eosc | RAISE

Research Analysis Identifier System

At a Glance

Title: Research Analysis Identifier System
Call identifier: HORIZON-INFRA-2021-EOSC-01
Topic: HORIZON-INFRA-2021-EOSC-01-04
Project number: 101058479
Total budget: € 4,836,125.00
Duration: 40 months
Start Date: 1 October 2022
Project Coordinator: Lab of Medical Physics & Digital Innovation, Aristotle University of Thessaloniki
Consortium: AUTH, VICOM, INTRA, UOWM, UOM, TECREANDO, WITA, CERTH, INNOV-ACTS, VILABS, CYCLOPT, KI, ATHENA, AINIGMA,UHASSELT, OPENAIRE, EUCENTRE, ENVE.X, SOTON, McGill.

Enabling Data Analysis without Data Sharing

- raise.science
- raise.science
- @RaiseScience
- info@raise-science.eu
- zenodo RAISE EOSC project

www.raise-science.eu

About RAISE project

The real value of **open data** for the **research community** is not to access them but to **process** them as conveniently as possible in order to **reduce time-to-result** and **increase productivity.** To that end, the **RAISE project** will provide the infrastructure for a **distributed crowdsourced data processing system**, moving from open data **to open access data** for **processing.** It will render the **mechanism** for sending the algorithm to the dataset instead of sending the data to the algorithm.

RAISE introduces the **Research Analysis Identifier** *(RAI)* and **services** that will support trustworthy research and data analysis services based on crowdsourced distributed resources along with local resources. Particularly, RAISE will produce the following **Outputs:**
1. RAI Certified nodes: A trustworthy crowdsourced network offering data storing and processing resources (community resources). The network will also provide the mechanisms for ensuring originality and authentication of data assets and research results.
2. RAI Cloud: Orchestrates the data sharing, processing and finding.

3. The Research Analysis Identifier (RAI): a unique identifier of any result along with the dataset information and the processing script (RAI Registration Service), without disclosing any source code or raw data.
4. Dataset plagiarism identification and dataset proof-of-origin services, to maximize the level of trust of the RAISE system.
5. RAI Synthetic Data Generator and the RAI SDK, to help researchers develop processing scripts by getting access to sample, small scale datasets which mock real data.

Project Objectives

- Enhance the proliferation and in-field demonstration of the **FAIR principles** of scientific research, providing innovative services to researchers and the wider research community.
- Develop a scalable distributed, crowdsourced network of processing nodes for **trustworthy sharing** and **on-demand provisioning of open science.**
- Enable easy and effective systematic **publishing** and **re-using of scientific results,** as well as ensure the accreditation of dataset collection and curation work in publications.
- Increase the available services for **European researchers** and offer new capabilities for data sharing and analysis in the **EOSC portal.**
- Engage end-users in the designing and real-life testing of the **RAI solution** in multiple science domains and disseminate the RAI system.

What's in it for the RAISE Community?

Researchers *(data providers & data processors)*	RAISE facilitates the data availability for researchers, by adopting a reproducible workflow that publishes the code and data in a form that allows other researchers to extend the approach to new applications with a minimum of effort.
Publishers & Open Data Repositories	RAISE offers a solution that enables the research community to publish their outcomes with evidence-based authenticity of the data-analysis.
EOSC Community	RAISE provides the effective tools that comply with the FAIR Guiding Principles for scientific data management and stewardship, in order to enhance the Open Science ecosystem and EOSC with services that enable FAIRness of data.
Industry *(companies with algorithms)* **& Potential Investors**	RAISE project develops innovative services that provide the mechanisms and long-term resources for seamless access, discovery, analysis and on-demand provisioning of open science data and tools and will comply with the EOSC Rules of Participation to be integrated with and offered through the EOSC.
General Public	RAISE offers a solution that enables the research community to publish their outcomes with evidence-based authenticity of the data analysis performed.

Application Domains

Within the broad domain of **Open Data Science**, RAISE will test and assess the effectiveness of **RAISE system** in real life use cases in **3 scientific domains** that are producing large amounts of data and facing major challenges on **data sharing**, and **sensitive data handling.**

Health:
Analysis of a variety of datasets ranging from clinical data and medical images, EEG signals and data collected from sensors for different conditions *(Heart failure, stroke)* but, also, different phases of care *(hospitalization, transitions, home care).*

Mobility:
Focus on mobility patterns and driving behavior. Investigation of the exploitation and combination of open datasets with non-open datasets that contain sensitive data.

Environment:
Analysis of environmental exposure from data collected from personal sensors. Investigation of how RAISE will tackle the computing power required as well as careful sharing of data sources while also maintaining privacy protections for personal data.

Cross-disciplinary:
The initial results of the real-life pilots will be combined in a multidisciplinary research study for the city of Thessaloniki in order to also investigate the reuse of code and results in the creation of new joint studies and publications.

Consortium