EOSC-EVERSE

Paving the way towards a European Virtual Institute for Research Software Excellence

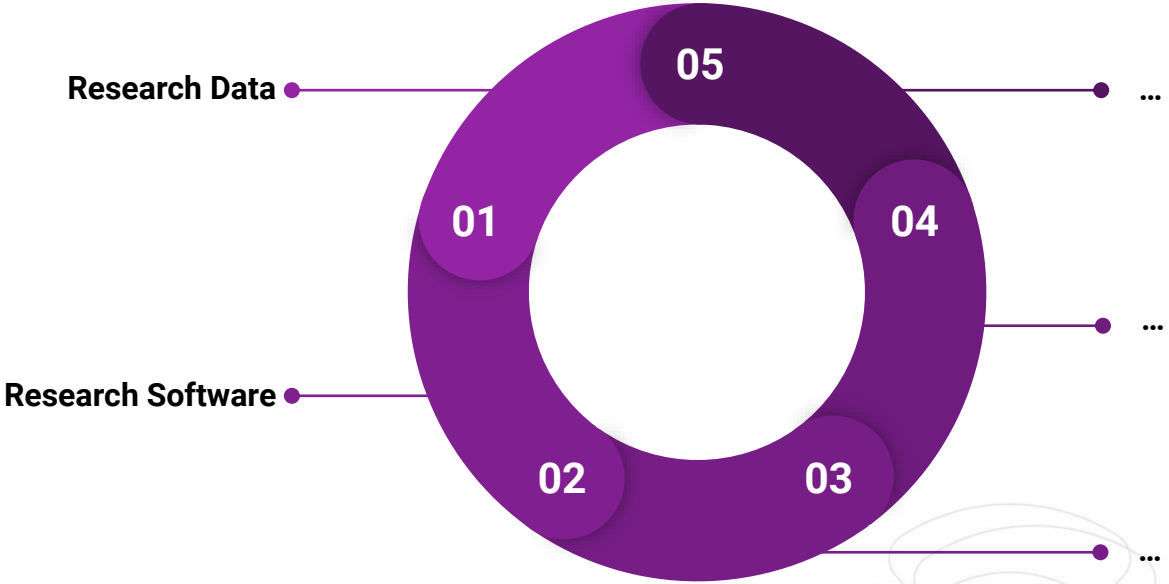
Funded by the European Union

This project has received funding from the European Union's Horizon Europe Programme under GA 101129744 — EVERSE — HORIZON-INFRA-2023-EOSC-01-02

Research Software: The backbone of research

- Software is ubiquitous in Research.
- Communities have created numerous software applications that are essential to progress in their fields.
- The reliability of these applications, how effectively they can be reused and their long-term sustainability, are critical aspects for future progress.
- For example:
  - software used in healthcare must be trustworthy to ensure the safety and well-being of patients.
  - reliable software ensures accurate weather forecasts that we all rely on for planning.
- This necessary trust in research/community software longevity requires a transparent display of good engineering and clear organisational processes that enable continuity.
- Ensure research software curation, quality, preservation and adoption of best practices tailored to developers at all levels.

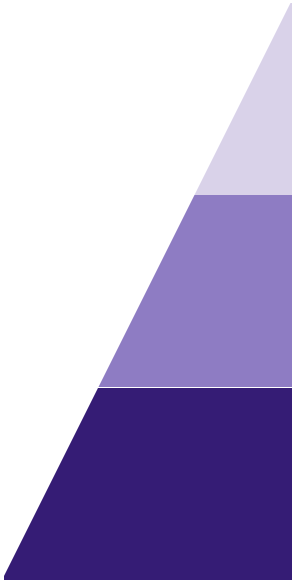
Making Research Software a first class citizen for the scientific endeavours



In a given programming language design, a first-class citizen[a] is an entity which supports all the operations generally available to other entities. These operations typically include being passed as an argument, returned from a function, and assigned to a variable.[1]

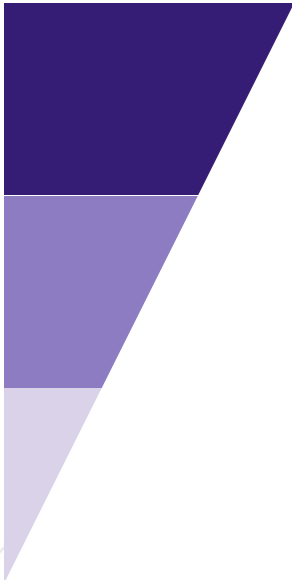
Wikipedia (2023/11/12). https://en.wikipedia.org/wiki/First-class_citizen

Not all software has the same level of importance …



abundance

1. **Research software infrastructure**
   It involves research software that captures more broadly accepted and used ideas, methods and models for use in research, and warrants close researcher involvement in their development.
2. **Prototype tools**
   It refers to research software that demonstrates a new idea, method or model for use by others outside the project within which it originated, often as a substantive intellectual contribution in its own right and often in the form of a proof of concept.
3. **Analysis code**
   It includes research software that captures computational research processes and methodology, and often occurs in the context of simulation, data generation, preparation, analysis and visualisation.



Importance

Foundational Software

REFERENCE TO THE 3-TIERS RESEARCH SOFTWARE MODEL

… but it is ubiquitous across data-intensive scientific domains

ESFRI cluster projects - Position papers on expectations and planned contributions to the EOSC
https://zenodo.org/record/367508

EVERSE

Paving the way towards a **E**uropean **V**irtual Institut**e** for **R**esearch **S**oftware **E**xcellence

EVERSE aims to create a framework for research software and code excellence, collaboratively designed and championed by the research communities, in pursuit of building a European network of Research Software Quality and setting the foundations of a future Virtual Institute for Research Software Excellence

- ✓ ensure research software curation, quality, preservation and adoption of best practices, by the Communities, for the Communities, build on collaboration with the five EOSC Science Clusters
- ✓ adopt a three-tier model for research software, i.e., analysis code, prototype tools and research software infrastructure, which captures the varying complexity of research software and its development, and can be used as a basis for research software excellence
- ✓ credit and recognition for both developers and software are essential components of our strategy to promote sustainable software practices
- Start: 1 March 2024 (36 months)

Partners, associates, and affiliated entities

Consortium:
15 full beneficiaries, 2 associated partners & 1 affiliated entities across 10 countries

Objectives (1/2)

Objective #1

Build a collaborative, community-led structure for **evaluating**, **verifying**, and **improving the quality** of **research software** and code, by actively involving researchers, software developers, and other stakeholders in the research community.

Objective #2

Leverage existing tools and resources to support the evaluation, verification and improvement of research software and code quality, based on **existing practices and standards** across **research communities** represented by the five EOSC Science Clusters.

Objectives (2/2)

Objective #3

Establish a **sustainable** and **collaborative ecosystem of stakeholders** across the research communities associated with the five EOSC Science Clusters to ensure research software and code quality assurance and support the advancement of **reliable and reproducible research**.

Objective #4

Provide a **framework** that will ensure appropriate **recognition**, **reward**, and **career development** for researchers and **RSEs** who implement research software and code quality assurance practices and policies

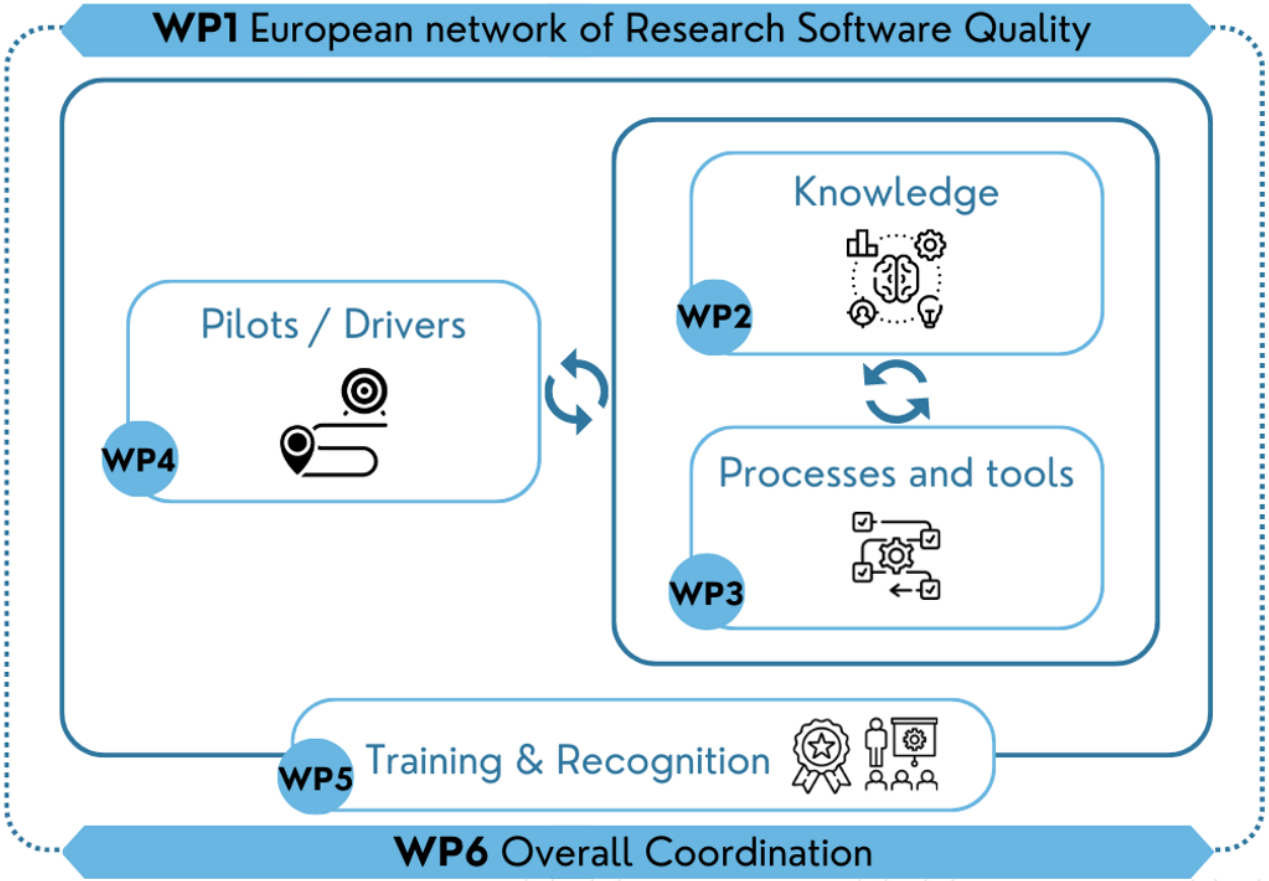
Expected Outcomes and Impact (1/3)

- A framework of **community curation** is established and promoted that ensures **quality of software and code** across the **different disciplines**.
- **Infrastructure, tools and services** are deployed that allow researchers to properly develop, describe with proper metadata, version, archive, share and reuse research software.
- The notion of **software quality** is defined in the context of EOSC and builds upon established practices by the FAIR and other communities.

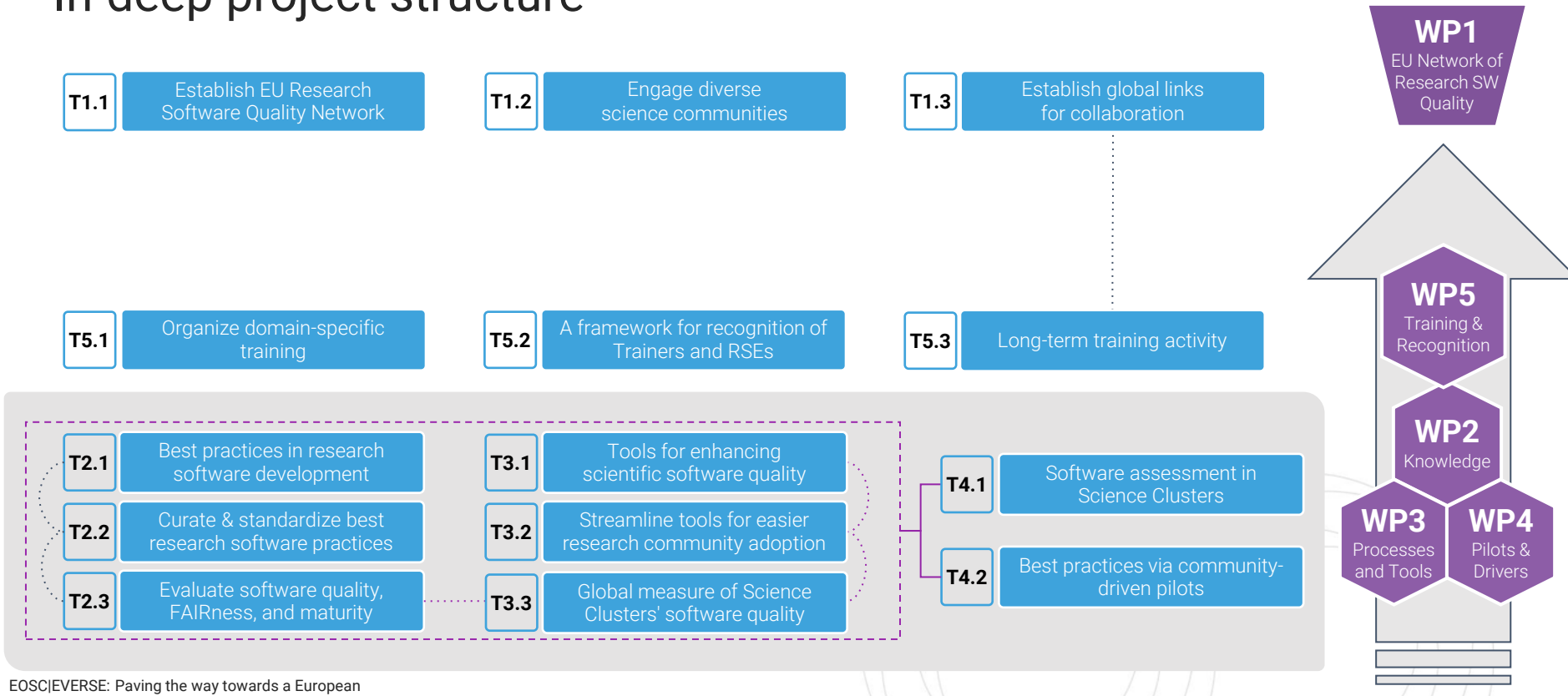
Expected Outcomes and Impact (2/2)

- Baseline **quality indicators** of “minimum quality” defined for the different types of digital objects targeted (software, code, etc), taking into account the concept of “**fit for purpose**”.
- The quality of research software (technical and organisational) improved, in general (e.g. software for data analysis) and in particular for software used in the services offered through EOSC.
- Software is developed in a **sustainable way and its reuse is maximized**.

Project Overview



In-deep project structure



Pilots & Drivers

Environmental Sciences: Integration of Science Cluster ENVRI through ENVRI-HUB

- Integrate EVERSE framework into the ENVRI-HUB Knowledgebase and Virtual Research Environment
- Apply to the development of the Essential Climate Variable computing program and cloud workflows

Life Sciences: Integration of Science Cluster EOSC-Life through ELIXIR

- Make RO-Crate actionable by incorporating the five safes concept into WfExS for secure and federated workflow orchestration
- Use of community-led standards for materialising research software packaged using container technologies and mobilising encrypted data whenever needed

Astronomy and particle physics: Integration of Science Cluster ESCAPE through the Dark Matter Test Science Project

- ML for scientific data compression (standalone code, python)
- A Common Tracking Software
- Choose an ATLAS trigger algorithm as an option for the collaboration

Proton and neutron science: Integration of Science Cluster PaNOSC through LEAPS/LENS
Transition software to high performance computing (HPC) and heterogeneous computing architectures

Social sciences: Integration of Science Cluster SSHOC
Develop a multilanguage textual analysis pipeline of tools that use a combination of open source tools and own code to create an integrated SotA tool capable of deploying locally or as a service

Project Ambition

Collaborative effort to produce and improve high-quality research software, code and other digital artifacts

- Knowledge Hub (RSQkit) containing:
  - Curated best practices.
  - Tools and processes to support the best practices and to assess software and code quality.
  - Guidelines and policies.
  - Catalogue of training materials and courses.
- Additional elements contributing to enhance quality.
- Examples for increasing researchers and RSEs recognition.
- Focus capacity building action through training to increase reach.

European Virtual Institute for Research Software Excellence

Thank you!

contact: everse-contact@lists.certh.gr

+info: everse.software/

Funded by
the European Union

This project has received funding from the European Union's Horizon Europe Programme
under GA 101129744 — EVERSE — HORIZON-INFRA-2023-EOSC-01-02