A glimpse into the ESCAPE project.

Patrick Fuhrmann(DESY) for the ESCAPE Collaboration

With contributions by G. Lamanna(LAPP), S. Campana (CERN), Z. Meyer (ASRON-NWO-I), K. Graf (FAU), S. Serjeant (OU), M. Allen (ASTRO), M. van Haarlem (ATRON-NWO-I)

EOSC Hub Week, Prague Czech Republic

10 April 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.

ESCAPE in a nutshell

ESCAPE convenes a large scientific community

- **31** partners (including 2 SMEs)
- **7** ESFRI projects & landmarks: CTA, ELT, EST, FAIR, HL-LHC, KM3NeT, SKA
- **2** pan-European International Organizations: CERN, ESO (with their world-class established infrastructures, experiments and observatories).
- **4** supporting ERA-NET initiatives: HEP (CERN), NuPECC, ASTRONET, APPEC
- **1** involved initiative/infrastructure: EURO-VO (Virtual Observatory)
- **2** European research infrastructures: EGO and JIVE-ERIC
- Budget: **15.98 M€**
- Started: **1/2/2019**
- Duration: **42** months (end date 31/7/2022)
- Coordinator: **CNRS (**Centre national de la recherche scientifique**) - LAPP - G. Lamanna**

Home page: https://escape2020.eu ; Twitter: @ESCAPE_EU

Funded by the European Union's Horizon 2020 - Grant N° 824064

ESFRI: Strategy Forum for RI's

ESFRI's mandate

- to support a coherent and strategy-led approach to policy making on research infrastructures in Europe
- to facilitate multilateral initiatives leading to a better use and development of research infrastructures
- to establish a European Roadmap for research infrastructures (new and major upgrades, pan-European interest) for the next 10-20 years, stimulate the implementation of these facilities, and update the roadmap as needed
- to follow-up on implementation of ongoing ESFRI projects after a comprehensive assessment, as well as the prioritisation of infrastructure projects listed in the ESFRI Roadmap

Funded by the European Union's Horizon 2020 - Grant N° 824064

The inevitable Partner Logo Slide

Funded by the European Union's
Horizon 2020 - Grant N° 824064

ESCAPE Science

- Multi Messenger Astronomy
  - Radio
    - SKA (Scare Kilometre Array)
    - JIVE VLBI (Very large Baseline Instrument)
  - Visible Light
    - European Extreme Large Telescope (ELT)
    - European Solar Telescope (EST)
  - Gamma Rays
    - CTA
  - Cosmic Rays: Neutrinos
    - KM3Net
  - Gravitational Waves
    - EGO-VIRGO
- High Energy Physics, Matter
  - HL-LHC
    - High Energy Particle
  - FAIR
    - High density exotic matter physics

Funded by the European Union's Horizon 2020 - Grant N° 824064

ESCAPE High Level Goals

- ESCAPE aims to address the Open Science challenges shared by ESFRI and pan-European research infrastructures in astronomy and particle physics.
- ESCAPE actions will be focused on developing solutions for the large data sets handled by the ESFRI facilities, by
  - connecting ESFRI projects to EOSC,
  - fostering common approaches to implement open-data stewardship,
  - and establishing interoperability within EOSC as an integrated multi-messenger facility for fundamental science.

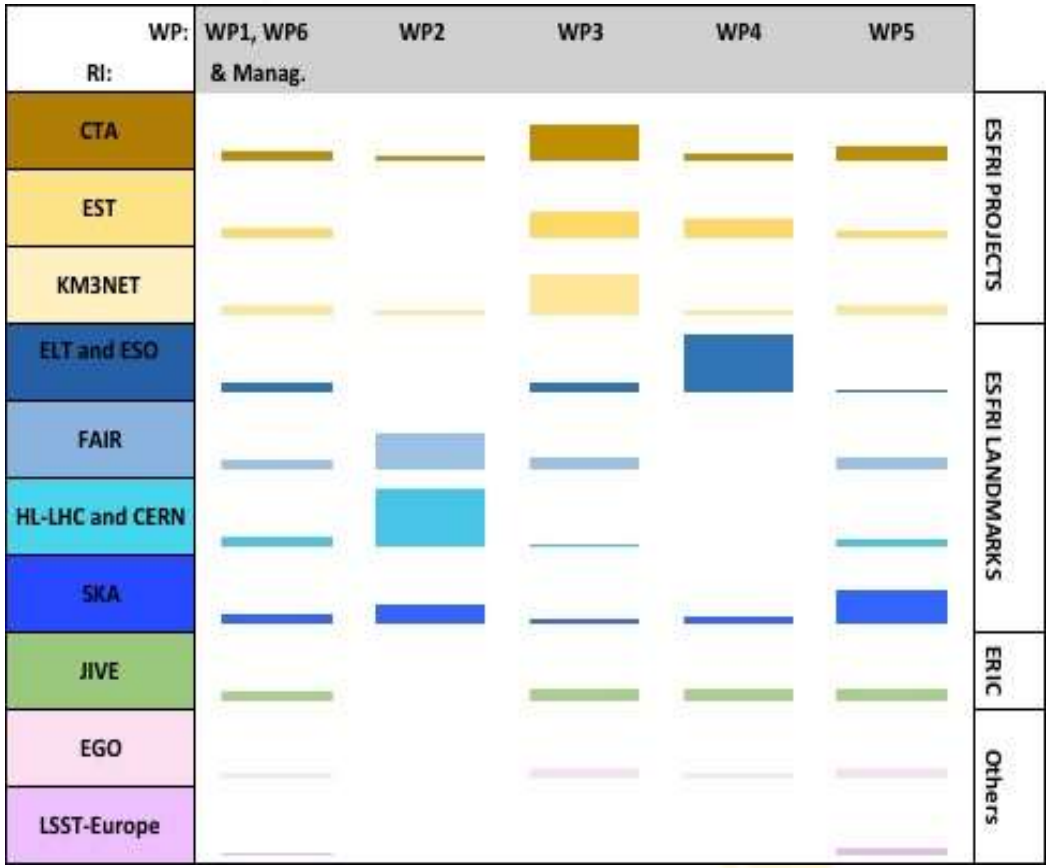
Funded by the European Union's Horizon 2020 - Grant N° 824064

The work package structure

- WP1 MIND: Leader: Giovanni Lamanna, LAPP-CNRS
  - Management and policy
- WP2 DIOS: Leader: Simone Campana, CERN
  - Contribute to the federation of global EOSC resources through an implementation of the Data-Lake concept (evolution of WLCG and other ESFRI RIs computing models) to manage extremely large volumes of data up to the multi-exabyte scale
- WP3 OSSR. Leader: Kay Graf, FAU
  - Support for "scientific software" as a major component of the ESFR-RI “data” to be stored and displayed in EOSC via dedicated community-based catalogues. Implementation of a community-based approach for the continuous development of shared software and for training of researchers and data scientists.
- WP4 CEVO. Leader: Mark Allen, CDS-CNRS
  - Extend FAIR standards, methods, tools of the Virtual Observatory to a broader scientific context; demonstrate EOSC's ability to include existing platforms.
- WP5 ESAP. Leader: Michiel van Haarlem, ASTRON-NWO (Deputy : Zheng Meyer)
  - Implementation of scientific analysis platforms enabling EOSC researchers to organize data collections, analyse them, access ESFRI's software tools, and provide their own customized workflows.
- WP6 ECO. Leader: Stephen Serjeant, Oxford Open University
  - Citizen Science, Open Science and Communication

Funded by the European Union’s Horizon 2020 - Grant N° 824064

Communties and workpackages

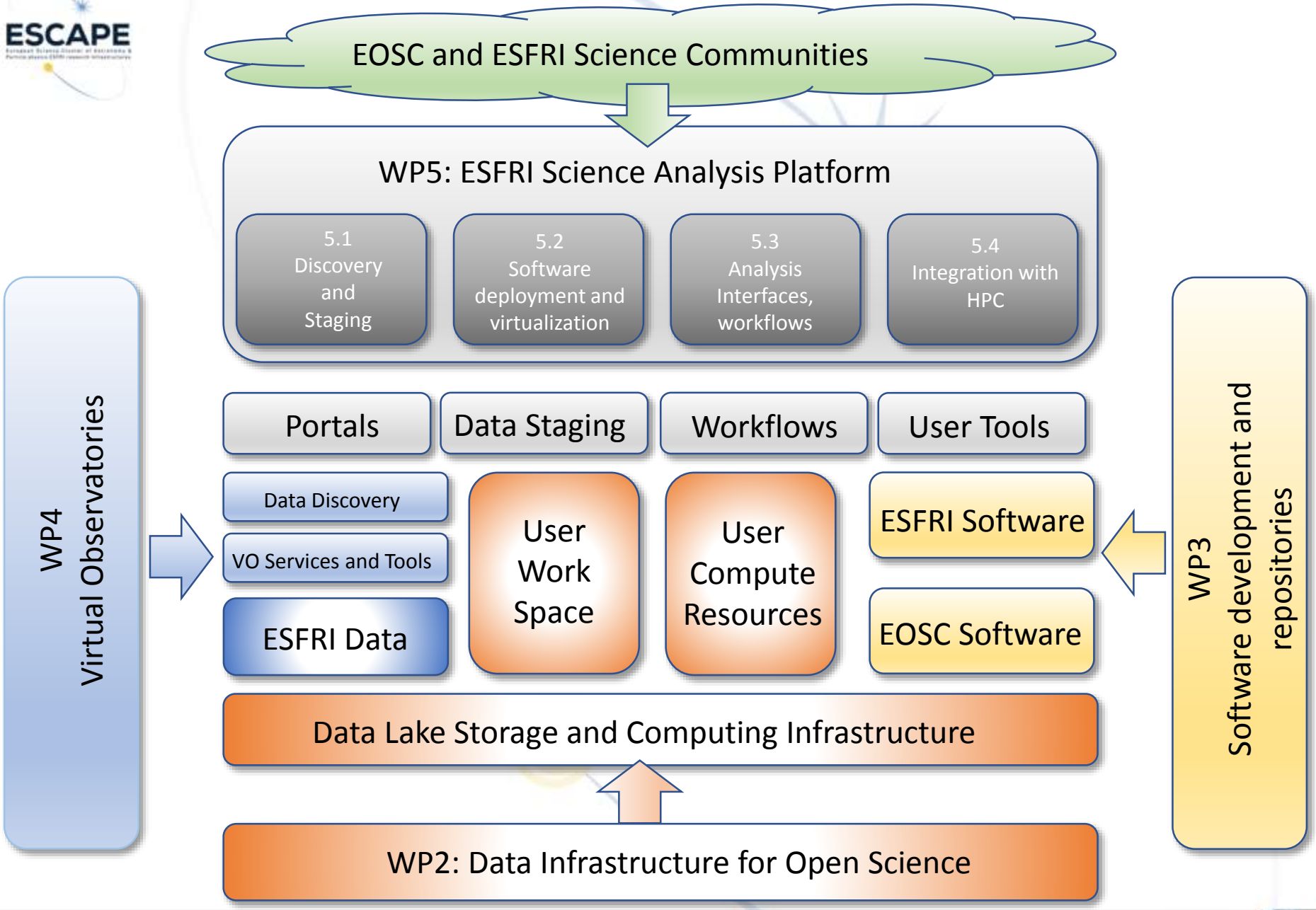


The allocated EU funded staff effort is proportional to the respective boxes' surface areas.

An heathy matrix:

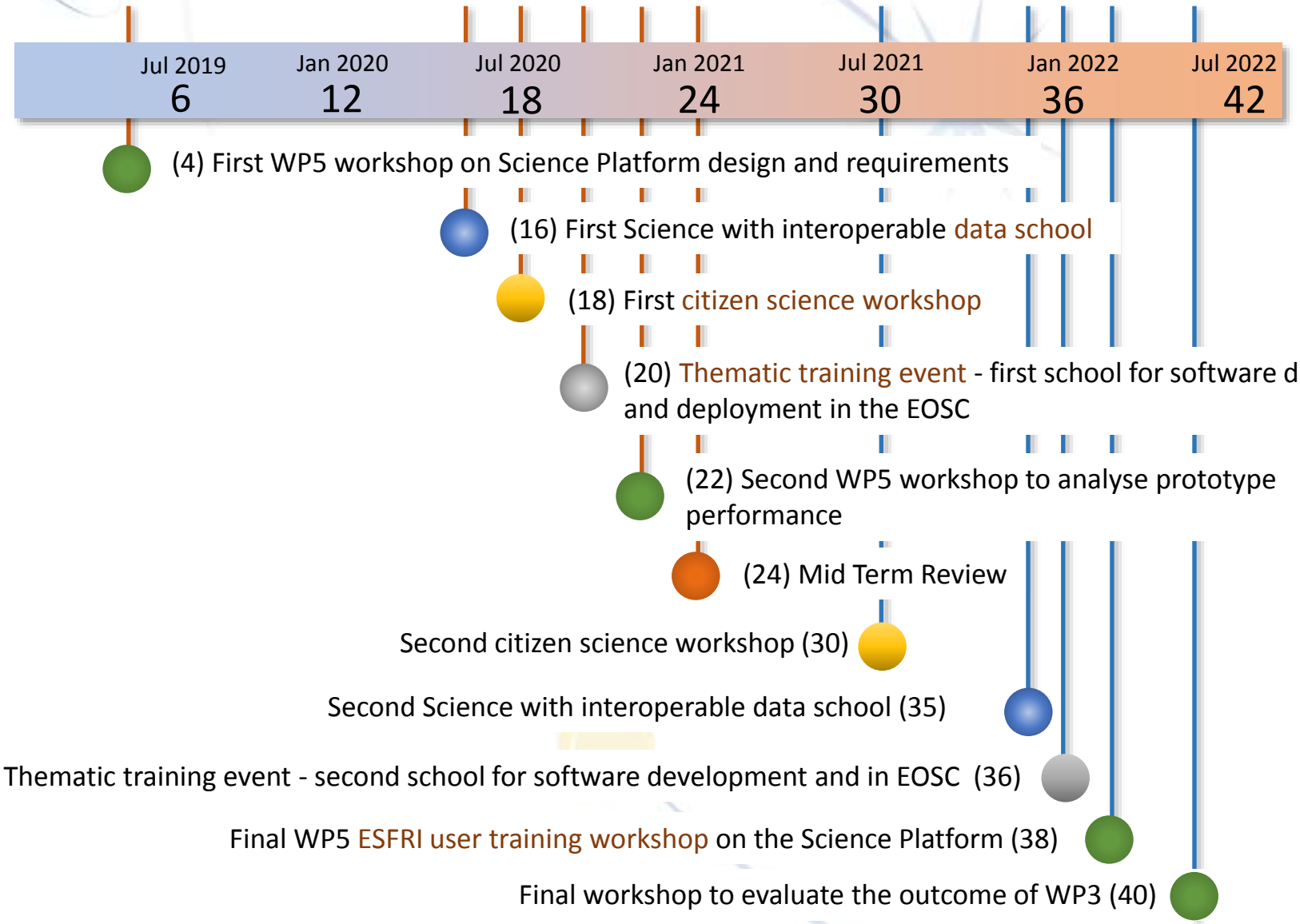
- Some clear priorities per each RI
- RIs' use-cases in almost all WPs
- Sub-sets of RIs driving a WP
- All RIs involved in the EOSC support

Funded by the European Union's Horizon 2020 - Grant N° 824064



Funded by the European Union's Horizon 2020 - Grant N° 824064

Workshops and Meetings



Funded by the European Union's Horizon 2020 - Grant N° 824064

WP1 : MINT

Management, Innovation, Networking and Dissemination Infrastructure for Open Science

CNRS (Centre national de la recherche scientifique**)** - LAPP

Management and policy

CNRS	**CERN**	**NWO-I**	**OU**	**FAU**

Funded by the European Union's Horizon 2020 - Grant N° 824064

WP1 Tasks

- Task 1.1: Governance, coordination and project management.
- Task 1.2: Dissemination, innovation and networking programme

Funded by the European Union's Horizon 2020 - Grant N° 824064

WP2 : DIOS

Data Infrastructure for Open Science

Contribute to the federation of global EOSC resources through an implementation of the Data-Lake concept (evolution of WLCG and other ESFRI RIs computing models) to manage extremely large volumes of data up to the multi-exabyte scale.

CNRS	CERN	NWO-I	FAIR	GSI	IFAE	INFN	SKA0	DESY	RUG	SURFSARA

Contributed by Simone Campana, CERN

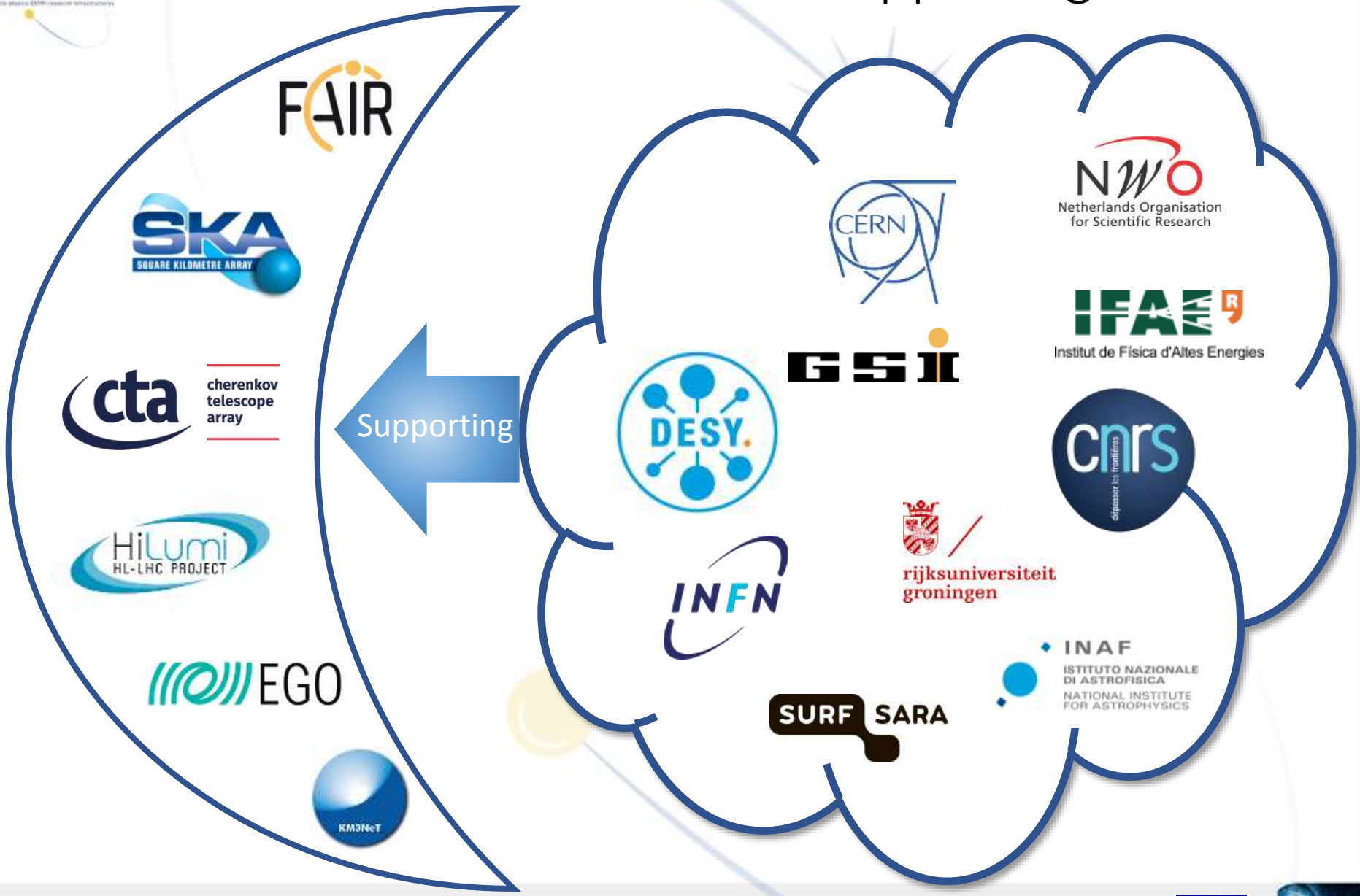
Funded by the European Union's Horizon 2020 - Grant N° 824064

Data Infrastructure for Open Science (DIOS)

- Goal: design, implement and operate a cloud of data services for open access and open science at the Exabyte scale
- The backbone of the Data Lake are well experienced large national data centers supporting the ESFRIs in ESCAPE
- The data lake will serve as underlying data infrastructure to manage and serve data to the user communities
- This solution will be proposed as key component of the future EOSC framework, supporting FAIR principles

Funded by the European Union's Horizon 2020 - Grant N° 824064

Involved sciences and supporting RI



Funded by the European Union's Horizon 2020 - Grant N° 824064

WP2- specific objectives

Prototype a reliable and scalable **federated data infrastructure.**

- **Stores and organizes scientific data (Findable) and enables the provisioning of data processing (Accessible)**
- Enables sciences to build open data repositories (Interoperable)
- In general, supports the world-leading data challenges of the Research Infrastructures in ESCAPE
- Ensure long term **data preservation** (Reproducible) at the infrastructure level.
- The Data Lake development leverages **collaboration** with and **integrates** the work from previous and ongoing frameworks
- **Computing** Interface and Scalability
- **Industrial** and **Commercial** involvement

Funded by the European Union's Horizon 2020 - Grant N° 824064

WP2- specific objectives (cont.)

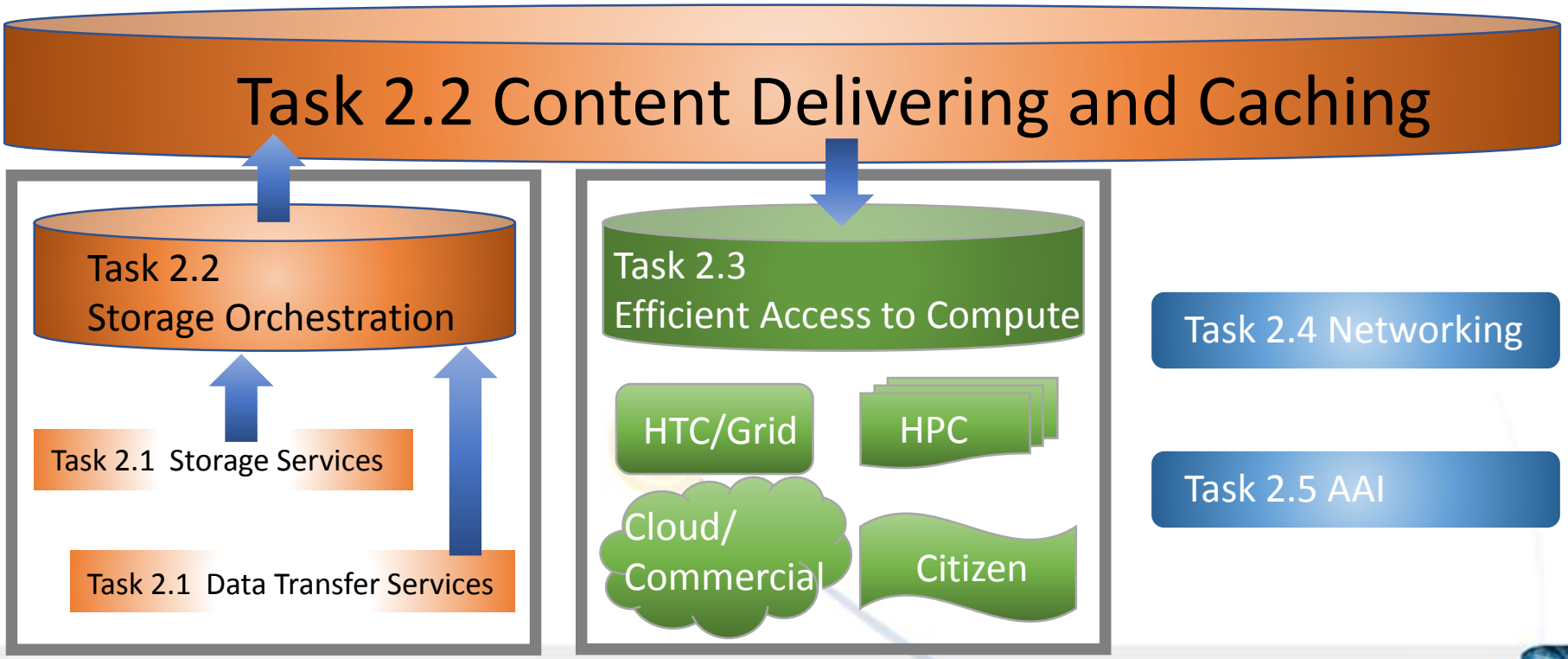
Create a cloud of data services, often referred to as a "*Data Lake*" by building on and integrating existing work from a variety of areas:

- Data services of ESFRI and other world-class RI.
- E-Infrastructures such as EGI, EU-DAT, PRACE, GEANT.
- Services from other H2020 E-INFRA projects
  - INDIGO-DataCloud, DEEP, eXtreme DataCloud.
- State of the art solutions in the appropriate areas.
- Particularly collaborating with advanced infrastructure projects like EOSC Hub.

Funded by the European Union's Horizon 2020 - Grant N° 824064

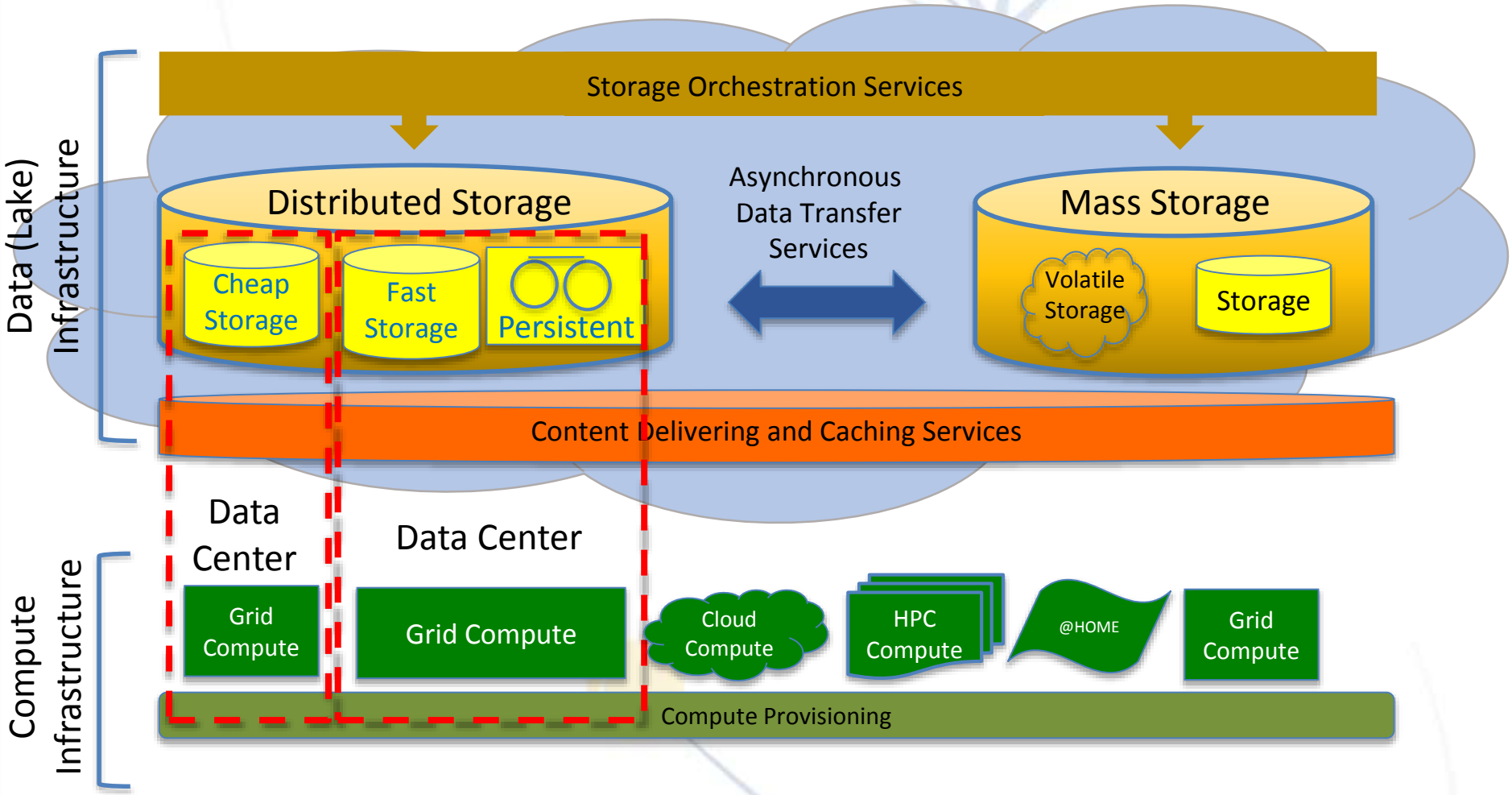
Structure of ESCAPE WP2

- Task 2.1 Data Lake Infrastructure and Federation Services. CERN (Xavier Espinal)
- Task 2.2 Data Lake orchestration service. DESY (Patrick Fuhrmann)
- Task 2.3 Integration with Compute Services. NWO-I-ASTRON
- Task 2.4 Networking. SKAO (Rosie Bolton)
- Task 2.5 Authentication and Authorization. INFN (Andrea Ceccanti)



Funded by the European Union's Horizon 2020 - Grant N° 824064

Data Lake strawman



Funded by the European Union's Horizon 2020 - Grant N° 824064

Collaboration on building the data lake and considerations on QoS in storage.

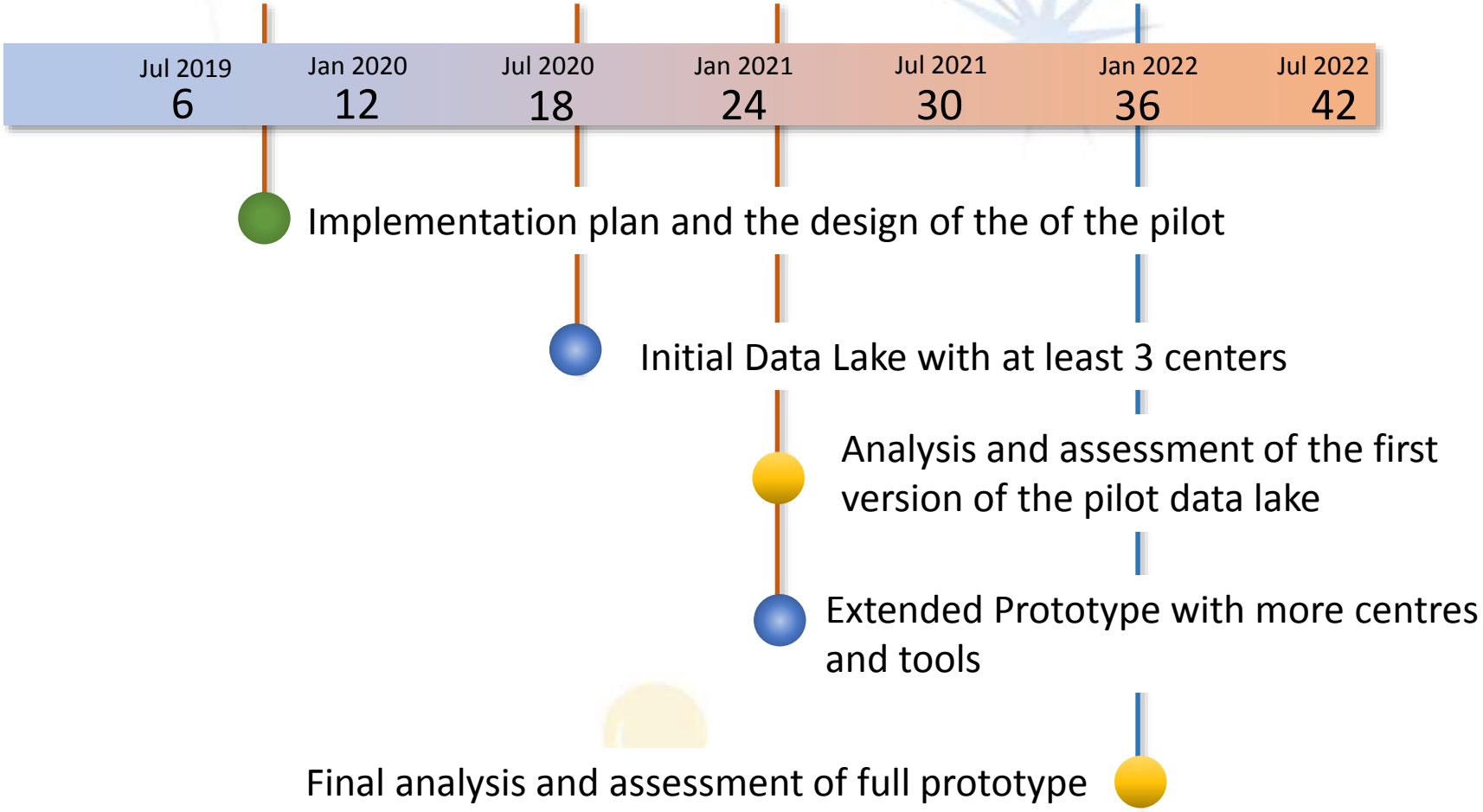
SNIA

RDA
RESEARCH DATA ALLIANCE

DOMA Working Groups
- Access
- QoS in Storage
- 3rd Party Transfer

Funded by the European Union's Horizon 2020 - Grant N° 824064

WP2 important milestones



Funded by the European Union's Horizon 2020 - Grant N° 824064

WP3 : OSSR

Open-source scientific Software and Service Repository

Support for "scientific software" as a major component of the ESFR-RI “data” to be stored and displayed in EOSC via dedicated community-based catalogues. Implementation of a community-based approach for the continuous development of shared software and for training of researchers and data scientists.

CNRS	CERN	NWO-I	FAU	AIP	GSI	IFEA	JIV EIRC	MPG
INFN	SKAO	UMC	CTA O	EGO	UNITOV	HITS	OROBIX	

Contribution by Kay Graf

Funded by the European Union’s Horizon 2020 - Grant N° 824064

E-OSSR Mission Statement

ESCAPE WP3 – E-OSSR:
Open-source scientific Software and Service Repository

Establish a foundation for **scientific software and services** as major components of the ESFR-RI **open science exploitation:**

- Implementation of a **community-based approach** for continuous development of shared software and services.
- Interoperability, software re-use and cross-fertilisation
- Open innovation environment for open standards, common regulation and shared (novel) software for multi-messenger & multi-probe data.
- Enabling **open science** via the exploitation of data across and beyond the participating communities following the FAIR principle.
- Training of researchers and data scientists.

Funded by the European Union's Horizon 2020 - Grant N° 824064

E-OSSR Work Program

- Partners from Astronomy, Astroparticle and Particle Physics
  - All ESFRI/RI participating in E-OSSR
- Work structured in five tasks:
  - Management activities, policy and support actions
    → link to EOSC entities
  - ESFRI Software and Services Collection
  - Common Approaches for Software and Services
  - Foundation of Competence for Software and Service Innovation
  - Repository Implementation and Deployment
- Exposition to and integration into the EOSC marketplace (catalogue of services) via **community-based repositories**;
  - Implementation of continuous deployment, exposure and preservation of software and services

Funded by the European Union's Horizon 2020 - Grant N° 824064

WP4 : CEVO

Connecting ESFRI projects to EOSC through VO framework

Extend FAIR standards, methods, tools of the Virtual Observatory to a broader scientific context; demonstrate EOSC's ability to include existing platforms.

CNRS	NWO-I	INAF	ESO	JIV ERIC	KIS	ORB
SKA O	UEDIN	UHEI	CTA O	EGO	INTA	

Contributed by Mark Allen

Funded by the European Union's Horizon 2020 - Grant N° 824064

WP4 Objectives

- Assess and implement the connection of **ESFRI** and other **astronomy research infrastructures** to the **EOSC** by the **Virtual Observatory**
- Refine and pursue implementation of **FAIR principles for astronomy** data via common interoperability standards - extending the VO to new communities
- Establish **data stewardship practices** for adding value to scientific content of ESFRI data archives
- Prerequisite: The inclusion of the data of the ESFRI facilities from astronomy and astro-particle physics in the VO is already well advanced thanks to the collaboration between ESFRI pathfinders and European VO teams in the ASTERICS Data Access, Discovery and Interoperability WP since 2015.

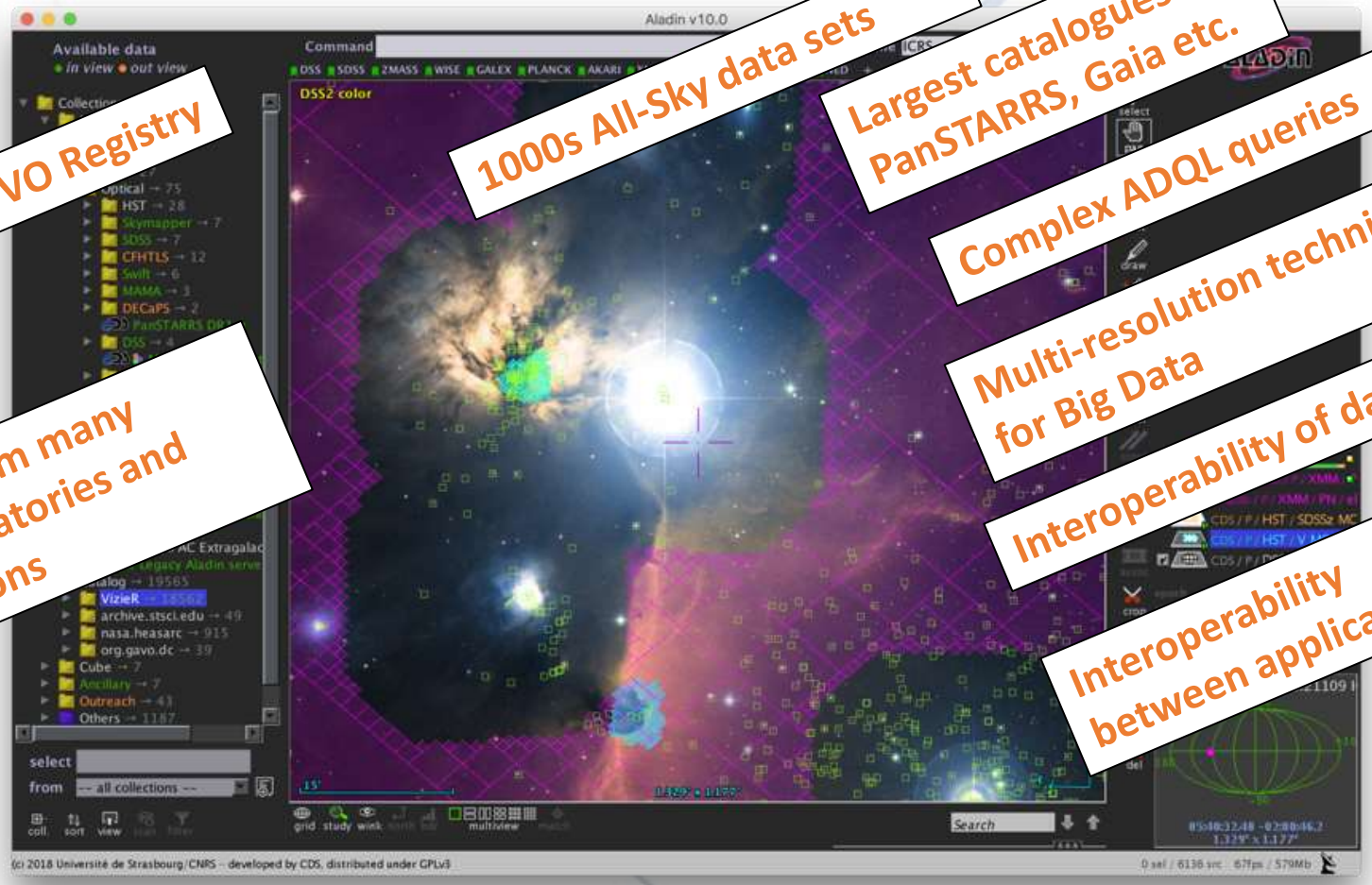
Funded by the European Union's Horizon 2020 - Grant N° 824064

What is the Virtual Observatory?

- **Operational framework** for interoperable access to astronomical data and services across all areas of astronomy
- Provides unique scientific capabilities, opening up new ways of using rich data in astronomy archives and services
- A pioneer of FAIR data sharing - **an existing global framework** – populated by major data providers (space and ground based) that is heavily used by the community *(e.g. Gaia data access is fully VO)*
- **Re-used and customized** by planetary science (EuroPLANET), atomic and molecular physics (VAMDC) and materials sciences (via RDA Working Group)

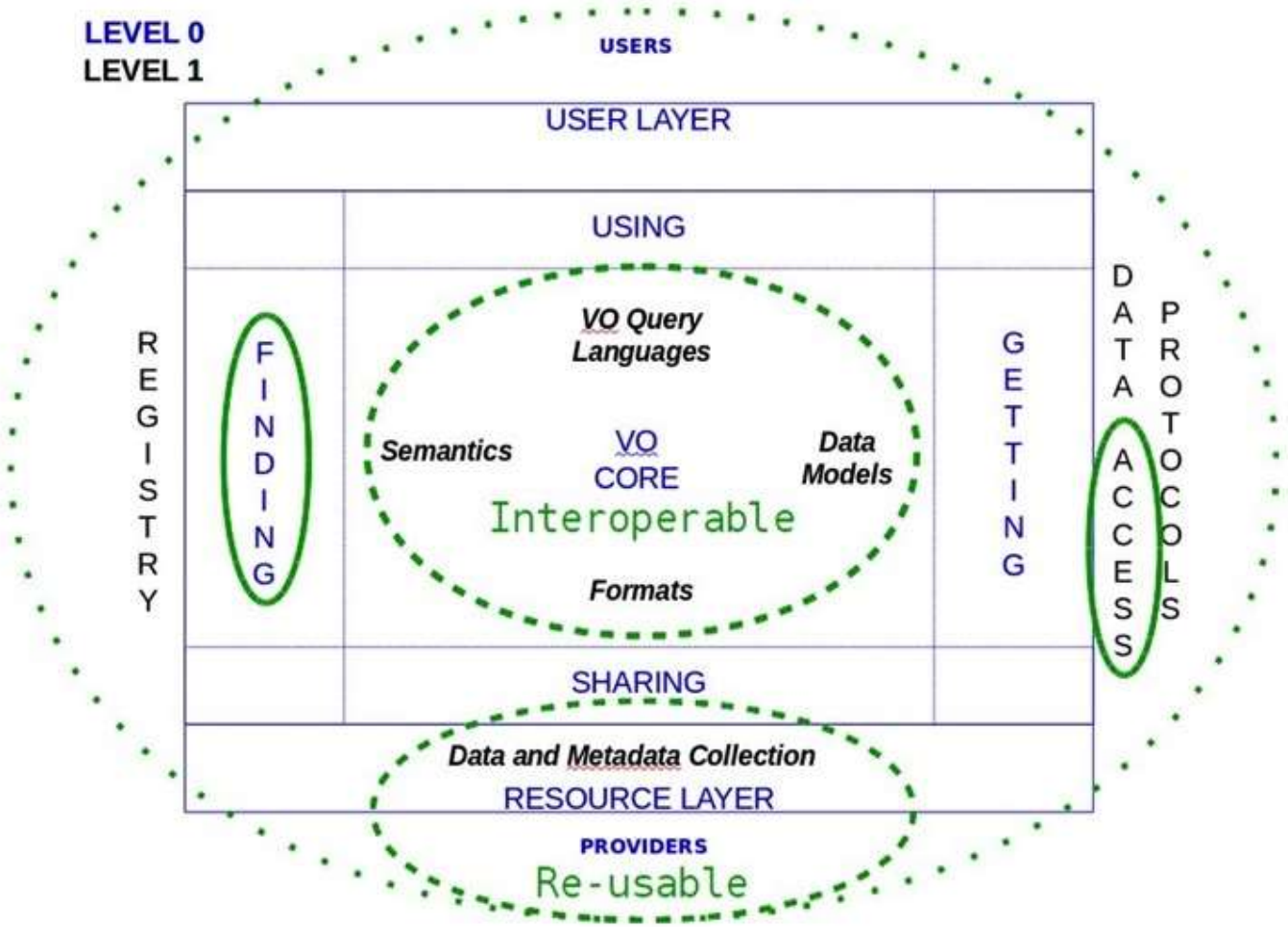
Funded by the European Union's Horizon 2020 - Grant N° 824064

One view of the VO from an application:



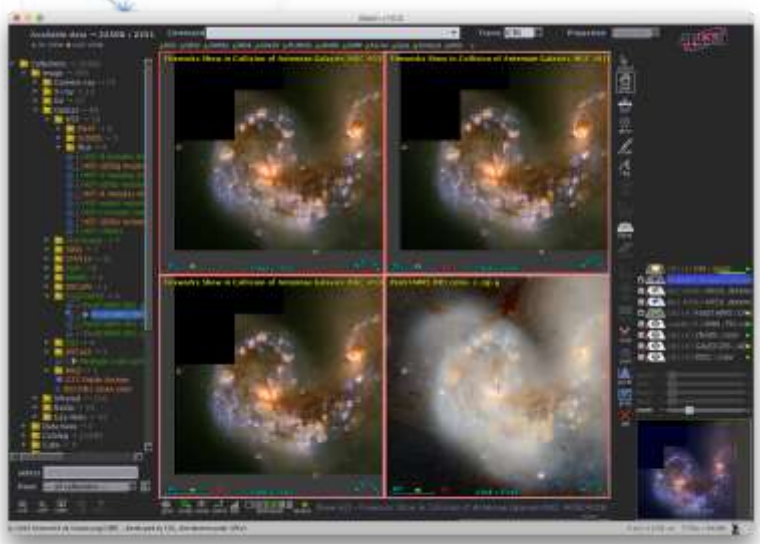
Funded by the European Union's
Horizon 2020 - Grant N° 824064

VO is FAIR

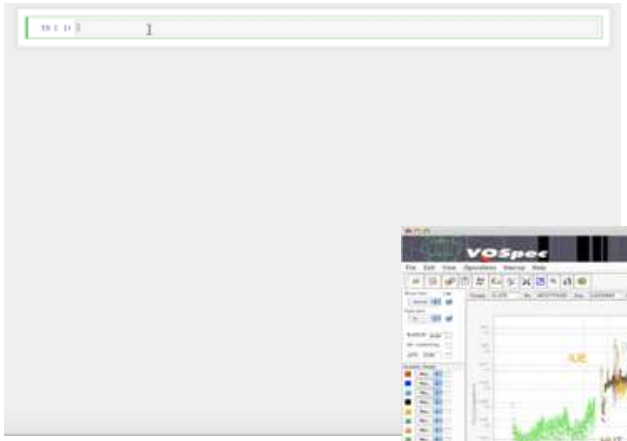


Funded by the European Union's Horizon 2020 - Grant N° 824064

Interoperable applications and services

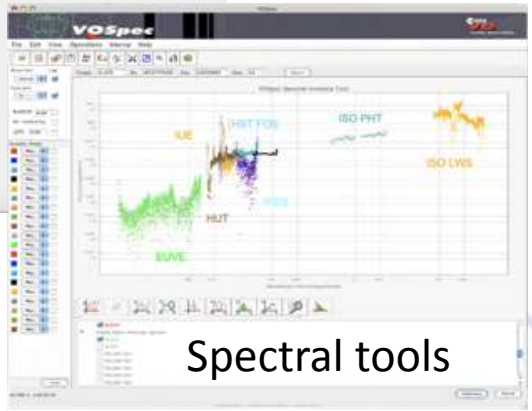


Aladin

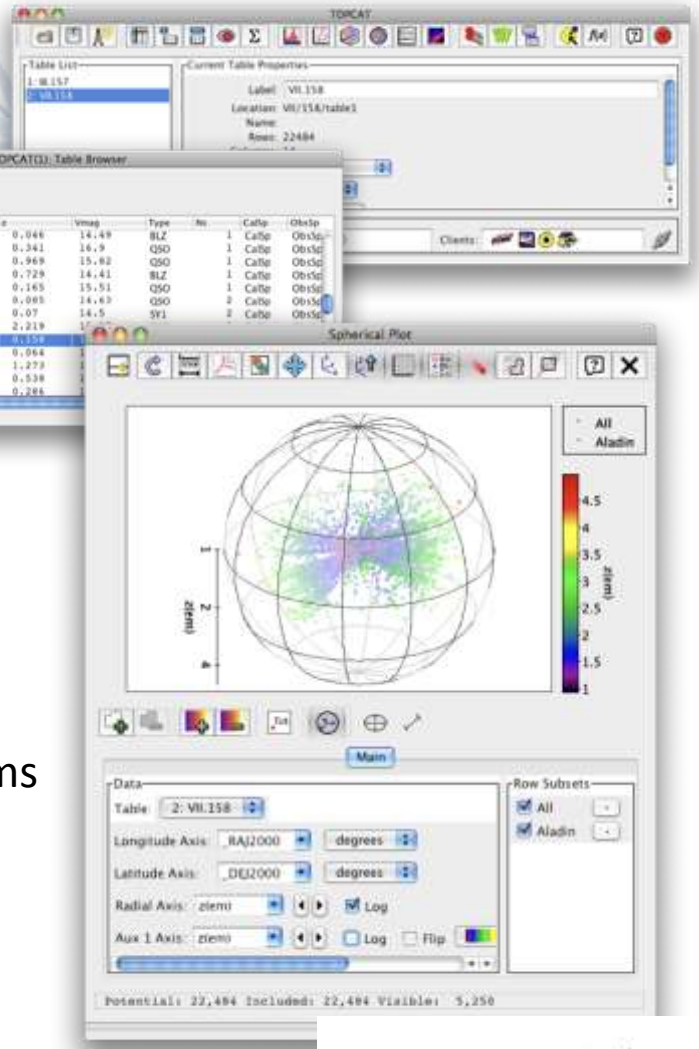


Notebooks

Your apps
& programs



Spectral tools



TOPCAT

Broadcast

Funded by the European Union's
Horizon 2020 - Grant N° 824064

Connecting ESFRI to the EOSC via the VO

In practice: ESFRI-VO-EOSC connection

- Inclusion of VO registry will be a key factor
- Implement FAIR principles via the use of common interoperability standards
- **VO next-steps** of connection to computing, and extension to new communities *(in particular EST)*
- Stewardship – technical and human. Adding value to the scientific content
- **Following all steps of EOSC evolution :**
  - EOSCpilot, EOSC-hub, eInfraCentral, …

Funded by the European Union's Horizon 2020 - Grant N° 824064

Expected Impacts

- Connection of the existing *(operational, heavily used, global)* VO framework to the EOSC
  - Astronomy specific services on EOSC
  - Contributing to the setting up of EOSC
- Training astronomers, ESFRI and RI on use of interoperable tools for science and data sharing
- Extension of the VO framework to new communities, and VO next-steps (computing, Bigger Data)
- Significant added value to ESFRI archive services *(e.g. machine learning classification)*

Funded by the European Union's Horizon 2020 - Grant N° 824064

WP5 : ESAP

ESFRI Science Analysis Platform

Implementation of scientific analysis platforms enabling EOSC researchers to organize data collections, analyse them, access ESFRI's software tools, and provide their own customized workflows.

Netherlands Organisation for Scientific Research

CNRS	CERN	NWO-I	FAU	INAF	FAIR	IFAE	JIV ERIC	KIS	SKA O
UCM	UEDIN	CTA O	EGO	CSIC	RUG				

Contributed by Michiel van Haarlem and Zheng Meyer

Funded by the European Union's Horizon 2020 - Grant N° 824064

WP5 Task Structure

- Task 5.1: Discovery and Staging
  - This task will provide users of the science platform with the capability to access and combine data from multiple collections and stage that data for subsequent analysis.
- Task 5.2: Software deployment and virtualization
  - This task will incorporate the work on the software repository described in WP3 and focus on tools and services to support the virtualization of relevant software packages and pipelines.
- Task 5.3: Analysis interfaces, work flows and reproducibility
  - The analysis interface task combines a number of elements to form the working surface for the user of the EOSC science platform.
- Task 5.4: Integration with HPC and HTC infrastructures
  - Once data for analysis has been located and staged, and workflows have been defined, either by accessing the EOSC software repository or by the user directly, the next step is to deploy those workflows on the underlying processing infrastructure.

Funded by the European Union's Horizon 2020 - Grant N° 824064

WP5 Specific Step

- Build prototype science analysis platform
  - Data gathering and discovery
  - Access to software & services
  - Customised processing & workflows
  - Interface with large-scale computing infrastructure
  - Adds analytics and visualisation
- Ready for future challenges
  - Increased scale of data volumes
  - Processing co-located with data

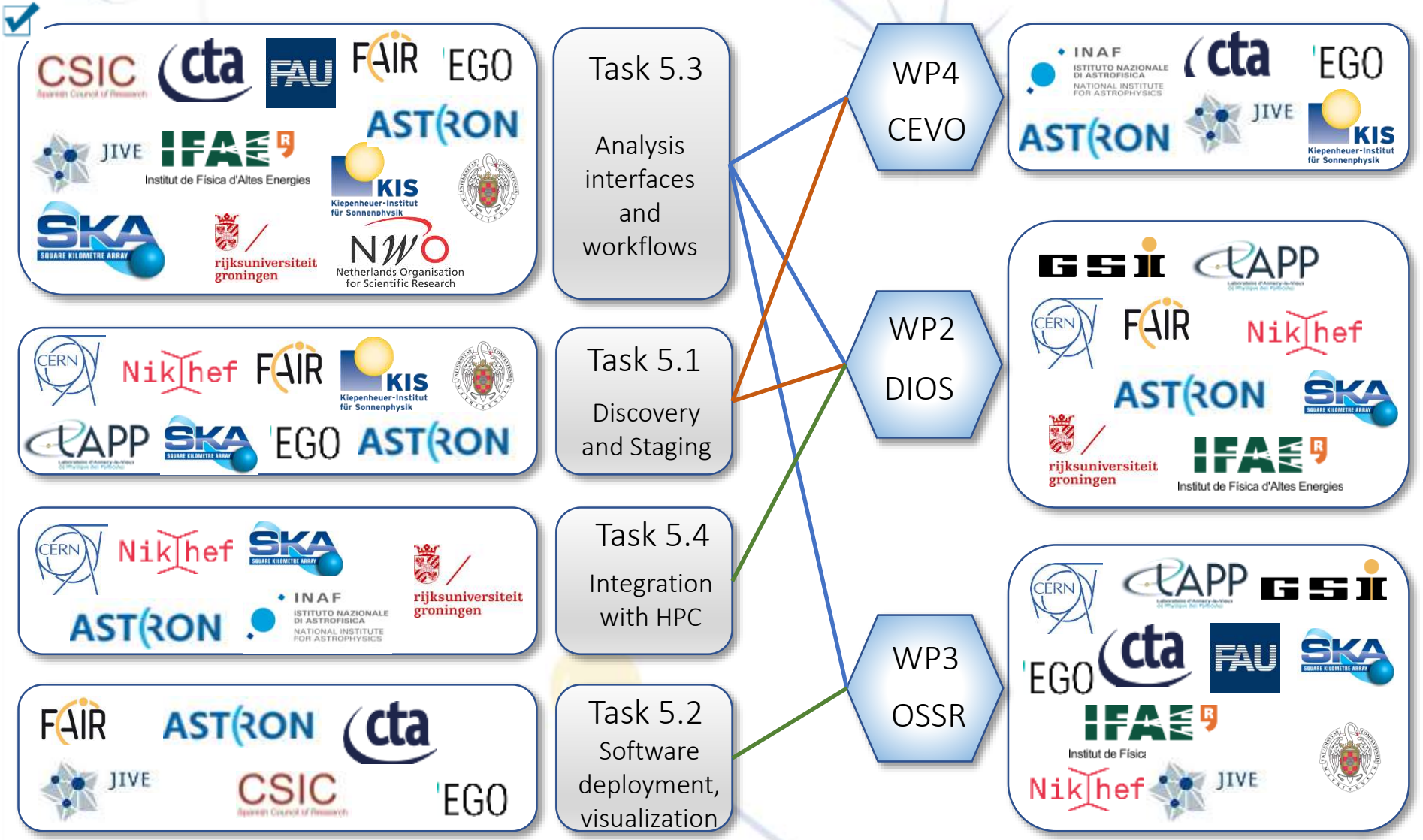
Funded by the European Union's Horizon 2020 - Grant N° 824064

Links with other ESCAPE WP's

- WP1 - General EOSC policy for services & infrastructure access
- WP2 - Integration with Data Lake - distributed computing & storage
- WP3 - Access to software & services in ESCAPE-EOSC catalogue
- WP4 - connect science platform with existing astronomical data archives and VO-enabled data collections

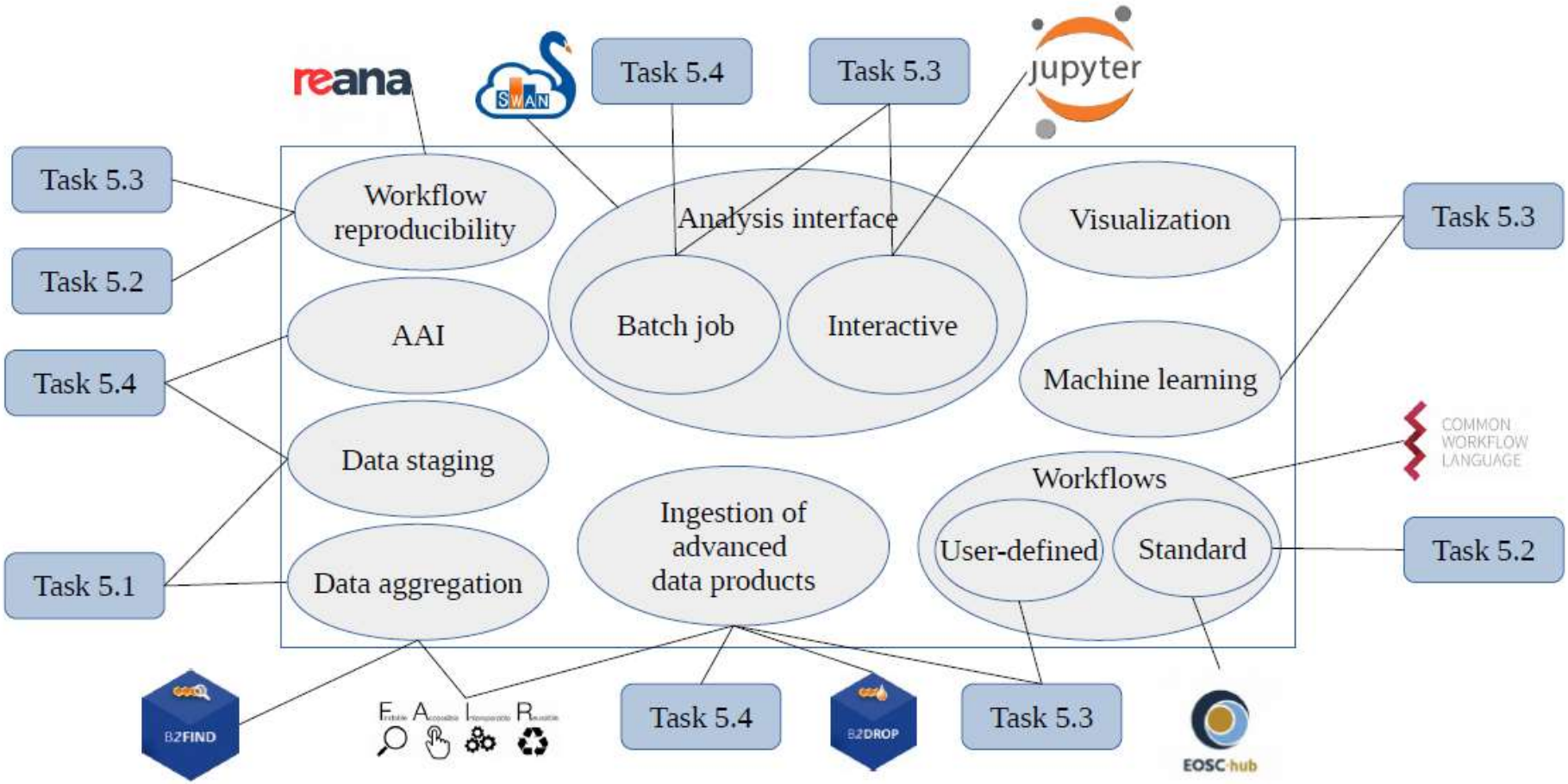
Funded by the European Union's Horizon 2020 - Grant N° 824064

Links with other ESCAPE WP's



Funded by the European Union's Horizon 2020 - Grant N° 824064

Detailed task interactions



Funded by the European Union's Horizon 2020 - Grant N° 824064

WP6 : ECO

Engagement and COmmunication

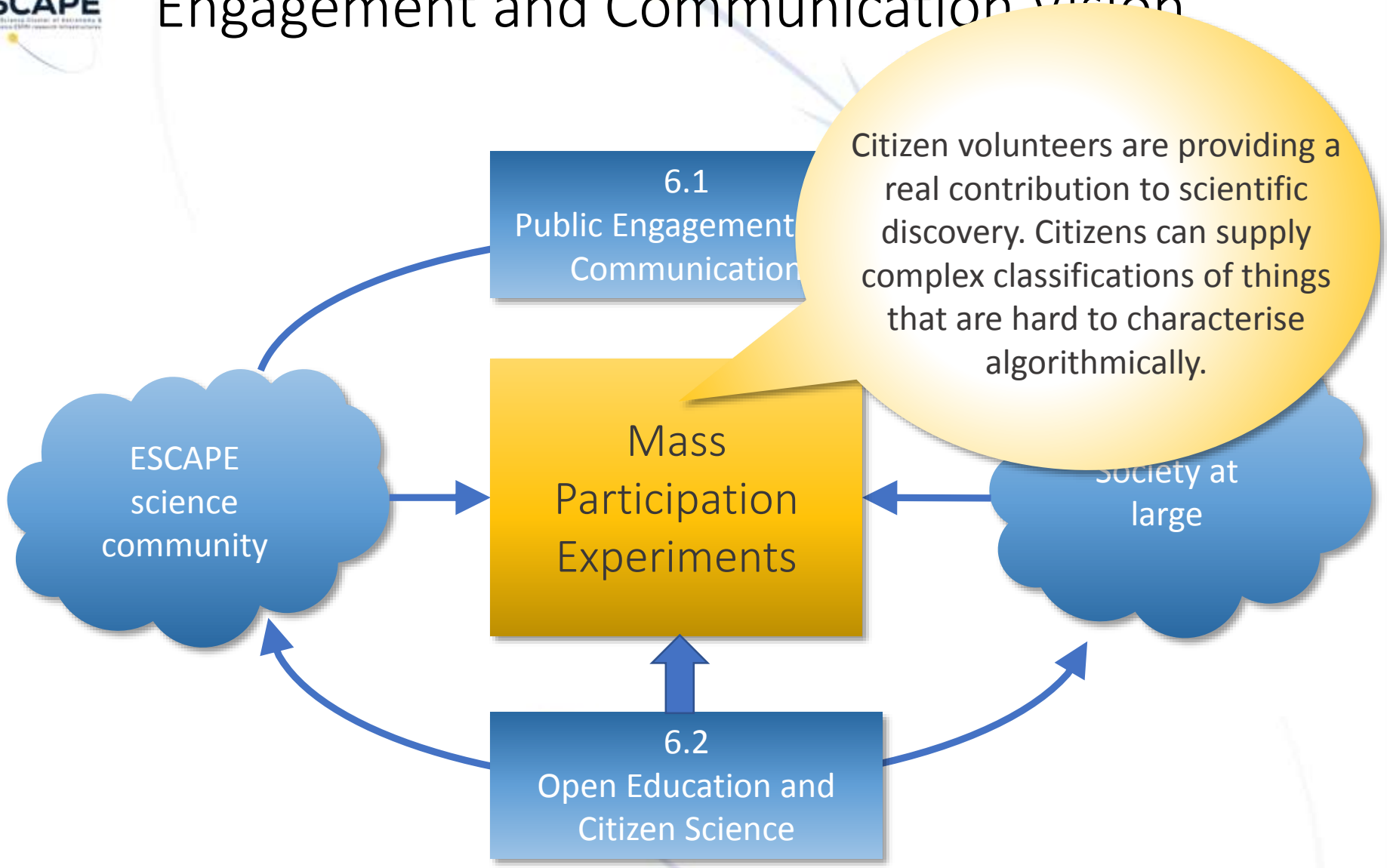
Citizen Science, Open Science and Communication

CNRS	OU	Trust IT

Contributed by Stephen Serjeant

Funded by the European Union's Horizon 2020 - Grant N° 824064

Engagement and Communication Vision



Funded by the European Union's Horizon 2020 - Grant N° 824064

ESCAPE Stakeholders

Engagement between ESCAPE and stakeholders, building on partners' networks.

STAKEHOLDERS

- e-Infrastructures
- ESFRI projects
- Industry (Namely SMEs)
- EOSC Governance
- Policy Bodies
- Pan-European Research Organisations

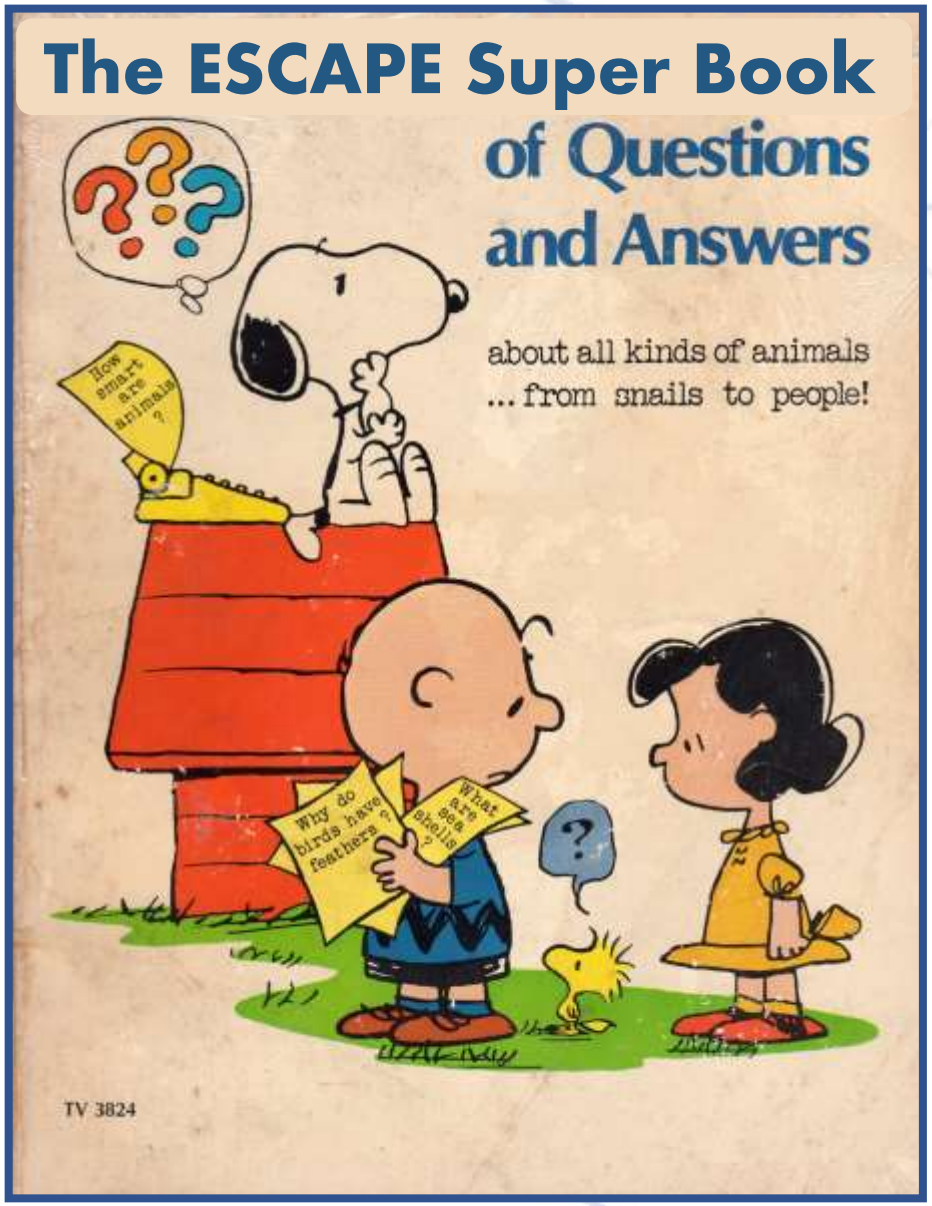
What is the value proposition for each of these stakeholders using ESCAPE's outputs?

How can ESCAPE's assets fit the EOSC offer for EU/ international researchers?

EUROPEAN OPEN SCIENCE CLOUD

Engagement with scientific, research, industrial & policy communities

Funded by the European Union's Horizon 2020 - Grant N° 824064



Funded by the European Union's Horizon 2020 - Grant N° 824064

What would we expect from EOSC (Hub)

- Establish a framework for a federated cloud infrastructure, including the interfaces to the scientific and IT/computing infrastructures.
- Enabling Open Science for a broad variety of users and use cases via:
  - re-use of data and services for increased scientific outcome;
  - ease of access and use of data and services across boarders of user types and user communities;
  - effective use of available resources for data and service providers.
- Cooperation for integration of the Astronomical Virtual Observatory framework into EOSC
  - Demonstrating the capacity of EOSC to include existing operational frameworks
- Visibility of astronomy resources in EOSC
  - Data and Services
  - First step is integration of VO Registry *(building on the experience of the VO Registry in EUDAT B2FIND)*
- Connection to computing resources
- Help to work with global EOSC and domain service portfolios
- An integrated AAI solution which provides a single point of entry instead of having to choose among different AAI solution providers.
  - Compatible with standard systems
- Ensure wider society included as stakeholder – easily overlooked
  - As a first concrete step, our citizen science platform, the Zooniverse, is being submitted as a service provider for the EOSC-Hub.

Funded by the European Union's Horizon 2020 - Grant N° 824064

Questions to EOSC (Hub)

- Will the EOSC (via EOSC-hub) provide computing resources and how will the federation be organised in this case?
- How deep will be a repository be integrated into the catalogue of services, who will take over maintenance and preservation of the repository?
- How will the interfaces between EOSC-hub und the EOSC clusters be established and how will they be defined (e.g. for AAI and usage of computing and hosting resources)?
- To which extent can science communities access compute and storage resources, and other services EOSC Hub provide?
- How can use cases of different user communities affect the practical design of EOSC Hub interfaces?
  - As the volunteer use case is quite different to that of a policy maker or a practicing scientist.

Funded by the European Union's Horizon 2020 - Grant N° 824064

How wants ESCAPE to support EOSC (Hub)

- Enriching the catalogue of services with an Astro, Particle, Physics community based repository of software and services building on the well established expertise of the partners.
- Defining the needs for the EOSC from the point of view of these communities.
- We bring in the experience of an operational domain specific interoperability framework
  - Lessons learned from VO pioneering activities
  - IVOA, Euro-VO projects, ASTERICS-DADI
  - Operational and services ready for wider interdisciplinary interoperability
- Re-usable tools
  - applicable beyond astronomy data (tables, images, cubes…)
- Re-usable standards
  - that can be modified and applied in different disciplines
- Data Stewardship practices for Open Science
  - Scientific, technical and human aspects of making data FAIR
- WP5 (analysis platform) will use as many existing EOSC services as possible to build the various SAP services. In this way, we will bring ESCAPE ESFRI user communities to use the underlying EOSC services.
- Low cost but high profile impact by including volunteer use cases.

Funded by the European Union's Horizon 2020 - Grant N° 824064

The END

Funded by the European Union's Horizon 2020 - Grant N° 824064

Deliverables

Month	D	Goal	Partner
5	D6.1	Project Website live	Trust IT
6	D4.1	Detailed project plan for WP4	CNRS
6	D5.1	Preliminary report on requirements for ESFRI science analysis use cases	NWO
9	D5.2	Detailed Project Plan	NWO
12	D6.2	Dissemination and exploitation plan	CNRS
14	D4.2	Intermediate analysis report on use for IVOA standards for FAIR ESFRI and community data	CNRS
16	D4.3	First Science with interoperable data school	INTA
18	D4.4	Intermediate analysis report on integration of VO data and services into the EOSC	INAF
18	D6.3	Brochure publication	Trust IT

Funded by the European Union's Horizon 2020 - Grant N° 824064

Deliverables

Month	D	Goal	
24	D5.3	Performance WP5 assessment of initial Science Platform Prototype	NWO
24	D6.4	Citizen science experiments with embedded educational resources (midterm)	OU
36	D3.8	Thematic training event - second school for software development and deployment in the EOSC	CNRS
36	D6.5	Promotional education animation videos	OU
38	D4.7	Final analysis report on integration of VO data and services into the EOSC	INAF
40	D1.1	Final Integration Event	CNRS
40	D2.3	Final assessment and analysis of the full prototypes, outlook for further development and deployment towards full production services within EOSC	CERN

Funded by the European Union's Horizon 2020 - Grant N° 824064

Deliverables

Month	D	Goal	
40	D4.8	Final analysis report on use of IVOA standards for FAIR ESFRI and community data, and best stewardship practices for value- added data	CNRS
41	D6.6	Citizen science experiments with embedded educational resources	OU
42	D5.4	Final assessment of the performance of the Science Platform prototype and plan for deployment of production version within the EOSC	NWO

Funded by the European Union's Horizon 2020 - Grant N° 824064