ASTRONOMY & PARTICLE PHYSICS CLUSTER

Research Data Management at GSI/FAIR workshop

Mohammad Al-Turany

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
European Science Cluster of Astronomy &
Particle physics ESFRI research Infrastructures

Five EOSC Clusters funded in H2020-INFRAEOSC-04-2018

- **EOSC-LIFE:** Life science RIs
- **ENVRI-FAIR:** Environmental Research Infrastructures
- **ESCAPE:** Astronomy and Particle Physics
- **PANOSC:** Photon and Neutron sources RIs
- **SSHOC:** Social Sciences and Humanities

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

ESCAPE in a nuthshell

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
- **2** European research infrastructures: EGO and JIV-ERIC
- Budget: **15.98 M€**
- Started: **1/2/2019**
- Duration: **42** months (end date 31/7/2022)
- Coordinator: **CNRS**

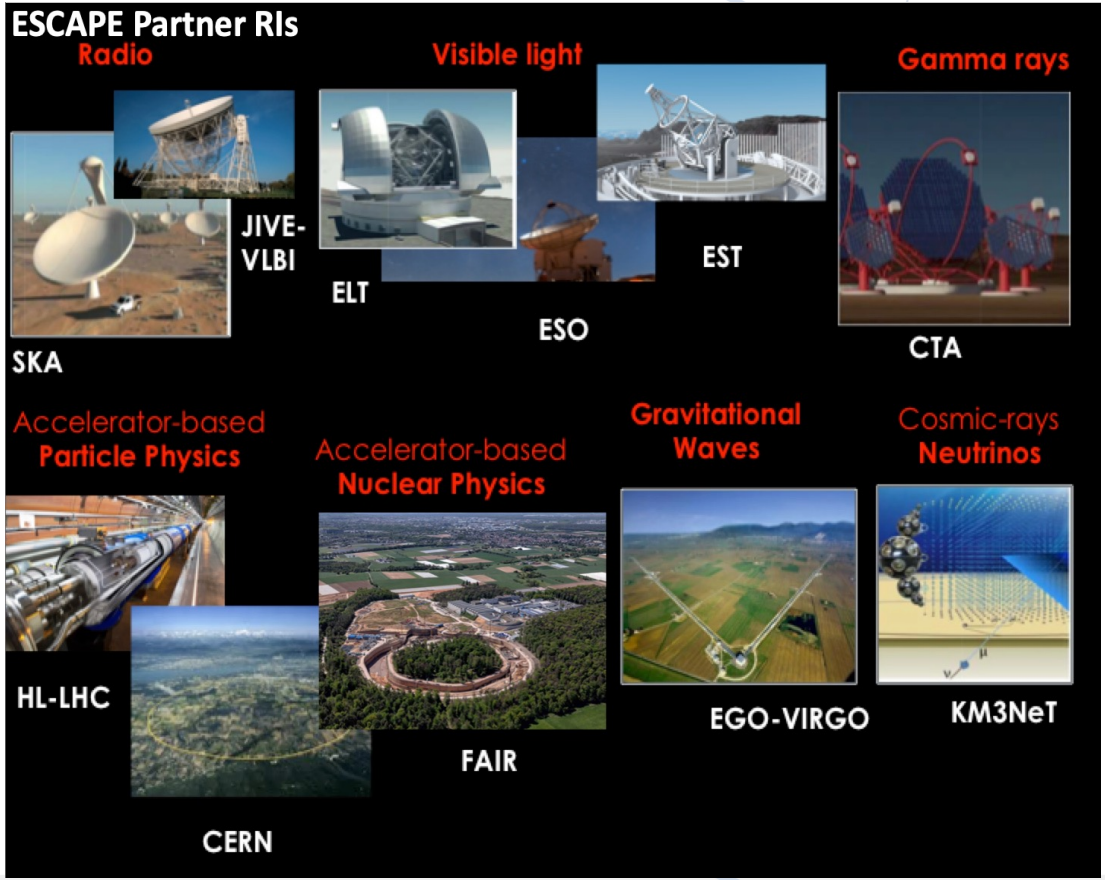
Home page: https://escape2020.eu ; Twitter: @ESCAPE_EU

SMEs: small and medium-sized enterprises

ESFRI: European Strategy Forum on Research Infrastructures

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

brings the astronomy, astroparticle and particle physics communities together



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

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

Astronomy and Particle Physics

ESCAPE is a step forward…

- The astronomy-related ESFRI projects and the accelerator-based particle physics ESFRI facilities will open together new paths towards the understanding of the Universe through a multi-probe approach.

- Enhance the coordination leveraging two major complementary excellences in data stewardship:
  - the astronomy Virtual Observatory infrastructure;
  - long-standing expertise of the particle physics community in large-scale distributed computing and big-data management.

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

ESCAPE goals

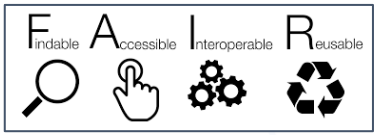
1. Implementing Science Analysis Platforms for EOSC researchers to stage data collections, analyse them, access ESFRIs' software tools, bring their own custom workflows.

2. Contributing to the EOSC global resources federation through a Data-Lake concept implementation to manage extremely large data volumes at the multi-Exabyte level.

3. Supporting "scientific software" as a major component of ESFRI data to be preserved and exposed in EOSC through dedicated catalogues.

4. Implementing a community foundation approach for continuous software shared development and training new generation researchers.

5. Extending the Virtual Observatory standards and methods according to *FAIR* principles to a larger scientific context; demonstrating EOSC capacity to include existing frameworks.



6. Further involving of small and medium-sized enterprises (SMEs) and society in knowledge discovery.

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

ESCAPE work programme

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
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 et Communication

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

ESCAPE work programme

PMs per WP and per each RI.
The allocated staff effort is proportional to the respective boxes' surface areas.

- 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

ESCAPE work programme

PMs per WP and per each RI.
The allocated staff effort is proportional to the respective boxes' surface areas.

- 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

We are part of:

WP1- MIND:
Management, Innovation, Networking and Dissemination

WP2 – DIOS:
Data Infrastructure for Open Science

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

WP5 – ESAP:
ESFRI Science Analysis Platform

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

WP2- 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

WP2- specific objectives

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

- Stores and organizes scientific data (**F**indable) and enables the provisioning of data processing (**A**ccessible)
- Enables sciences to build open data repositories (**I**nteroperable)
- In general, supports the world-leading data challenges of the Research Infrastructures in ESCAPE
- Ensure long term **data preservation** (**R**eproducible) at the infrastructure level

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

WP3 - Open-source scientific Software and Service Repository

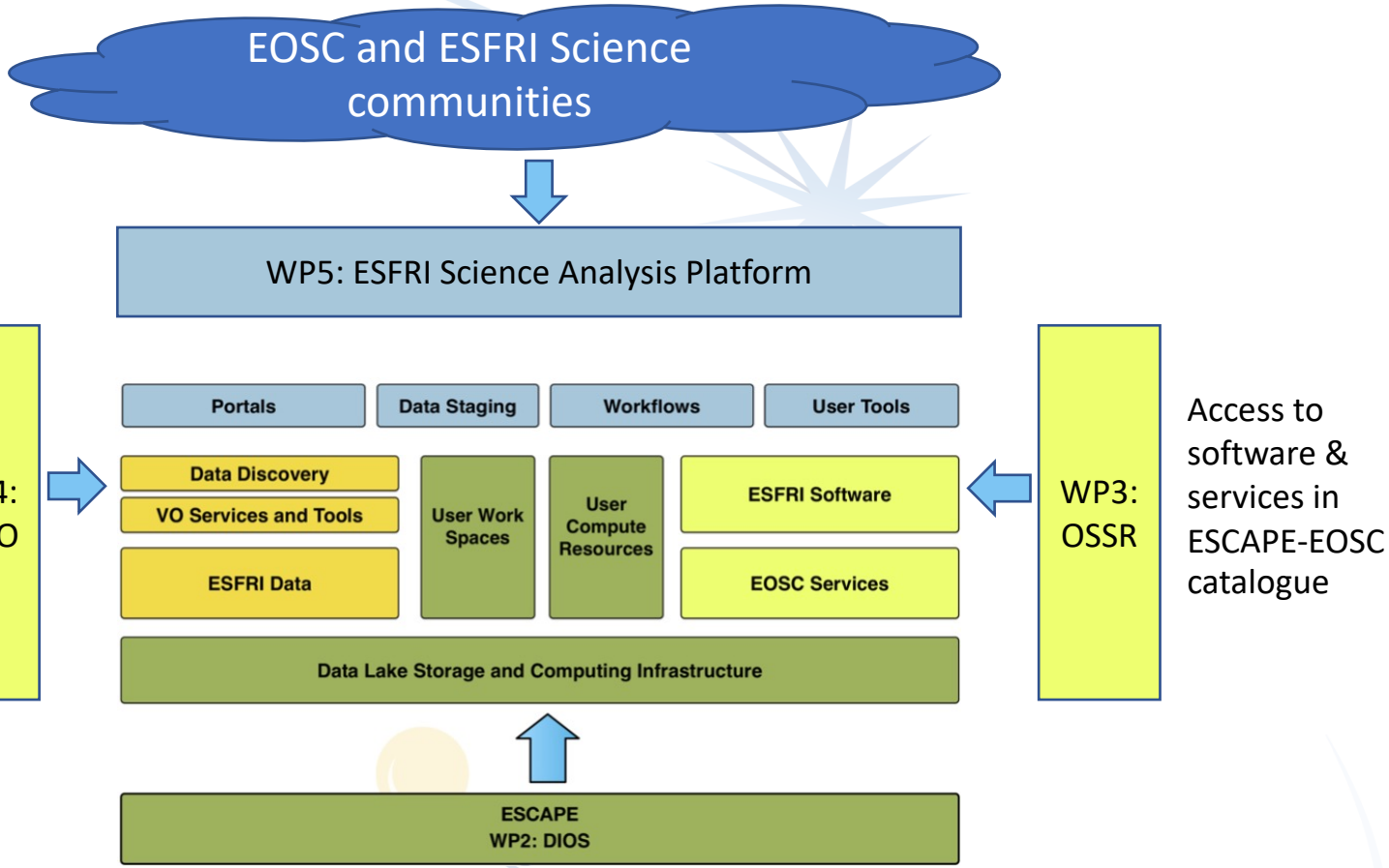
- Aim: expose the tools of the ESCAPE (ESF)RI projects in a repository under the EOSC catalogue of services
- Objectives:
  - continuous development, deployment, exposure and preservation of software/tools/services
  - interoperability, software re-use and cross-fertilisation
  - open innovation environment for open standards, common regulation and shared (novel) software for multi-messenger&multi-probe data
- All objectives follow the community-based approach

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

WP5 - ESFRI Science Analysis Platform

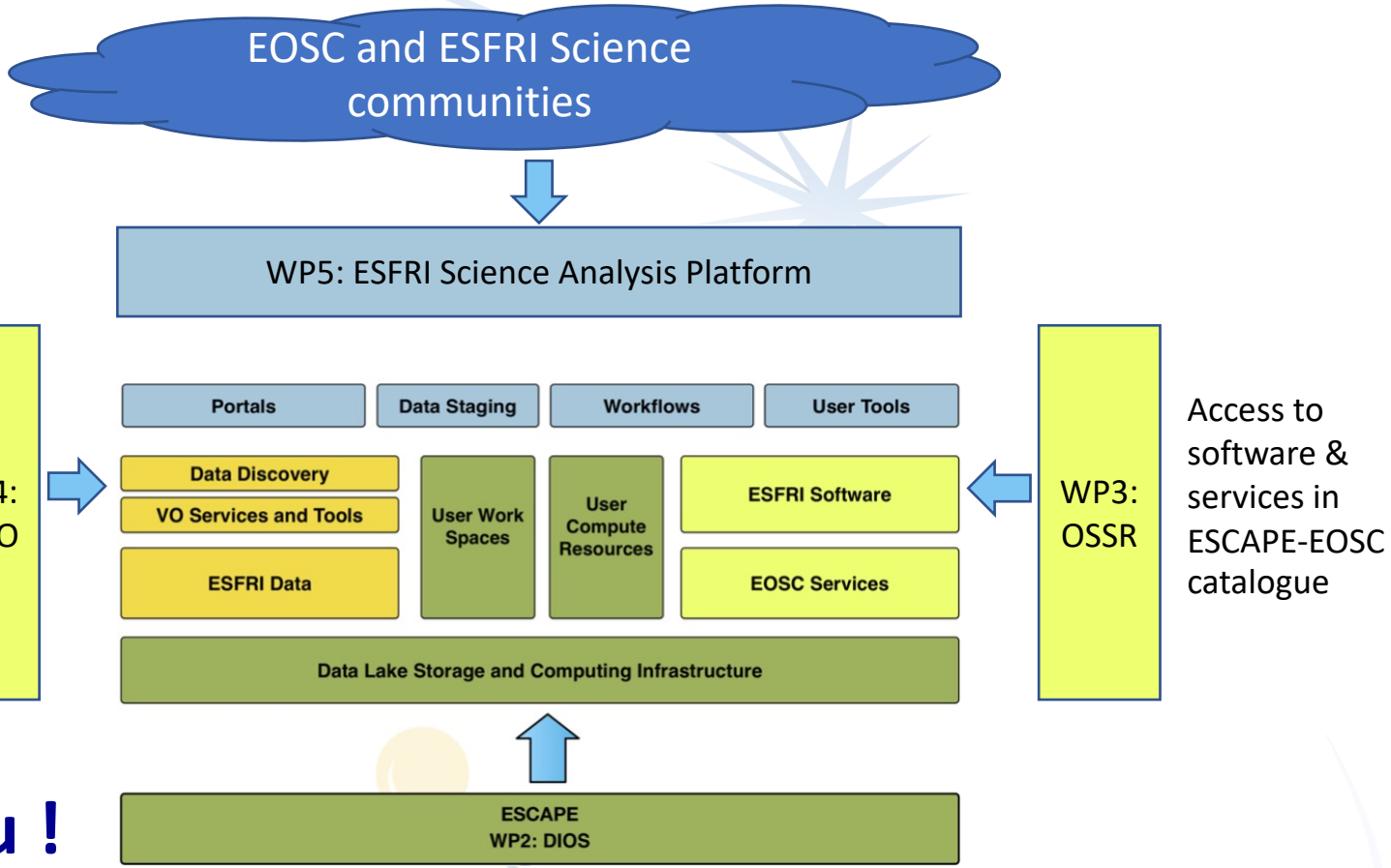
- Implement Flexible and Expandable Science Platform
  - make EOSC a working interface
  - bring analysis to data
- Support users to:
  - identify & stage existing data collections
  - tap into software tools & packages developed by ESFRIs
  - bring own custom workflows
  - take advantage of available HTC and HPC infrastructure
- Focus on core common functions to support two communities
  - look to expand to other domains in future
  - flexibility rather than single platform for all users

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



Integration with Data Lake - distributed computing and storage

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



Thank you !

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