ESCAPE OSSR
WG Meeting, 07/04/2022

Kay GRAF

ECAP, Erlangen Centre for Astroparticle Physics, Friedrich-Alexander-Universität Erlangen-Nürnberg

for the OSSR Team

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

Review Report

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

ESCAPE EOSC cell

ESCAPE
European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

ESCAPE OSSR

Catalogue & Repository of resources

Datasets
Software & services
Tutorials
Training
Publications

TSP's

RI-Specific Science Platforms

ESCAPE VO Virtual Observatory

Astronomy Data centres

VO Registry

VO Registry
Analysis Tools
VO Services

ESCAPE SAP Science Platforms

Workflows, notebooks, deployment platforms, packaging

ESCAPE CS Citizen Science

ESCAPE DIOS Data Lake

FAIR data management
Content discovery and delivery

HPC

PRACE

EuroHPC Joint Undertaking

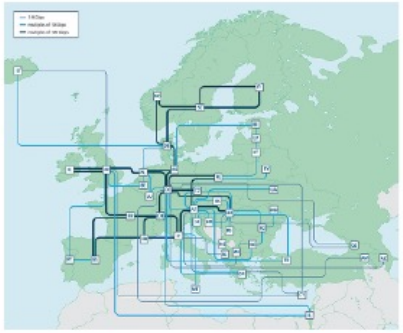
HTC

Grid clusters, etc

Private/public clouds

Commercial clouds

GÉANT



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

OSSR Aims and Objectives

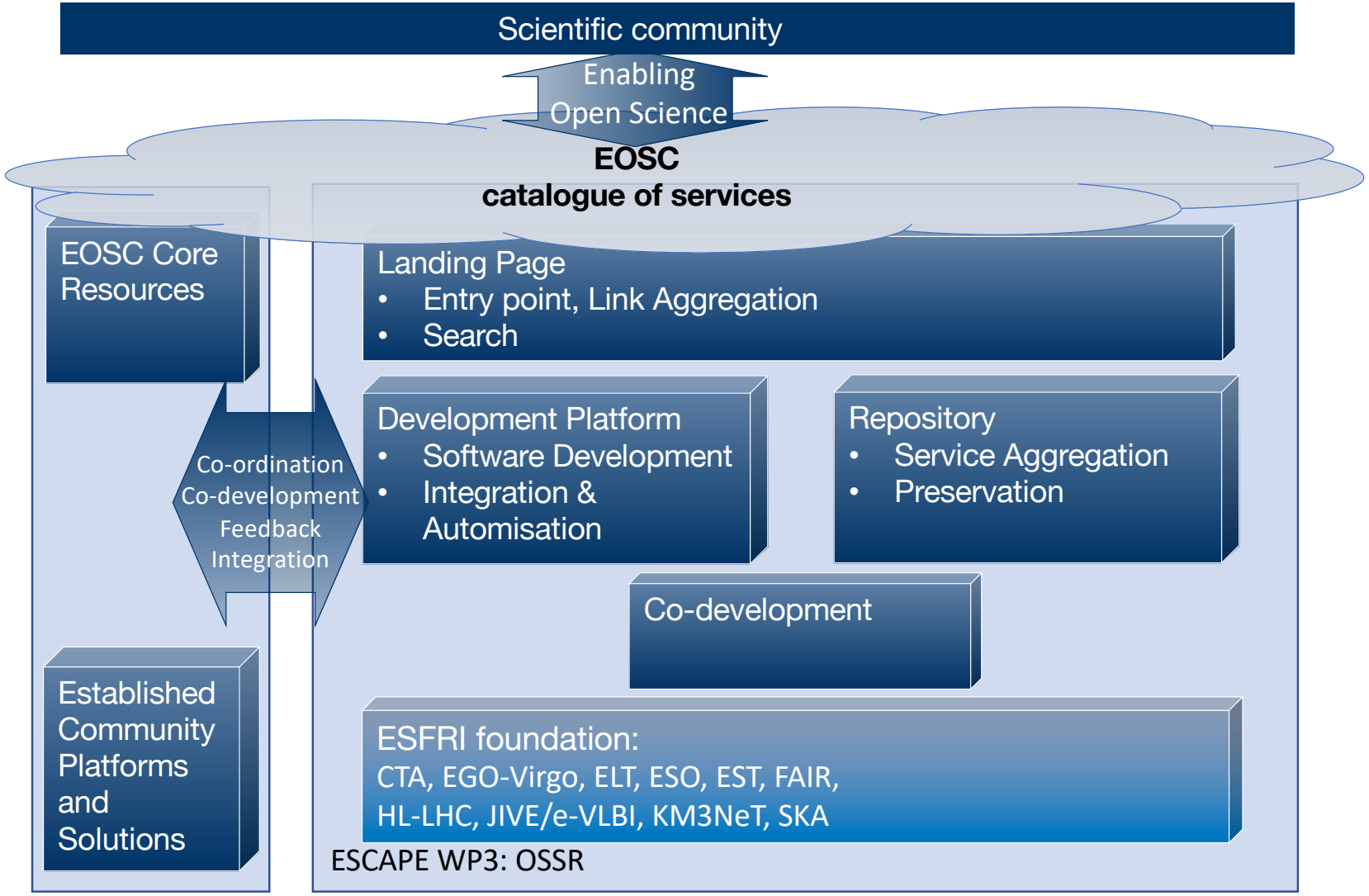
ESCAPE OSSR
Catalogue & Repository of resources

Datasets
Software & services
Tutorials
Training
Publications

- Aim:
  shared open science software and services based on FAIR principles
- Objectives:
  - Facilitate and support continuous **development, deployment, exposure and preservation** of partners' software/tools/services
  - Foster **interoperability, software re-use and cross-fertilisation** between ESFRIs (e.g. simulation)
  - Offer an **open innovation environment for open standards** (e.g. workflows, data-formats), **common regulations** and **shared (novel) software** for multi-messenger & multi-probe data
- All objectives follow:
  - Paradigm of **enabling open science** – with **software as "first class citizen"**
  - a **community-based and inclusive** approach
  - the **FAIR principles** for open science resources – software and derivatives
  - **Federation** of available resources

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

OSSR Overview



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

Work Organisation – 5 Tasks and Focus Groups

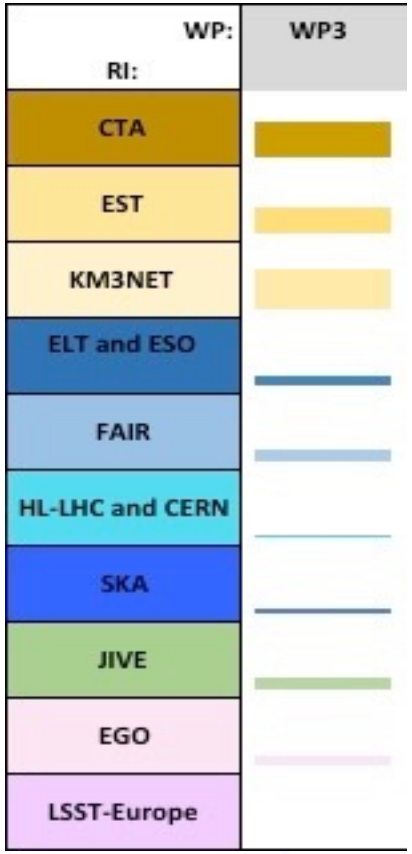
- Tasks
  - 3.1: Management Activities, Policy and Support Action (MAPS) –
    Providing assistance and support for work package related activities
  - 3.2: ESFRI Software and Services Collection (ESSC) –
    Systematically list available software and requirements to facilitate integration into the repository
  - 3.3: Common Approaches: Software and Services (CASS) –
    Fostering common developments and facilitate sharing of software
  - 3.4: Foundation of Competence for Software and Service Innovation (COSSI) –
    Establish competence group for mainstreaming new approaches to data analysis (e.g. deep learning)
  - 3.5: Repository Implementation and Deployment (RIAD) –
    Setting up a demonstrator for a common software repository
- Focus groups
  - Focus group 1: Collecting Software requirements (related to Task 3.2)
  - Focus group 2: Technical implementation of the repository (related to Task 3.5)
  - Focus group 3: Innovative workflows (related to Task 3.4)
  - Focus group 5: Common approaches to CORSIKA (related to Task 3.3)
  - Focus group 6: Common Data Formats (related to Task 3.3)

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

OSSR Overview and Organisation

- Wiki as general entry point for information Organisation
- **Tasks** formulate the main objectives of the work package
- **Focus groups** facilitate day-to-day work with ~ bi-weekly meetings (~ 150 to date)
- Central communication tools and bi-yearly written updates from partners
- Partners commonly responsible for reaching the aims and objectives

Budget Distribution



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

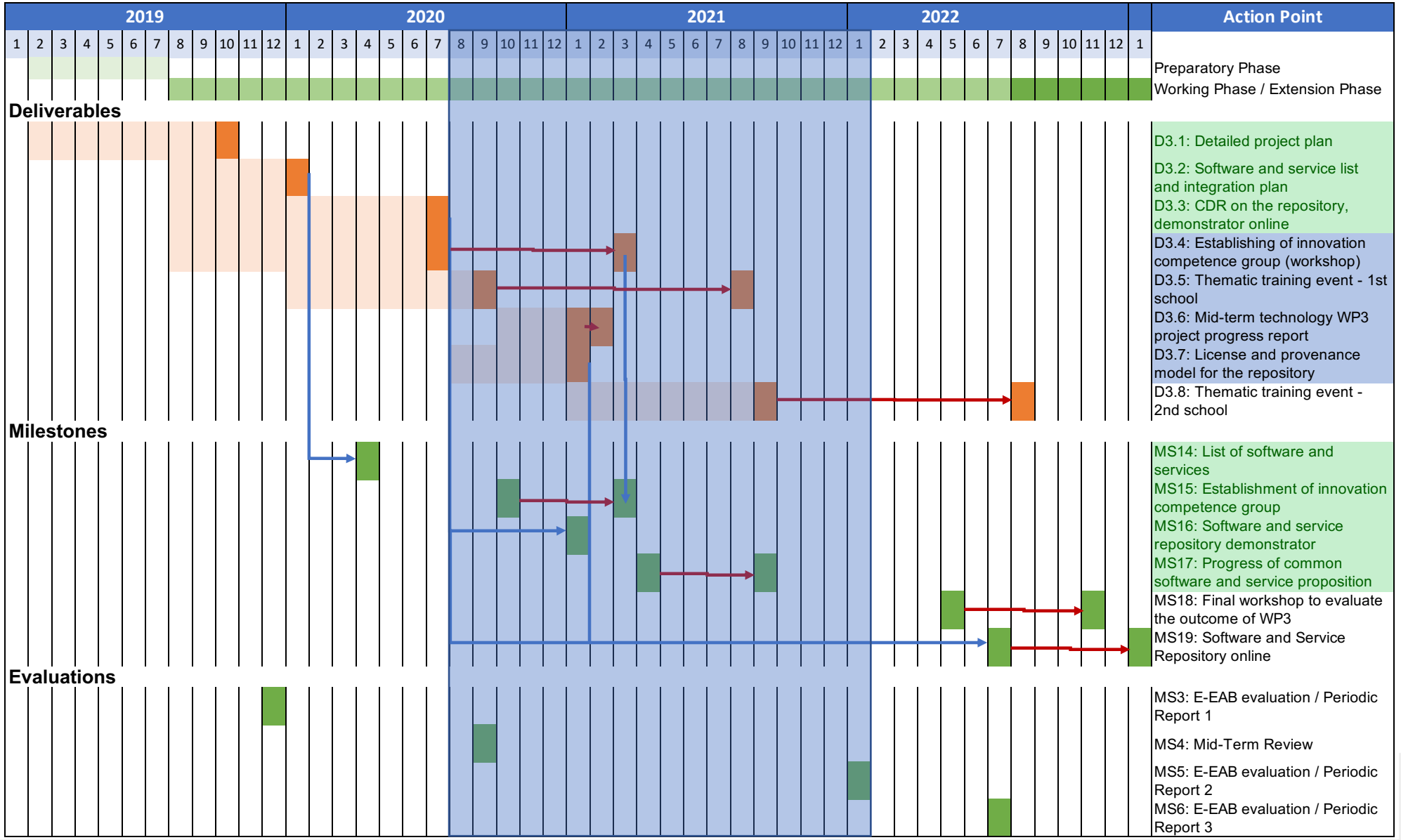
Partner Contribution

Institutes/SME	ESFRI/RI	PM	Task 3.1	Task 3.2	Task 3.3	Task 3.4	Task 3.5
CNRS-LAPP	CTA	39	6	6		6	21
CTAO	CTA	6		6			
IFAE	CTA	20			20		
MPG-MPIK	CTA	36			36		
UCM	CTA	18			12	6	
EGO	EGO-Virgo	30				30	
HITS	ELT	12				12	
AIP	EST	36				36	
NWO-I-CWI	EST	26				26	
UNITOV	EST	14				14	
GSI	FAIR	36		18	18		
CERN	HL-LHC, CERN	6		6			
JIVE	JIVE	36		36			
CNRS-CPPM	KM3NeT	24			24		
FAU	KM3NeT	54	30	9	6		9
INFN	KM3NeT	36			24	12	
NWO-I-Nikhef	KM3NeT	12		12			
SKAO	SKA	12					
OROBIX	SME	15				15	
TOTAL		**468**	**36**	**93**	**140**	**157**	**30**

OSSR: Deliverables & Milestones

accepted

submitted



Task 3.1

- Management Activities, Policy and Support Action – K. Graf (FAU)
  - IT services and helpdesk for the WP
  - Implement workflows on project platform
  - Establishing links
    - within ESCAPE
    - towards the EOSC working groups/task forces and other projects
    - EOSC Future and national activities
  - **Organisation** of workshops, questionnaires, meetings, training events
- Open Points:
  D3.8 (2nd school) and MS18 (final workshop)

17 February 2021

Webinar: ESCAPE OSSR | Enhancing science through sharing software - benefits & use cases

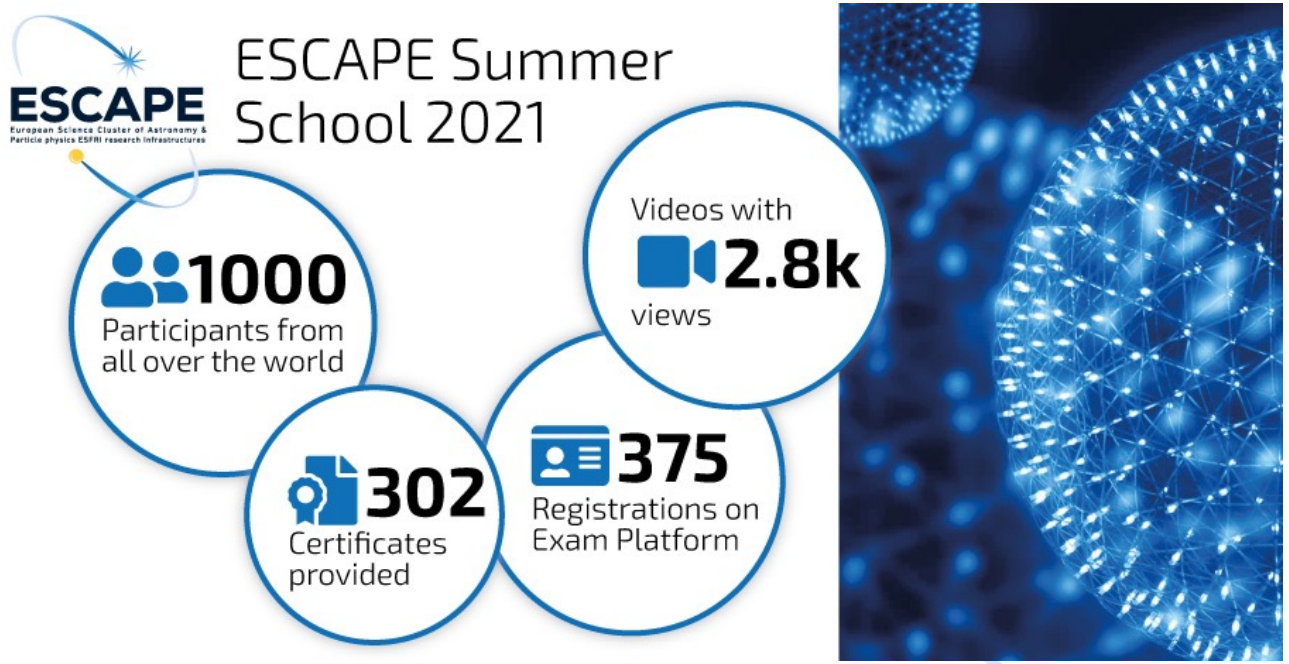
Virtual

When: 17th February 3pm CET. In the webinar "ESCAPE OSSR Enhancing science through sharing software - benefits & use cases" we will show the ESCAPE OSSR developments and achievements towards a FAIR multi-messenger data-driven cooperative approach.

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

T3.1 - Outreach and Training – D3.5

- Software as first class EOSC citizen
- Enable software custodians
- All lectures/materials online: https://escape2020.github.io/school2021/



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

Task 3.2

- ESFRI Software and Services Collection – K. Graf (FAU):
  - Development, benchmarking and **onboarding of software** within and across partners;
  - Gathering of **common practices** and know-how towards the definition of best practices to be shared with the community (WOSSL workshop)
  - **Workflow and co-ordination** for software onboarding
  - Software to become part of the repository collected
    - partners started to add and prepare the software for repository integration;
  - Co-operation with HGF DMA project
- Open points: finalise onboarding of partner software; onboard test science project software

23 July 2020 to28 July 2020

ESCAPE Workshop on Open-Source Software Life Cycles

Virtual

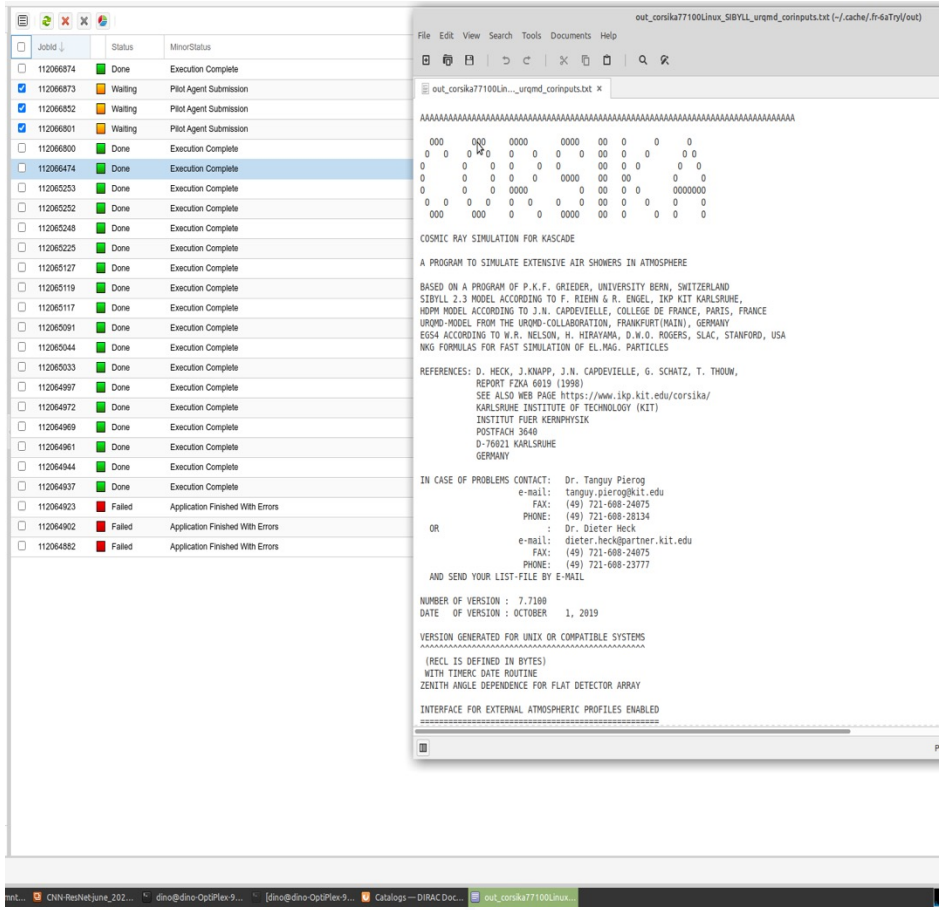
Software development is an integral part of modern science, gaining knowledge from data. All ESCAPE partners develop and

[...]

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

Task 3.3

- Common Approaches:
  Software and Services – C. Bozza (INFN):
  - Coordinative work for **common data formats** and software tools between the partners and the ESF/RIs they represent (together with CEVO);
  - Development and production of **CORSIKA** – an air shower simulation program - turnkey containers for various use cases and functional development of CORSIKA for specific purposes;
  - Usage of the DIRAC interware for large scale simulation productions is pursued.
- Open points: integration with **ESAP backend**, continue data format specification



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

Task 3.4

- Foundation of Competence for Software and Service Innovation – E. Cuoco (EGO)
  - **Machine learning approaches** to simulation and experiment data adapted and benchmarked;
  - Definition of data formats and different **deep-learning approaches** pursued
  - Exchange of experience and harmonisation of approaches for **innovative workflows**; gathering of workflows between different partners
  - IWAPP workshop → innovation working group formed
  - Establishment of use cases for multi-messenger analysis workflows connecting several ESFRIs
- Open Point: follow up in test science projects

08 March 2021 to12 March 2021

IWAPP Workshop - Innovative Workflows in Astro and Particle Physics

Online

The objective of this workshop is to bring together the scientists' communities of Astrophysics, Astroparticle Physics and Particle Physics who are leading the development of Innovative Workflows within their domain.

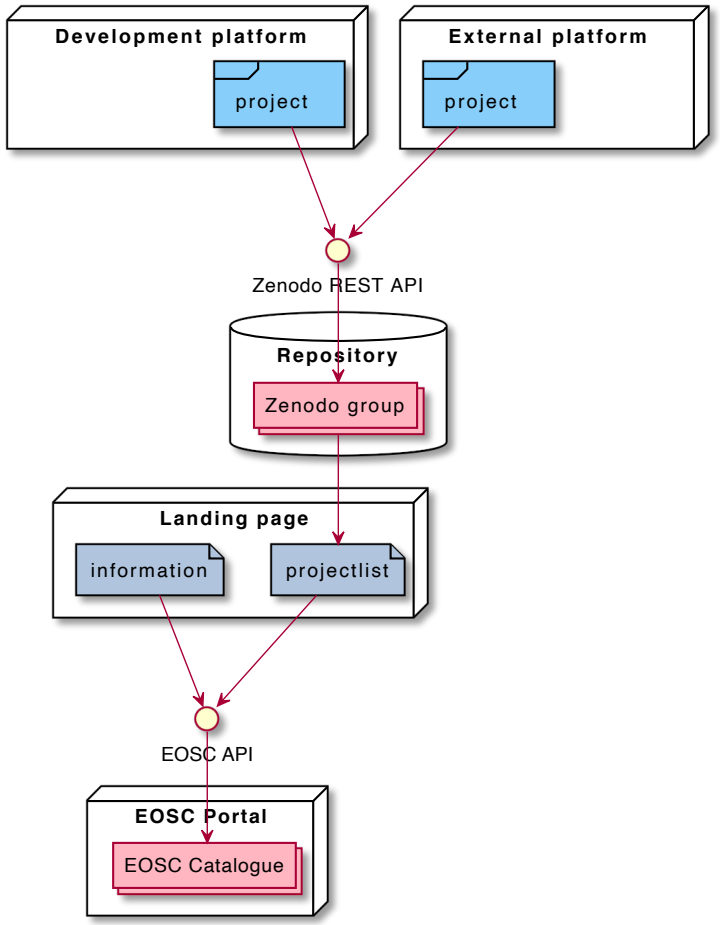
Funded by the European Union's Horizon 2020 - Grant N° 824064

Task 3.5

- Repository Implementation and Deployment – T. Vuillaume (CNRS-LAPP):
  - **partner feedback** for the repository gathered;
  - design of the repository and the definition of technical solutions for its **implementation**;
    - eossr library: implementation of OSSR developments
    - OSSR guidelines and (metadata) definition (CodeMeta)
    - Continuous integration
  - prototype **set up and expanded**
  - Onboarding procedure and technical implementation hardened during a sprint in 02/2022 – aim: user-friendliness, sustainability (also beyond ESCAPE)
- Open Points:
  - continue **implementation** tasks, **EOSC catalogue** integration
  - Follow up of **Virtual Research Environment** (link to DIOS and ESAP)

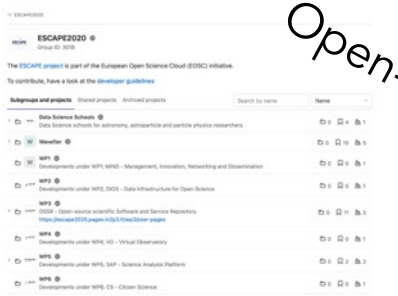
Funded by the European Union's Horizon 2020 - Grant N° 824064

OSSR Schematic



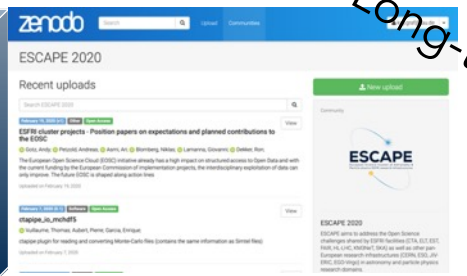
Development Platform
- Software Development
- Integration & Automization

Open-source



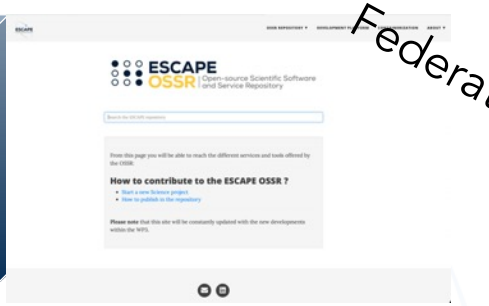
Repository
- Service Aggregation
- Preservation

Long-term



Landing Page
- Entry point, Link Aggregation
- Search

Federation



Same structural approach as [CERN Open Data Portal]

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

eOSSR library

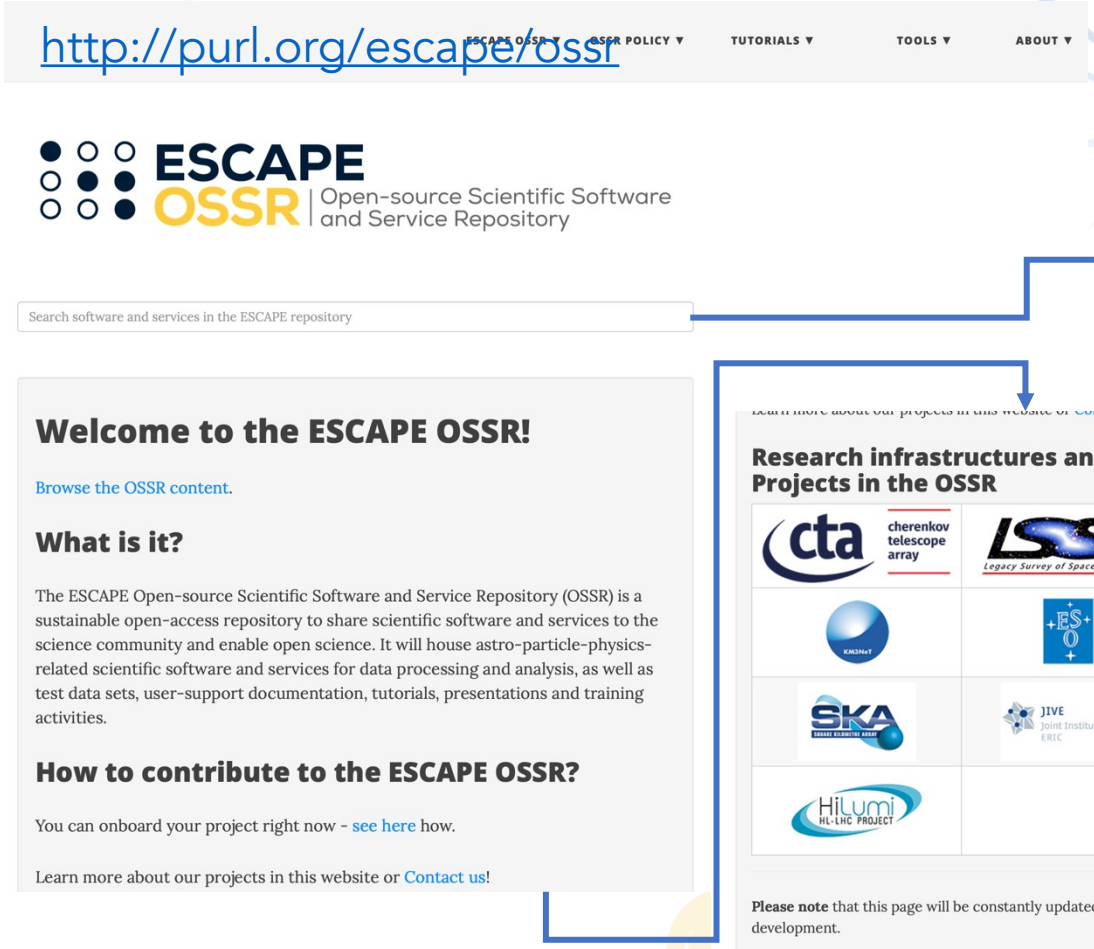
DOI 10.5281/zenodo.5524912

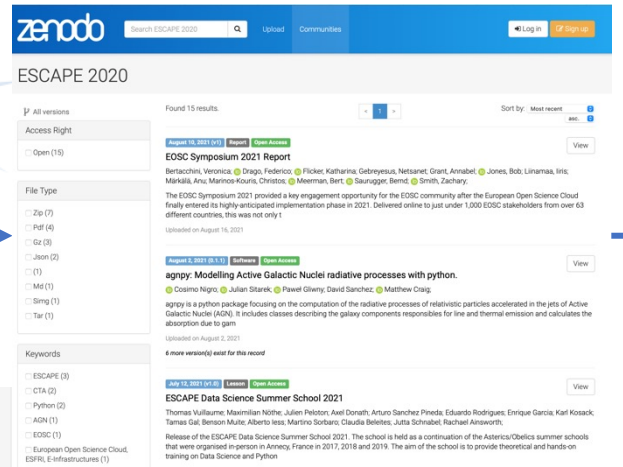
- *eOSSR* library https://gitlab.in2p3.fr/escape2020/wp3/eossr
  - incorporates all OSSR developments, based on the commonly defined practices and standards
  - python-based
  - OSSR API : send request to the OSSR, find and filter software and services, upload new entries, update existing entries
  - CI : automated upload / update using gitlab CI
  - Metadata : schema definition, crosswalk between CodeMeta and Zenodo, generator and validator available (implementation of D3.6 and D3.7)
- Used by **ESAP** to get OSSR records

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

OSSR - Current Landing Page

http://purl.org/escape/ossr





OpenAIRE | EXPLORE

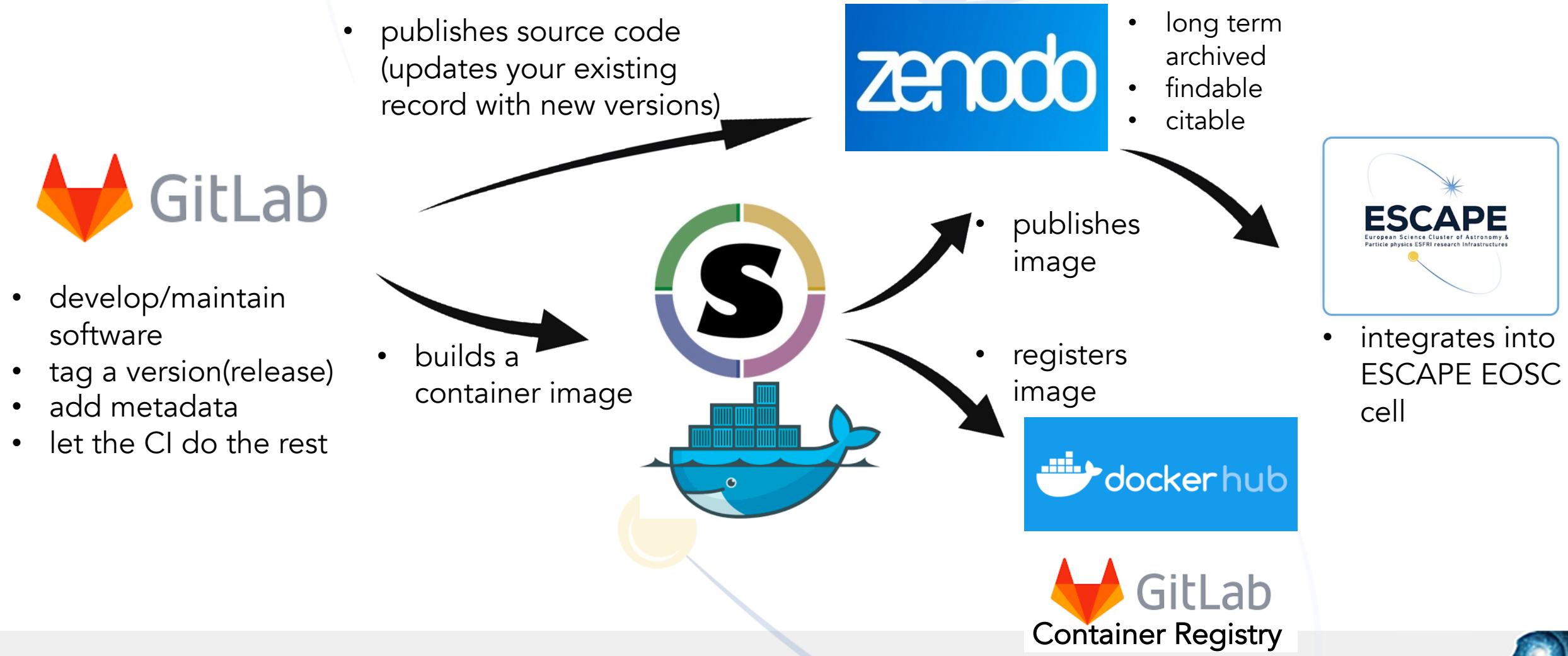
openAIRE

EOSC integration

+ add related projects / collections

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

ESCAPE OSSR Workflow



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

Onboarding Procedure and Status

- Onboarding workflow defined
  - Tracked on project platform, supplied by eOSSR library
  - Necessary:
    Onboarding presentation & documentation
    (e.g. technical report)
    → entry on the landing page
  - To be extended to external partners and streamlined
    - Add EOSC terms of use
- OSSR progress workshop (2021/09) to update the status of all 19 partners
  - all planned contributions (D3.2) will be integrated into OSSR by the end of the ESCAPE project.

Status:

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

Joint cross-WP Activities supported by Technical Coordination Team

WP2 – WP3 Topics

- **Datalake as a Service** onboarding
- Linkage of Software to Datalake

WP3 – WP4 Topics

- Software **metadata** in the IVOA & OSSR
- **Onboarding strategy** of IVOA software
- Including IVOA in key list of OSSR
- Use cases with IVOA software
- **VHE data level format** (CTA, KM3NeT, …) development: upcoming CEVO workshop

WP3 – WP5 Topics

- Integrating **Jupyter notebooks** from OSSR
- **Adding workflows** to OSSR
- **Finding software** through ESAP
- Use of **containerized software**
- How to treat **complex workflows**?

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

Coordination with EOSC and Extension of Scope

- National projects:
  - **Punch4NFDI** (German national scientific data management project - OSSR, data-lake and analysis platform discussion on PunchLunches)
  - **HGF** (German Helmholtz Institute funding) together with *FAIR*
    → extend OSSR to other communities
- EOSC Future
  - **Test Science Projects** (Extreme Universe and Dark Matter)
    - Support for technical implementation: software onboarding and virtual research environment
  - **EOSC Cluster Projects**
  - OSSR in EOSC: EOSC Provider Agreement / harmonising ToU
- EOSC Association: KG member of Software Infrastructure TF

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

Risks and Mitigation

- Unforeseen risks:

Nr.	Description of Risk (Level)	Proposed risk mitigation measures
5	Difficulty to find suitable contract staff (Medium)	Advertisements for contractual positions are to be distributed as widely as possible using the channels of the ESCAPE project.
6	Hired contract staff depart prematurely (Medium)	The partners will manage their contract personnel and indicate any risks if they arise to the WP management (Task 3.1).
7	Work and organisation restrictions due to COVID-19 pandemic	Delay of affected deliverables and milestones, extension of the grant agreement.

- Risk mitigation

Nr.	Did you apply risk mitigation measures?	Did your risk materialize?	Comments
5	Yes	partly	The risk mitigation did help,
6	Yes	Yes	Managed internally by partners and by result documentation
7	Yes	Yes	The severity will be governed by the further COVID-19 pandemic situation; funding transferred from offline events into organisation of online ones.

Deviations from and Implementation of the DoA

- OSSR provided input to the general ESCAPE dissemination and data management plans, no updates needed
- Implementation:
  - All the tasks have been fully implemented, all of the foreseen critical objectives have been achieved.
  - Partners on track to fulfil the agreed software and services to be onboarded; minor HR changes (re-hiring)
  - Otherwise all resources are available and deemed adequate until the end of the project
- There is no significant difference between the actual and planned resources up to this point.
- E-AB and RP review feedback part of work plan:
  - Resources listed in D3.6, cross-WP work intensified
  - OSSR landing page visible
  - User engagement via TSPs started; all software necessary will be tracked and listed
  - Intensify engagement with HSF: planned for this year

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

Summary and Outlook

- Project generally **in updated timeline** (6 month delay) with DoA
- Lessons learned:
  - the concept and implementation status of EOSC needs constant dissemination and explanation
  - Status and landscape of software diverse within the community, building the awareness and best practices needs compromises → community approach necessary and effective
- Open Points to be finalised in the remainder of the project:
  - Partner contributions
  - OSSR implementation:
    - Service onboarding
    - MetaData definition and eossr library
  - Enlargement of Community engagement
  - EOSC onboarding
  - Sustainment of OSSR
- Verification via TSPs and interactions with community and partner RIs

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

Feedback / Report

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

Feedback – WP3-related

- Objectives: WP3: E-OSSR
  The ESCAPE Open-source Scientific Software and Service Repository (OSSR) is now fully operational but has started to be populated with software and services. The last part of the project will be critical as a significant increase in the number of scientific software and services are expected to be added: some adjustments are expected, in particular the documentation on how to contribute. Advanced search (beyond zenodo search) to exploit the metadata and facilitate the reusage of software and services would be beneficial. Training material and online training event have been very successful.
- General Comment:
  The Open-Source Scientific Software and Service Repository is also a service that could be deployed/reused by other scientific domains: the guide for on-boarding new project (scientific software, container images, or repositories with full analyses environments) is very good; easy to follow and foster best software practices. What is unclear is if additional metadata are added. Also, it is recommended to put more emphasis on FAIR software practices. The search functionality is rudimentary.

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

Feedback – WP3-related

- Recommendations:
  - *FR2. Improve the way the training materials are "presented" online by i) provide an overview of the training strategy (from foundational skills for moving to Open Science to bespoke trainings on ESCAPE & EOSC services for the community) and detail the learning objectives and key concepts taught; ii) highlight the outcomes of the WP5 workshops, including addition of the (mini-)project topics of the participants and their final presentations. This could also lead to a more systematic Train the Trainer program through "ambassadors" of ESCAPE community Open Science services.*
    → set up a training concept web page?
  - *FR5. Consider more systematic approach to the continuous measurement of the effectiveness of your engagement with the user community (e.g., surveys / post-use questionnaires / resource consumption stats)*
    → follow up of first school
- Implementation:
  Moreover, the dissemination activities are somewhat unbalanced. Emphasis is put on their own website and the presence in Social media (647 posts). Several conference (13) and workshop (37) participations are reported, but no conference or workshop has been organised. This propagates in the estimates of persons reached.
  -> IWAPP, WOSSL not taken into account?

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

Other topics

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

Dates

- Next general assembly on 25/05/2022
- 2nd ESCAPE Data Science School – 20 to 24/06/2022
- ESCAPE Closing Event (general) on 25/10/2022
- Closing Meeting (WP3) – 11/2022?

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

Onboarding to EOSC Portal

- OSSR is a "data provider" to the EOSC
  - Question: how should all OSSR entries be registered with the EOSC portal
- Coordination meeting with openAIRE on 05/04/2022
  - openAIRE will establish a full link with the EOSC portal by 09/2022; so it is sufficient to register with them only; all OSSR entries will then also go to the EOSC portal → *visibility to partners*
  - So far there is no way for openAIRE to have a logical link of a catalogue (=OSSR) and a data provider (=Zenodo) – they will work on that to provide *visibility to OSSR*
- Tasks for us:
  - Provide a vocabulary of OSSR entry types to openAIRE
  - Register OSSR as data provider with https://re3data.org/ or https://fairsharing.org → to be picked up by openAIRE

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