType*).

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

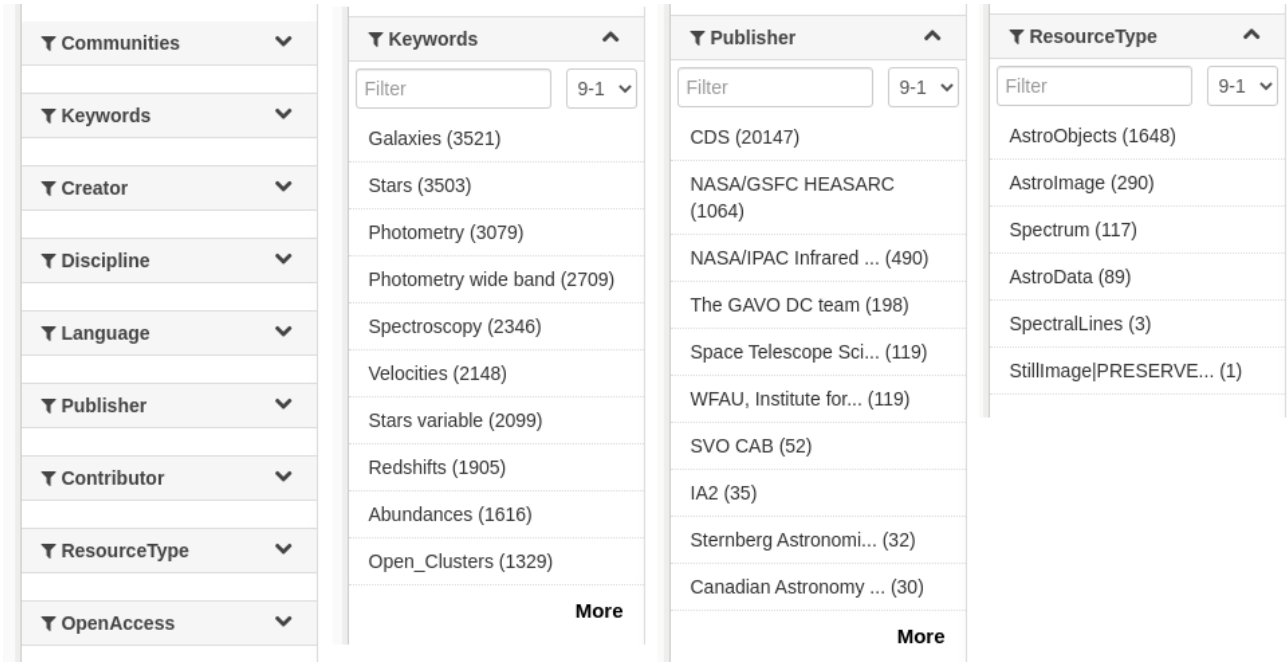

Figure 3: The leftmost panel shows the features that B2FIND allows filtering upon. The other three panels, Keywords, Publisher and ResourceType, show the available values for each ot the three features expanded (including the count of resources for each value).

The filtering applied by B2FIND resembles the basic requirements answered by the IVOA Registry – the fact that the metadata schema of the IVOA Registry includes the Dublin Core is an asset in that respect, but does not allow for the subsequent machine-actionable interaction that the standardised interfaces provided by the IVOA standards allow. This is, of course, a known barrier when matching domain and general purpose repositories, because the depth of insight given by the descriptive metadata depends a lot on the semantic content of the metadata documents. Part of this is taken into account by the *Metadata Access* direct reference to the IVOA Registry full record (Figure 4).

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

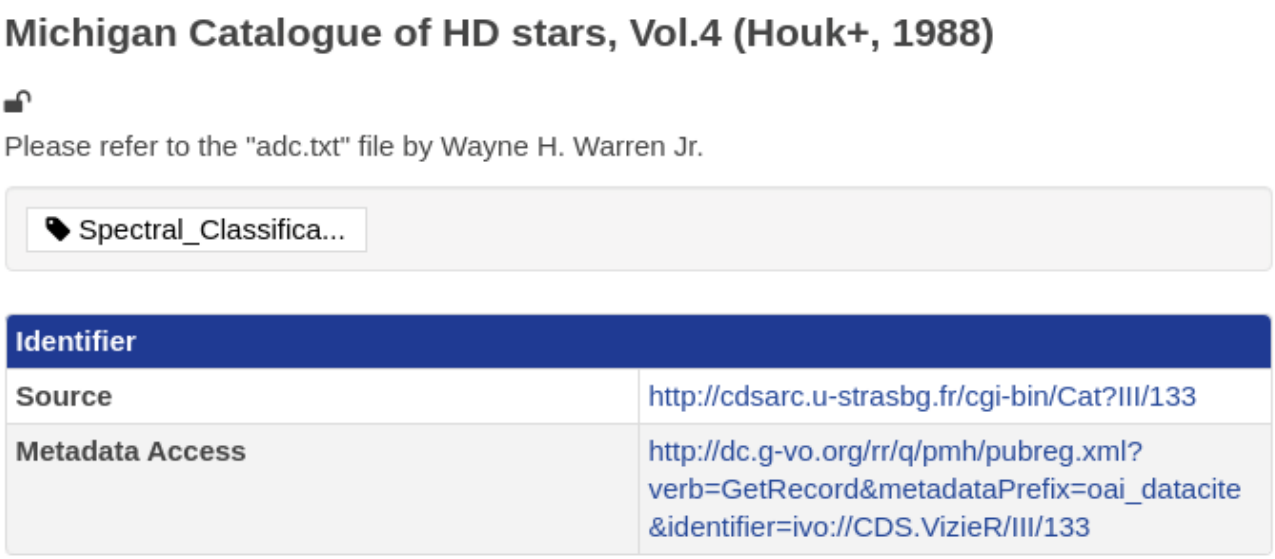

Figure 4: B2FIND, single resource description. The Identifier metadata shows the Metadata Access that connects to the IVOA Registry content.

Further work towards cross-disciplinary integration has involved a study of cross domain user stories for resource discovery. The latter is a challenging task, which is not directly included in the ESCAPE work plan and its follow up will depend on potential connections to other cluster projects.

Service on-boarding

Besides proving the technical feasibility of integrating the IVOA Registry content within the base repository structure of EOSC, our activities have also started to compare the concept of *service onboarding*, as described by EOSC, with the IVOA scenarios for service description and the deployment architecture. The "Adding a service to EOSC" report [7], presented at the CEVO Technology Forum 1 by André Schaaff (CNRS-ObAS), was a first attempt in this direction. It was used to report the experience in the EOSC-Enhance survey. While continuously checking on the EOSC updates about service onboarding, for instance in the EOSC-Hub Week 2020 report,[2] a further attempt will be made through cross-project connection with H2020 NEANIAS[3].

Semantics

Another set of resources to be tested for inclusion in the EOSC services are the IVOA-standardised

[2] https://www.eosc-hub.eu/eosc-hub-week-2020/agenda/service-onboarding-catalogue-services
[3] https://www.neanias.eu/

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

semantic vocabularies[4]. The activity of checking how this integration could be done has already started, however progress on the EOSC side is currently not defined well enough to provide a clear connection at this stage of the project.

2.2. VO Services portfolio

This sub-section reports on the status of activities for actually providing resources to create a portfolio of services dedicated to astronomy within the EOSC.

These activities are at an early stage, given the complexity of the EOSC architecture. The EOSC-Hub Week 2020 showed an inclusion plan for national, cluster and domain managed repositories. In CEVO we are following the evolution of that harvesting scenario while keeping awareness of other astronomy efforts, like the Europlanet VESPA containerised solution using the DaCHS resource and service software.

The idea of operating a custom marketplace for astronomy is probably better pursued in connection with ESCAPE WP3. This solution will be followed up.

2.3. EOSC Hybrid Cloud contribution

Another activity meant to integrate the VO architecture with the EOSC is through the federation of Research Infrastructure (RI) and connecting them to the EOSC itself. Here, the work has progressed in connection with ESCAPE WP5 ESFRI Science Analysis Platform in two main directions.

On the one hand the UEDIN partner who leads the interactive data analysis work in WP5, is developing components to discover and access data analysis facilities based on agreed standardised APIs between the facility providers. To do so they are working on developing use cases, notebooks and containers for the WP5 platform. Specific requirements for astrophysics sub-domains (time domain, radio astronomy, …) are collected and worked upon by all WP4 partners.

On the other hand work is progressing for a possible implementation of the IVOA VOSpace standard [6], the IVOA interface to distributed storage, to contribute as a service to the ESCAPE/EOSC Cloud solution, possibly on top of the ESCAPE Data Lake (WP2). A solution has been identified, both by INAF and UEDIN, to test VOSpace on top of RUCIO. It will be discusssed with WP2.

[4] https://ivoa.net/rdf/

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

2.4. Service containerisation

A final activity taken up to integrate and contribute to the EOSC cloud investigated the containerisation of web applications and other VO software to be integrated into the ESCAPE software repository. This activity will see a follow-up with WP3 Open-source scientific Software and Service Repository, and the harmonization of these containers with the science platform will see an interaction among WP4, WP5 and WP2 in the second half of the project.

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

3. ESCAPE cross-WP interactions

Progress on the integration activity described above has involved building connections between the WPs of the ESCAPE project.

Activities led to provide software packages and services to the community through a FAIR repository solution, which complement CEVO Task 4.1 to connect services directly to EOSC through the VO, are managed in ESCAPE WP3 Open-source scientific Software and Service Repository. Task 4.1 got in connection with WP3 to contribute requirements to the ESCAPE repository. That repository, as shown during the EOSC-Hub Week 2020 (Figure 5), should flow into the general EOSC repository. The WP3-WP4 collaboration should facilitate the contribution of VO resources to the ESCAPE repository and requirements gathering on media type annotations from the VO scenario.

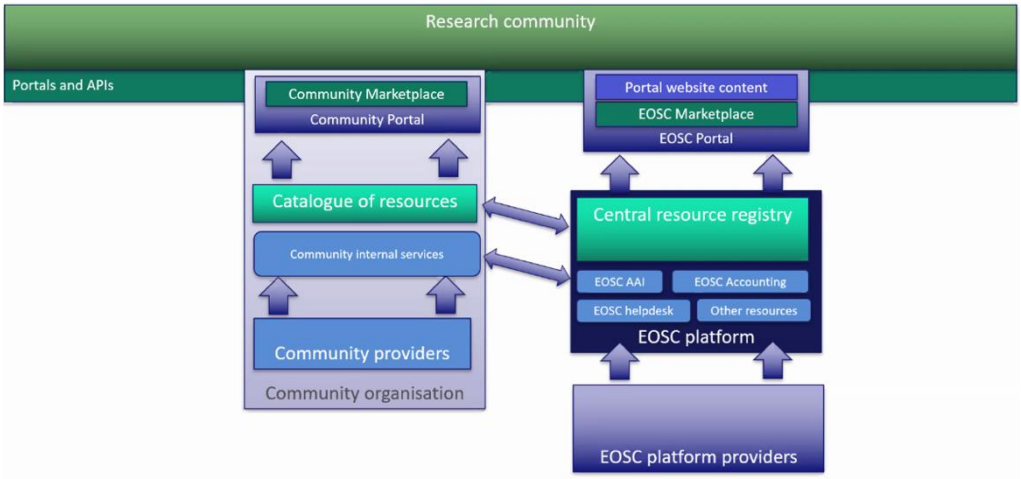

Figure 5: EOSC architecture's foreseen catalogue integration. O. Appleton presentation at the EOSC-hub Week 2020.

Another point of connection is between CEVO EOSC integration activities and WP5 on containerisation and hybrid cloud contributions. On the CEVO side this means porting IVOA derived technologies and having astronomy RIs implement or use them, while on the WP5 (ESAP) side it touches the integration of the Science Platform with the data resources. A specific aspect connected to this integration is the discussion of the AAIs (Authentication and Authorization Infrastructures) and how they affect the existing landscape of VO resources and services.

Discussions held during the CEVO Technology Forum 1 and the ESCAPE Progress Meeting proved useful not only to make the above connections clear, but also to understand how the ESCAPE efforts in federating its RI resources could connect with the EOSC platform.

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

4. Interaction with EOSC-related projects

In addition to connecting internally with other ESCAPE WPs and following the evolving FAIR best practices (e.g. through the RDA FAIR Data Maturity Model WG), our activities for integrating the VO expertise within the EOSC landscape have also involved the interaction with other EOSC-related projects, particularly FAIRsFAIR and FREYA.

Mark Allen (CNRS-ObAS) is one of FAIRsFAIR Champions. He participated actively in the Secong FAIRsFAIR Synchronisation Force Workshop, which was held as a series of sessions in April-June 2020, during which he represented ESCAPE. The meeting was aimed at identifying the on-going activities related to the Action Plan of the "Turning FAIR into Reality" report, and to identify gaps and possible additions to the Action Plan. Another FAIRsFAIR activity in which CEVO participated actively is the two workshops on metadata catlogue integration (9 September and 11 October 2020), and its follow-up DDI-CODATA Workshop (20 November 2020). This allowed us to explain the IVOA metadata framework, and to participate in the discission on the possible generic solutions and how disciplinary frameworks could interface with them.

CEVO participated in a meeting organised by FREYA on persistent identfiers in research disciplines (5 August 2020), which allowed us to present the way the IVOA defines and uses identifiers.

One also has to cite here the active participation of CEVO in the tests and discussions of the RDA FAIR Data Maturity Model. The tests were performed jointly by Task 4.1 and Task 4.2: they enabled an assessment of how the IVOA framework was aligning with the proposed RDA Maturity Model and to identify discrepancies. The conclusion is that we can live with the maturity criteria, but that there are issues with the priorities: in particular, there is no "Essential" priority on Interoparbility, whereas it is an essential element for astronomers, hence in the VO. The tests and the conclusions from them, were reported during the Technology Forum and the IVOA meeting in May 2020. This work brings useful information for the integration of the VO-enabled resources in EOSC, since the FAIR Metrics proposed by the EOSC FAIR WG are a subset of the RDA FAIR Data Maturity Model.

Attendance at the EOSC related meetings has been valuable both to showcase astronomy interoperability solutions, to present the developing astronomy requirements for EOSC, to understand the requirements and solutions from other research domains, and to get knowledge of the work performed by other projects and how astronomy could fit in.

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

5. VO architecture integration analysis

Considering the progress shown in Sections 2 and 3, the integration of the IVOA-based open architecture of standards for astronomy into the connected ecosystem of data resources and services in the EOSC is identified as being mostly feasible, but there are still challenges that need to be tackled to actually bring this integration to a mature level.

A definitely positive conclusion regards the high level description of data holdings and browser based services, leveraging the heritage and work of communities like the Open Archive Initiative, foundations like DataCite, DOI, ORCID and similar. The careful initial choice of the IVOA community to base its registry of resources on the generic OAI-PMH protocol, and to include the Dublin Core as mandatory elements of the metadata schema, has made it relatively simple to map the basic annotations and concepts of the IVOA into metadata models and universal identifiers. This made feasible the integration of the IVOA Registry into the EOSC repository provided by EUDAT that also relies on OAI-PMH and DataCite.

The challenge is more on the machine actionability of the discovery and access services attached to the data resources. The VO Registry uses standardised URIs to annotate the standard interfaces as well as to identify the relationships among resources. It has discipline-specific metadata for the coverage (e.g., in spatial and temporal coordinates) and scope of the data resources. EOSC, being the portal environment meant to be more general and abstract in terms of research domains, has a more abstract approach to data access and thus makes it difficult to reproduce the machine actionability level available in the IVOA. On the other side, IVOA, having developed its own URI schema for identifiers (IVOID) lacks somewhat on the take up of more general identifiers for its resources, i.e. the DOI nowaday more commonly used (although DOIs are commonly implemented by data providers, in particular for citation in publications). Another example of a difference emerging from the domain versus general repository behaviour are the tesselletion solutions (HiPS [8] and MOC [9] IVOA standards, based on HEALPix) provided by the VO for sky positional and time coverage of the observations. Those are a valuable support in filtering the growing number of registered available resources and could be a technical feature to be reported alongside the other metadata also in a general portal solution like the EOSC one, being the natural counterpart of the geolocation of general web or earthbound resources.

A more challenging integration task has been comparing the user authentication and authorization (A&A) solutions available within the astrophysical community and the interface architectures made available by the projects developing the EOSC.

On the one hand, the standardisation efforts within IVOA regarding A&A have not seen a high level of implementation. This is due to the *public* approach to data and service delivery that

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

represents the majority of VO registered holdings, as well as to the lack of a safely funded, globally operating identity provider. Moreover, IVOA standards, and VO contents consequently, deal more with resource filtering, operations on catalogues (i.e. metadata annotated listings of astrophysical sources) and direct access to datasets, with a relatively minor support for providing computing operations. This means that, besides being only partially concerned by data access policies because of the *public data and resources* approach, the VO community has, up to now, not been concerned with accounting, because storage and computational resources made available to provide astrophysical data holdings for the community are usually enough to make them available and useable.

On the other hand, FAIRness and Openness requirements in the VO architecture were, since the beginning, answered by interface protocols that were meant to be machine consumable. Thus the A&A stress was more on the credential delegation technical aspects than the actual identification of the user and the permissions given to them.

For these reasons, the AAI (Authentication and Authorization Infrastructure) solutions being developed by the EOSC projects were initially not well aligned with the requirements of the VO data providers: because on the one hand they pointed towards a licensing system where the only possible solution for most heritage data was a plain CC-0 (and that also felt difficult to attach in some cases), and on the other hand they lacked the machine actionability that full credential delegation would allow.

Things have progressed in that respect since the start of the ESCAPE project, which can be seen by comparing the outcomes of EOSC Symposium (November 2019)[5] with those of the EOSC-Hub Week 2019 (April 2019)[6] on the AAI aspects. Convergence and interoperability of the main A&A solutions provided by the EOSC services (EGI Check-In, EUDAT B2ACCESS, GEANT eduTEAMS) has been provided to the benfit of the users. First discussions and awareness of machine-to-machine credential delegation were raised.

There is also a difference in the way data resources and services are included in the VO Registry (and thus made available to VO-aware users and tools) with respect to what happens currently in the EOSC onboarding scenario. On the IVOA side a data/service provider has to be aware of the available typed-resources and how to describe them. The steps to perform then to publish a resource can be summarized as:

- identify (or implement) a *publishing registry* and create an *authority* resource in it. This means providing an identifiable (it will have its own IVOID) root resource under which all the resources provided will be placed and could be uniquely identified by their ID;

[5] https://www.eoscsecretariat.eu/eosc-symposium
[6] https://www.eosc-hub.eu/events/eosc-hub-week-2019

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

- generate the appropriate (VO)Resource descriptions for its data collections and services and publish them through the registry.

The *publishing registry*, exposing its OAI-PMH interface, will be harvested by the *full registries* that form the actual resource repository known as the IVOA Registry. The data collections and services are made available to the VO community when the publishing registry is harvested. The above steps can be fully automated and human interaction is more on the operational side to check the health status of the Registry as a whole, following up on corrupted records, invalid resource descriptors, and unreachable resources.

Figure 6 illustrates the way the IVOA Registry works. It shows how the OAI-PMH interface allows for resource harvesting among the different registry instances and how the dedicated RegTAP solution can be used by applications (and proves more actionable). RegTAP [5] defines an interface for s earching the resource metadata qtored in the registry using the IVOA Table Access Protocol. The same RegTAP interface is the one that facilitated the IVOA Registry mapping to the B2FIND repository.

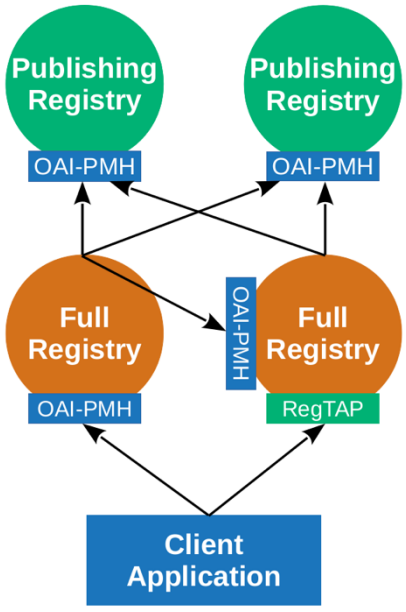

Figure 6: IVOA Registry ecosystem: publishing, harvesting, cross-domain interoperability and domain specific interface solutions.

On the EOSC side, the service onboarding solution is based on human interaction through web forms and contact points until the resource/service is validated and released through the portal.

This difference might be a temporary one while EOSC tools and architecture become more mature. It cannot be considered a challenge from the technical point of view because, as reported above for the registry integration, the metadata models and FAIR approach are comparable. It is, however, a potential hurdle which could produce a slower take up of astrophysical resources within the EOSC

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

ecosystem when progress will have been made on the service onboarding in EOSC.

Another challenging difference between the VO landscape and the EOSC proposed architecture is related to how service provision is handled. On the IVOA side, the contribution of data collections, services and other resources, is provided by the deployment of resources and services following the agreed standards (and often by autonomous systems). This leads to an ecosystem where the validation of the deployed resources in the VO is an operational activity focusing only on providing reliable resources to the end user.

In the EOSC architecture, where the governance is different, the data providers have to go through a different process, where metadata standards and interfaces are not automatically a means of creating an interoperable solution, and thus the onboarding process fills this gap by a preliminay validation of the resources that tends more to a preservation of the content and access points to information holdings, rather than being a check of actionable interoperability.

The challenge here is to find a way to automatise as much as possible the on-boarding process for a resources that are already registered in the IVOA, to reduce the amount of additional effort required on the providers side: the domain-specific technical one for the VO landscape, the general-purpose EOSC repository controlled validation on the other.

Last but not least important, a topic where integration is required is the interpretation on the FAIR principles applied to data resources and services already living in the VO landscape. This is an activity that crosses borders in CEVO between tasks 4.1 and 4.2 and a direct connection to FAIR initiatives at the EOSC, but also global, level.

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

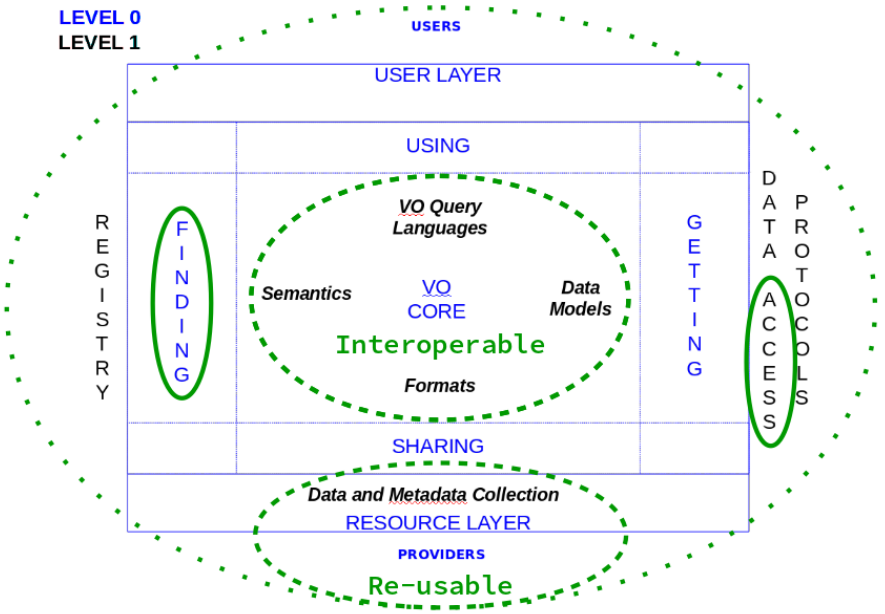

Figure 7: A basic mapping of the FAIR principles main concepts on top of the IVOA architecture.

The IVOA vision, which has been developed starting in 2001-2002, was already following the FAIR principles, since its aim was to enable astronomers to find, access and interoperate data, a reuse capability being provided by the usage of a common format including data and metadata, and to provide seamless access to astronomical resources. of the concept of openness of data resources and services and always workied keeping in mind machine actionnable interoperability. It can be seen from the IVOA architecture schema shown in Figure 7. IVOA developed a disciplinary framework of FAIR practices which fulfills the community science needs but does not follow step by step all the fine grained details of the FAIR principles, as for instance defined in the RDA FAIR Data Maturity Model [10].

The activities to be brought on, in connection to EOSC and the other EOSC-related project will require testing the VOFAIR-ness with respect to EOSC FAIR Metrics and to continue on the work already started for the RDA FAIR Data Maturity Model.

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

6. Open questions and vision

The integration of the VO architecture into the EOSC seems not an impossible task. Still it requires to tackle a few challenges that are currently discussed within the EOSC development bodies and projects.

From the point of view of the astrophysical community the main concern seems to be about how the existing machine-actionable and data providing semi-automated solution can be replicated in the EOSC without having to duplicate efforts in management and research technologies.

The effect of facing a multi-disciplinary environment has proven a challenge mostly because the astrophysical community concentrated on its science needs with respect to data sharing and reuse, which led it to skip hurdles like the authentication and authorization of users to access discovery and retrieval service. This allowed for a direct path in maturing a re-usable and interoperable paradigm of the VO that consequently brought into its architecture the findability, accessibility and interoperability aspects, thus completing the FAIR principles scenario in an operational framework at an early stage.

Within ESCAPE, considering the progress made so far, the goal for the second reporting period of the project would be to investigate a solution for ESFRIs and RIs within the EOSC (and/or a domain specific, ESCAPE provided, cloud) to build their data resources and services adopting a VO aware vision and, as seamlessly as possible, find those same resources available in EOSC itself and able to use the other resources provided within the EOSC such as computing resources. To do so the challenges to tackle are currently the following:

- attach data holdings to resource descriptions (not only PIDs for metadata);
- provide relationships among data resources and attached services;
- test, and make easier, the integration of VO-enabled data providers managed resources in the portal;
- deal with current AAI services and solution without preventing public content access;
- work on actionability for domain specific services and annotations through the EOSC portal;
- evaluate the position of VO-enabled data with respect to EOSC FAIR Metrics.

All the above challenges, building on the activities reported in this deliverable, will be tackled during the remaining part of the project. How far they proceed will depend on the actual status of the EOSC development and finalisation, and on the amount of resources necessary with respect to the tasks to perform.

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

Appendix: Participation in EOSC related meetings

Participation in EOSC related meetings.

Meeting	Location	Date
EOSC-hub Week 2019	Prague	10-12 April 2019
Research Data Alliance Plenary 14	Helsinki	23-25 October 2019
EOSC Symposium	Budapest	26-28 November 2019
Reasearch Data Alliance Plenary 15	(online)	18 March-10 April 2020
FAIRsFAIR Second Synchronisation Force Workshop	(online)	Series of workshops April-June 2020
EOSC-hub Week 2020	(online)	18-20 May 2020
FREYA Workshop “Persistent identifiers in Research Disciplines”	(online)	5 August 2020
FAIRsFAIR Workshop “Metadata catalogues integration for Interdisciplinary Research”	(online)	11 September 2020
Second ESFRI RIs-EOSC Workshop "Research Infrastructures shaping EOSC"	(online)	6-7 October 2020
FAIRsFAIR Workshop “Metadata Catalogue Integration	(online)	9 October 2020
EOSC Governance Symposium	(online)	19-22 October 2020
Research Data Alliance Plenary 16	(online)	9-12 November 2020
DDI-CODATA Workshop	(online)	20 November 2020

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union’s Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

References

[1] ESCAPE Deliverable 4.1 “Detailed WP4 Project Plan”, M. Allen & al., 12.07.2020, https://projectescape.eu/deliverables-and-reports/d41-%E2%80%93detailed-wp4-project-plan

[2] “Integrating the VO Framework in the EOSC”, M. Molinaro & al., ADASS XXIX proceedings, volume in preparation, pre-print version available at https://arxiv.org/abs/1911.08205

[3] ESCAPE Milestone 20, “Progress & priotities at IVOA - 1”, M. Allen & al., https://projectescape.eu/Milestone%2020

[4] ESCAPE Milestone 21, “Progress & priorities at IVOA - 2”, M. Allen & al., https://projectescape.eu/deliverables-and-reports/milestone-21-report-progress-and-priorities-ivoa-2

[5] “IVOA Registry Relational Schema Version 1.1”, M. Demleitner & al., IVOA Recommendation, 2019, https://ui.adsabs.harvard.edu/abs/2019ivoa.spec.1011D

[6] “VOSpace Version 2.1”, M. Graham & al., IVOA Recommendation 2018, https://ui.adsabs.harvard.edu/abs/2018ivoa.spec.0621G/abstract

[7] “Adding a service to EOSC - A short study about how to propose its service(s) through EOSC”, André Schaaff, talk at the ESCAPE WP4 Technology Forum 1, https://indico.in2p3.fr/event/20005/#18-adding-a-service-to-eosc

[8] “HiPS - Hierarchical Progressive Survey Version 1.0”, P. Fernique & al., IVOA Recommendation, 2017, https://ui.adsabs.harvard.edu/abs/2017ivoa.spec.0519F

[9] “MOC - HEALPix Multi-Order Coverage map Version 1.1”, P. Fernique & al., IVOA Recommendation, 2019, https://ui.adsabs.harvard.edu/abs/2019ivoa.spec.1007F

[10] “FAIR Data Maturity Model: specification and guidelines”, FAIR Data Maturity Model Working Group, 2020, Research Data Alliance, https://doi.org/10.15497/RDA00050

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union’s Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.