ESCAPE OSSR

Thomas Vuillaume, on behalf of all contributors:

Enrique Garcia, Jutta Schnabel, Kay Graf, Tamas Gal, Mark Kettenis, Christian Tacke, Marjolein Verkouter …

29-11-2022, ESCAPE OSSR workshop, Erlangen

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.

Todays' Menu

OSSR overview…

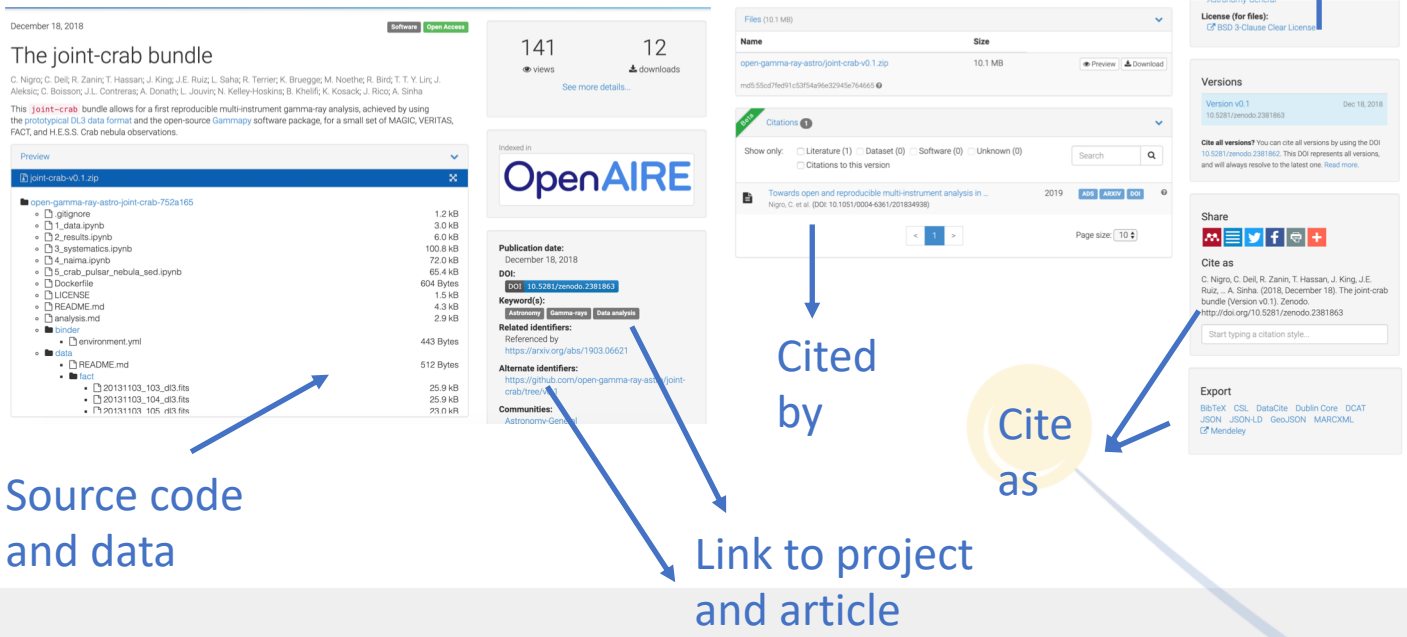
… To the future

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

A step back in time… Bruxelles, February 2020

An example of open science project : The Crab bundle

The Crab multi-instrument gamma-ray analysis with MAGIC, VERITAS, FACT and H.E.S.S.
https://github.com/open-gamma-ray-astro/joint-crab/tree/v0.1
https://zenodo.org/record/2381863#.XkxcD5NKhhA



Now imagine this as a standard,

in an integrated environment - with a single login,

allowing you to (re-)run (part of) any analysis, with another dataset,

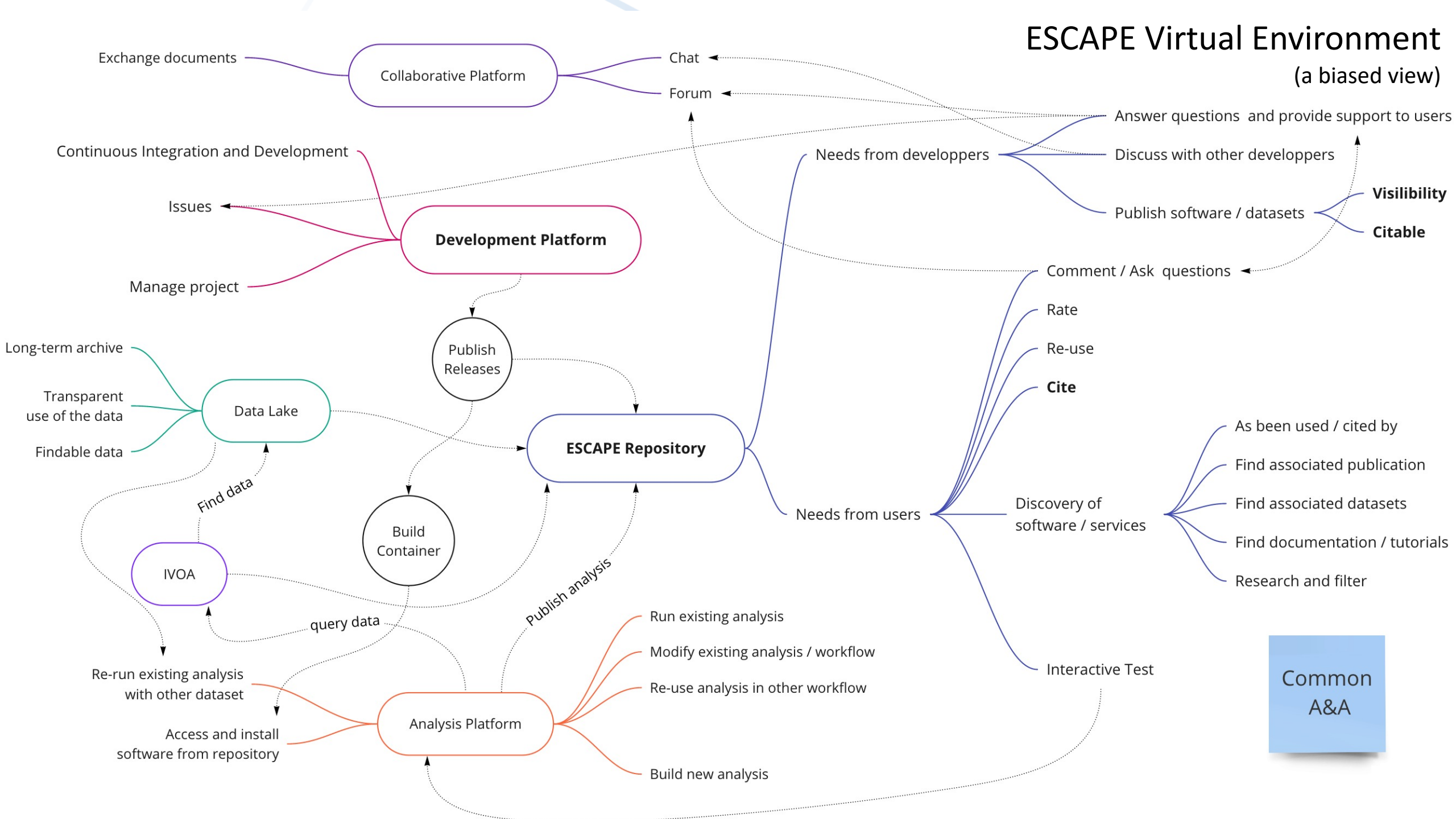
and easily publish your new results,

automatically giving credit to original analysis, datasets, workflows…

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

ESCAPE Virtual Environment

(a biased view)



Community Driven Repository, Organised by Science Projects

- Starting a new Science projet (through validation) sets up a complete virtual environment
- Researchers / Institutes contribute to the science project by publishing software / workflow / data
- The contributions are validated by science project curators
- Users can search the repository or explore it through the science projects

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

Policy & recommendations

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

OSSR policy

- Find it… in the OSSR! https://doi.org/10.5281/zenodo.6757113

- TLDR
  - License !
    - Permissive and open-source if possible (MIT, BSD-3…)
  - Add metadata
    - As a CodeMeta file
    - The more (accurate) metadata, the better
  - Version
  - Follow modern development practices (git, environment files, dockerise…)

- Final/updated version before ESCAPE ends?

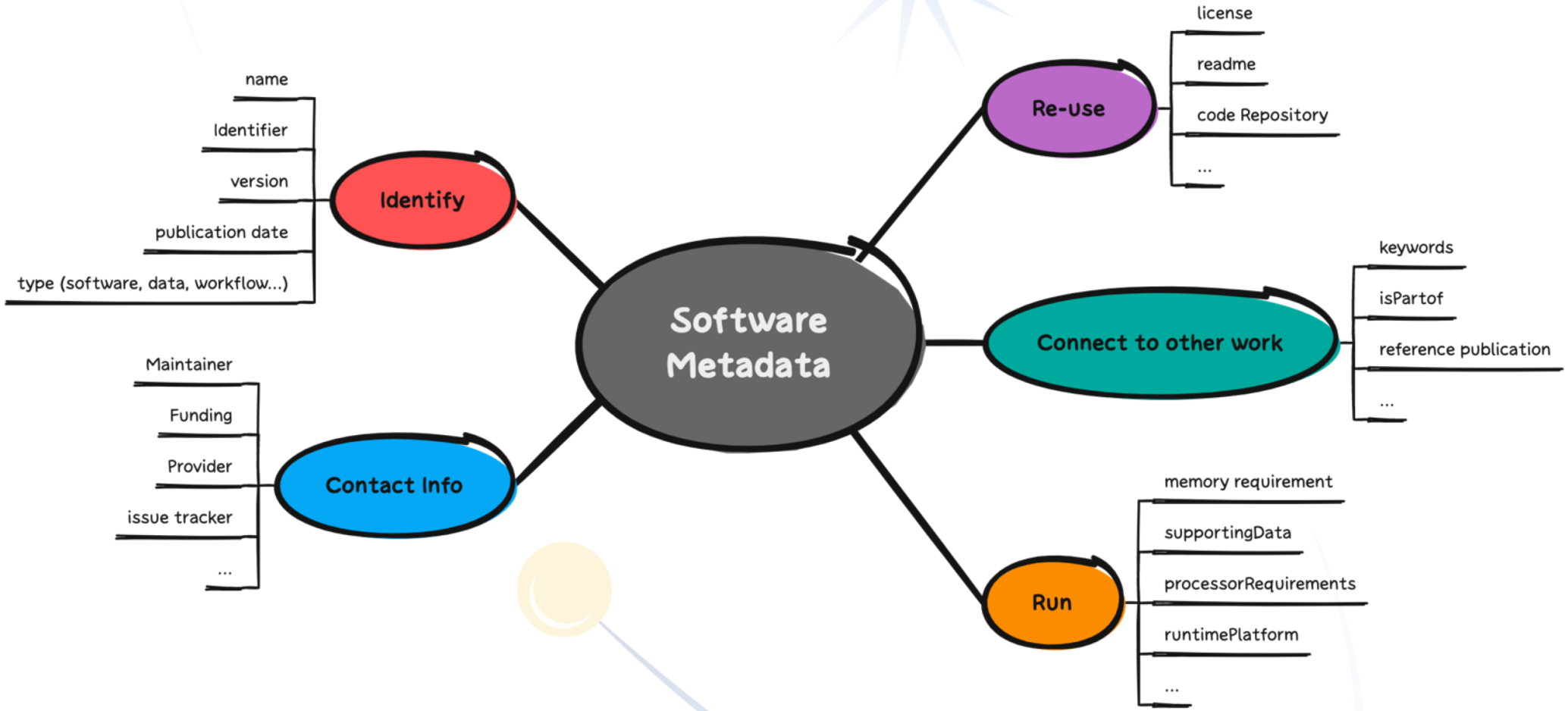
Funded by the European Union's Horizon 2020 - Grant N° 824064

MetaData & FAIRness

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

Metadata & FAIRness

- MetaData are the heart of the FAIR principles



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

Metadata & FAIRness

- It also allows services to install a software or run an analysis
  - by finding it: see e.g. the jupyter-notebook keyword in the OSSR allowing the ESAP to identify records that include and interative analysis
  - by running it: e.g. knowing the CPU or memory requirements
- For FAIRness, **MetaData must be included in a file included in each record.**
- ESCAPE implementation choice for software: CodeMeta (as a codemeta.json file)
- CodeMeta generator and validator online

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

An example of codemeta.json

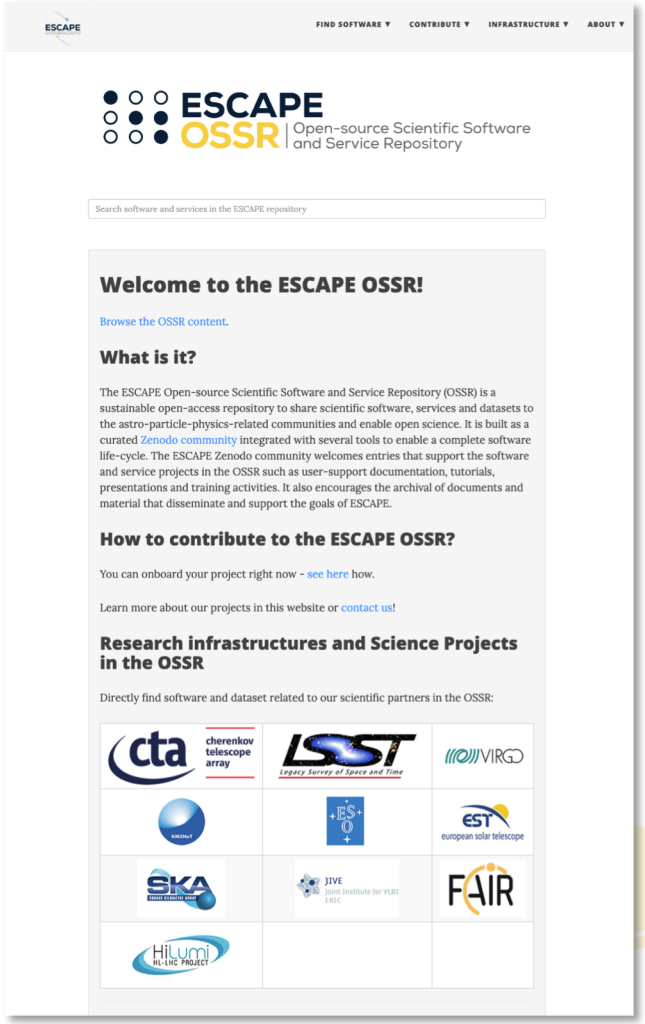

```
{
    "@context": "https://doi.org/10.5063/schema/codemeta-2.0",
    "@type": "SoftwareSourceCode",
    "name": "eossr",
    "description": "<p align=\"left\"><img src=\"docs/images/eossr_logo.png\" width=\"400px\" ></p><h1>The ESCAPE OSSR library</h1><p>T
    "license": "https://spdx.org/licenses/MIT",
    "version": "v0.6.2.dev34+g75395d9",
    "softwareVersion": "v0.6.2.dev34+g75395d9",
    "codeRepository": "https://gitlab.in2p3.fr/escape2020/wp3/eossr",
    "developmentStatus": "active",
    "isAccessibleForFree": true,
    "isPartOf": [
        "https://gitlab.in2p3.fr/escape2020",
        "https://projectescape.eu/"
    ],
    "contIntegration": "https://gitlab.in2p3.fr/escape2020/wp3/eossr/-/pipelines",
    "buildInstructions": "https://gitlab.in2p3.fr/escape2020/wp3/eossr/-/blob/master/README.md",
    "issueTracker": "https://gitlab.in2p3.fr/escape2020/wp3/eossr/-/issues",
    "readme": "https://gitlab.in2p3.fr/escape2020/wp3/eossr/-/blob/master/README.md",
    "programmingLanguage": [
        {
            "@type": "ComputerLanguage",
            "name": "Python",
            "url": "https://www.python.org/"
        }
    ],
    "softwareRequirements": [
        {
            "@type": "SoftwareApplication",
            "identifier": "requests",
            "name": "requests",
            "softwareVersion": ">=3.6"
        },
        {
            "@type": "SoftwareApplication",
            "identifier": "pytest",
            "name": "pytest",
            "softwareVersion": ">=5.4.2"
        }
    ],
    "keywords": [
        "jupyter-notebook",
        "zenodo"
    ],
    "runtimePlatform": "Python 3",
    "downloadUrl": "https://gitlab.in2p3.fr/escape2020/wp3/eossr/-/archive/v0.6.2.dev34+g75395d9/eossr-v0.6.2.dev34+g75395d9.zip",
    "releaseNotes": "",
```


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

Technical Implementation

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

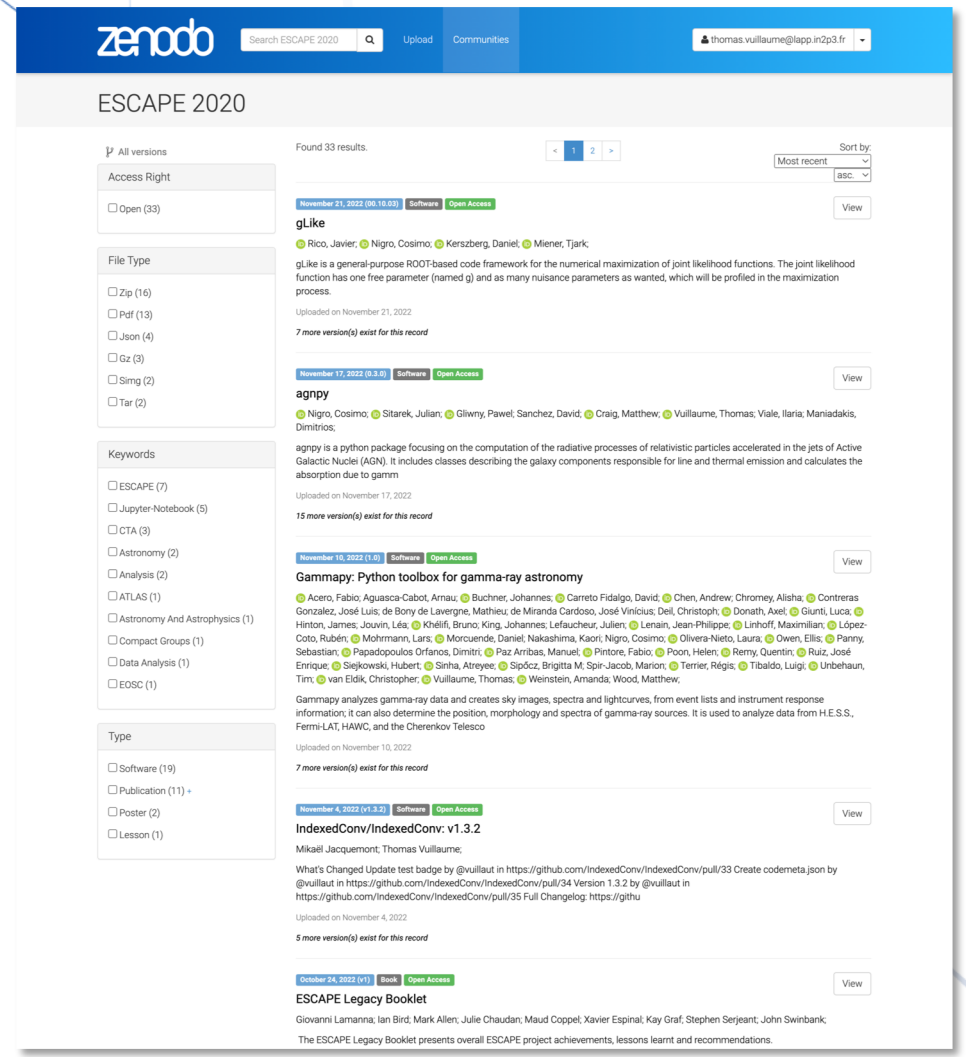
OSSR entry point: https://purl.org/escape/ossr



- Browse the OSSR
- Guidelines & policy
- Tutorials
- Contact

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

OSSR core: Zenodo & the *escape2020* community



- [Current stats]:

[3]: **OSSR statistics generated the 2022-11-30**

[4]:
```
md(text)
```

[4]: **There are 19 records in the OSSR.**

downloads: 4806

unique downloads: 2042

views: 19073

unique views: 6964

[6]:
```
md(text)
```

[6]: **Note that there are also 14 records in the ``escape2020`` community that are not software or datasets.**

You may find them directly on Zenodo

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

eOSSR

- *eOSSR* library
  - Dev: https://gitlab.in2p3.fr/escape2020/wp3/eossr
  - Doc: https://escape2020.pages.in2p3.fr/wp3/eossr/
  - OSSR: https://doi.org/10.5281/zenodo.6826881
  - **Gathers all OSSR developments and technical definitions**

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

eOSSR

- *eOSSR* library
  - Dev: https://gitlab.in2p3.fr/escape2020/wp3/eossr
  - Doc: https://escape2020.pages.in2p3.fr/wp3/eossr/
  - OSSR: https://doi.org/10.5281/zenodo.6826881
  - **Gathers all OSSR developments and technical definitions**
    - **OSSR API:** send request to the OSSR, find and filter software and services, upload new entries, update existing entries

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

eOSSR

- *eOSSR* library
  - Dev: https://gitlab.in2p3.fr/escape2020/wp3/eossr
  - Doc: https://escape2020.pages.in2p3.fr/wp3/eossr/
  - OSSR: https://doi.org/10.5281/zenodo.6826881
  - **Gathers all OSSR developments and technical definitions**
    - **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

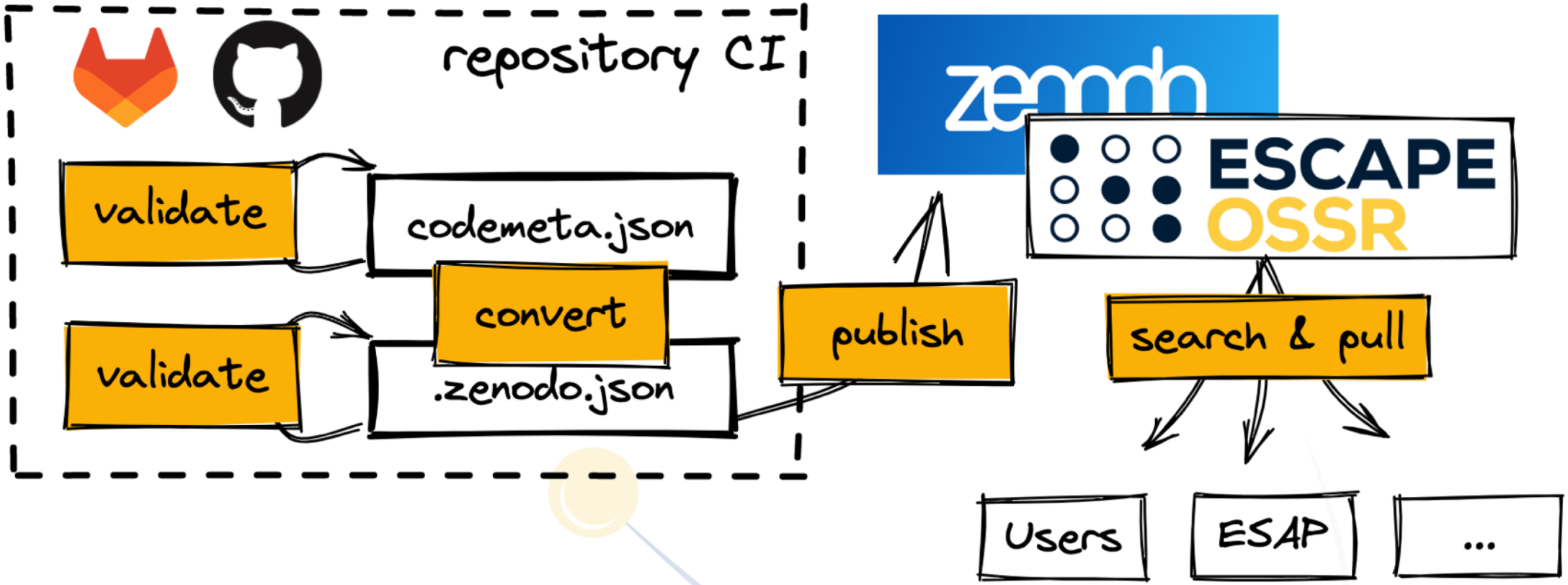
Funded by the European Union's Horizon 2020 - Grant N° 824064

eOSSR

- *eOSSR* library
  - Dev: https://gitlab.in2p3.fr/escape2020/wp3/eossr
  - Doc: https://escape2020.pages.in2p3.fr/wp3/eossr/
  - OSSR: https://doi.org/10.5281/zenodo.6826881
  - **Gathers all OSSR developments and technical definitions**
    - **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, validation

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

eOSSR



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

eOSSR proceedings

eOSSR presented as ADASS 2022 poster

proceedings: https://hal.archives-ouvertes.fr/hal-03876630v1

ESCAPE brings together the astronomy, astroparticle and particle physics ESFRI and pan-European communities with aligned challenges of data-driven research and demonstrated capabilities in addressing various stages of data workflow and concerned with fundamental research through complementary approaches.

eOSSR

Authors: Thomas Vuillaume*, Enrique Garcia, Christian Tacke, Tamas Gal, for the ESCAPE project
*thomas.Vuillaume@lapp.in2p3.fr

The aim of the ESCAPE WP3-OSSR (Open-source Scientific Software and Service Repository) is to provide the tools necessary for the communities to share their science products in a harmonized way respecting the FAIR principles, promoting open science and maximizing cross-fertilization by software re-use and co-development. One of the key components to achieve this goal is the software and service repository.
For its concept implementation, the ESCAPE repository is using Zenodo web service through the curated escape2020 community integrated with several tools to enable a complete software life-cycle. The ESCAPE Zenodo community welcomes entries that support the software and service projects in the OSSR such as user-support documentation, tutorials, presentations and training activities. It also encourages the archival of documents and material that disseminate and support the goals of ESCAPE.
We developed a Python library called eOSSR in order to allow an automated integration of Zenodo with other tools and platforms forming the OSSR as well as providing an integrated environment to external users.
The library is open-source and has been published in the OSSR itself.
You may find it at https://doi.org/10.5281/zenodo.6826881.
You may find documentation and running examples online: https://escape2020.pages.in2p3.fr/wp3/eossr/

Metadata: The OSSR has chosen codemeta.json as schema and format for its software metadata. The definition of this schema has been integrated within the eOSSR, thus allowing:
- An automated verification of the metadata.
- A converter between Zenodo metadata schema and the OSSR metadata

API: For other tool to communicate with the OSSR, an API was necessary. The eOSSR takes advantage of Zenodo's API to propose a set of high-level functionalities through Python functions or command-lines, such as:
- Requesting software in the OSSR via wide or specific searches, using plain text or recognized metadata such as keywords or file type.
- Uploading software to the OSSR (or Zenodo) via the CLI.
- Search for Zenodo's communities, and supported licenses, funders and grants.

JOIN OUR COMMUNITY

REGISTER ON THE WEBSITE! www.projectescape.eu
@ESCAPE_EU linkedin.com/in/escape-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 No. 824064.

The eOSSR library

Thomas Vuillaume[1], Enrique Garcia[1,2], Christian Tacke[3], and Tamas Gal[4]

[1]*Univ. Savoie Mont-Blanc, LAPP, CNRS, Annecy, France;* thomas.vuillaume@lapp.in2p3.fr

[2]*IT Department, CERN - 1211 Geneva 23 - Switzerland*

[3]*GSI Helmholtz Centre for Heavy Ion Research GmbH, Darmstadt, Germany*

[4]*IT Erlangen Centre for Astroparticle Physics, Erlangen, Germany*

Abstract.
The astronomy, astroparticle and particle physics communities are brought together through the ESCAPE (European Science Cluster of Astronomy and Particle Physics ESFRI research infrastructures) project to create a cluster focused on common issues in data-driven research. Among the ESCAPE work packages, the OSSR (ESCAPE Open-source Scientific Software and Service Repository) is a curated, long-term, open-access repository that makes it possible for scientists to exchange software and services and promote open science. It has been developed on top of a Zenodo community, connected to other services. A Python library, the eOSSR, has been developed to take care of the interactivity between Zenodo, services and OSSR users, allowing an automated handling of the OSSR records. In this work, we present the eOSSR, its main functionalities and how it's been used in the ESCAPE context to ease the publication of scientific software, analysis, and datasets by researchers.

1. The ESCAPE OSSR

The aim of the ESCAPE OSSR (Open-source Scientific Software and Service Repository) is to provide the tools necessary for the communities to share their science products in a harmonized way respecting the FAIR principles, promoting open science and maximizing cross-fertilization by software re-use and co-development. One of the key components to achieve this goal is the software and service repository. For its concept implementation, the ESCAPE repository is using Zenodo web service through the curated escape2020 community integrated with several tools to enable a complete software life-cycle. The ESCAPE Zenodo community welcomes entries that support the software and service projects in the OSSR such as user-support documentation, tutorials, presentations, and training activities. It also encourages the archival of documents and material that disseminate and support the goals of ESCAPE.

1

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

MetaData generator and validator

https://escape2020.pages.in2p3.fr/wp3/codemeta-generator/

eOSSR ESCAPE OSSR Open-source Scientific Software and Service Repository

[Run online](#)

Validate and convert your metadata

This notebook will help you validate your metadata for an upload to the ESCAPE OSSR.

To do so, upload your codemeta metadata, either using an URL pointing to the `codemeta.json` file, uploading a `codemeta.json` file or copying the metadata in the text box below.

Note that you can generate your ESCAPE codemeta file using the online generator: https://escape2020.pages.in2p3.fr/wp3/codemeta-generator/

`Load codemeta from a json file`

Upload (0)

`Load codemeta from an URL`

URL:

Load

codemeta:

Validate !

Convert to .zenodo.json

Add ESCAPE metadata

Implementation into the OSSR environment

From a single click

- Publishes source code (updates your existing record with new versions)

zenodo

- Long term archived
- Findable
- Citable

GitLab

1. Make a new tag (release)
2. Let the CI do the rest

- builds images

- publishes to OSSR

- Publishes on registries

GitLab registry

docker hub

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

Curation

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

Curation

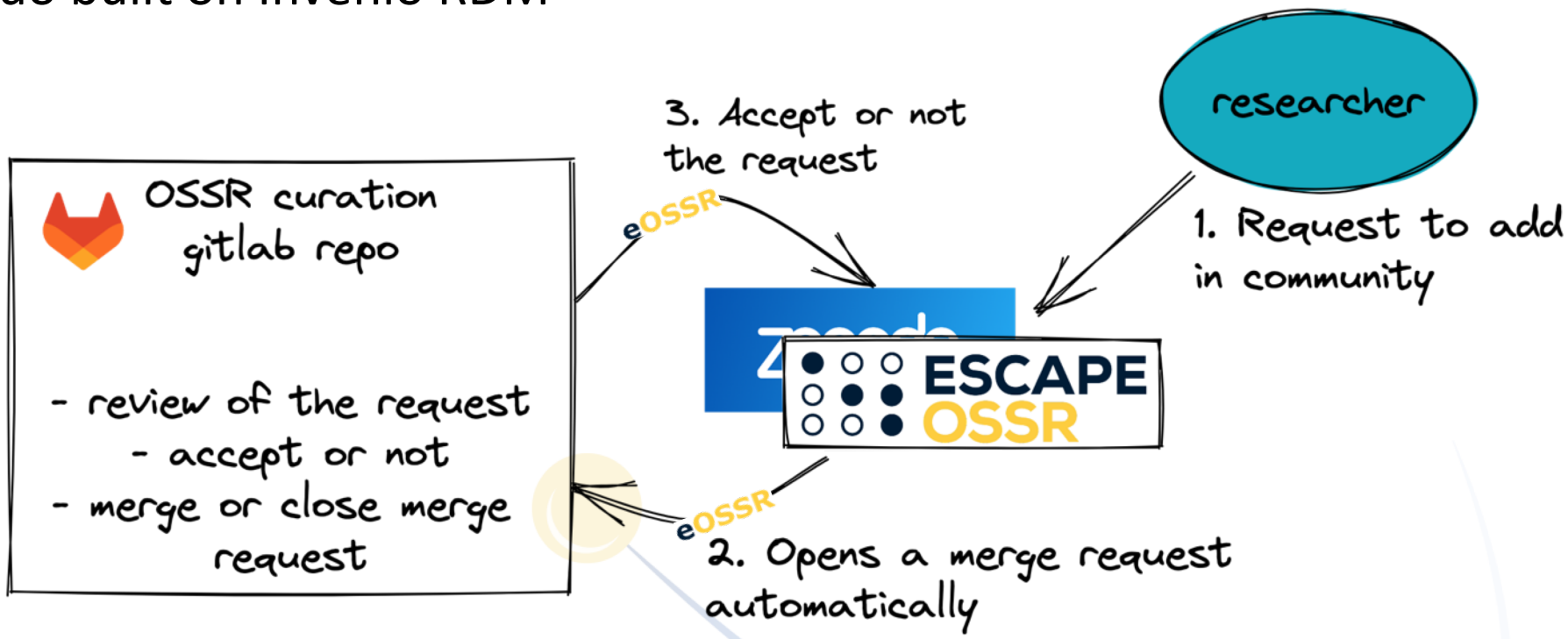
- Each entry in the OSSR is reviewed to:
  - Ensure that it matches our policy and requirements
  - Make recommendations
  - Help implement them

- It happens through:
  - A recorded presentation of the software, analysis or service
  - A formal final review and acceptance through ossr-curation

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

Curation: ossr-curation

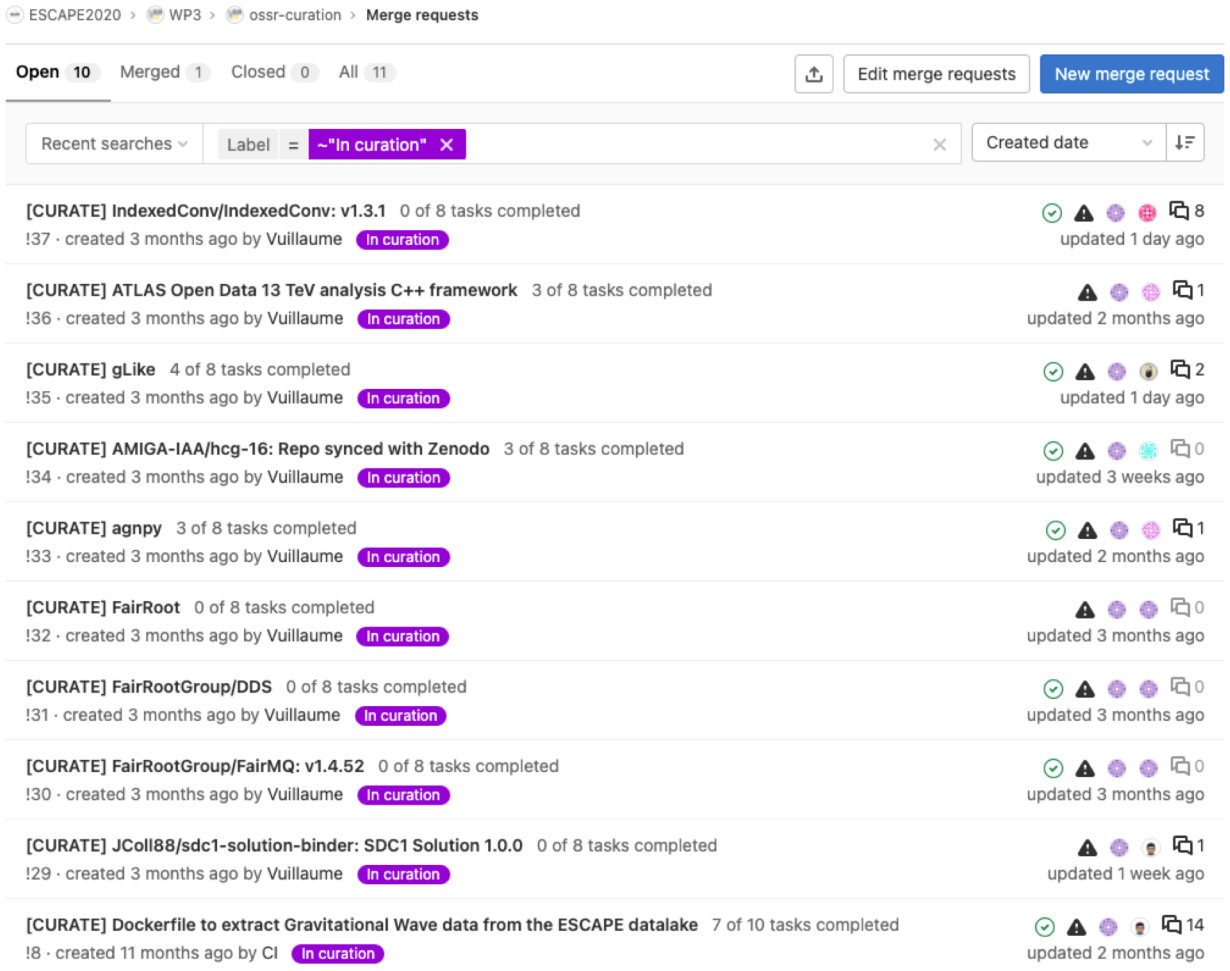
- Issue: Zenodo currently supports a single community maintainer and reviewer
- OSSR curation: a solution to collaboratively and openly review requests in a Zenodo community. Note that it might become obsolete in a future version of Zenodo built on Invenio RDM

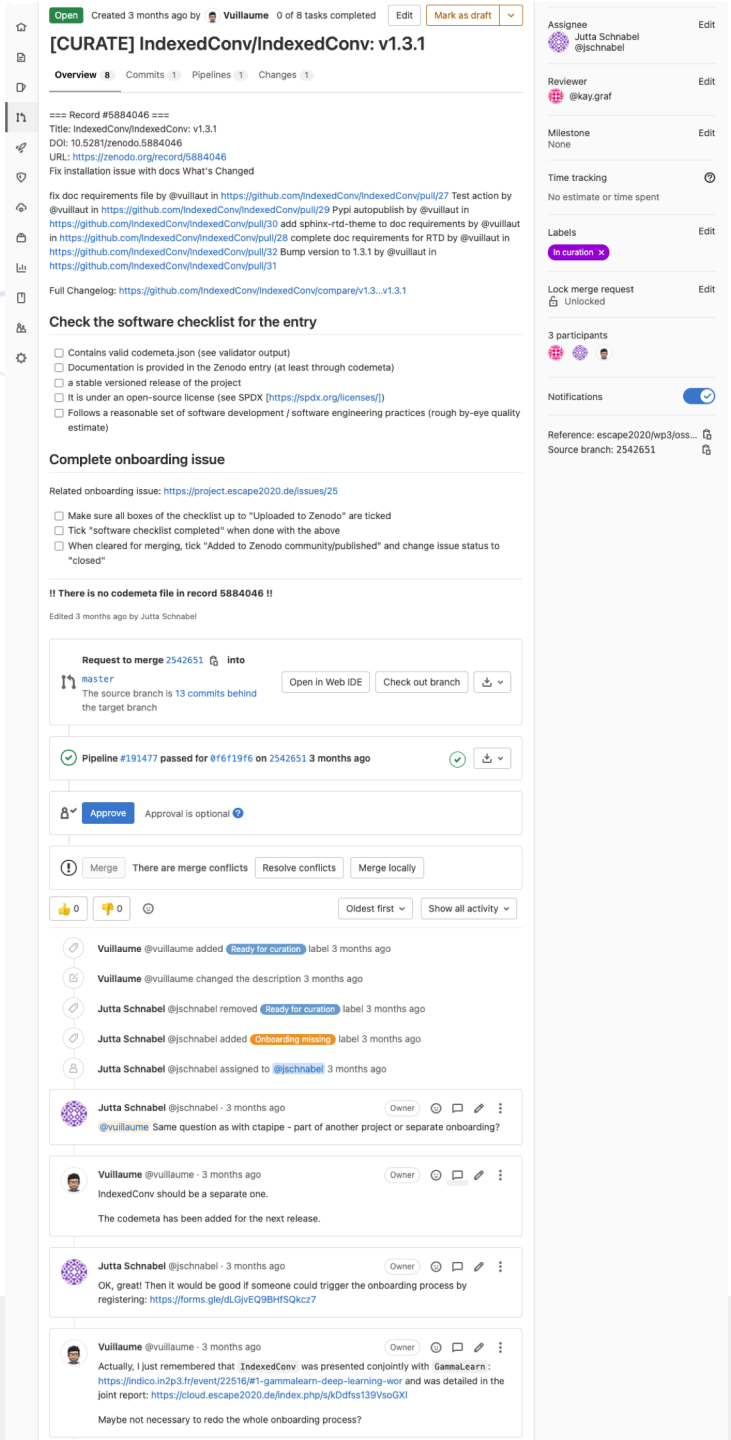


Funded by the European Union's Horizon 2020 - Grant N° 824064

Curation: ossr-curation

https://gitlab.in2p3.fr/escape2020/wp3/ossr-curation





How to publish in the OSSR

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

How to publish in the OSSR

- Instructions online: https://escape2020.pages.in2p3.fr/wp3/ossr-pages/page/contribute/onboarding/

1. Fill a form to request a presentation
2. Prepare your repository: see requirements and metadata
   - Validate your codemeta.json using the eOSSR in your CI
3. Publish to Zenodo: ease your life using the eOSSR
4. Add escape2020 as a community (this comes for free if you use the eOSSR)
   - If your software was already in Zenodo, you can edit the metadata and add the escape2020 community through the web portal
5. Final review and acceptance → ossr-curation

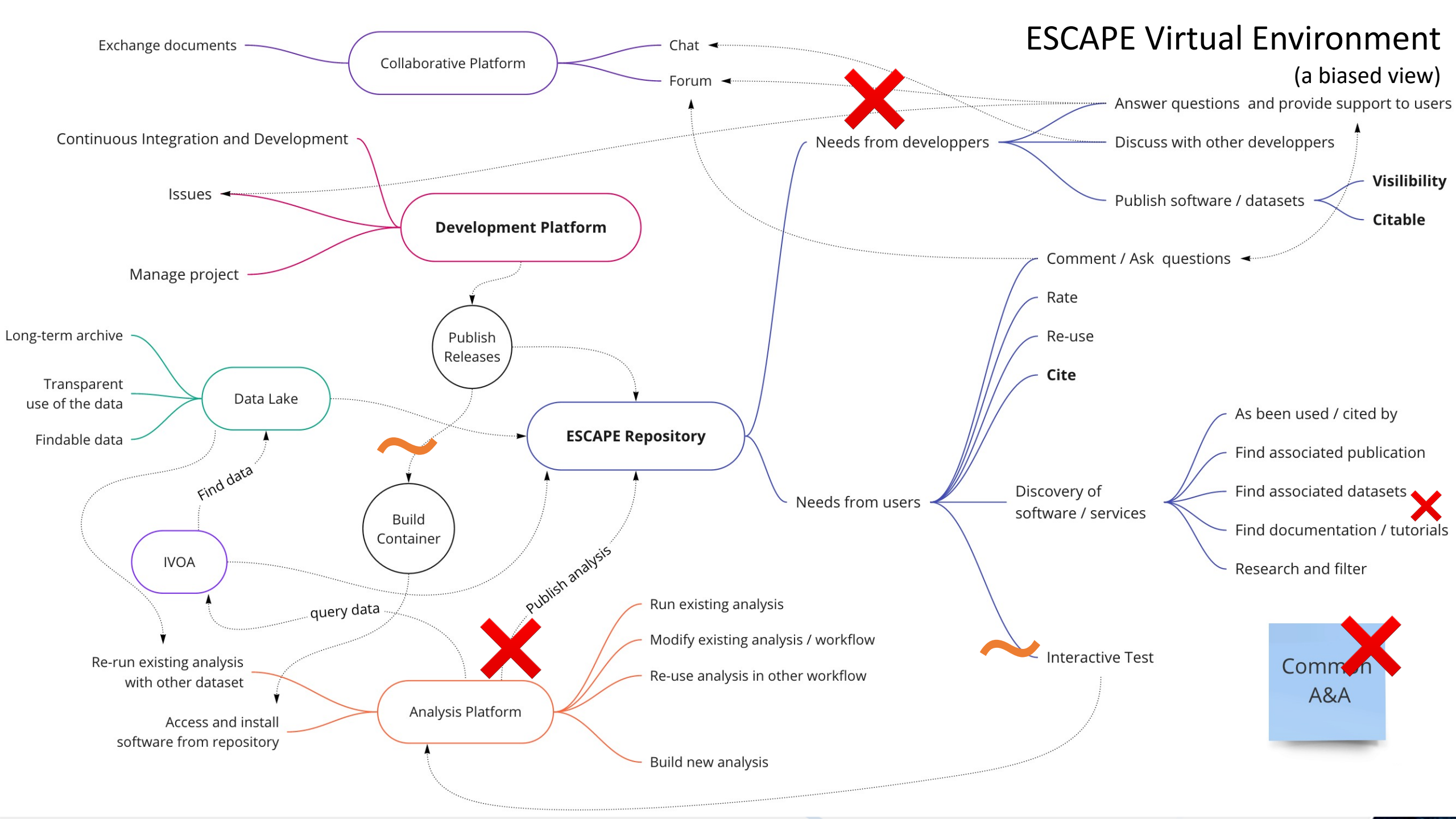
Funded by the European Union's Horizon 2020 - Grant N° 824064

Things we could improve
(a.k.a. I wished we did better)

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

ESCAPE Virtual Environment

(a biased view)



- Find associated software / dataset / analysis…
  - Possible in metadata but still not clearly defined and in the hands of developers
  - In Zenodo, ***related identifiers*** (continues this upload, is cited by this upload, etc…)

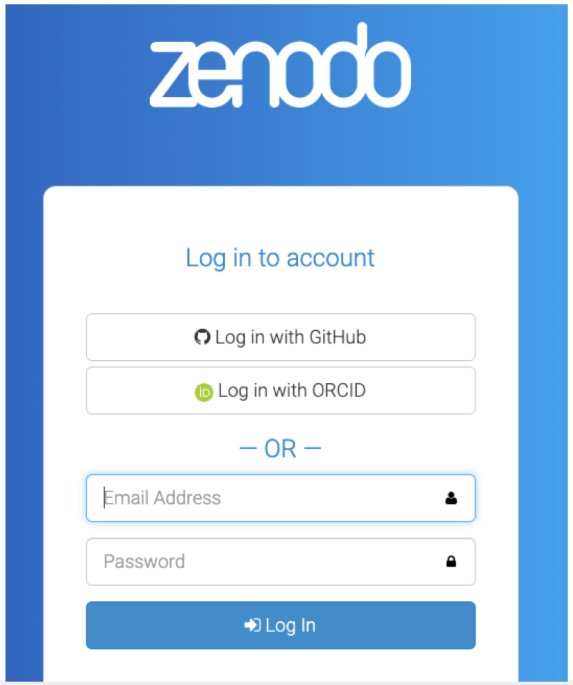
- Publish analysis from analysis platform
  - Possible using eOSSR but requires an integration between ESAP and the development platforms

- CodeMeta
  - Not automated enough
  - Not integrated enough => some hope with InvenioRDM ?
  - Still too much a burden for developers

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

- Containers and registry
  - Some attempts: https://gitlab.in2p3.fr/escape2020/wp3/ossr-registry
  - We rely on developers to do their job

- Common A&A
  - Not entirely in our hands for the OSSR
  - Will improve with InvenioRDM?



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

The future of the OSSR ?

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

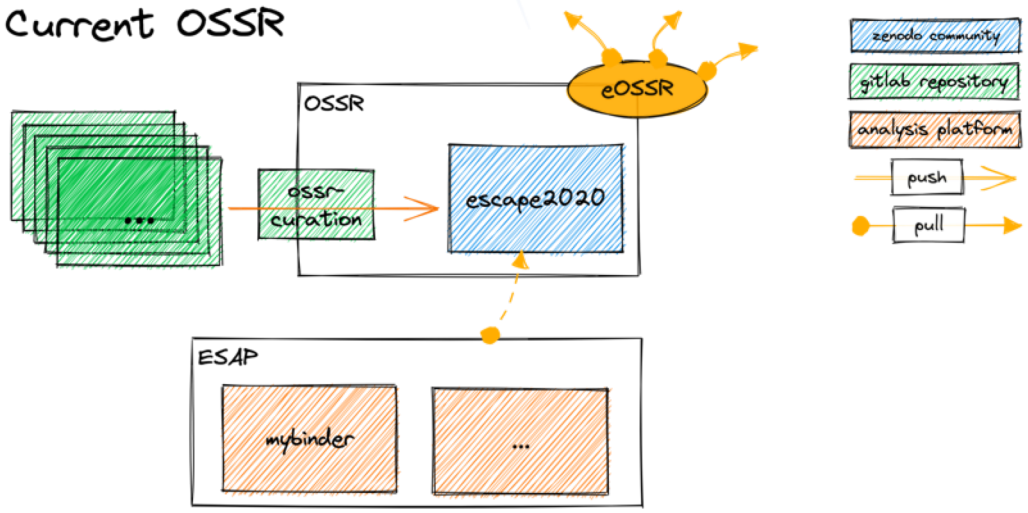
Some things I'd like to finish before ESCAPE ends

- eOSSR v1.0 release – see milestone in GitLab
- make clear recommendations on how to maintain codemeta.json in a repository
  - Based on the work done recently for gammapy and other repositories
- Finish all on-going curations
- OSSR paper gathering these technical developments and recommendations

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

The future of the OSSR: federating the decentralized?

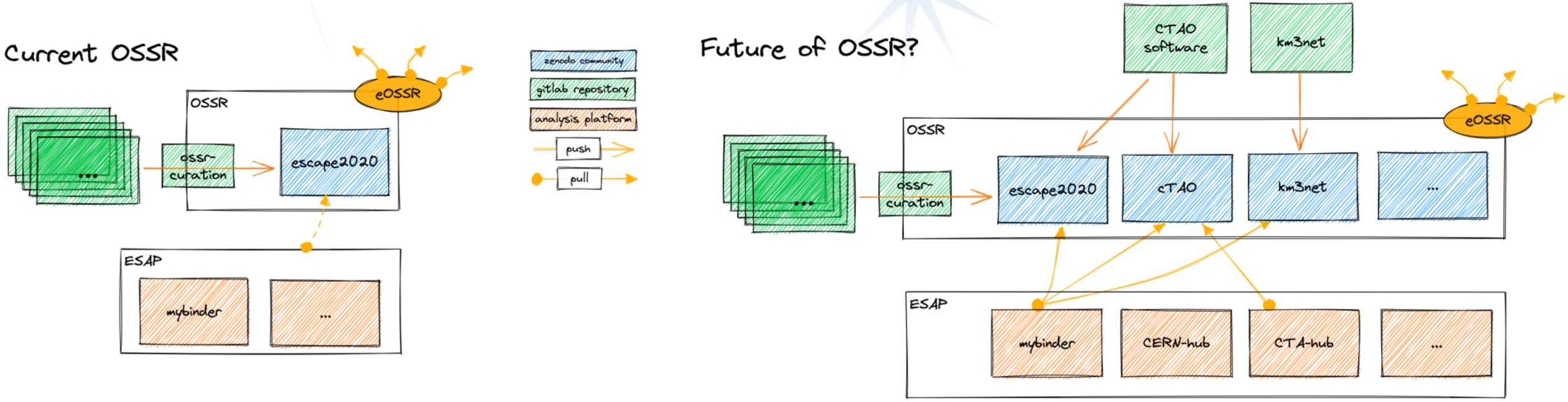
- ESCAPE as a funded project is ending, but the OSSR can stay (though in different forms)



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

The future of the OSSR: federating the decentralized?

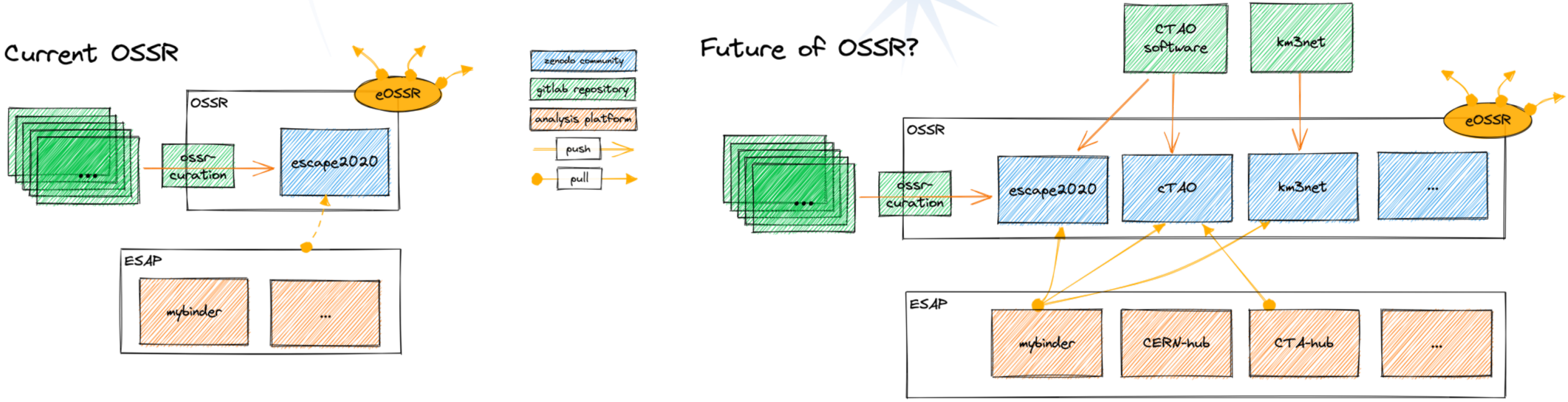
- ESCAPE as a funded project is ending, but the OSSR can stay (though in different forms)



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

The future of the OSSR: federating the decentralized?

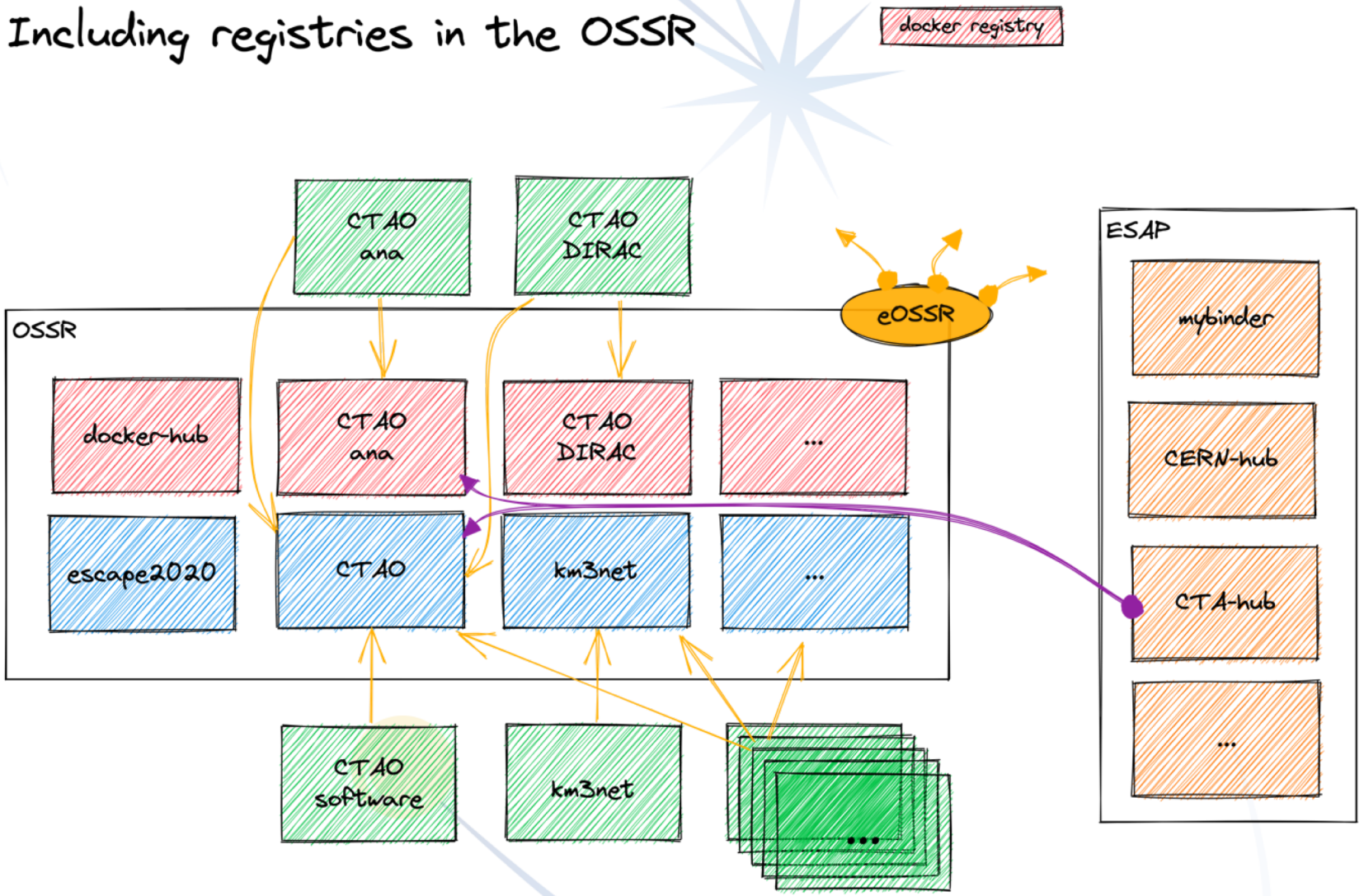
- ESCAPE as a funded project is ending, but the OSSR can stay (though in different forms)



- Decentralization powered by the eOSSR, serving the API to fetch the federated OSSR
- Each community has to commit to follow a minimal set of requirements to be part of the OSSR
- Domain specific platforms will have different requirements, that can be requirements of each Zenodo community

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

The extra mile: community specific **registries** could also be integrated the same way in the OSSR



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

Curation

- The curation is a central part of a trusted repository
  - part of it is automated (e.g. checking for correct metadata)
  - but human curation is essential
- Curation is (sometimes) domain specific
- For the OSSR to keep growing after ESCAPE ending, we need curators

⇒ A commitment from the research infrastructure is needed to assign experts able to do software curation in their domain

⇒ In the decentralized version, RI commit to maintain and follow a common base of requirements, and do the curation internally (with potentially more requirements)

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

Curation

- The curation is a central part of a trusted repository
  - part of it is automated (e.g. checking for correct metadata)
  - but human curation is essential
- Curation is (sometimes) domain specific
- For the OSSR to keep growing after ESCAPE ending, we need curators

⇒ A commitment from the research infrastructure is needed to assign experts able to do software curation in their domain

⇒ In the decentralized version, RI commit to maintain and follow a common base of requirements, and do the curation internally (with potentially more requiremer

⇒ In line with the Open Collaboration agreement signed in Brussels

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

Recommendations to ESFRIs

- A clear set of guidelines and requirements to publish any software and analysis in collaborations
  - Following OSSR requirements
- A curation team dedicated to software and analysis code
- A software/analysis publication process for analysers
- A software/analysis review process for curators
- A set of templates for software and analysis for members to use

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

Conclusion: mission accomplished!

- There are still things to improve
- We did a good job and achieved our goals and more
- Let’s make sure it does not all disappear with ESCAPE ending

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

Thank you for your attention

Questions?

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

Finally, some inspiring projects

A list of conference publications:

- ESCAPE Open-source Scientific Software and Service Repository
- The ESCAPE Data Science Summer School 2021 ADASS XXXI Poster
- The eOSSR library

Some inspiring projects:

- https://deepnote.com/ - collaborative analysis platform
- https://codeocean.com/ - an analysis platform with and integration publication process
  - eOSSR demo: https://codeocean.com/capsule/6789403/tree
- https://biocontainers.pro/ - an example of workflows registry in the bio community

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