

IRB, 26th of April 2023

Innovation programme

On-going H2020 & HE projects:

-BICIKL (Jose Benito Gonzalez)

- Finished the record batch uploader tool for Zenodo and tested the tool with a dataset from Biodiversity Literature Repository (200+ records). Used for importing datasets provided to BLR by third-parties.

-CS3MESH4EOSC (Jakub Moscicki)

- CS3 Conference in Barcelona (6 - 8 Mar) - the project was represented through a [presentation](#) and [demo session](#). Recordings are available.
- The Project had its third in-person meeting in Barcelona, on 9-10 Mar. The month of June was confirmed as the target date for Production. A plan for the preservation of the results of the Project was put in motion.
- Nextcloud and ownCloud have now merged the changes which allow for the ScienceMesh plugins to work, into their stable release channels. This means that partners are now able to move to production with their deployments.
- Discussion is now ongoing on the first subset of partner sites which will be moving to production in June and on the timeline for those which will follow.
- We were very saddened to hear that External Advisory Board member Andrew Cormack (Jisc) passed away. The StC acknowledged Andrew's contribution to the Project at its last meeting. The Project Coordinator sent a message of condolences to Jisc.

-EGI-ACE (Romain Wartel, Bob Jones, Xavier Espinal)

- There will one full day dedicated to the EGI-ACE project outcomes at the upcoming EGI 2023 conference in Poznań, Poland, from June 19th until June 23rd, 2023.
- Deviations
 - The IT personnel contributing to the cyber security activities for the remainder of the project are subject to confirmation by the Chief Security Officer.

-EOSC-Future (Xavier Espinal, Romain Wartel, Bob Jones)

March 2023:

- The Virtual Research Environment is being developed via Terraform and Flux and hosted on a public [Github](#) repo. A new [documentation](#) is also under construction.
- The work was presented in the [IT Technical Forum at CERN](#) and sparked discussion about connecting and coordinating the work done between the VRE team, the Swan team and the CS3MESH4EOSC team.

- The VRE Data Lake is a new fully working Rucio instance, with storage elements at INFN-CNAF, DESY, IN2P3-CC, CESNET and CERN accepting both authentication with token and x509.
- Continuous testing of transfers between RSEs is in place and can be seen from the [Gafana monitoring dashboard](#).
- The [VRE Reana](#) cluster is being used by an ATLAS workflow and will soon be accessible with users with an IAM account.
- The Jupyterhub notebook interface is developed, the Rucio extension on it is working apart from some caveats that need to be solved for the token implementation.
- The implementation of a Dask server on the nodes of the cluster to provide an alternative to run and distribute workflows is under construction.
- The proceedings paper following the [ACAT](#) conference was successfully submitted for publication.
- The VRE work was [presented in Aachen](#).

April 2023:

- The outputs from the project are being used as input by the EOSC Association to define the access policies to the services to be made available to researchers for the on-going 35M euro procurement organised by DG CNECT.
- The General Assembly 2023 will be held on 3 May, 14.00-17.00 CEST in Seville.
- Deviations
 - The IT personnel contributing to the cyber security activities for the remainder of the project are subject to confirmation by the Chief Security Officer.

-FAIRCORE4EOSC (Lars Holm Nielsen)

- Implemented the first part of the technical specification, namely: software-specific metadata fields, import of software-specific vocabularies, software fields in the deposit form and record landing page, export formats for software.
- We were not able to attend the project meeting due illness.

- interTwin (Maria Girone)

March 2023:

- T4.2: Finalizing our use case's technical requirements. Discussions between partners from WP7 around the structure of the deliverables concerning the reporting of requirements consolidation. Activities have been conducted in terms of preparing input for the different sections of D7.2: report on requirements and thematic modules definition for the physics domain.
- T7.7: Studying existing tools for hyperparameter optimization. Work on exploring alternatives to GAN models, more specifically a Variational AutoEncoder. Works continued on GAN-based model integration with Monte Carlo framework (Geant4).
- T6.5 has made progress in its requirement analysis, resulting in an updated version of the requirements table. Discussions with T6.1 and T6.2 have helped to better define their preliminary design, allowing T6.5 to draft some viable interfaces. The proposed solution is for T6.5 to be deployed as a standalone container and interact with T6.1 via Kubernetes-like

APIs. T6.2 is able to test/validate AI/ML workflows as a microservice. The draft architecture for T6.5 will be included in the deliverable T6.1, along with the requirements analysis from the use cases.

April 2023:

- T6.5: Contribute to the D6.1 deliverable, concerning the analysis of use cases' requirements and architecture design. Discussion with T6.1 resulted in the agreement on using Common Workflow Language (CWL) as the try point of digital twin (DT) workflows. This allows us to define a draft interface with the workflow manager. Furthermore, planned internal delivery of a minimal working prototype for T6.5, by the end of the month.
- T4.2: Discussions between partners involved in WP7 around the structure and content of the deliverables concerning the reporting of the thematic modules' requirements. Organizing and writing use case contribution to the D7.2: report on requirements and thematic modules definition for the physics domain.
- T7.7: Hyperparameter optimization of a Variational AutoEncoder in terms of network size, loss function terms contribution to the total loss, learning rate, integrating alternative pre-processing steps.

-OpenAIRE-NEXUS (Jose Benito Gonzalez)

- Completed last updates for the Zenodo pitch slides and Zenodo book to close up the task.

-RAISE (M. Girone)

March – mid April 2023

- The work on MLPF continues with studies on CLIC-based datasets, with new versions of the datasets being developed and generated in rapid iterations. Hyperparameter Optimization (HPO) has been performed for both the GNN- and the Transformer-based MLPF models and a performance comparison based on the results from this HPO are underway.
- A TCB meeting was held on the 15th of March, discussing data and I/O challenges related to Exascale. · The CERN openlab Technical Workshop was held March 16-17, where two presentations based on work done in CoE RAISE were given, one on enabling AI on HPC and one on HPC benchmarking.
- A seminar on HPO of Deep Learning models using HPC was given at the CERN EP-IT Data Science Seminar on April 5th. The seminar included an introduction to HPO as well as results from work done in T4.1 of CoE RAISE during the last couple of years.

- OpenWebSearch (Andreas)

- WP5: Federated data infrastructure:
 - iRODS instance up and running, opened in firewall. Data federation established with IT4I and DLR.
 - Setting up MinIO with iRODS and experimenting with exposing the MinIO dashboard. Currently testing MinIO with S3 volumes in our OWS instances.

- Prepared and published Milestone report M11: "Requirements for data structure needs & co-design V1"
- WP6 - Open Web Search Ecosystem and Sustainability:
 - We participated in a meeting where we discussed different funding models.
- OpenWebSearch.EU Consortium Meeting, 17-19 April 2023, Graz, Austria.
 - OpenWebSearch.EU Partner Meeting 17-19 April 2023, in Graz, Austria.
 - Presented our past/current activities in WP 5
 - ELSA Workshop (Ethical, Legal and Social Aspects) 19-20 April 2023:
 - Workshop to discuss Legal and Ethical Aspects of creating an Open Web Index (OWI)

- **SYCLOPS** (Maria Girone)

March and April 2023:

- First press release of the project was prepared.
- IT fellow joined the project after the easter break, ramping up now.

- **Proposals for upcoming Horizon Europe funding calls**

Projects accepted for funding:

- HORIZON-ZEN - A two years project coordinated by CERN (single beneficiary) managed by

Lars Holm, with 429.536 EUR allocated to CERN, was selected and invited to grant preparation. Evaluation score 13.5. Project likely to start June 1st. Kick-off meeting with EC participation tentatively set for June 12th-13th.

Submitted proposals on the 9th of March under the Research Infrastructure Program:

- AARC3 – Hannah Short (IT-PW-IAM): 174.460 EUR, Lump sum grant, and 24 months duration
- SPECTRUM – Maria Girone (IT-GOV): 503.196 EUR, Lump sum grant, and 30 months duration
- OSCARS – Xavier Espinal (IT-GOV): 351.090 EUR, 48 months duration
- EVERSE – Stefan Roiser and Jose Benito (IT-GOV & IT-CA): 556.559 EUR, CERN-EP is also involved with 351.489 EUR and 36 months duration
- EOSC Beyond – Luca Mascetti (IT-SD-PDS): 318.487 EUR, Lump sum grant, and 36 months duration