OSSR Implementation concept & status

Enrique GARCIA – LAPP/CNRS

27/07/2020 – ESCAPE WOSSL workshop

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.

Outline

- Repository implementation Goals

- Development platform
  - GitLab in2p3 instance
- Long term repository
  - Zenodo
- Connection of services
  - Containerization
  - CI/CD pipeline between Dev platform and repository
  - Metadata

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

OSSR Goals

- Converge to common Open Source and FAIR practices that adapt and evolve to the community

1. Provide an open-source **development platform** available for all partners and beyond;

2. A long-term **repository** insuring the implementation of the FAIR principles and the **connection** with **other** ESCAPE **services** (ESAP, EOSC portal/marketplace)

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

Development Platform

- Environment where to:
  - Develop software.
  - Test, modify, debug, release... software

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

GitLab in2p3 instance

- GitHub is not OS
  - Bought by Microsoft 2008
- GitLab is hosted within a ESCAPE institution
  - https://gitlab.in2p3.fr/escape2020
- No intention of forcing communities to change habits !
  - Make available an environment if anybody needs it !
  - Feel free to create new projects here !

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

(Git)hub/Lab as a service/software

	+	-
Github or gitlab as a service	- Free for open source - No setup - Integration with lot of services - Huge community	- Data stored outside the E.U. - No control over the data & conditions of use - No control of accounts
Gitlab as a software	- Private projects - Control - Open source	- Setup, install, maintenance of the service - Costs - Less integration with other services

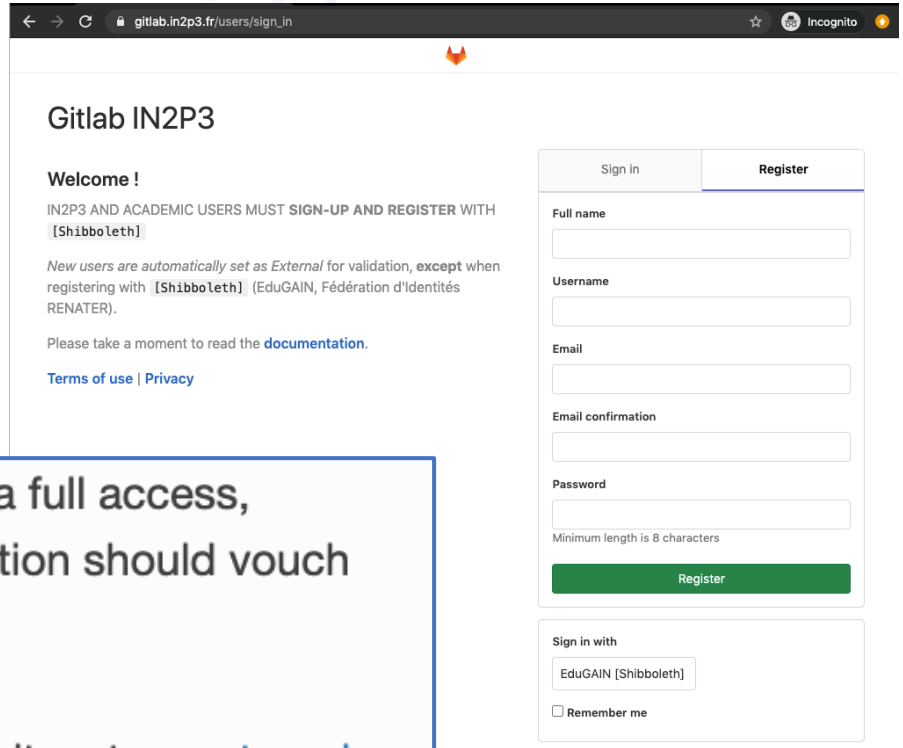
- Slide from T. Vuillaume (24/07)

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

GitLab in2p3 instance - FAQs

- `I don’t have a EduGAIN/Shibboleth account’

  (thanks Jutta !)



For non-eduGAIN users who wish to get a full access, someone belonging to a partner organization should vouch for you to user support.

New non-eduGAIN users will be, by default, set as external. Validators will unlock the account after verification.

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

Long term repository

- A environment where to store digital artefacts.

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

Long term repository

- A environment where to store digital artefacts.

Software **Repos**itory != **Develop**ment Platform
Stable release != Minor changes in a project

- ESCAPE needs → General purpose;
  - Software
    - Software containers
  - Documents, presentations, multimedia…
  - Full projects
  - Data sets *

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

Zenodo

- General purpose repository

- FAIR principles
  - Findable
    - Reference/Identification by DOIs + metadata
  - Accessible
    - Archival + Human/machine readable metadata
  - Interoperable
    - Accepts different metadata schemas
  - Reusable :
    - License and provenance

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

Zenodo as a service/software

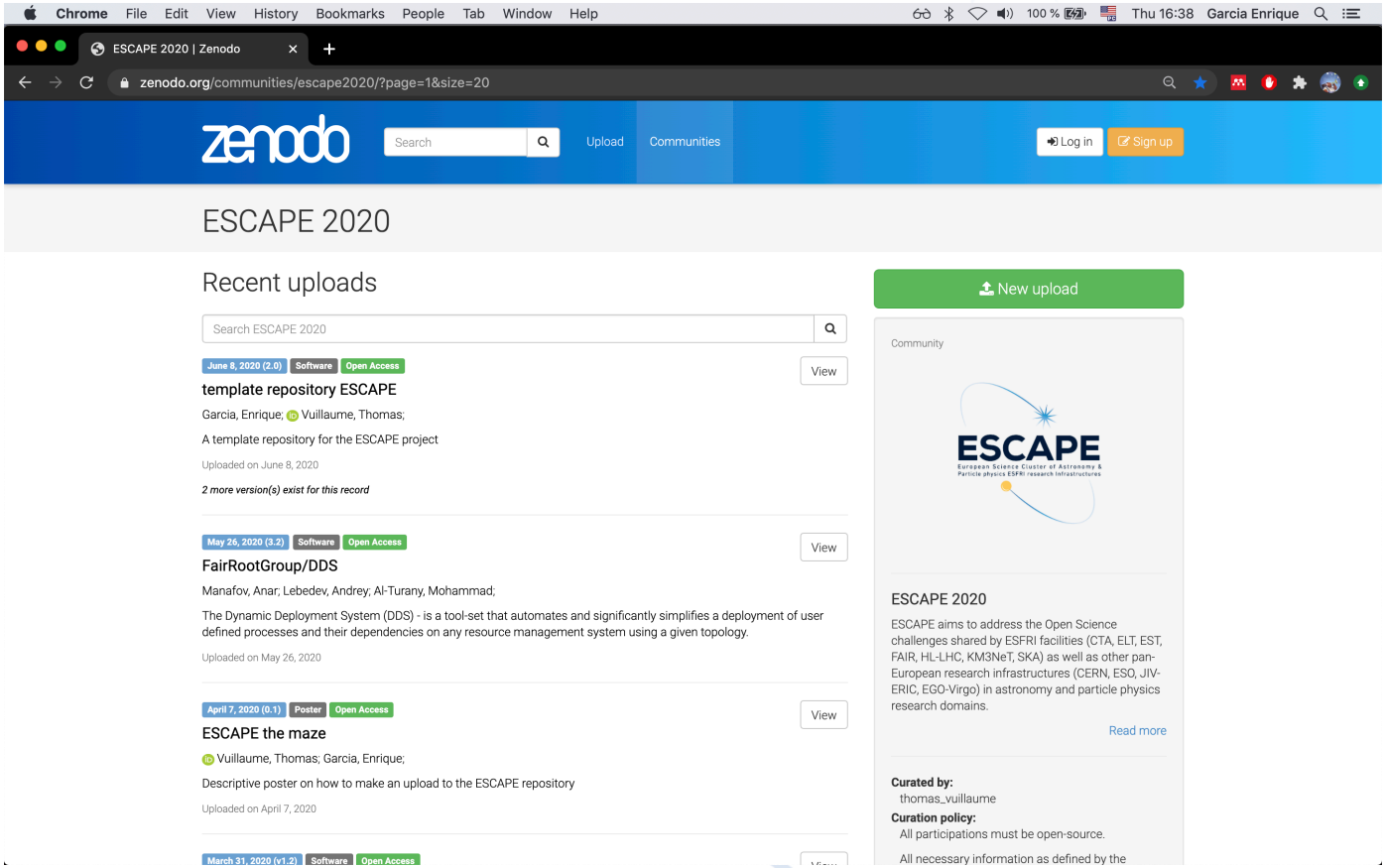
	+	-
Zenodo as a service	- Free - No setup - CERN support - Secure, **very** long time, archive - Always up to date - Integration with OpenAIRE	- Size limitation to 50GB per upload (an agreement is possible for bigger datasets) - No multiple ownership as of today - No multiple curators as of today - Limited customization
Zenodo as a software (later InvenioRDM)	- Better branding - Own the data - Custom front-end - Custom A&A	- Need time and money to dev/setup/install - Need computing infrastructure - Need maintenance - Future ?

- Slide from T. Vuillaume (24/07)

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

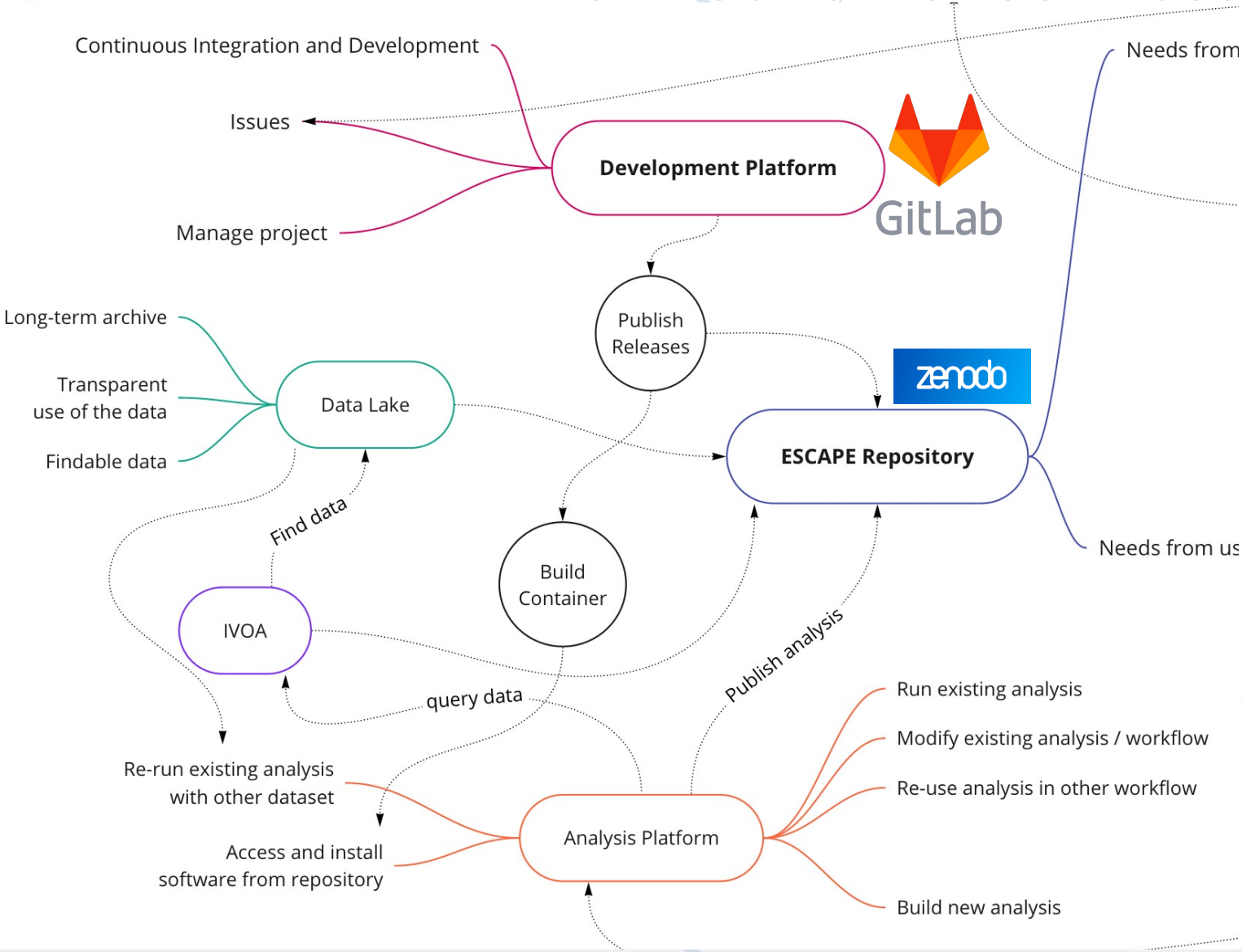
ESCAPE 2020 Community

- https://zenodo.org/communities/escape2020/



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

Connection of services



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

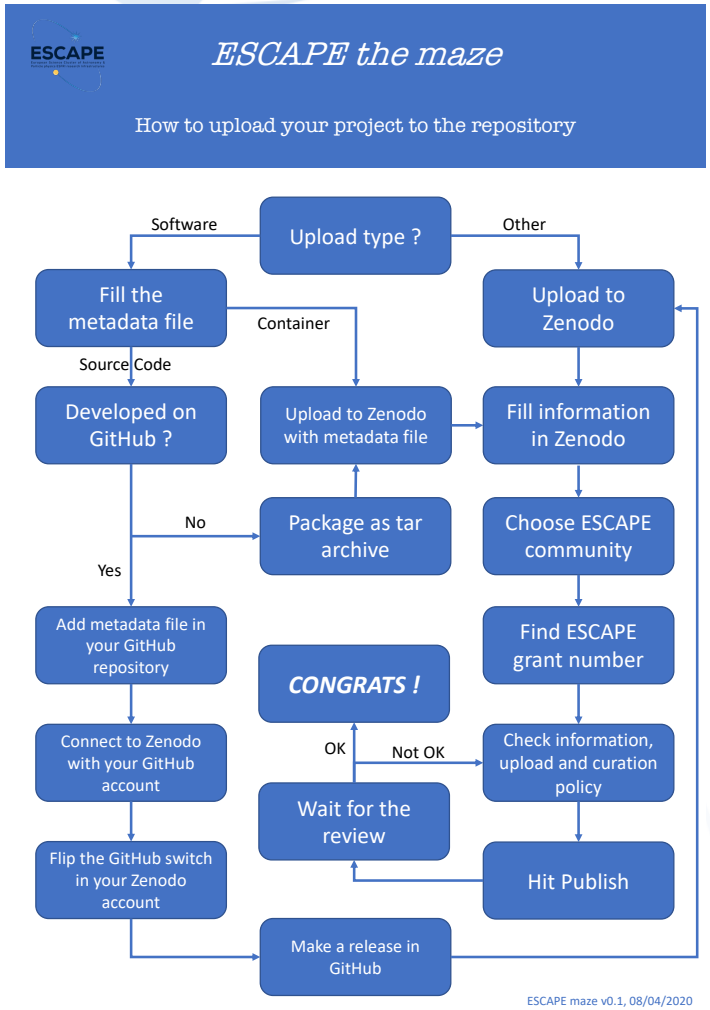
Connection of services. Containers

- Satisfy all the FAIR principles.
- Allow an easy and fast reproducibility of very complex environments.
- The expected/proposed way of reproducing studies in the analysis platform.

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

Connection dev platform - repository

- Upload a project to Zenodo:
  - *ESCAPE* the maze
  - zenodo.org/record/3885172



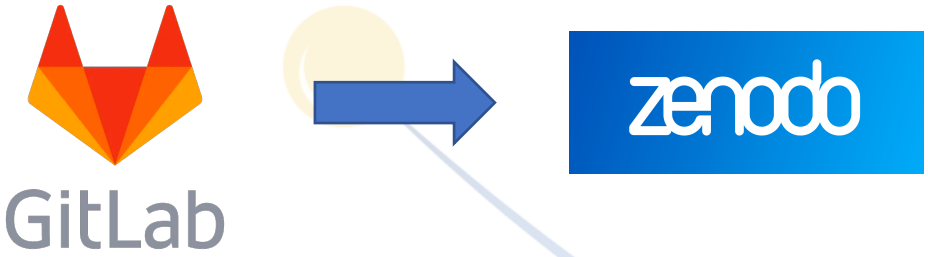
Funded by the European Union's Horizon 2020 - Grant N° 824064

Automatic connection GitLab-Zenodo

- Development platforms provide very useful tools for continuous integration and continuous delivery/deployment (CI/CD).

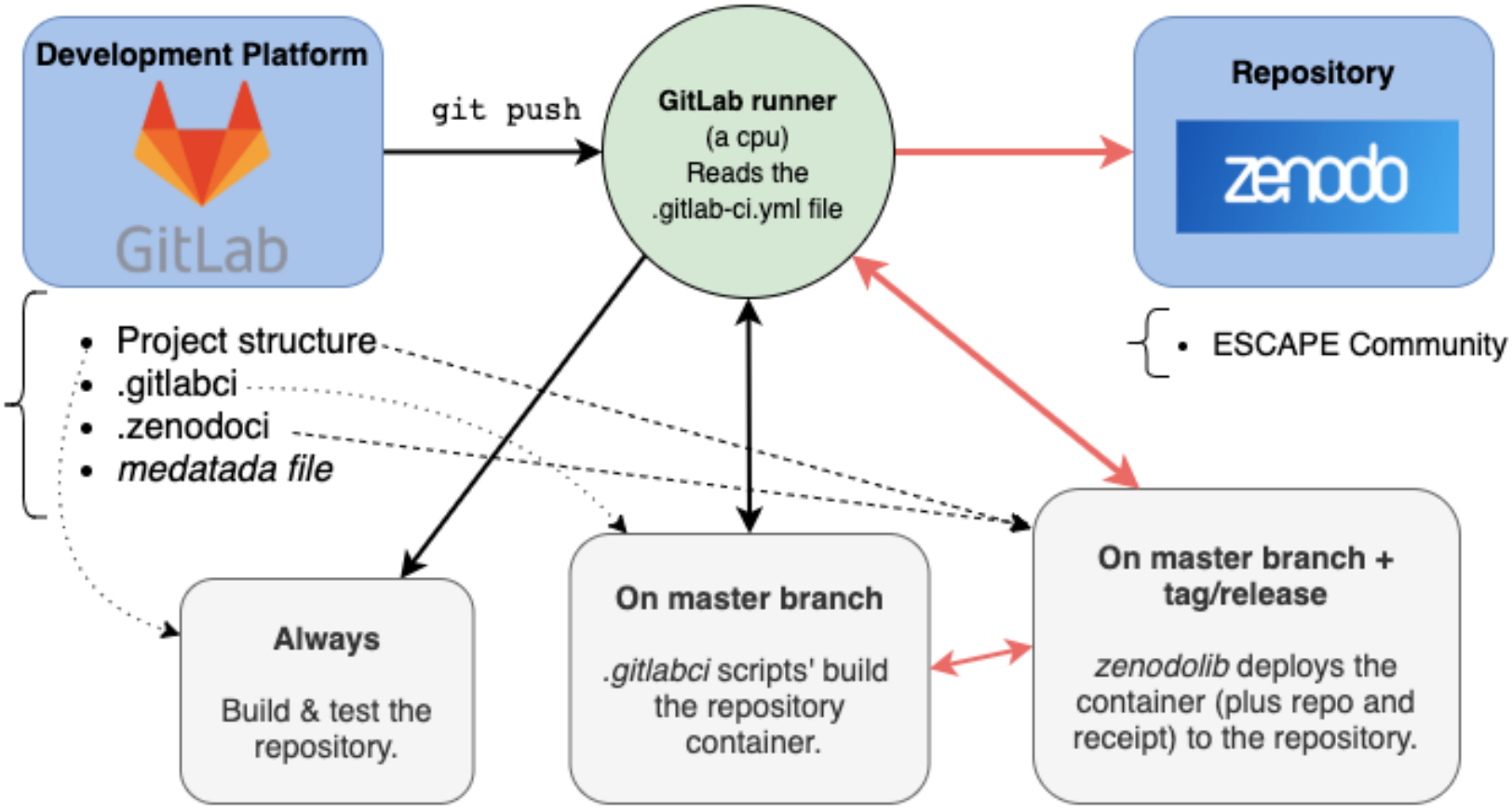
https://gitlab.in2p3.fr/escape2020/escape/template_project_escape

- Use the development platform to develop, build and test,
- and deploy a container/project to the repository (or endpoint).



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

Automatic connection GitLab-Zenodo



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

Automatic connection GitLab-Zenodo (I)

- Following FAIR principles:
  - Singularity-CI builders

1. Invokes DockerHub
2. Creates a Singularity container
3. Uploads the container to an endpoint (+ dev plat)

- https://github.com/singularityhub/singularity-ci
- Tutorial : https://vsoch.github.io/2018/gitlab-singularity-ci/

Singularity-CI Builders

This is an entrypoint to show you how you can build Singularity containers on different Continuous Integration Services. The templates provided are served by Gitlab or Github, depending on where the CI services are commonly used.

General Templates

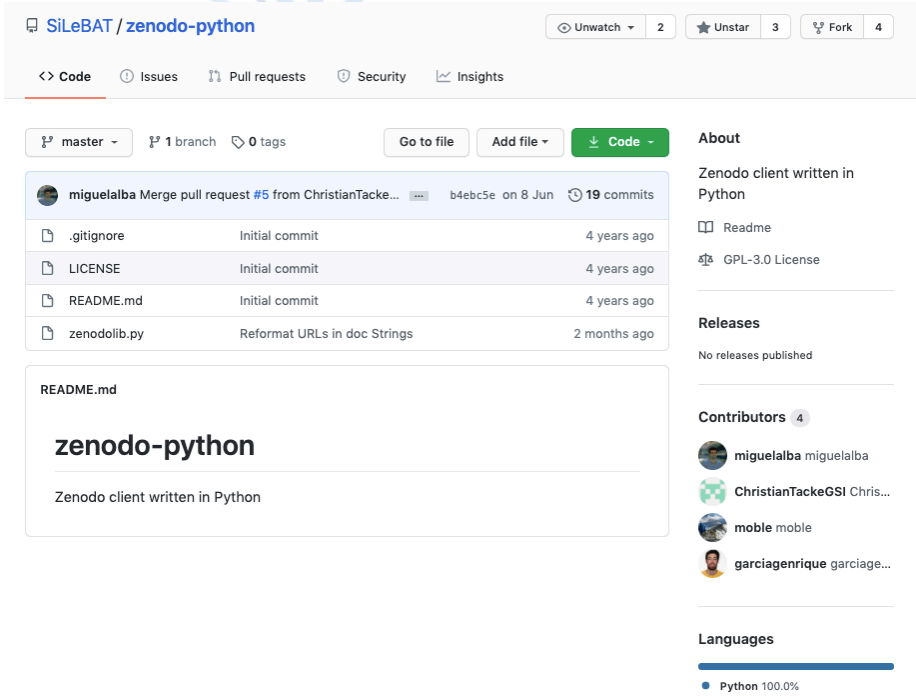
The following are general builders that you can customize to build on Travis, CircleCI, GitHub, or GitLab.

Service	Status	Template	
GitLab		singularityhub/gitlab-ci	
TravisCI	build passing	singularityhub/travis-ci	
CircleCI	PASSED	singularityhub/circle-ci	
GitHub		singularityhub/github-ci	

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

Automatic connection GitLab-Zenodo (II)

- Following FAIR principles:
  - *zenodo-python* repository
    - Provides the *zenodolib* library
    - Manages upload to Zenodo through their API
    - Please use Zenodo sandbox for tests ! https://sandbox.zenodo.org/deposit



- https://github.com/SiLeBAT/zenodo-python/
- REST API tutorial: https://developers.zenodo.org

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

Automatic con. GitLab-Zenodo (wrap up)

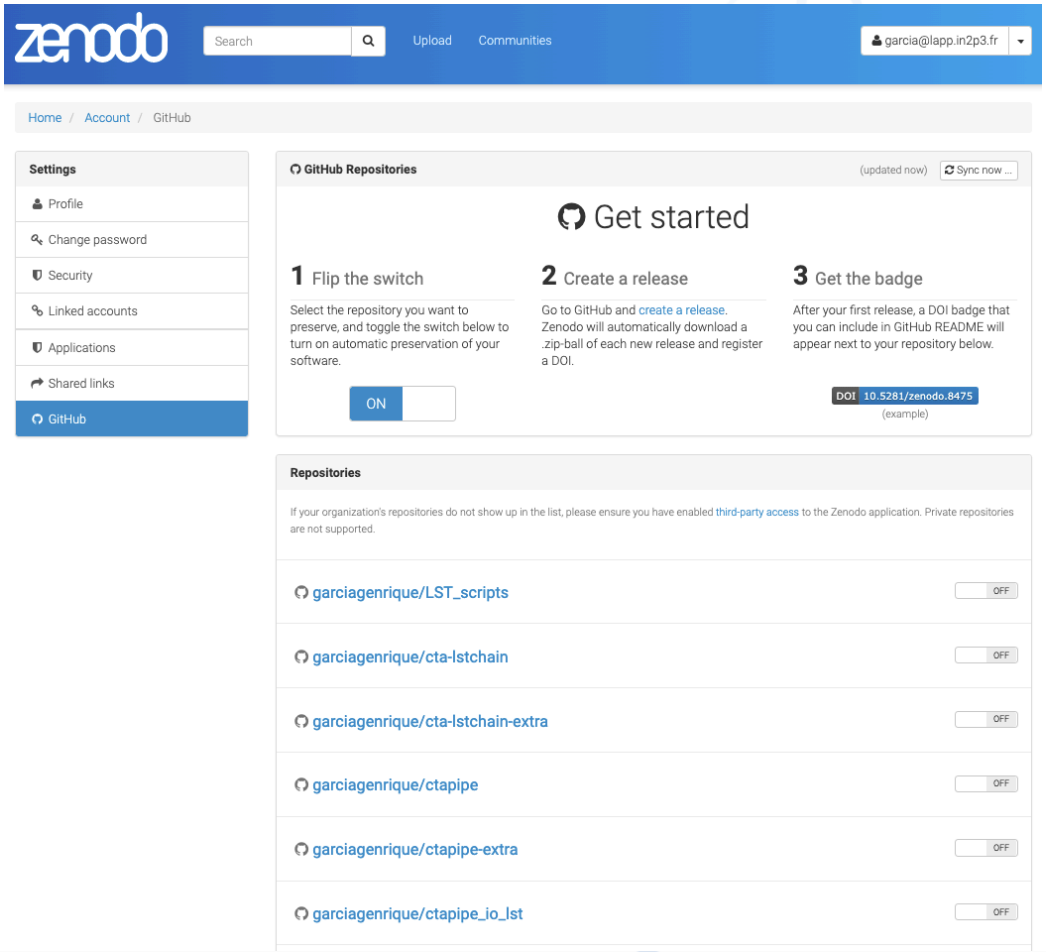
1. Clone/Fork the template repository
   https://gitlab.in2p3.fr/escape2020/escape/template_project_escape
2. Create an access token to communicate with the Zenodo API
3. Configure the .gitlab-ci.yml file
4. Configure of the GitLab CI/CD environment variables
5. Adapt/update the uploading/new version scripts
6. Create a release

Follow the **OSSR on-boarding demonstrator** at 13h15 (27/07/2020)

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

Automatic connection GitHub-Zenodo

- Zenodo → Account → Settings



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

Connection of services. Metadata (I)

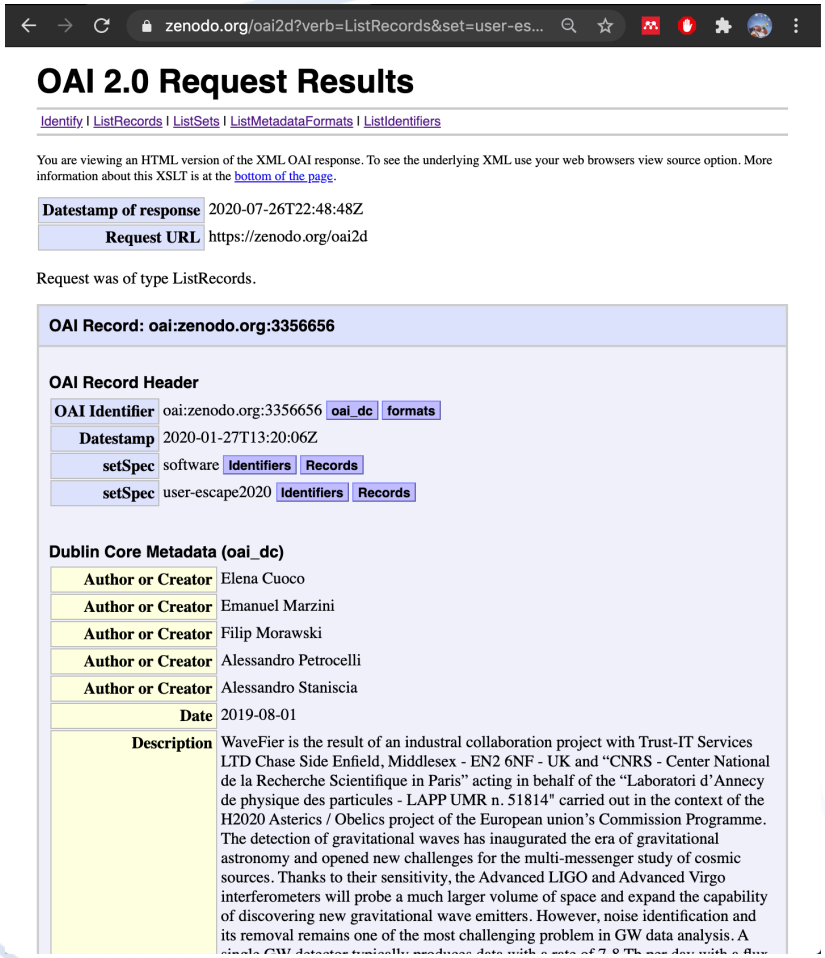
- Allows fulfilling all the FAIR principles in a single standard
  - Findable
  - Accessible
  - Interoperable
    - **Will allow the connection with other services ! (ESAP, Portal)**
  - Reusable :
    - License and provenance

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

Connection of services. Metadata (II)

Zenodo:

- harvest metadata using the OAI-PMH
  - Open Archive Initiative – Protocol of Metadata Harvesting
- accepts and export different metadata formats:
  - oai_datacite(3, 4,…)
  - marcxml
  - oai_dc
  - dcat
  - marc21

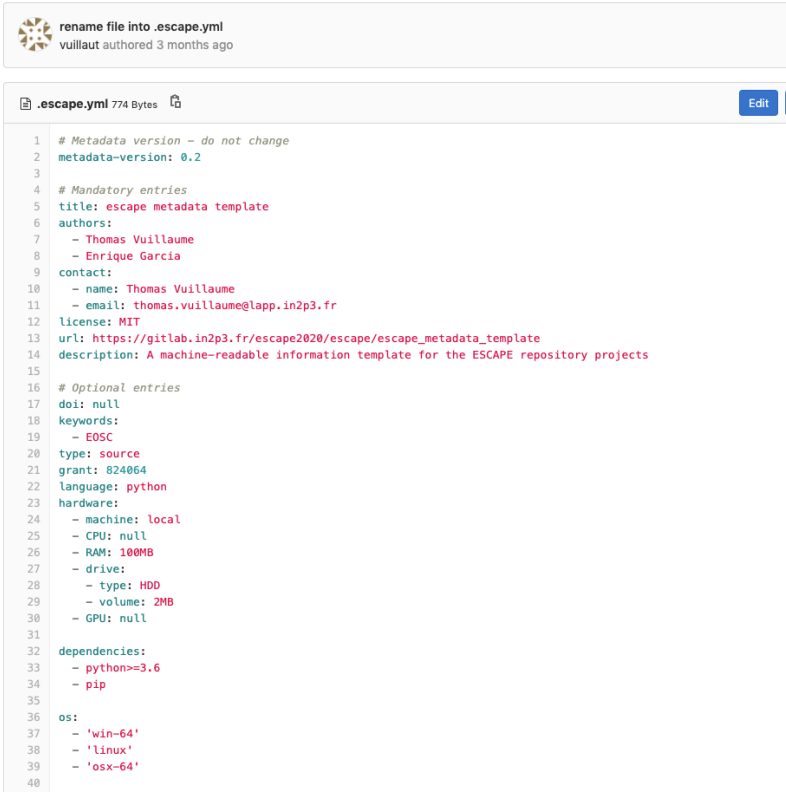


Funded by the European Union's Horizon 2020 - Grant N° 824064

Connection of services. Metadata (III)

- Started discussion (Mark Kettenis' talk – 24/07)
  - Metadata + I/o info + computing resources + media types
- ESCAPE metadata template **basic proposal**

  https://gitlab.in2p3.fr/escape2020/escape/escape_metadata_template
  - Gathers all the mentioned info
- Need to converge to a
  - Personalised schema ?
  - DataCite/CodeMeta – based ?



Repository License

- FAIR principles might lead to confusing situations
  - BSD 3-Clause → Singularity-CI builders
  - GPLv3 → zenodo-python repository
- Both licenses are compatible*, however GPLv3 is copyleft
- *It can be combined code released under the other license with code released under the GNU GPL

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

OSSR ToDo list

- Improve OSSR onboarding / uploads
  - More exhaustive documentation
- Implement ESCAPE ecosystem-like metadata schema
- Aim for an integrated virtual environment

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