Ref. Ares(2024)2353805 - 29/03/2024

Geosphere Infrastructures for Questions into Integrated Research

D5.3 WP5 Status of research infrastructures at M18

Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call

Recommended citation:

Christadler, I., Gabriel, A., Lorito, S., Romano, F., Babeyko, A. Y., Rodríguez Muñoz, L., Esposti Ongaro, T., Lanucara, P., Kalligeris, N., Scala, A., Murphy, S. (2024): Geo-INQUIRE Horizon Europe project. D5.3 WP5 Status of Research Infrastructures at M18, Potsdam: GFZ Helmholtz Centre for Geosciences, 47 p., Report

DOI: https://doi.org/10.48440/ZC54-K388

Reference information

Deliverable No.	**D5.3**
Deliverable Name	WP5 Status of research infrastructures at M18
Author(s) (Affiliation)	Iris Christadler (LMU), Alice Gabriel (LMU), Stefano Lorito (INGV), Fabrizio Romano (INGV), Andrey Babeyko (GFZ), Lucia Rodriguez Munoz (CINECA), Tomaso Esposti Ongaro (INGV), Piero Lanucara (CINECA), Nikos Kalligeris (NOA), Antonio Scala (UNINA), Shane Murphy (IFREMER), with contributions from: Laura Sandri (INGV), Chiara Montagna (INGV), Noemi Romano (ETHZ), Hannes Vasyura-Bathke (GFZ), José Manuel Gonzáles (UMA), Finn Lovholt (NGI), Steven Gibbons (NGI), Otilio Rojas (BSC), Francesco Martinelli (UNINA)
Lead Beneficiary	LMU
Due Date (month)	M18
Type	R – Document, report
Dissemination level	PU – Public
Reviewer(s)	Mateus Litwin Prestes (GFZ)

Versions

Version No	Date	Changes / Comments / Status
1.0	28/03/2024	Final version submitted to the EU
0.5	20/03/2024	Internal Review
0.3	11/03/2024	Updates from installations and task-leaders

Table of contents

Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call

1 Introduction

This WP5 deliverable highlights the advancements being made between M6 (March 2023) and today (March 2024). It is a follow-up deliverable on D5.1 “Status of research infrastructures at M6”. It is the first deliverable to report general KPIs and important information to WP6 for the unification of VA-KPIs across WPs as defined in D5.2 “Research infrastructures coordination progress report with definition of KPIs” from M12. Results on all KPIs will be reported in a unified way in the M24 deliverable D5.4 “Research infrastructures coordination progress report with access and KPI monitoring”

During the last twelve months of the project the following Milestones have been met:

- MS27 “Standards for SDL addressed” (M12)
- MS28 “TA TEB formed and draft call prepared” (M6)
- MS29 “KPIs identified for access monitoring” (M12)
- MS30 “Level-3 Gauss-Toolbox included in the portfolio” (M12)
- MS31 “Second TA draft call” (M18)

The 1st Transnational Access (TA) call has just been closed on 15 March 2024. The evaluation process started immediately to shorten any delays. Following some interactions of potential applicants with the Geo-INQUIRE PMO and TA providers, aimed to the clarification of formal call aspects, which included a public Q&A session, 19 applications were submitted for TA to 7 WP5 installations. Most of the applications include requests for computational resources provided by the two High-Performance Computing (HPC) installations. The 2nd TA call is about to open mid-April, to be advertised at EGU in Vienna.

The draft white paper for the Simulation Data Lake (SDL) has been started with D5.2. A final version, including recent developments both regarding the new SDL platform and the effort for standardization of simulation data and metadata is under preparation and is expected to be submitted for publication in May/June 2024. CINECA is currently coordinating the work on finalizing the data formats and the metadata scheme with input from all SDL partners, and in collaboration with the DT-GEO project regarding the workflow metadata schema harmonization. While b2share was available as interim solution the first S3-based implementation of the final SDL by CINECA will be available within the next two months.

A first set of KPIs was defined in D5.2, with some “WP5-special” KPIs (especially WP5-VA-KPIs). After some meetings with WP6 to understand the needs of both WPs, it has been decided that WP5 should concentrate on the core set of KPIs. It is however possible that WP5 will maintain some specific KPIs, as the potential for their future implementation was declared by VA providers (Table 1 in D5.2). WP5 had provided also one specific TA KPI metric regarding the core-hours consumed by the TAs. This deliverable provides updated information on how the different service providers are planning to measure the (VA-)KPIs. This will be used as input for WP6, so that final numbers could be reported in D5.4.

The only VA service that was not foreseen to be available right from the start ("TS-Gauss") has been put into operation. More details are provided in Chapter 2.3 "SaaS" and in the template for TS-Gauss in Appendix 1.

This deliverable reports the progress on inter-WP topics: The Implementation Level Matrix (ILM) and Key Performance Indicators (KPIs) (with WP6), FAIRness assessment (with WP7) and documentation, trainings and workshops (with WP9). A full description of all trainings efforts in the first half of the project will be reported in D5.5 "Training summary report at M24" (M24).

The deliverable is structured as follows: *Chapter 2* contains updates from each task of WP5: Coordination activities in T5.1 with a focus on the TA Call preparation and a homogenization task for Tsunami workflows, advances on the Simulation Data Lake (SDL) in T5.2, documentation and training together with KPI assessment in T5.3, an overview of the T5.4 workflows that participated in the TA call, and porting efforts to the HPC systems in T5.5. *Chapter 3* highlights the cross-WP activities with WP6 "Access and harmonization of integrated research infrastructures services for next generation data, services, and products", WP7 "FAIR data management, legal compliance, and impact assessment" and WP9 "Fostering cross-disciplinary research of the Earth through training and workshops". Even if some of the activities described in Chapter 3 pertain to T5.1 because they are coordination activities, they are reported in a separate chapter for the sake of consistency with the corresponding deliverables at M18 of the other Geo-INQUIRE WPs. *Chapter 4* gives a conclusion. Templates with short reports and information from all services are given in Appendix 1. Appendix 2 contains a copy of the KPIs as defined in D5.2 to make the document self-contained.

2 Status of research infrastructure at M18

2.1 Update of Task 5.1 "Access Coordination"

2.1.1 TA Call preparation coordination

The major topic for at least the last six months was the preparation for the 1st Geo-INQUIRE TA Call. Initially, there was a prolonged interaction between the WP and Task leaders with the TA providers, to clarify the details of the offer, regarding the overall budget and budget rules, the definition of the maxima regarding the physical stay number and duration, and regarding both the in-kind and budgeted core-hours to run the numerical simulations.

The WP5 leader then collaborated with the responsible person appointed by WP4, with the WP8 leader, and with the PMO to set up the call rules, to prepare the call documents and forms, to set up the evaluation rules and the evaluation forms, to prepare a homogeneous description for the Geo-INQUIRE web site of the TA offer for the different TA installations. Some of these activities, for example those regarding the installation of the numerical codes on the supercomputers, are reported within the section dedicated to the specific tasks.

Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call

The last steps in preparation for the TAs were interactions regarding technical and formal aspects between potential applicants, the PMO and the TA providers, with the help of the WP leaders. One important milestone of this interaction was the Q&A public event organized on 1 March 2024 which was attended by a significant number of potential applicants.

After the deadline, the WP5 installations have received 19 applications (including five from non-EU countries) for seven different WP5 installations, plus the two HPC installations One installation was not open for applications in this call. Two installations did not receive applications, but their "mirror" applications did (PTHA and CyberShake). The evaluation of the applications is just starting. The numbers just presented are quite satisfying considering that this was the first time that an offer like this in WP5 was proposed. We will make sure that the experience gained in the first call will help us to improve the 2nd call and there will be the possibility to advertise the 2nd call at the EGU General Assembly next month in Vienna.

2.1.2 Interoperability of VA Services

WP5 offers VA to a significant number of tools for modelling geophysical phenomena (earthquakes, tsunamis, volcanoes). Some of these VAs are also distributed by some Thematic Core Services in EPOS. However, they are mostly provided independently from each other. One of the purposes of Geo-INQUIRE is to improve the interoperability of its installations. We have started this process by considering the interoperability between earthquake and tsunami codes in WP5, as the former can be used as an input to the latter, and eventually with the sea level data measurements provided in WP2.

Rationale

Numerical modeling of earthquake-generated tsunamis and validation of outputs through field observations requires a series of steps by the user. Depending on the application, this process may involve multiple models/software among those offered as VA in WP5, each of which has specific input and output (I/O) formats. It is of great benefit to the tsunami community for the models used in the numerical modeling chain to be interoperable and thus to minimize pre/post-processing before and after the application of each model. This is particularly relevant to operational/automated tsunami simulations run in HPC environments, or even for hazard forecasting, for which minimum intervention and inputs from users is of paramount importance to meet operational time constraints.

This networking activity under Task 5.1 involves the coordination between service providers within WP5 to agree on the I/O formats and ensure interoperability between models involved in the numerical modeling of earthquake-generated tsunamis.

Conceptual scheme

A conceptual scheme for integration of inputs for numerical modeling of earthquake-generated tsunamis is shown in Figure 1. This scheme was proposed to and discussed with the WP5 participants during several meetings.

The first step is the definition of the tsunami-generating earthquake scenario, given input parameters of hypocenter, focal mechanism and magnitude, which are all the necessary information to start a numerical tsunami simulation. These input parameters can be defined through a user-defined input file, or sourced for real events from a webservice such as FDSN. The next step depends on the application, and it can be either to:

I. Obtain stochastic slip distributions for planar/non-planar faults using models k223d/VA4-532-5 and ANTI-FASc/VA4-532-6 (step 2a); stochastic slip distributions are then passed as ASCII files to numerical models in steps 3a and 3b to compute the seafloor deformation;
II. Obtain the seafloor deformation using earthquake simulation tools such as Pyrocko/ VA4-534-1 or SeisSol/ VA4-531-1 (step 2b) and pass it on to numerical models in steps 3a and 3b;
III. Obtain the seafloor deformation directly from the numerical simulation models (Tsunami HySEA/VA4-532-1 and TS-Gauss/VA4-532-3-4 in steps 3a and 3b) by employing earthquake scaling relations and the Okada analytical equations.

An additional desirable step on which some Geo-INQUIRE partners are presently working is shown in Figure 2 for translating the seafloor deformation into sea-surface displacement, for which a filter for the water column response can be applied, the horizontal component of co-seismic displacement factored by the bathymetric gradient can be considered in the sea-surface displacement, and finally for considering the contribution to the sea floor displacement of the horizontal movement of an oceanic slope and of the anelastic deformation of the material surrounding the fault zone (e.g. the sedimentary wedge near the trench in a subduction zone). Among these, only a uniform sea-depth approximation of the so-called Kaijura filter is presently implemented in Tsunami-HySEA.

Step 4 in the numerical modeling chain involves accessing sea-level observations from IOC/UNESCO's Sea Level Station Monitoring Facility (SLSMF, VA1-25-1), hosted by VLIZ, through the REST service, while step 5 comprises the model/data comparison through a user visualization tool.

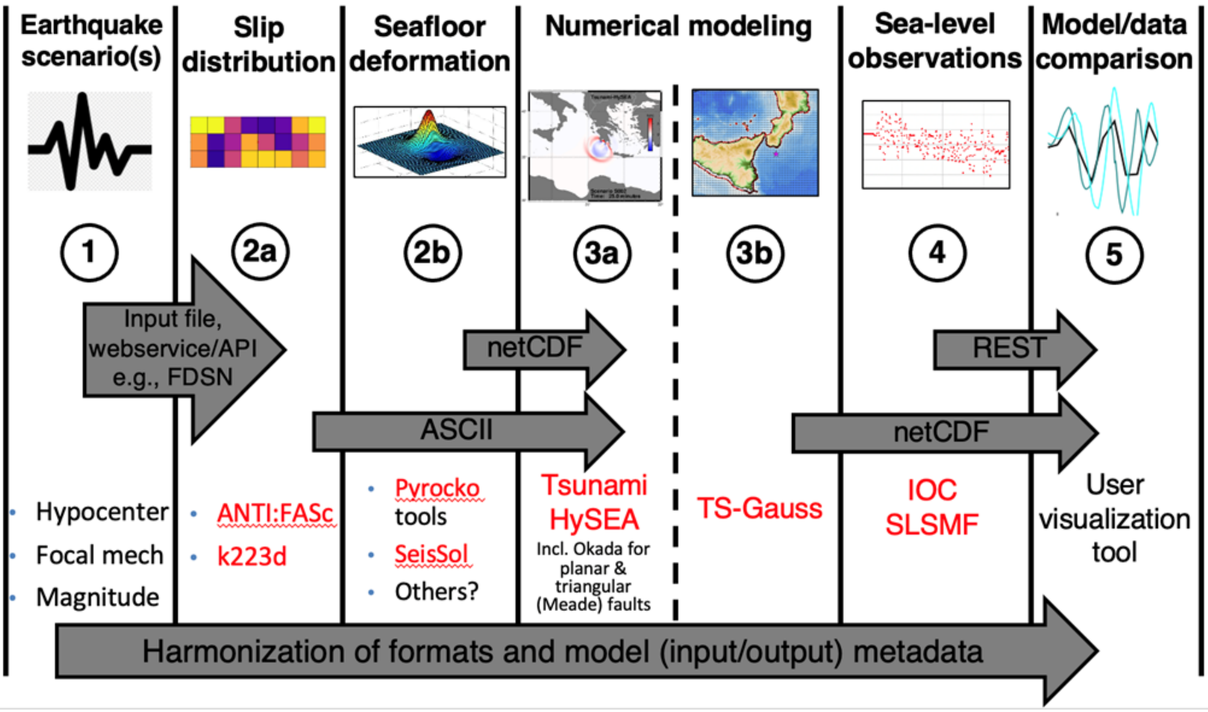

Figure 1: A conceptual scheme for the integration of inputs and validation of outputs in the numerical modeling chain for earthquake-generated tsunamis. The related models/services within WP5 are listed in red

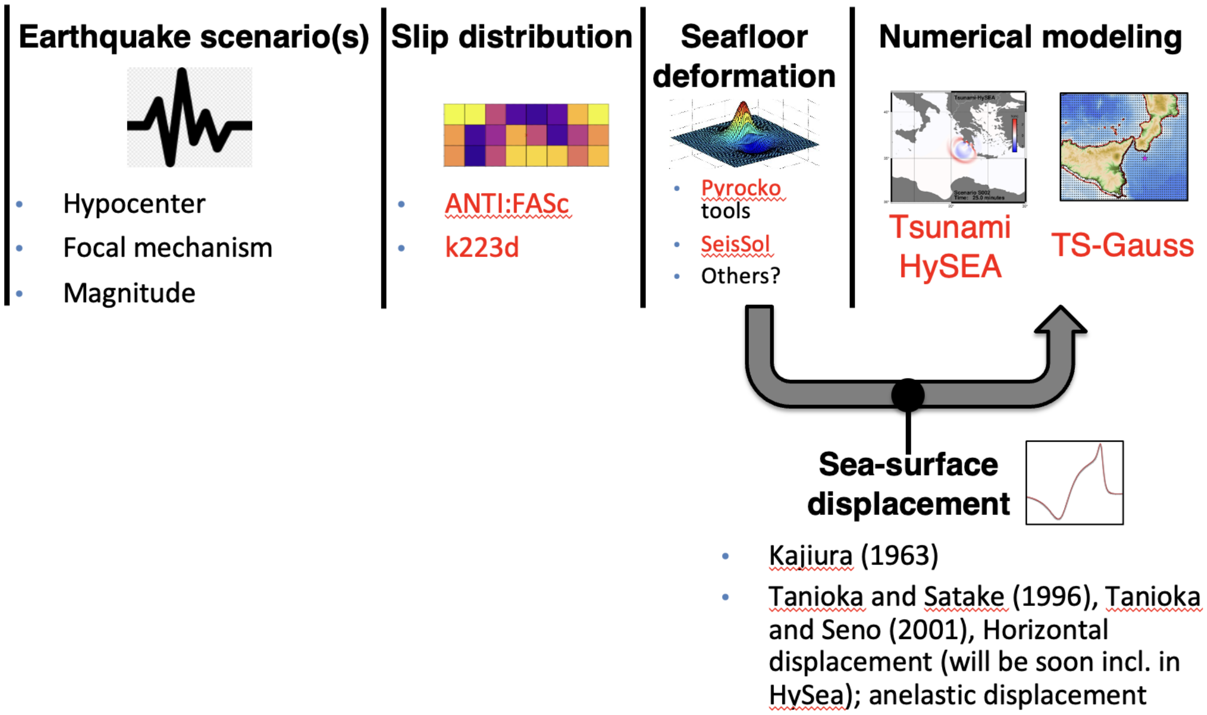

Figure 2: An additional step in the conceptual scheme shown in Figure 1 for translating the seafloor deformation to sea-surface displacement.

Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call

Landscape analysis

An initial landscape analysis was performed to identify the initial steps to improve the interoperability between models, starting with the definition of the I/O of each model/service involved in the numerical modeling chain for earthquake-generated tsunamis shown in Figure 1. Preliminary information on model I/O obtained by the service providers through a specific survey are outlined in the table below.

Table 1: Inputs and outputs of model/services involved in the numerical model chain

Inputs	***Outputs***
AntiFASc	
• The geometry of the slabs can be provided either in grd and/or xyz formats available at the webpage of the project Slab 2.0, or (if a mesh discretization is already available) through ASCII files containing the geographical coordinates of nodes and a table of the nodes being the vertices of each face; • Magnitude and hypocenter of the event are set in the general input json configuration file; • The magnitude binning and the earthquake scaling relations used are set in another configuration file in json format; • The focal mechanism: strike and dip angles are determined by the modelled geometry of the slab (or of the fault surface), while the possibility to allow users to set the rake is currently under construction.	The slip distributions are provided in a format compatible with the Input scheme of Tsunami HySEA, that is a file ASCII with the following 11 columns: LON1 LAT1 DEPTH1 LON2 LAT2 DEPTH2 LON3 LAT3 DEPTH3 RAKE SLIP , where 1, 2, 3 indicate the 3 nodes comprising each subfault of the mesh.
K223d	
Input data for the programme is an ascii file format linked with abaqus containing the mesh and an ascii input file. At present a Fortran programme provides an example of how to generate a simple planar mesh as an input. The input is the process of being updated to an ascii vtk format with jupyter notebook tutorials in preparation that demonstrate (a) how to generate a mesh from open source seismogenic fault databases and (b) choose a sub-section of the fault mesh that will be used as an input for k223d (based on a user choice of epicentre and magnitude).	Slip distribution is being updated from .inp (Abaqus format) to an ascii vtk format which is readable in Paraview. The file format is in cartesian coordinates.
Pyrocko-GF	
The focal mechanism (moment tensor) or fault geometry and slip (distribution) of an earthquake. Receiver locations, requested quantity to be simulated (displacement, velocity, acceleration), name of the Greens Function database. Everything in yaml/JSON format. Examples will be provided under https://pyrocko.org/examples	Simulated time-series of the requested quantity (displacement, velocity, acceleration) at requested receiver positions as generated by the provided source parameters. MSEED format.

Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call

SeisSol	
SeisSol input files are structured along a main parameter file (see e.g. https://seissol.readthedocs.io/en/latest/parameter-file.html) and yaml file readable by the easi library, describing spatially variable parameters. SeisSol also requires a mesh file and files describing fault and surface receivers. Many examples can be found at https://github.com/SeisSol/Examples	Velocities (variables v1 to v3) and ground displacements (u1 to u3) are outputted as a surface representation in an unstructured grid of triangle (XDMF/Hdf5 or XDMF/binary files, which can be opened in ParaView, or e.g. with the seissolxdmf python module). The surface mesh is given in the mesh coordinate system, which can be straightforwardly transformed back to the geographic coordinate system (latitude, longitude) given the coordinate projection description, with, e.g., the proj4 library. When running fully-coupled simulations (with seismic, acoustic and gravity waves), sea-floor (acoustic and elastic sides), sea-surface and normal free-surface cells are all written in the same file. The quantity "locationFlag" allows filtering the cells (see documentation at https://seissol.readthedocs.io/en/latest/free-surface-output.html).
Tsunami HySEA	
Mandatory inputs: • Text (ASCII) file containing all the input parameters; • A NetCDF grd file with the bathymetry/topography. • Optional inputs: • Initial sea surface displacement in NetCDF grd format; • A text (ASCII) file with the Okada parameters, either using standard or triangular faults; • One or more NetCDF files containing the deformations to be applied to the sea floor, one file for each deformation; • A NetCDF file grd containing a dynamic deformation; • A text (ASCII) file with the points in which to obtain the time series; • One or more NetCDF files containing variable frictions, one file for each grid.	• One NetCDF file for each grid used in the simulation; • One NetCDF file containing the time series data, if time series is used.
GAUSS	
• Either multi-Okada fault parameters as ASCII file with each line representing a sub-fault, or • Surface elevation grid in various formats (NetCDF, Golden Software, xyz).	• Time series at coastal POIs in tabular text form (ASCII): rows- time of a snapshot (typically 30 sec interval), columns- wave heights); • An aggregate file (ASCII): file with hmax and optionally with the results of mareogram analysis: polarity, dominant wave period, hmax.
SLSMF	
Utilize the IOC SLSMF API to retrieve real-time sea-level data. Inputs such as station codes, temporal parameters, and desired data formats are essential for the data retrieval process to fit the requirements of rapid simulation workflows.	Outputs, conforming to various formats including JSON and XML, ensure compatibility across diverse modeling tools.

Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call

Additional model/service updates relevant to the integration of inputs are outlined here:

- **k223d**: is presently working on linking the input for this code with open-source geological fault database such as the European Databases of Seismogenic Faults (VA2-33-1). This is being done through the generation of Jupyter Notebook tutorials which show how to mesh a fault plane based on the fault database and generate a vtk file that can then be read in by k223d. These tutorial notebooks will be realised by M24.
- **k223d**: a new version of the code is being developed where the I/O can be read by the free software Paraview. It is envisaged that it will be necessary to vary the output to include formats that are compatible to the HySEA and SeisSol codes, which can be seen as the next step to connect in the chain.
- **SLSMF**: By harmonizing the input/output mechanisms with real-time data accessibility, the IOC SLSMF API significantly enhances the capacity for early tsunami detection and risk assessment, underlining its value in supporting advanced research and mitigation strategies in the Tsunami geohazard field. Looking ahead, the development of a specialized research data endpoint is anticipated, which will extend the data availability beyond the current 30-day limit, incorporating automated quality control features and the innovative automatic single stream creation capability. This will enable the selection of the optimal sensor from multiple data streams at any given day, alongside the introduction of spatial filtering options. Such advancements will further empower researchers with enhanced data precision and flexibility, paving the way for more accurate and efficient tsunami simulation and analysis in HPC environments.
- **SeisSol**: An interface module allowing to convert SeisSol surface output to HySEA netcdf input has been developed in the DT-Geo project. Link: https://github.com/dtgeoeu-wp6-tsunamis/Interface-module

The next step towards improving model interoperability involves coordination between the WP5 service providers to agree on the feasible and desirable level of standardization of the model I/O formats to be achieved during the rest of the Geo-INQUIRE project lifetime.

2.1.3 Design of a new e-Infrastructure for VA to Scientific Software

One of the objectives of Task 5.3 is the development of a e-infrastructure providing Software as a Service (SaaS) to the geophysical community. The novel infrastructure is designed to ease the implementation of Pilot Demonstrator Applications (PDA) during the course of Geo-INQUIRE and beyond, in a sustainable perspective. In the reporting period, the following activities have been undertaken:

1. Definition of different categories of SaaS (as relevant for Geo-INQUIRE scientific and technological community).
2. Involvement of the Geo-INQUIRE software developers in the identification of the "functional requirements".
3. Design of a web application, based on REST architectural style, supporting interoperability between different computers.

Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call

4. Implementation of the Web Processing System (WPS), with three prototypes PDAs.
5. Starting of testing phase.

The service aims to provide APIs to make software portable and accessible as a Web Processing Service (WPS). The WPS can ingest input with a defined syntax, execute code on a dedicated server and return results asynchronously. WPS requires simple interactions between the server (hosting the service) and the client (requesting the service); no specific installation is required, neither by the developers of the scientific applications, nor by the users. The adopted solution includes the definition of API (Application Programming Interfaces) that formally specify how to access (i.e. "make requests to") the web service, potentially allowing other applications to directly access the service (Figure 3).

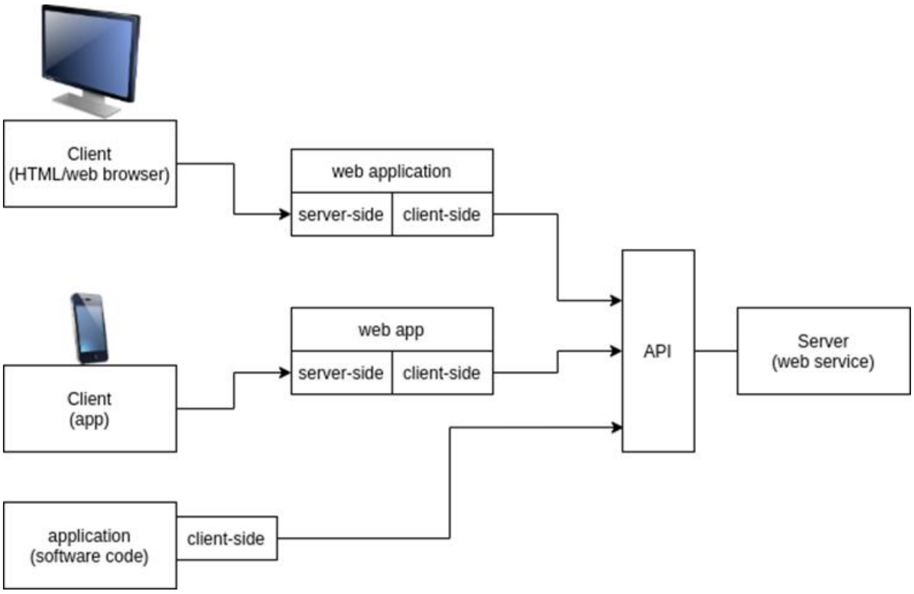

Figure 3: Sketch of the web service

The web service is based on the REST architecture. The server does not maintain the status of the communication; every request is independent by the previous and contains all the necessary information to provide the result; the requests to retrieve data (i.e. using GET HTTP method) are idempotent and cacheable. Therefore, the server requires less resources, is faster and simpler, and more scalable. The web service has been implemented in Flask, a micro web framework written in Python, and supports scientific applications written in several programming languages, both interpreted (e.g. Python, shell) or compiled (e.g. Fortran, C), possibly requiring specific libraries and installations.

The service is still in testing mode with three PDAs and will be probably released on the web in the next year. While the web service solution is designed to support interoperability, it does not limit the possibility to have a user-friendly interface, but rather leave the task of making the user interface to a different layer of code, possibly separate from the service; also, different code for different user interfaces (e.g. apps for mobile) can be developed without changes to the service.

2.2 Update of Task 5.2 “Simulation Data Lake”

Task 5.2 focused its activity on the design, development, and deployment of the Simulation Data Lake (SDL). Since M6, significant advancements have been achieved.

As mentioned in D5.2, an initial and temporary SDL@CINECA VA service was released. This temporary solution relies on the EUDAT Collaborative Data Infrastructure (EUDAT CDI) and its B2* service ecosystem. To enhance data sharing and dissemination efforts, a dedicated Geo-INQUIRE B2SHARE community was created. Additionally, seamless integration with B2ACCESS, B2HANDLE, and B2FIND services enhances the SDL infrastructure's capabilities in data access, management, and discovery, respectively.

CINECA is dedicating great efforts to the development of the updated solution for the SDL@CINECA installation, which will be hosted on ADA Cloud. The SDL@CINECA is being built upon a modular architecture, featuring components that work harmoniously to facilitate efficient data management and interaction. The backbone of the SDL is its REST API Service, providing developers seamless access to metadata and data resources. A robust metadata catalogue enables sophisticated queries such as full-text search and geospatial analysis. The SDL relies on S3 compatible object storage for storing datasets and data products, ensuring scalability and access through standard HTTP methods. A user-friendly Web Interface and a Command-Line Interface (CLI) tool will offer users versatile means of interaction with the system, supporting, e.g., interactive prompts and script integration. Complementing these features, a centralized Role-Based Access Control (RBAC) layer, will allow to manage access policies and permissions assignments to users. This comprehensive architecture of the SDL@CINECA system will empower researchers within the computational geoscience domain with advanced tools for data exploration and management.

The first version of the SDL@CINECA service will be released by the end of M18 and will provide an initial and basic set of functionalities. The subsequent gradual improvements of the service will rely on the first data uptake use cases, already in preparation.

As part of the development of the SDL@CINECA service, CINECA is working, in collaboration with the simulated data providers of WP5, on the definition of data standards, and in particular, an optimal metadata schema for Geo-INQUIRE simulation data (Figure 4). It is worth noting that the datasets stored in the SDL should bring along the information regarding the workflow which generated them. Hence, the metadata definition is being addressed in collaboration with the DT-GEO project, which is developing similar metadata schemas for the workflows. The metadata schema in DT-GEO describes workflows often including numerical simulations. The workflow description which accompanies a simulation experiment (a collection of simulation runs) whose results are stored on the SDL represents a specific realisation of potentially the same workflow, with the fields describing for example the format and the types of the input parameters taking actual values when used to describe a simulation experiment or a single simulation run.

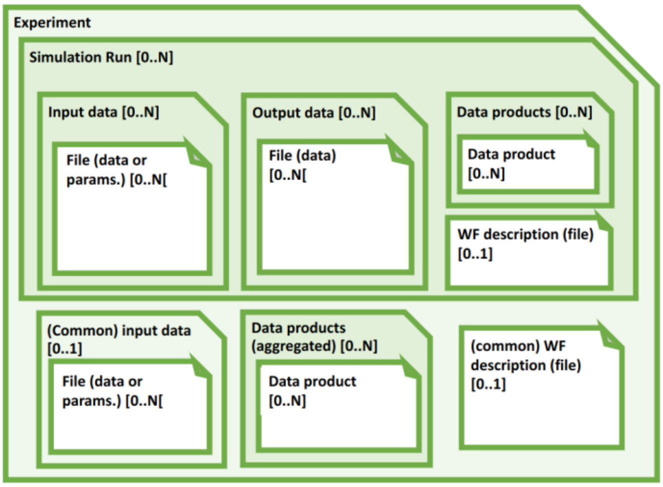

Figure 4: Graphical representation of a generic SDL dataset or experiment.

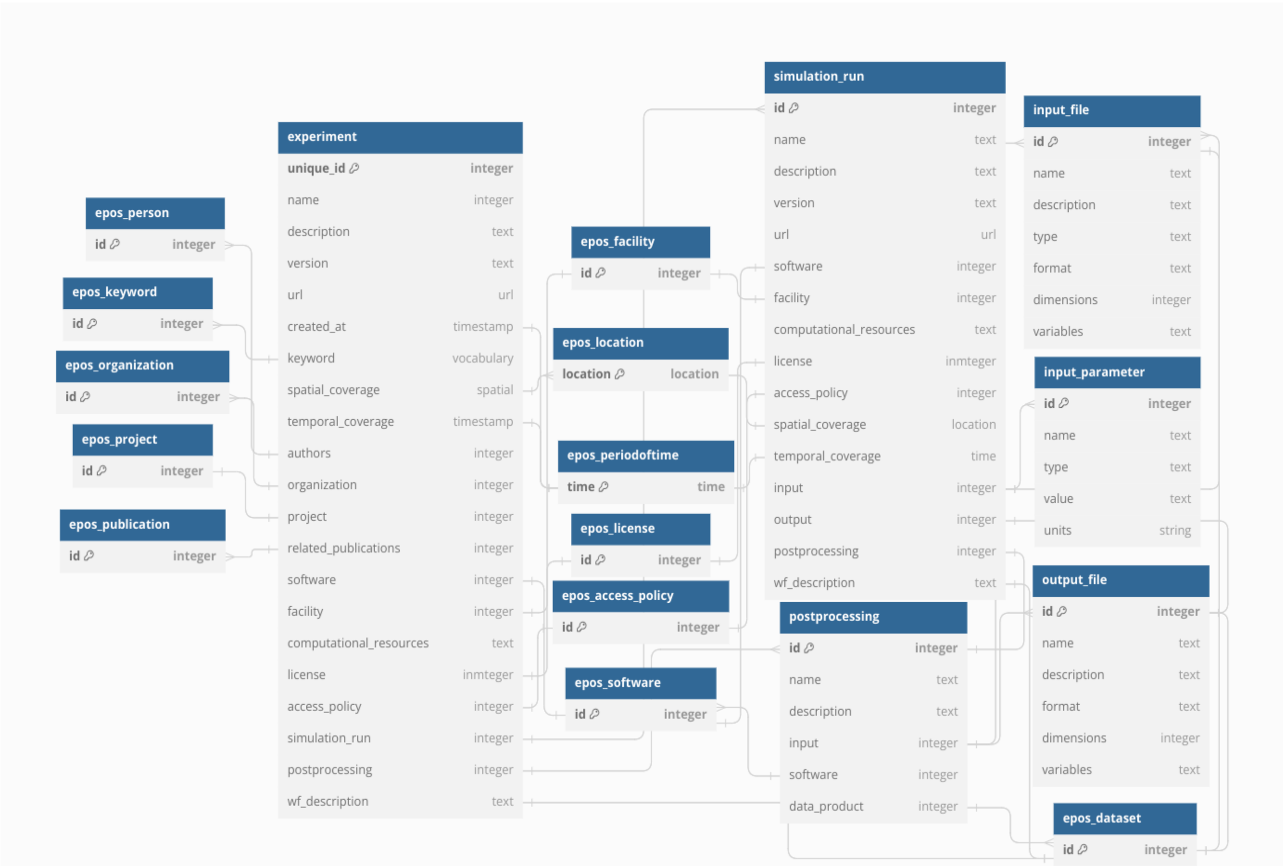

Figure 5: UML diagram of the Geo-INQUIRE simulation data metadata schema. The figure represents the current status of the schema definition rather than its final shape.

Our approach relies on a meticulous process of data modelling. We prioritized the creation of a comprehensive schema that could effectively capture the intricacies of simulation data and processing details. Through iterative development and feedback loops with researchers and domain experts, we refined our metadata schema to ensure it aligns closely with the specific needs and requirements of Geo-INQUIRE, in particular, the FAIR principles the project is committed to. This included defining standardized metadata fields for each entity identified and establishing controlled vocabularies. We identified and characterized the entities experiment, simulation run, input, output, and post-processing (see Figure 4).

In line with our commitment to interoperability among services and infrastructures in the geoscience community, we have established a collaboration with EPOS and -as noted above- DT-GEO, that aims to converge towards a common metadata schema, under the auspices of Geo-INQUIRE WP6-7. The ongoing efforts will materialize in a unified proposal for the extension of EPOS-DCAT-AP (https://epos-eu.github.io/EPOS-DCAT-AP/v3/). The collaboration articulates through regular meetings and data modeling sessions to harmonize our metadata schemas and vocabularies. Figure 5 displays the relationships between Geo-INQUIRE identified new entities and classes included in the EPOS-DCAT-AP v3.0 schema. Currently, an additional mapping of the entities to those proposed by DT-GEO is being carried out.

The shared scientific community and commitment to FAIR principles render this collaboration among projects of strategic importance. We are not only creating an optimal metadata schema for Geo-INQUIRE simulation data but also contributing to the development of a robust and standardized framework for the broader geoscience community.

In our ongoing effort to provide the most comprehensive and relevant insights into the SDL and its adherence to data standards, we made the strategic decision of delaying the publication of the White Paper drafted by M12 (see D5.2). By waiting for the latest significant progress to be included in the white paper, particularly regarding the metadata schema and SDL architectural design, we aim to release a more valuable and useful publication, thus maximizing its impact.

2.3 Update of Task 5.3 “Software as a Service”

The standardization and metadata preparation effort conducted for the SDL in cooperation with that regarding the DT-GEO workflows, and in coordination with EPOS will be also applicable to the software provision. Some of the software (the tsunami software, for example), is in fact already under distribution from the portal of the Tsunami TCS in EPOS. The application/extension to the EPOS-DCAT-AP metadata schema will make the distribution FAIRer.

All HPC-enabled codes (SeisSol/ExaHyPE (VA4-531-1, TA2-531-1) (see Figure 6), Fall3D (VA4-531-2) and HySEA (VA4-532-1, TA2-532-1)) will be made available on CINECA Tier-1 or Tier-0 systems G100/Leonardo booster. The efforts to unify their access by providing a Geo-INQUIRE module environment have been completed for two of the codes offered as TAS and are being completed for the other two (see Chapter 2.5). They have, in fact, been offered through the 1st TA Call, together with the OpenFOAM volcanology codes and BingClaw, which do not need that as they run at INGV

Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call

and NGI, respectively. The HPC-enabled code provision as spack files, which could be re-used by other supercomputing centers, is being achieved in collaboration with ChEESE-2P.

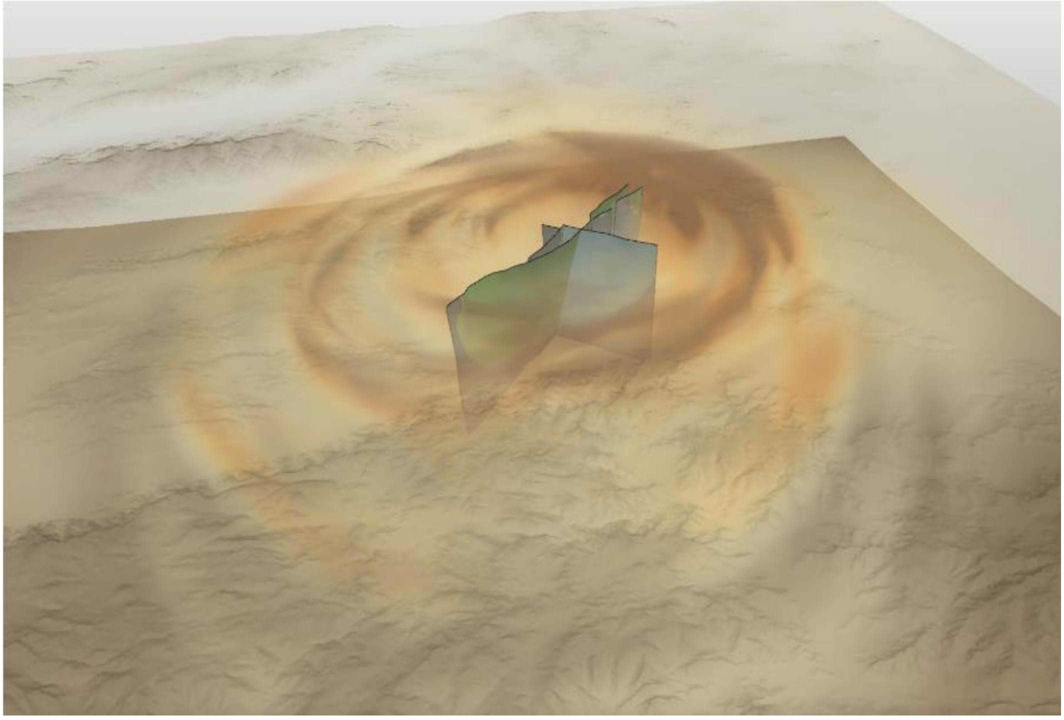

Figure 6: Volume rendering of SeisSol simulations (educational visualization showcase at the SC'23 supercomputing conference)[1]

Some of the computational less intensive services can directly be accessed via the web (BingClaw (VA4-532-1, TA2-532-1), Pyrocko(VA4-534-1), AmpFact(VA4-532-2), TS-Gauss (VA4-532-3, VA4-532-4), VO-CIV(VA4-533-1)). To make sure that all codes provide easy access for the users:

- We provided DOIs for all services, together with documentation on how to install and use the codes.
- For the codes that are more complex to install, docker containers could be a work-around to give users the opportunity to use these codes out-of-the-box for trainings or first test runs with little overhead. This approach has been successfully used in many sessions of the "High Performance Seismology (HPS) Training"
- For several codes training materials already exists (examples, slides, video recordings, etc.). These will be further improved and homogenized during Geo-INQUIRE.

[1] Abram et al., SC'23, Conversing Faults: The 2019 Ridgecrest Earthquake, https://sc23.supercomputing.org/proceedings/sci_viz/sci_viz_files/svs102s3-file1.pdf
The link to the full SeisSol video is here: https://sc23.supercomputing.org/proceedings/sci_viz/sci_viz_files/svs102s3-file2.mp4

Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call

- In close collaboration with WP9 several trainings and workshops have been held or are planned. The recordings of those could be used for either (self-paced) trainings or for advertisement of VA and TA calls

In December 2023, a new “Level 3” software+data service called “TS-Gauss” has been started. TS-GAUSS (the “Gaussian Toolbox”) provides datasets and tools for fast – within a few seconds – tsunami simulations based on pre-computed tsunami Green’s functions. Tsunami waveforms at pre-defined coastal locations along the Mediterranean Sea coasts can be calculated for an arbitrary sea-surface elevation. The calculation is a linear superposition, according to the given sea bottom displacement, of the elementary tsunamis produced by a set of gaussian-shaped unit tsunami sources, evenly distributed across the sea surface in a regular pattern (Figure 7). This service is now static (download-based), but is being upgraded into a web-application. The developers are in close-contact with T5.2 to use the capabilities of the SDL for this service.

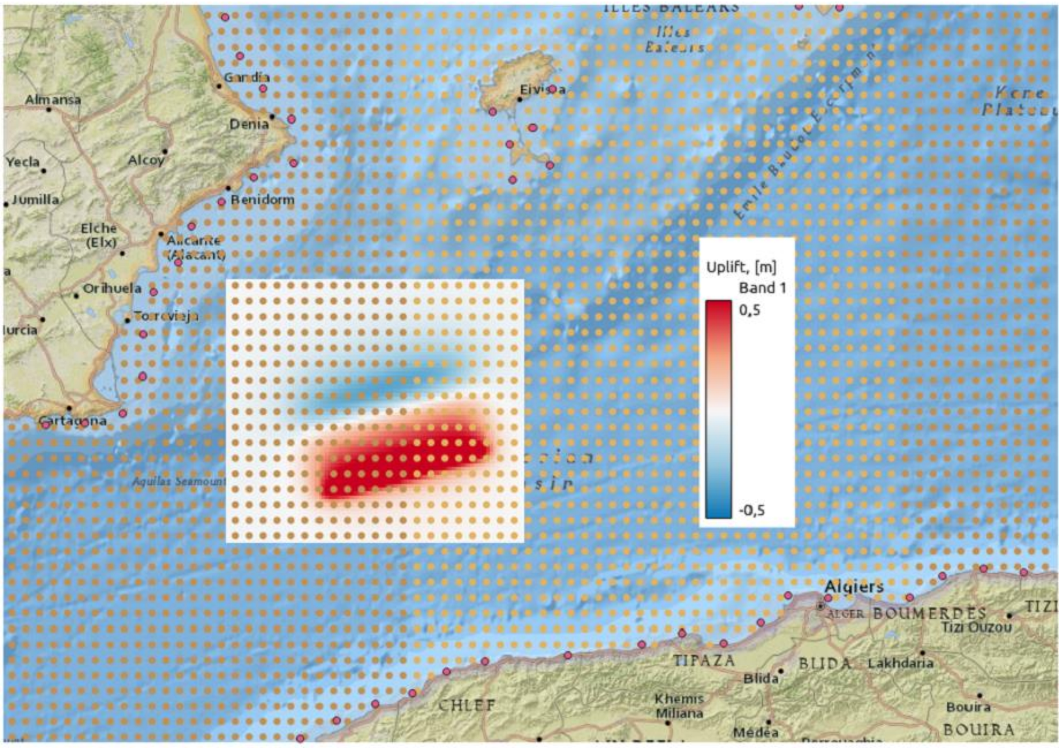

Figure 7: Example of initialization of the TS-Gauss tsunami calculations.

2.4 Update of Task 5.4 “Workflows as a Service”

The coordination activities in T5.4 have been focused on the definition of the access call, selection criteria and planning for Transnational Access. In particular, TA in the framework of T5.4 has specificities that needed to be coordinated:

Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call

1. some (most) of the e-infrastructures (workflows) are distributed in nature, i.e., they are installed in multiple instances, each of which has a specific focus (software development, scientific application, operational environment, supercomputing execution);
2. some (most) of the workflows require intensive HPC (High Performance Computing) resources provided by EuroHPC European-level installations (CINECA, BSC, LRZ) that need to be planned and secured for the TAs. Coordination of such activities has led to the 1st TA call.

Personalized training (Training-Through-Research, TTR) will be provided to grant owners during the development of TAs. Training material will be made publicly available through Zenodo for ChEESE installations (T5.4.1), and through the EFHER platform for T5.4.2.

One TA example for which TTR will happen is the Probabilistic Tsunami Hazard Analysis (PTHA) workflow TA2-541-3/4/5/6/8 for tsunamis of seismic origin (Figure 8). As any PTHA for seismic sources, the first step includes the preparation of the input seismic probabilistic source model, which defines a large ensemble of tsunami scenarios to be simulated, in this case with the Tsunami-HySEA software (also distributed in Geo-INQUIRE as both VA and TA outside of this TA on PHTA). For this to happen, it is necessary to assist the awardees regarding the geophysical and statistical aspects of the model preparation. It is then necessary to train them regarding the usage of Tsunami-HySEA and the computational resources, and to post-process the simulation results to eventually aggregate them with the source probabilities to obtain the hazard model.

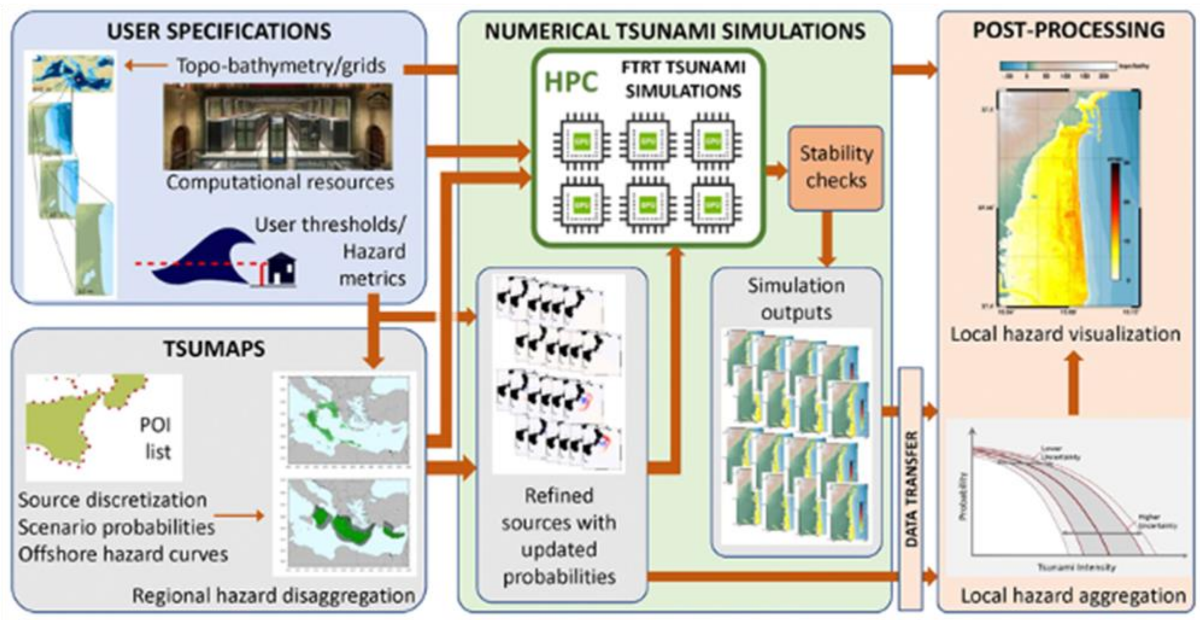

Figure 8: Conceptual scheme of the PTHA TA workflow. Modified after Gibbons et al., 2020, Front. Earth Sci. 8:591549. doi: 10.3389/feart.2020.591549.

In parallel, each installation provider has cured the maintenance and upgrading of the software, and of the execution environment, and tested the workflows on the new target HPC architectures

Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call

for the TAs. For T5.4.1, this activity has been carried out in collaboration with the ChEESE-2P Centre of Excellence for Exascale in Solid Earth Sciences.

Task T5.4 will only report KPIs for Transnational Access, i.e., concerning the actual success of the first TA call, in terms of number of applications and interaction with the providers of each installation, and of the scientific impact of the TAs.

2.5 Update of Task 5.5 "Access to HPC Systems"

Task 5.5 of WP5 in Geo-INQUIRE is related to the use of HPC CINECA Tier-0 and Tier-1 architectures for selected Transnational Access (TA) activities. The two CINECA HPC systems available to Geo-INQUIRE are:

- **Galileo100** (G100, https://www.hpc.cineca.it/systems/hardware/galileo100/) and
- **Leonardo booster** (ranked No.6 in the latest top500 list from November 2023, Figure 9 from https://www.hpc.cineca.it/systems/hardware/leonardo below)

Figure 9: Screenshot about Leonardo booster

In the framework of T5.5, a deployment campaign was implemented as a close cooperation between the installation teams of the service providers and the user-support team at CINECA. The status of the installations is shown in the table below (end of February 2024):

Table 2: Status of installations on February 2024

Software	Status	Machine(s)
SeisSol	on-going	G100/Leonardo booster
ExaHyPE	on-going	Leonardo booster
FALL3D	tested	G100/Leonardo booster
HySEA	tested	Leonardo booster

Some comments on the work done in the reporting period for Task5.5:

- Every code is able to efficiently use A100 accelerators of Leonardo booster. Two codes (SeisSol and FALL3D) are able to use the CPU part (Intel) of G100 machine. In both cases, CINECA user-support team have checked tools and dependencies needed for the correct and efficient execution on a CPU (G100) and CPU+GPU (Leonardo booster) environment.
- The local installations try to reinforce the use of spack environment (which is the natural choice on both machines). The original installation procedure has been, eventually, modified to take into account the use of a spack environment.
- For every code, a module in the Geo-INQUIRE profile has been created to simplify the further use of the code(s) in the TA activity.

Next steps:

- Finalize the installations on both machines for the entire set of codes.
- Refine and improve the spack-awareness in the framework of TA activity.
- Eventually add optional features (e.g. monitoring of the performances, containerisation, in coordination with the ChEESE-2P project).

2.6 Update on in-kind CPU-hours

In addition to Galileo100 and Leonardo booster the two systems at CINECA, several other sites will provide in-kind CPU/GPU-resources for the TAs:

Table 3: In-kind contributions of CPU/GPU resources for TAs.

TA	Code	System	Computing Center
TA2-531-1	SeisSol	Heisenbug / SupermucNG	LMU / LRZ
TA2-531-2	OpenFoam	Vesuvio cluster	INGV-Pisa
TA2-531-3	HySEA	Mercalli	INGV-Roma
TA2-541-9/10	Cybershake	MareNostrum4	BSC
TA2-542-1	EFEHR	Local cluster / Euler IX	ETHZ

3 Coordination activities with cross-cutting WPs

3.1 Outcomes of the WP6/7 workshops from Jan/Feb 2024

In October 2023, during the yearly Geo-INQUIRE meeting in Rome, the outcomes of the M12 deliverables on KPIs were presented to WP6 and the general audience. With a clear focus on virtual access to high-end software packages instead of highly integrated web services (like those in WP2 and WP3) WP5 defined an additional set of KPIs in D5.2 at M12. These were KPIs for the SDL, several WP5-specific KPIs for VAs and also one for TAs (computing time budget used). To make this document self-contained, we included the list of KPIs as described in deliverable D5.2 in Appendix 2. Figure 10 shows the status of KPIs as identified by the different service providers in D5.2. It shows that none of the VA services is currently able to locate users geographically (VA-04), a KPI that was introduced relatively late in the process and is only measurable by webserver statistics. All services are planning to measure VA-03, the number of users served. VA-02, the number of data or metadata requests is only a useful measure for those services that provide data packages for download in addition to the software (e.g. TS-Gauss). It was discussed if the software downloads could also be considered as data, but then VA-02 would just be the same as VA-03. VA-01, the percentage of service usage logging systems in place is in general 100%.

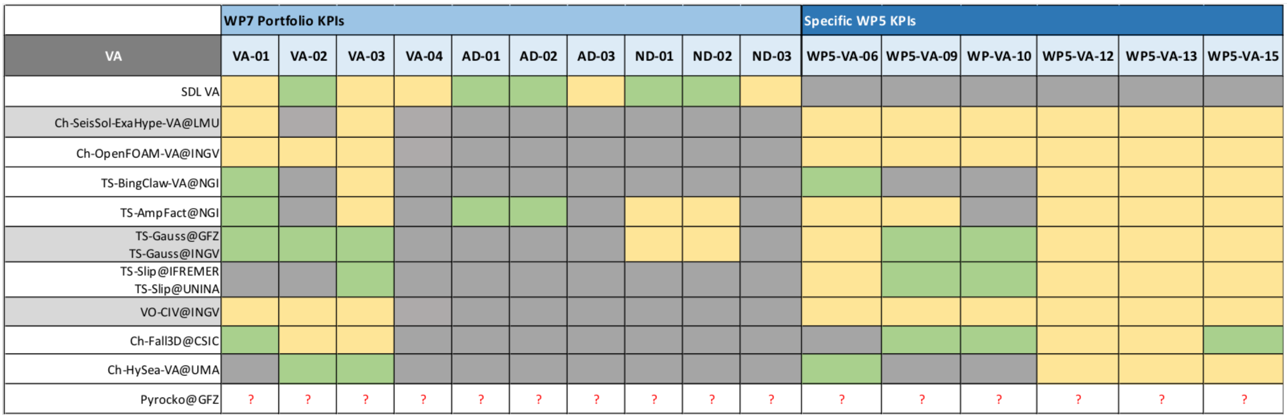

Figure 10: WP5 survey on VA-KPIs from D5.2. Green: Already available, yellow: planned, grey: not expected/not feasible.

When this was presented a fruitful discussion emerged. After the peculiarities of WP5 became clearer to WP6 and WP7, they explained their clear goal of unifying the set of KPIs among all WPs. In order to achieve this unification, the idea of slightly adapting the meaning of the basic set of KPIs to fit the needs of the WPs came up.

On late November 2023, this approach was further elaborated during a T5.3 meeting that was joined by WP6. T5.3 is responsible for all WP5 VAs except for the SDL that has its own unique set of KPIs. So, the codes in T5.3 were the ones for which the additional WP5-VA-KPIs were drafted in the first place. The T5.3 service providers showed examples how they plan to measure their KPIs, explained in more detail the WP5-specfic KPIs, their meaning and the way they could be measured. WP6 provided ideas how to adapt the core set of VA-KPIs for WP5. A technical overview of how WP5 could measure both the core set and the WP5-specific VA-KPIs can be found in Chapter 3.2. This is meant as input for WP6 for their homogenization efforts that will lead to the M24 deliverables with KPI figures.

After further discussions across WP leaders in the PMB meetings, it was decided that WP6/7 would organize workshops with each of WP2, WP3, WP4 and WP5. The coordination workshop with WP5 was held on 16 January 2024 among the WP and task leaders. The agenda covered was:

1. Implementation Level Matrix
2. Service in installations
3. FAIRness evaluation

3.1.1 Implementation Level Matrix (ILM)

Besides a check of all contact points for the VA services in the ILM, a few points were discussed in more detail: The meaning of the column "RI integration" and the specialties of the SDL. The SDL consists of several Vas. The main VA implements the service itself at CINECA, the partners could be considered as sub-services that deliver data to the SDL (similar to other services in WP3).

The link to the ILM has been forwarded to all service providers with a request to check and update the information. The information there is up-to-date and WP6 will report on this in more detail in their M24 deliverable. With the information from the ILM the four SSRI-KPIs can be given:

- **SSRI-01**: *Percentage of installations with full RI integration*
  3/12 services are fully integrated, the remaining 9 services are partially integrated
- **SSRI-02**: *Percentage of installations fully operating*
  12/12 services are fully operating
- **SSRI-03**: *Percentage of installations under implementation operating*
  0/0
- **SSRI-04**: *No. of services running and reachable publicly*
  12/12, all services are up and available

3.1.2 Key Performance Indicators (KPIs)

As explained above, after consultation with WP6 and WP7, it was decided that WP5 should concentrate on the VA-KPIs designed for all services at least for the M24 deliverables. This will mean that WP5 will slightly "interpret" the VA_KPIs to grasp the core purpose of each. It will also mean that WP5 will prepare to measure those KPIs. The numbers might not be as comparable to other WPs (WP2,3,4) that mostly use webserver statistics, since WP5 will have to rely on zenodo or GitHub statistics, but it will enable a unified set of KPIs for all services.

A more, provisional, technical analysis how to measure VA-KPIs in WP5 can be found in Chapter 3.2. We decided to provide in-depth information how each of the services is planning to collect the KPIs to have a solid information base for WP6/7 for the M24 deliverables. The Information on how the VAs plan to collect KPIs can be found in Appendix 1.

The four VA-KPIs and their adaptation to WP5 are:

- **VA-01:** *Number (and/or % of installations) of Service Usage Logging systems in place*
  This will be 100% of all installations, since zenodo DOIs are provided by everyone
- **VA-02:** *Number of data or metadata requests*
  This is one of the trickiest KPIs for WP5. Only very few VAs deliver data itself, examples are the webservices (e.g. Pyrocko, VO-CIV), a few will provide data packages together with their software. For the remaining codes It was decided that software downloads could count as data requests. For those codes VA-02 will be equal to VA-03. (VA-02 == VA-03)
- **VA-03:** *Number of users served*
  As explained above, finding the right number might be tricky. Since the codes differ very much in their target user base, it might be best to collect GitHub clones where possible but we should understand that some codes might only be able to report zenodo views for this.
- **VA-04:** *Number of systems capable of geographically locate users*
  This will only be possible for those codes that are webservices (e.g. Pyrocko, VO-CIV). Technically, it will also be possible for local GitHub/GitLab services with a tracking system but this would cause a lot of overhead and is certainly not a solution for all codes.

For TAs the situation is different, because their KPIs depend on the number of calls, applicants, etc. The first TA call closed 15 March 2024 so we have first information on some of the KPIs (other will only be available after some analysis, after the selection process finishes or after the TAs itself have been carried out).

3.1.3 FAIRness

To assess FAIRness, the project will rely on the F-UJI-Tool (www.f-uji.net). This assessment will be carried out by WP7 on its own. WP7 already provided guidelines on how service providers could support that tool (e.g. by providing license information in json format). They will use the collection of DOIs provided by the service providers to WP7 and run an automated assessment. A feedback loop with the service providers is foreseen.

Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call

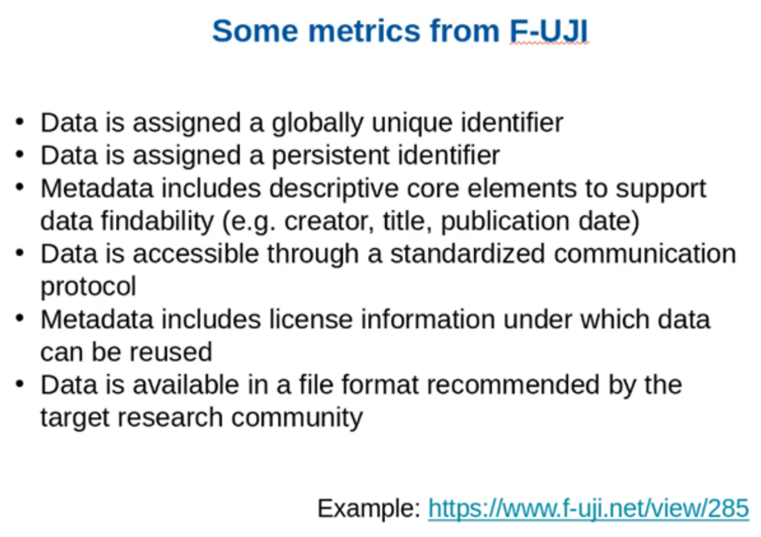

Figure 11: Examples of the metrics from F-UJI tool.

3.2 Different ways to measure VA-KPIs in WP5

This chapter will show different ways how WP5 partners could measure VA-KPIs for the M24 deliverable (both the core VA-KPIs and the WP5-specific KPIs). This chapter is meant as input to WP6 to decide on the actual measurements. We show possibilities and limitations of different approaches and put into relation two different codes that represent the "extremes" of WP5:

- Slip@UNINA is a code with one main developer, that you can simply download and use on a standard computer.
- SeisSol is an example of an open-source software project with many contributors that runs on high-end HPC machines but can take weeks to be ported to new systems because of its dependencies.

All installations provide DOIs for their codes through zenodo. Zenodo (see Figure 13 and Figure 14) is an easy tool to use to count the number of views and downloads. However, zenodo views and downloads are not a very representative way of measuring "data access" in WP5.

The main access to the codes is through their GitHub or GitLab repositories. GitHub "Insights" for example provides some information on the codes. The numbers that best matches VA-03 ("number of users served") is probably the number of clones (see Figure 12).

The number of GitHub "clones" is also provided to developers with push rights to the repository. They can use the GitHub CLI to query:

```
gh api \
  -H "Accept: application/vnd.github+json" \
  -H "X-GitHub-Api-Version: 2022-11-28" \
  /repos/<owner>/<repo_name>/traffic/clones
```

{"count":2897,"uniques":132,"clones":[
{"timestamp":"2024-02-27T00:00:00Z","count":327,"uniques":28},
{"timestamp":"2024-02-28T00:00:00Z","count":239,"uniques":71},
{"timestamp":"2024-02-29T00:00:00Z","count":256,"uniques":17},
{"timestamp":"2024-03-01T00:00:00Z","count":389,"uniques":5},
{"timestamp":"2024-03-02T00:00:00Z","count":1,"uniques":1},
{"timestamp":"2024-03-03T00:00:00Z","count":2,"uniques":1},
{"timestamp":"2024-03-04T00:00:00Z","count":435,"uniques":11},
{"timestamp":"2024-03-05T00:00:00Z","count":50,"uniques":5},
{"timestamp":"2024-03-06T00:00:00Z","count":179,"uniques":7},
{"timestamp":"2024-03-07T00:00:00Z","count":373,"uniques":5},
{"timestamp":"2024-03-08T00:00:00Z","count":79,"uniques":3},
{"timestamp":"2024-03-09T00:00:00Z","count":3,"uniques":3},
{"timestamp":"2024-03-10T00:00:00Z","count":3,"uniques":2},
{"timestamp":"2024-03-11T00:00:00Z","count":485,"uniques":6},
{"timestamp":"2024-03-12T00:00:00Z","count":76,"uniques":1}
]}

This will then return the number of clones over the last two weeks and the number of unique clones. So, it is a good way to measure VA-03 for software projects but in order to measure constant numbers, GitHub actions have to be used.

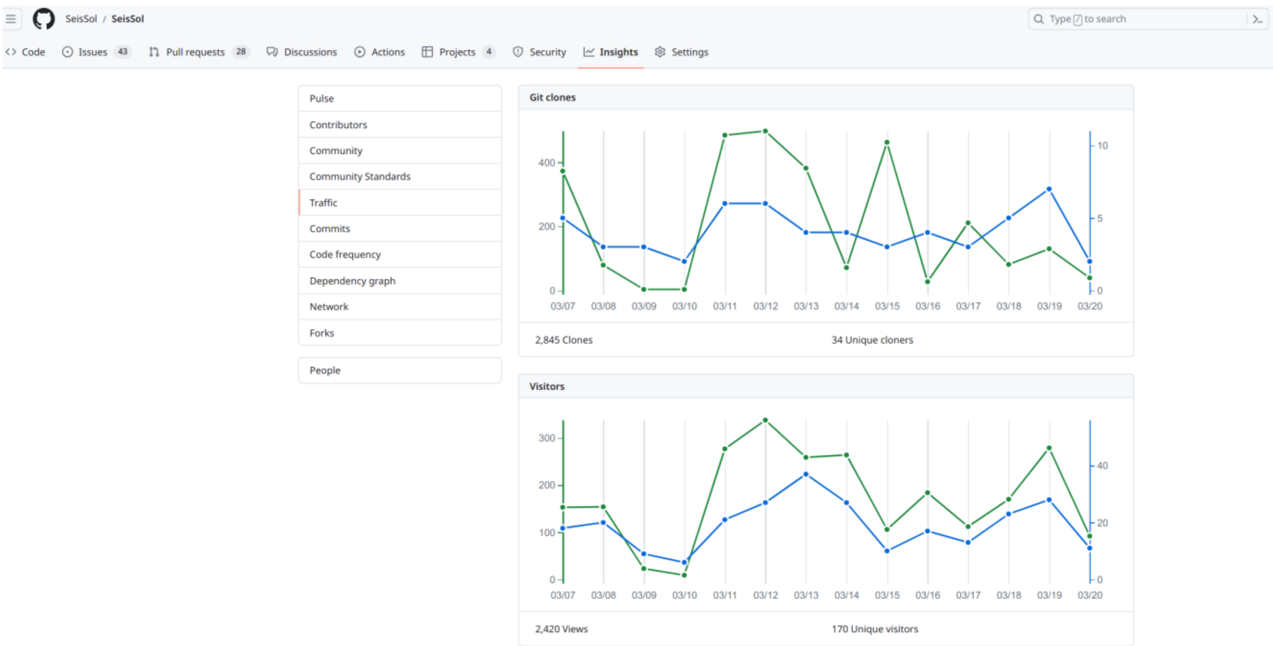

Figure 12: GitHub "Insights" on clones and visitors for SeisSol

Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call

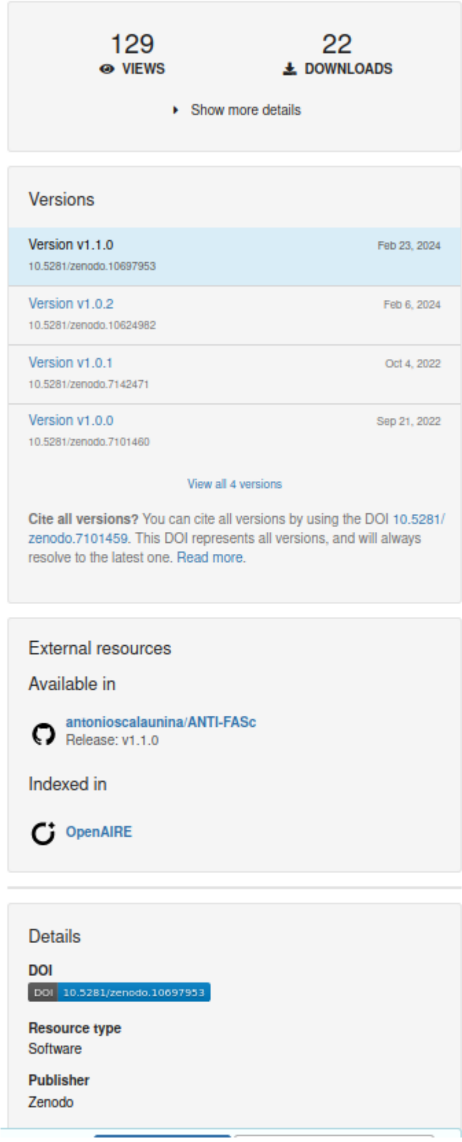

Figure 13: Zenodo figures for ANTI-FASc (Slip@UNINA)

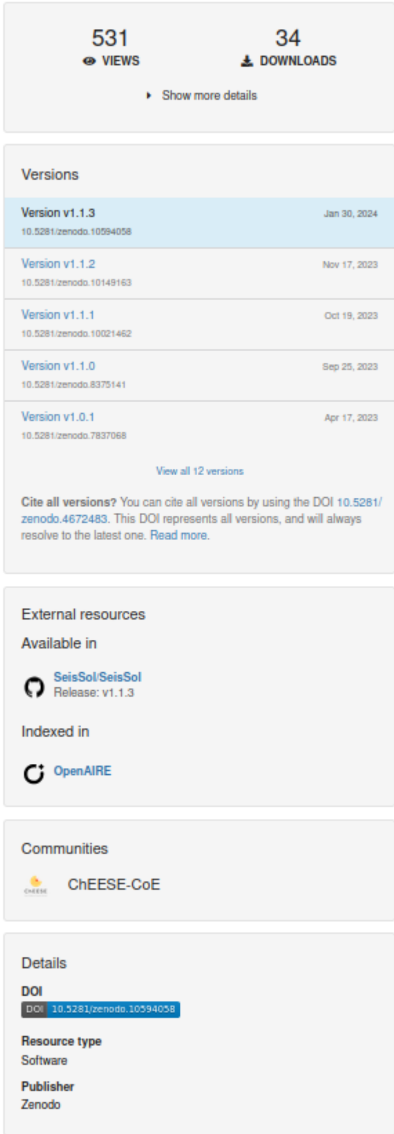

Figure 14: Zenodo figures for SeisSol

Repo-tracker gives a nice visual summary of GitHub repos and is able to show many of the WP5-specific VA-KPIs (see Figure 15 and Figure 16). Repo-tracker will only give you one month of insight unless you register your repo, but that could be easily done if WP6 decides to include the WP5-specific KPIs. By doing this exercise of analyzing and comparing the numbers it also became clear that, to some extent, those WP5-VA-KPIs measure more the performance of software projects than the actual benefit those projects bring to the community.

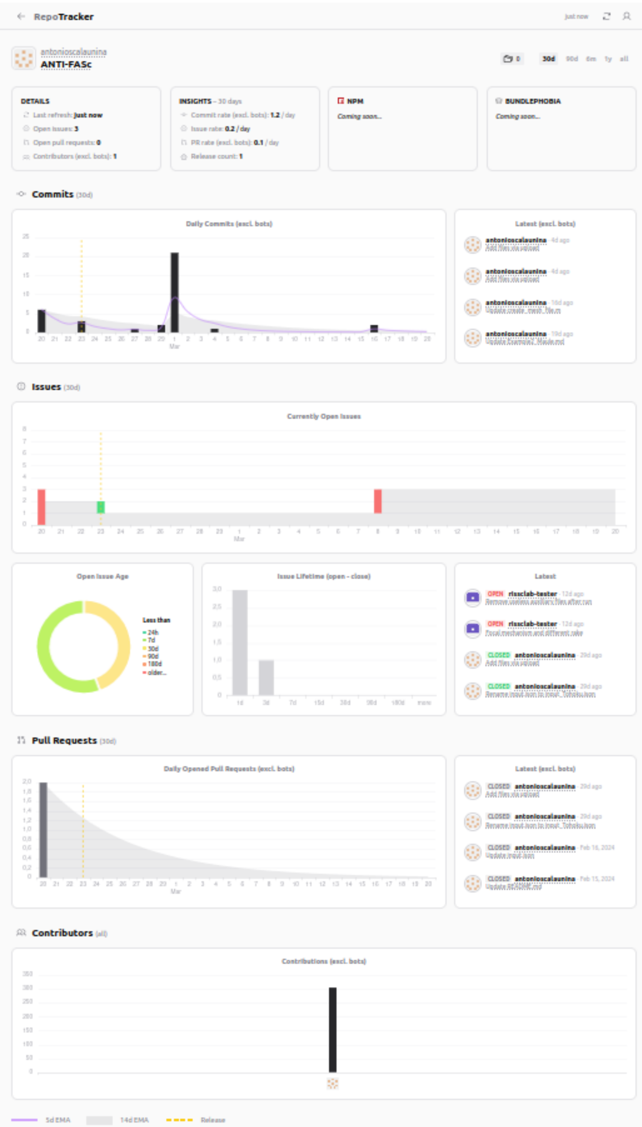

Figure 15: Repo-tracker for ANTI-FASc (Slip@UNINA)

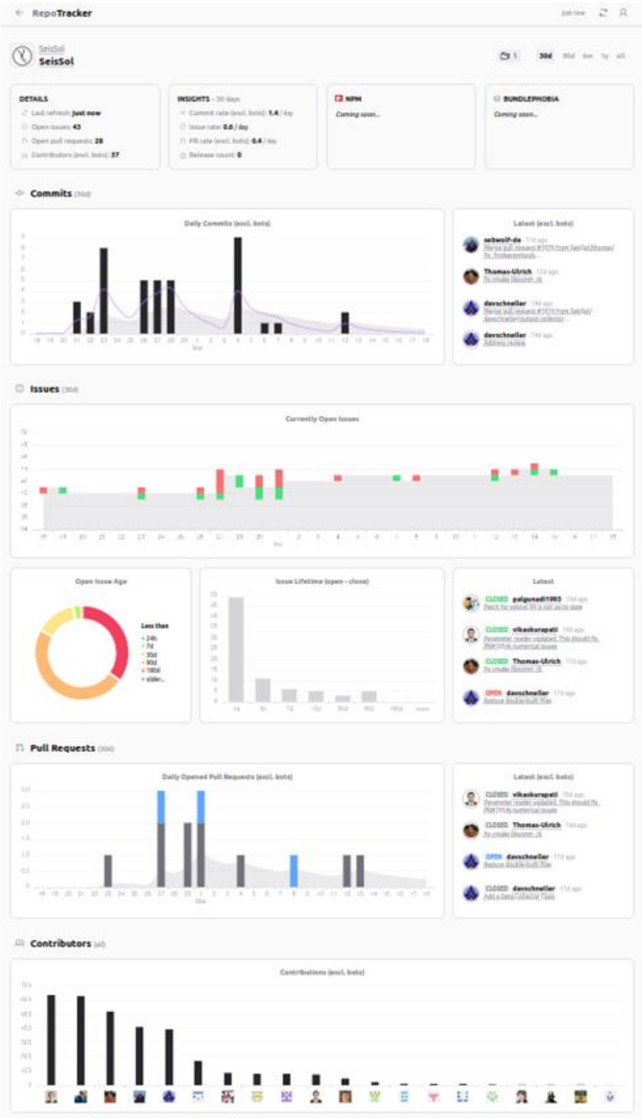

Figure 16: Repo-tracker for SeisSol

3.3 Dissemination Activities (WP9)

WP5 is tightly connected to WP9 through IFREMER and the LMU T5.3 co-leader which is also WP9 co-leader. During the first 18 months of the project WP9:

- updated the trainings & workshop plan (T9.1),
- prepared the first summer school in autumn 2024 in Greece (T9.2),
- started a service catalogue to identify cross-disciplinary opportunities (T9.3),
- looked for installations that are willing to participate in personalised training (T9.4)
- identified conferences and workshops for networking events (T9.5)

Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call

WP5 will mostly contribute to trainings and workshops and some installations volunteered to participate in the personalised training program (call will open in spring/summer 2024). In addition, many WP5 installations work on improving their documentation and training materials.

All TAs have a dedicated training plan that will be tailored to the needs of the successful TA applicants. We expect that the experience gained from "Training-Through-Research" will help each installation to improve their documentation, i.e. identify missing parts, update information, add examples, etc. Something that has not yet been decided (because it would need the involvement of TA users) would be some kind of "results workshop", e.g. an online event, where applicants could show their results and report on their experiences. This could be used both as input for reporting but also as advertisement for future calls.

To better understand what the VA services are working on, WP5 asked each installation what they already provide and what they plan to develop/improve during the project:

Table 4: Training and/or user support provided by WP5 installations

Code	*Training*	*Documentation*	*Online Support*
SeisSol	Joint HPS Workshop 05/2023, Hackathon Rome 2025 (Tsunami-SeisSol-Link)	x	x
ExaHyPE	Maybe in the future	x	x
Fall3D	Joint Cheese/Geo-I Workshop 03/2024	x	x
OpenFoam	-	x	x
HySEA	Online Training 04/2024, Hackathon Rome 2025	-	-
BingClaw	Hackathon Rome 2025	-	-
AmpFact	Hackathon Rome 2025	-	-
TS-Gauss	Hackathon Rome 2025	x	x
Slip@IFREMER	Online Training in late 2024	(use recording of training)	-
Slip@UNINA	-	(video course)	x
VO-CIV	(-)	x	x
Pyrocko	Evtl. Online training end of 2024	x	-

WP5 is one of the main contributors to T9.1 "Training & Workshops". One big event was already held in 2023 and materials and recordings are available online:

- **09 to 12 May 2023 (4-day training): *"High Performance Seismology Cybertraining"***
  Alice Gabriel (LMU/Scripps) and Marine Denolle (Univ. Washington)
  Joint Workshop with input from Geo-INQUIRE, ChEESE and DT-Geo, online

All trainings and workshops for 2024 are listed below. Those that happen before September 2024 will be reported in detail in D5.5 "Training summary report at M24".

Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call

- **11 to 14 March 2024: *"Computational Volcanology: Modeling volcanic eruptions using European Flagship Codes"*,** at BSC, Barcelona, Spain
  Arnau Folch (CSIC), Tomaso Esposito, Laura Sandri (INGV)
  joint ChEESE and Geo-INQUIRE Workshop

- **23 and 24 April 2024 (2-day training): *"Tsunami HySEA course: Introduction and tips on installation"*,** online
  Jorge Macías (UMA)
  Day 1: Introduction to Tsunami HySEA and virtual computational platform
  Day 2: HySEA Tsunami installation: tips and tricks

- **02 to 04 September 2024: *"Toward Data Lakes for Recorded and Simulated Earthquake Ground Motions"*,** at GFZ in Potsdam, Germany
  Fabrice Cotton & Ssu-Ting Lai (GFZ), Alice Gabriel (LMU/Scripps)

- **Autumn 2024: *"Stochastic slip training course"***
  Shane Murphy (IFREMER) & Andre Herrero (INGV)

Through **WP9 Personalized Training** funding is available to successful applicants (students, young scientists or other colleagues) to fund visits to VAs or TAs. The visits are meant to take a few days to perform a "mini-project" (data access, algorithm test, etc.). Up to two visits are possible, and they can be accompanied by online meetings. The first call for participation is planned for M18, the second for M30.

These WP5 installations expressed interest to participate in the first call for personalized training are (I) Ch-SeisSol/ExaHyPE@LMU, (ii) TS-Slip@UNINA and (iii) TS-Slip@IFREMER.

4 Final considerations

WP5 has used the last 12 months to tightly grow together in its endeavour to provide access to the next generation of software to the community.

Task 5.1 is our coordination task, where task- and WP leaders decide on common issues. T5.1's intra-WP coordination activities have been reported in Section 2.1, and the inter-WP coordination activities in Section 3, for consistency with other deliverables. The coordination work in the reporting period regarded: the preparation of the TA calls, the interoperability among WP5 VAs, and the design of a concept for VA access to software. Moreover, WP5 has closely worked with WP6 on KPIs and integration in large RIs (e.g., EPOS), WP7 on FAIRness and WP9 on trainings and workshops. Finally, there are active links (mostly through joint meetings on common aspects and key personnel) with the following projects:

Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call

- DT-Geo (on the SDL and its metadata scheme, on workflows, containers for digital twins and on using CompSs)
- ChEESE-2P (several HPC codes: ChEESE is actively working on improving those codes and their performance, as well as adopt well-known SW principles (e.g. CI/CD development, mini-apps, etc)

Task 5.2, the Simulation Data Lake, has used b2share as an interim data lake and is now very close to a final agreement on the metadata schema for its first own implementation (including input from DT-Geo and EPOS to offer a harmonized solution at the European level). The implementation is work-in-progress and should soon go live with a test version based on the new schema and defined functionalities.

Task 5.3 “SaaS” has progressed in training and documentation as well as in offering the HPC codes at CINECAs Tier-0 system, available through EuroHPC to all earth sciences. Many codes provided first information how they are planning to measure VA-KPIs to ensure that there is time for another feedback loop by WP6 before the first numbers will be reported in the M24 deliverables. All four TA services participated in Geo-INQUIRE’s 1st TA Call. As envisioned, only one service was not available right from the start of the project: “TS-Gauss” has meanwhile been implemented and is available through GitLab.

Task 5.4 “WaaS” has provided very valuable input to the documents and the selection process of the TA call. Nearly all workflows participated in the TA call and the next months will be a busy time both for the hosts and the receivers of the TA grants. T5.4 also helped in shaping the TA KPIs and will provide those for the M24 deliverables.

Task 5.5 “HPC@CINECA” ported -in close collaboration with the service providers- first codes to their platforms. Those are already available through the use of the module command and also spack files, that should simplify porting the codes to other Tier-1 and Tier-0 systems. Many TAs rely on computing budget at CINECA and their user support for carrying out the transnational access projects.

Appendices

Appendix 1: Status Updates of all WP5 VAs & TAs

T5.2 Simulation Data Lake

Installation ID	VA4-52-1 to VA4-52-8
Installation Name	**SDL@CINECA,** SDL@CINECA/GFZ, SDL@CINECA/INGV, SDL@CINECA/CSIC, SDL@CINECA/NGI, SDL@CINECA/LMU, SDL@CINECA/BSC, SDL@CINECA/UMA
Hosting Institution	**CINECA Consorzio Interuniversitario, Italy** GFZ, INGV, CSIC, NGI, LMU, BSC CNS, UMA
Description	The Simulation Data Lake (SDL) provided by CINECA (SDL@CINECA infrastructure) is engineered to ensure the enduring preservation, accessibility, and insightful exploration of results from numerical simulations and in-silico experiments. Committed to the FAIR principles (Findable, Accessible, Interoperable, and Reusable), this innovative service not only guarantees the reproducibility of analyses but also opens avenues for pioneering research while nurturing collaboration across diverse geoscience domains. SDL@CINECA comprises multiple installations. At the core lies the actual storage server at CINECA, serving as the Virtual Access (VA) provider. Complementing this, the sub-installations refer to the effort that different institutions will dedicate to curating existing simulation datasets and integrating new ones from various Virtual Access (VA) and Task Areas (TAs) within the Geo-INQUIRE WP5 framework int the SDL. CINECA is actively developing a state-of-the-art SDL hosted on its cloud infrastructure, envisioning it as a strategic asset within the geo-science domains. Upholding FAIR principles, the future SDL@CINECA will efficiently manage vast and intricate simulation datasets, fostering advancements in geoscience research. In preparation for the full release of SDL@CINECA, a temporary initial VA service is now available. This temporary solution relies on the EUDAT Collaborative Data Infrastructure and its B2* service ecosystem. To enhance data sharing and dissemination efforts, a dedicated Geo-INQUIRE B2SHARE community has been established. Additionally, seamless integration with B2ACCESS, B2HANDLE, and B2FIND services enhances the SDL@CINECA infrastructure's capabilities in data access, management, and discovery, respectively.
Update	CINECA is deeply invested in developing the SDL@CINECA infrastructure, which boasts a modular architecture designed for efficient data management. Key components include a REST API Service for seamless data access, a robust metadata catalog enabling sophisticated queries, a user-

Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call

	friendly Web Interface, The SDL utilizes S3 compatible object storage for scalability and accessibility. Finally, efforts have also been dedicated to the creation of a Command-Line Interface that will offer versatile user interaction, and the development of a Role-Based Access Control layer to ensure secure access control. This comprehensive system will empower geoscience users with powerful data exploration and management tools. In parallel, our journey towards an optimal metadata schema for Geo-INQUIRE simulation data involved meticulous data modeling and collaboration with EPOS ERIC and DT-GEO. Through iterative development, we refined our schema to meet project needs. Ongoing discussions aim to converge towards a common future metadata schema, potentially the updated version of EPOS-DCAT-AP. This collaborative effort fosters data interoperability, enhancing research efficiency and collaboration within the geoscience community. The progress achieved has been integrated into the drafted White Paper presented in D5.2.
KPIs	The current temporary SDL solution, based on the EUDAT infrastructure, employs key performance indicators (KPIs) to gauge its effectiveness and usage. On the one hand, it tracks the number of times datasets or records within the Geo-INQUIRE B2SHARE community have been viewed by users. On the other hand, it measures the total number of file downloads from the community, which indicates the extent to which researchers are accessing and utilizing the data available within the system. The future SDL@CINECA will adopt and track the VA KPIs presented in D5.2. A detailed presentation of the metrics will be presented in upcoming deliverables.
Links	
DOI	-
URL	https://b2share.eudat.eu/communities/Geo-INQUIRE
GitHub/GitLab/or other	-
Training & Documentation	
Training Material	-
Documentation	-
Training or Workshop	22 June 2023 - 3rd Geo-INQUIRE online Seminar: SDL - Simulation Data Lake by Alice Gabriel and Gabriella Scipione

T5.3 Software as a Service

Ch-SeisSol/ExaHyPE (VA & TA)

Installation ID	VA4-531-1, TA2-531-1
Installation Name	Ch-SeisSol-ExaHype@LMU
Hosting Institution	LMU
Update	A big SeisSol training ran in May 2023 (also covering Cybershake, HPC Usage, FAIR principles and Open Software best practices). The training material is available online, as well as the recordings. For the training a SeisSol aptainer was created to be used on Frontera and the docker containers were updated. SeisSol took part in the TA call. We are working together with TUM and CINECA on the SeisSol port to Leonardo. A first SeisSol SDL dataset has been uploaded to the test server of b2share and SeisSol input for the metadata scheme has been submitted. Our SeisSol simulations of the 2019 Ridgecrest earthquakes (Taufiqurrahman et al., 2023) served as an educational visualization showcase at the SC'23 supercomputing conference (Abram et al., SC'23 https://sc23.supercomputing.org/proceedings/sci_viz/sci_viz_files/svs102s3-file1.pdf) Link to the SeisSol video is here: https://sc23.supercomputing.org/proceedings/sci_viz/sci_viz_files/svs102s3-file2.mp4
KPIs	VA-01: 100% VA-02: Number of SW downloads via Zenodo: 509 views, 34 downloads (as of 11.3.2024) via GitHub: 132 clones in 2 weeks (as of 12.3.2024) VA-03: Number of users served (= VA-02) WP5-VA-KPIs: see repo-tracker screenshot in Chapter 3.2
Links	
DOI	https://doi.org/10.5281/zenodo.10594058
URL	http://www.seissol.org/ http://www.exahype.org/
GitHub/Gitlab/or other	https://github.com/SeisSol/SeisSol https://gitlab.lrz.de/hpcsoftware/Peano
Link to repo-tracker or webservice statistics	https://repo-tracker.com/r/gh/SeisSol/SeisSol
Training & Documentation	
Training Material	https://github.com/SeisSol/Training https://hub.docker.com/r/seissol/training https://hub.docker.com/repositories/peanoframework
Documentation	https://seissol.readthedocs.io

Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call

Training or Workshop	9.-12. May 2023: Alice Gabriel (LMU/Scripps), Marine Denolle (Univ. Washington) et. al. "**High Performance Seismology Cybertraining"** (joint online Workshop with input from Geo-INQUIRE, ChEESE and DT-Geo)

Ch-Fall3D (VA)

Installation ID	VA4-531-2
Installation Name	Ch-FALL3D@CSIC
Hosting Institution	CSIC - INGV
Update	New version release (8.3), SQAaaS passed with gold badge, deployed on several EuroHPC systems (Leonardo, LUMI, Vega), container version almost ready
KPIs	The KPIs applicable to the VA have been described in D5.2.
Links	
DOI	https://doi.org/10.5281/zenodo.6343786
URL	https://www.geo-inquire.eu/virtual-access/software
GitHub/GitLab/or other	https://gitlab.com/fall3d-suite
Training & Documentation	
Training Material	
Documentation	https://fall3d-suite.gitlab.io/fall3d/
Training or Workshop	Performed in the frame of the ChEESE CoE

Ch-OpenFoam (VA & TA)

Installation ID	VA4-531-3, TA2-531-2
Installation Name	Ch-OpenFOAM-VA@INGV
Hosting Institution	INGV Pisa
Update	A new version (August 2023) of the OpenPDAC-11 software has been released on GitHub and Zenodo for open-source access and reference.
KPIs	No repo-tracker yet, but planned. Otherwise, standard KPIs (number of downloads, number of views, number of commits, bug fixed) cannot be tracked yet on the long period with the free version of GitHub. Zenodo software pages report 70 downloads (68 + 2), 439 views (377 + 62 + 9)
Links	
DOI	10.5281/zenodo.7701702 10.5281/zenodo.5031824 10.5281/zenodo.10639208
URL	https://zenodo.org/record/7701703 https://zenodo.org/record/5031825 https://zenodo.org/records/10721203
GitHub/GitLab/or other	https://github.com/demichie/OpenPDAC-10

Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call

	https://github.com/demichie/OpenPDAC-11
Training & Documentation	
Training Material	-
Documentation	10.5281/zenodo.5031824
Training or Workshop	-

Ch-HySEA (VA & TA)

Installation ID	VA4-532-1, TA2-532-1
Installation Name	T-HySEA@UMA
Hosting Institution	Universidad de Málaga (UMA)
URL	There is a DOI for version 1.1.0 on Zenodo 10.5281/zenodo.7812152
Update	The Tsunami-HySEA software has been made available on the open repository https://github.com/edanya-uma/TsunamiHySEA and releases from this repository are automatically archived under Zenodo with a DOI. The most recent release was June, 2023. The GitHub has been updated with both improved documentation and notes on compilation and use examples. In our recent VA access offer we will introduce a new version of the documentation and compilation for newer Linux systems.
KPIs	
Links	
DOI	https://doi.org/10.5281/zenodo.7812152
URL	https://edanya.uma.es/hysea/
GitHub/GitLab/or other	https://github.com/edanya-uma/TsunamiHySEA
Training & Documentation	
Training Material	The README.md file on the GitHub page is under constant renewal. A revision of existing documentation is ongoing.
Documentation	A revision of existing documentation is ongoing.
Training or Workshop	A VA training activity consisting on two parts: compiling T-HySEA and Intro to T-HySEA is being organized in collaboration with WP9. This activity will be held online in April 2024.

TS-BingClaw (VA & TA)

Installation ID	VA4-532-1, TA2-532-1
Installation Name	TS-BingClaw-VA@NGI
Hosting Institution	Norwegian Geotechnical Institute (NGI)
URL	There is a DOI for version 1.0.1 on Zenodo https://dx.doi.org/10.5281/zenodo.8354763
Update	The BingCLAW software has been made available on the open repository https://github.com/norwegian-geotechnical-institute/BingCLAW_5.6.1 and

	releases from this repository are automatically archived under Zenodo with a DOI. The most recent release was September 18, 2023. The GitHub has been updated with both improved documentation and notes on compilation and, in January 2024, an extra directory was added with a visualization example. The code for this example was modified following an in-person seminar and the updates are now uploaded to the repository. A docker container has been generated in which the process will run and documentation and official release of this containerized method of use is in preparation.
KPIs	The KPIs that were already adopted for VA4-532-1 (c.f. D5.2) are VA-01 (Number of installations and Service Usage Logging in place) and WP5-VA-06 (Number of consultation requests: total and per month). A logging system is in place for accesses both from GitHub and from Zenodo, although we have no way of measuring how many clones lead to actively used installations. A total of 3 direct consultation requests have been made and these resulted in a series of email conversations and an in-person meeting and demonstration at the NGI premises in Oslo. All of the requests came from young MSc candidates. The KPIs labelled for future adoption were VA-03 (Number of users served), WP5-VA-12 (Number of services running and reachable publicly), WP5-VA-13 (is the software packaged, containerized, and downloadable. Ease of deployment.), and WP5-VA-15 (number of publications citing software). We are now monitoring VA-03, see the short explanation below. Regarding VA-03, we have the number of access to GitHub and Zenodo as an indication. Regarding WP5-VA-13, the software provided on the GitHub/Zenodo can be deployed and compiled using only openly available third-party software; a containerized version is in preparation. Regarding WP-VA-15, one publication from 2024 cites the BingCLAW software (Walker et al., 2024, https://doi.org/10.1016/j.quascirev.2023.108433)
Links	
DOI	https://doi.org/10.5281/zenodo.8354763
URL	https://www.ngi.no/en/research-and-consulting/natural-hazards-container/tsunamis/model-for-simulating-dynamics-of-cohesive-landslides
GitHub/GitLab/or other	https://github.com/norwegian-geotechnical-institute/BingCLAW_5.6.1
Link to repo-tracker or webservice statistics	The following links provide a day-by-day breakdown of clones and views of the https://github.com/norwegian-geotechnical-institute/BingCLAW_5.6.1 since June 2023. https://github.com/norwegian-geotechnical-institute/BingCLAW_5.6.1/blob/github-repo-stats/norwegian-geotechnical-institute/BingCLAW_5.6.1/ghrs-data/views_clones_aggregate.csv https://github.com/norwegian-geotechnical-institute/BingCLAW_5.6.1/blob/github-repo-stats/norwegian-geotechnical-institute/BingCLAW_5.6.1/latest-report/report.pdf They show that between June 6, 2023, and February 27, 2024, there were 159 unique visitors to the site and a total of 415 clones.

	On Zenodo, https://doi.org/10.5281/zenodo.8354763, the native access tracker shows that as of 2024-02-27 there have been a total of 172 views and 24 downloads.
Training & Documentation	
Training Material	The README.md file on the GitHub page is under constant renewal and training can be provided in-person, via email, or online. A revision of existing documentation is ongoing.
Documentation	A comprehensive description of a use case (the simulation of the Storegga landslide) is provided in the following Open Access publication: https://doi.org/10.1029/2018JC014893 *(Jihwan Kim, Finn Løvholt, Dieter Issler, and Carl Fredrik Forsberg, Landslide Material Control on Tsunami Genesis—The Storegga Slide and Tsunami (8,100 Years BP), JGR Oceans, vol 124(6), June 2019, Pages 3607-3627)* Additional documentation is available on the GitHub link above.
Training or Workshop	An in-person workshop for compiling and running BingCLAW was held at NGI on February 22, 2024, for students from Oslo Metropolitan University (OsloMET) who had accessed the online resources. Similar trainings can be arranged online for remote users. Additional training for successful TA applicants will take place in-person during the stay and through electronic communication and online meetings before and afterwards.

TS-AmpFact (VA)

Installation ID	VA4-532-2
Installation Name	TS-AmpFact@NGI
Hosting Institution	Norwegian Geotechnical Institute (NGI)
URL	https://doi.org/10.5281/zenodo.10732728
Update	The amplification factors are now systematized into a single file covering all areas globally where amplification factors are computed for 12634 points along the 50 m depth contour. The new file is now found in the open repository norwegian-geotechnical-institute/global-tsunami-amplification-factors: Global tsunami amplification factors (github.com). Still some coastal areas and islands are lacking, these will be included in future updates.
KPIs	The KPIs that were already adopted for VA4-532-2 (c.f. D5.2) are VA-01 (Number of installations and Service Usage Logging in place). A logging system is in place for accesses and downloading both from GitHub and from Zenodo. The KPIs labelled for future adoption were VA-03 (Number of users served), ND-01 and ND-02 (New datasets), WP5-VA-06 (No of consultation requests), WP5-VA-09 (Number of commits on public repository), WP-VA-12 (Number of services running and reachable publicly), WP5-VA-13 (is the extended software for computing own amplification factors as well as tools for computing the MIH from timeseries), and WP-VA-15 (number of

	publications citing the data and software). We are now monitoring VA-03, see the short explanation below (see "webservice statistics").
Links	
DOI	https://doi.org/10.5281/zenodo.10732728
URL	https://www.ngi.no/en/research-and-consulting/natural-hazards-container/tsunamis/amplification-factors-for-tsunami-run-up-estimation/
GitHub/GitLab/or other	https://github.com/norwegian-geotechnical-institute/global-tsunami-amplification-factors
Link to repo-tracker or webservice statistics	An access counter was installed on the GitHub link in February 2024. It is too early to report on accesses registered by this mechanism. The first version of the TS-ampfactors was available on Zenodo on https://doi.org/10.5281/zenodo.7127777 and this repository had, as of March 1, 2024, 144 views and 38 downloads.
Training & Documentation	
Training Material	The README.md file on the GitHub page is under constant renewal.
Documentation	A revision of existing documentation is ongoing.
Training or Workshop	--

TS-Gauss (VA)

Installation ID	VA4-532-3, VA4-532-4
Installation Name	TS-Gauss
Hosting Institution	GFZ, INGV
URL	https://git.gfz-potsdam.de/babeyko/ts-gauss
Update	A git repository has been created, containing the code to perform fast tsunami simulations based on pre-computed Gaussian elementary sources (currently stored in a Google Drive repository). The git repository also provides a user guide for TS-Gauss.
KPIs	The service is just started, hence there are still no statistics
Links	
DOI	https://doi.org/10.5880/GFZ.2.5.2024.002
URL	https://git.gfz-potsdam.de/babeyko/ts-gauss
GitHub/GitLab/or other	https://git.gfz-potsdam.de/babeyko/ts-gauss
Training & Documentation	
Training Material	In preparation
Documentation	https://git.gfz-potsdam.de/babeyko/ts-gauss/-/blob/main/TS-Gauss.Readme.pdf?ref_type=heads
Training or Workshop	-

Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call

TS-Slip@UNINA (VA)

Installation ID	VA4-532-6
Installation Name	TS-Slip@UNINA
Hosting Institution	UNINA
Update	During the last year of project from M6, the software has been used by some research groups for scientific activities whose results are going to be soon published on scientific journals. These researchers provided some feedbacks either through personal communications to the developers or posting issues on the GitHub platform. Following these feedbacks, the provided service has been improved focusing on the following aspects: 1 Some bugs, concerning the meshing module, the rigidity variations with depth, the GEO-UTM coordinates conversion, have been either fully or partially fixed. The documentation has been updated accordingly. 2 The output slip distributions are now provided also in a format compatible with the input format of HySea, a commonly used tsunami simulator which is provided as a service (both as VA and TA) within G-I 3 The Tohoku and Maule earthquake guided examples have been modified accordingly to the code improvements. Moreover, a fully guided example for the computation of an ensemble for PTHA has been added to the released documentation. Such example has been developed using the Makran subduction zone. 4 Several pre-computed mesh discretizations for 50 slab interfaces around the world have been now made available by the service. The geometrical features of these interfaces have been obtained from Slab1.0 and Slab2.0 projects outcomes (https://www.usgs.gov/tools/slab20-interactive-map). The pre-computed meshes and the other input files are stored both through GitHub and the zenodo service, as now described in the updated documentation. Beyond that a Wiki documentation is currently under construction and will be completed by M24 (October 2024)
KPIs	We measure VA-01 (100%). We measure VA-02 and VA-03 in two ways. 1 Through the Insights/Traffic function on GitHub the number of clones and unique clones can be used as a proxy of no. of SW downloads (VA-02) and no. of users served (VA-03). However, this function only allows to view the clones related to the last two weeks and does not contain information about users directly downloading zip file of the repository. As example we report 40 clones and 20 unique cloners in the two weeks from 17th February to 1st March. 2 Through the zenodo service, the total number of download (for all the released version) can be monitored. Until the 2024, March the 1st, we measured 22 total downloads through this service For what concerns specific KPIs for SaaS:

	• WP5-VA-06: can be monitored through personalized user support provided, the views and the "unique" visitors available at the Traffic/Insights function of GitHub (only last 2 weeks available, e.g. 421 views and 9 unique visitors in the period from 17th February to 1st March) • WP5-VA-09: Monitored on the GitHub repository (Traffic/Commits on GitHub insights for the last year and through RepoTracker for the last month, see attached screenshots) • WP5-VA-10: Monitored on the GitHub repository (on the Repo-Tracker at the moment currently available for the last month, see attached screenshots) • WP5-VA-12: The software is a single service. A database of precomputed input files is now available on zenodo and downloads can be monitored, as already described for the software • WP5-VA-13: Software not packaged • WP5-VA-15: Documentation contains a reference to cite the SaaS in publications. However, it is quite difficult to monitor.
Links	
DOI	https://doi.org/10.5281/zenodo.7101459
URL	https://zenodo.org/doi/10.5281/zenodo.7101459
GitHub/GitLab/or other	https://github.com/antonioscalaunina/ANTI-FASc
Link to repo-tracker or webservice statistics	https://repo-tracker.com/r/gh/antonioscalaunina/ANTI-FASc
Training & Documentation	
Training Material	We have planned to prepare a brief video-course (between 2 and 4hrs) to show the functioning of the code and the execution of the guided test-case examples. The link to the video-course will be included in the documentation. This course will be presumably available between M24 and M30
Documentation	A brief documentation is available at the GitHub repository webpage along with the guided examples: https://github.com/antonioscalaunina/ANTI-FASc A more detailed wiki documentation (currently under construction) is available at the following link: https://github.com/antonioscalaunina/ANTI-FASc/wiki It will be presumably completed by October 2024 (M24)
Training or Workshop	Video-course material (see above)

TS-Slip@IFREMER (VA)

Installation ID	VA4-532-5
Installation Name	Slip (k223d)
Hosting Institution	Ifremer
Update	During last year of the project, work has gone on offline on a major update to the programme which will see the inclusion of rupture velocity in the

	stochastic slip model, while this work is outside of the Geo-INQUIRE project it impacts on the development of version 1 of the programme which is currently available. This new version will be released before M24 with a publication submitted also. This work over the last year has focused on: The input format for the programme has been update to take a standard vtk format file which will include information about the rupture velocity on the fault plane, the calculation of the rupture time on the fault plane based on a user defined nucleation location and the rupture velocity. The output will be the slip distribution and the rupture time across the fault in a file with a vtk format. Documentation is currently being updated to coincide with this update. Tutorial Jupyter Notebooks are currently being prepared to show how k223d can be used to generate slip distributions on realistic faults. To do this, the notebooks will show how to take data from open source fault databases (e.g. Seismofaults.eu, Slab 2.0), import them into jupyter environment, mesh them and then generate slip distributions using the k223d.These notebooks and an updated user guide will be realised by M24 of the project.
KPIs	VA-01 (100%) is measured VA-02 and VA-03 is measured in two ways. 1 In GitHub, the Insights/Traffic function providing information on number of downloads over the last 2 weeks. For the period from the 21st February and 5th March there have been 1 clone and 1 unique users. 2 On Zenodo there have been no downloads since 11th January 2023 when the tracking was first started. For what concerns specific KPIs for SaaS: • WP5-VA-06: on GitHub there have been 23 views and 5 unique visitors between 21st February and 5th March. • WP5-VA-09: Monitored on the GitHub repository (Traffic/Commits on GitHub insights for the last year and through RepoTracker since 30/01/24). As a major update is being planned, there have been no commits on this version of the code in • WP5-VA-10: Monitored on the GitHub repository (on the Repo-Tracker since 30/01/24). Again, as a major update will be released soon, no bug reports/fixes have been provided. • WP5-VA-12: The software is a single service. Documentation will be updated by M24 • WP5-VA-13: Not containerised but is planned in the future • WP5-VA-15: GitHub readme contains a reference to publication but this has not been
Links	
DOI	https://doi.org/10.5281/zenodo.7525449
URL	(see below)
GitHub/GitLab/or other	https://github.com/s-murfy/k223d

Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call

Link to repo-tracker or webservice statistics	https://repo-tracker.com/r/gh/s-murfy/k223d
Training & Documentation	
Training Material	Brief description of code is currently provided on GitHub. This will be updated soon to give a more extensive information on the programme.
Documentation	https://github.com/s-murfy/k223d
Training or Workshop	Autumn 2024: Shane Murphy (Ifremer), Andre Herrero (INGV) **Stochastic slip training course** *Online Training Course*

VO-CIV (VA)

Installation ID	VA4-533-1
Installation Name	VO-CIV@INGV
Hosting Institution	INGV Pisa
Description	The Volcano Dynamics Computational Centre (VDCC - www.pi.ingv.it/progetti/eurovolc/) is an infrastructure providing Virtual Access to volcanological software through an interactive web framework. It allows the user to run volcanological applications online, but it is not easy to maintain and to expand because it requires a significant workload to implement new software in the web framework. The EPOS Computational Infrastructure for Volcanology (CIV - civ.pi.ingv.it) is a software catalogue made available by EPOS. The service offers a description and some documentation about volcanological software, and a catalogue of metadata accessible through the EPOS portal, but does not offer the possibility for online execution of the applications. The Geo-INQUIRE VO-CIV@INGV service being implemented aims to combine (and to overcome the limits of) the EUROVOLC VDCC VA infrastructure and the EPOS CIV, to provide APIs to make volcanological software portable and accessible as a Web Processing Service (WPS). The VO-CIV infrastructure is designed to ease the implementation of Pilot Demonstrator Applications during the Geo-INQUIRE project, in a sustainable perspective.
Update	To include VA computational capabilities to the volcanological model catalog EPOS-CIV, a new web processing service (WPS) has been designed and implemented from scratch. The WPS is able to ingest input with a well-defined syntax, execute code on a dedicated server and return results asynchronously. The web application, based on REST architectural style, is designed to support interoperability between different computers. The service requires simple interactions between the server hosting the service and the client requesting it; no specific installations are required by the developers of the scientific applications. The server does not need to maintain the status of

Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call

	the communication; every request is independent and contains all the necessary information to provide the result. Therefore, the server requires less resources, it is faster, simpler and more scalable. The web service has been implemented in Flask, a micro web framework written in Python. The web service is currently in a testing phase. Prototype Pilot Demonstrator Applications (PDAs) have been developed and tested within the web framework. During the current reporting period, three PDAs have been integrated in VO-CIV: • SOLWCAD – Fortran code to compute the saturation surface of H2O-CO2 fluids in silicate melts of arbitrary composition (Papale et al., Chem. Geol. doi:10.1016/j.chemgeo.2006.01.013, 2006). • PLUME-MoM-TSM - a FORTRAN90 code designed to solve the equations for a steady-state integral volcanic plume model, describing the rise in the atmosphere of a mixture of gas and volcanic ash during an eruption (de' Michieli Vitturi et al., Geosci. Model. Dev. doi:10.5194/gmd-14-1345-2021, 2021). • PyBetVH - Long-term volcanic hazard assessment using the Bayesian Event Tree for Volcanic Hazard (BET_VH) model (Tonini et al., Comput. Geosci. doi:10.1016/j.cageo.2015.02.017, 2015). During the next six months the VO-CIV will be exposed on the web. Later development will include developing user-friendly interfaces. Specific layers of code, separate from the service and targeted to different needs (e.g. mobile apps) will be developed to this aim.
KPIs	No repo-tracker as the codes are executed online. VDCC service has a tracking system based on Google Analytics, featuring 264 downloads and 136 online runs for the current reporting period. EPOS-CIV does not have a logging system, but it is being implemented in the more general EPOS portal. A dedicated logging system will be developed for the new VO-CIV service. Standard KPIs (number of downloads, number of service requests, number of views, number of email requests) will be tracked.
Links	
DOI	n/a
URL	https://www.pi.ingv.it/progetti/eurovolc/ https://civ.pi.ingv.it/
GitHub/GitLab/or other	
Training & Documentation	
Training Material	Solwcad: https://www.improve-etn.eu/index.php/mixture-thermodynamics/
Documentation	(see above)
Training or Workshop	none

Pyrocko (VA)

Installation ID	VA4-534-1
Installation Name	Pyrocko, seismosizer-webengine
Hosting Institution	GFZ
Update	The architecture and software design have been discussed, i.e. which libraries and packages are to be used for the implementation. A repository has been setup and package are installable in >=python3.8. The serialisation of objects for remote communication has been implemented.
KPIs	Since the package is not finished yet nothing can be tracked, but will be in the future.
Links	
DOI	-
URL	-
GitHub/GitLab/or other	https://github.com/pyrocko/pyrocko https://github.com/hvasbath/seismosizer-webengine
Training & Documentation	
Training Material	Needs to be created once the package is running. This will be a website and a software hosted on https://pyrocko.org.
Documentation	-
Training or Workshop	-

T5.4 Workflows as a Service

Ch-PVHA (TA)

Installation ID	TA2-541-1/ TA2-541-2
Installation Name	PVHA Probabilistic Tsunami Hazard Assessment TA
Hosting Institution	INGV – CSIC
Update	The announcement has been finalized (see URL below) and has been advertised actively on international volcanology-related contexts such as at Cities On Volcanoes 12 conference (just held in Antigua Guatemala) and with ALVO (Latin-American Association of Volcanology)
KPIs	The KPIs applicable to the TA have been described in D5.2. The only one applicable prior to the TA visit is TA-03, i.e., the number of calls – currently the TA hosts have received exploratory questions from 4 potential applicants.
Links	
DOI	https://zenodo.org/records/6375497#.ZCPzZ-xBwlU
URL	https://www.geo-inquire.eu/transnational-access-offer/ta2-541-1/ta2-541-2
Training & Documentation	

Training Material	Documentation will be provided prior to and during the visit
Documentation	A comprehensive overview of the procedure and application to the Campi Flegrei caldera (Ily) is provided in the Open Access publication Martinez-Montesinos et al (2022) https://www.frontiersin.org/articles/10.3389/feart.2022.941789/full (doi: 10.3389/feart.2022.941789, title "On the feasibility and usefulness of high performance computing in probabilistic volcanic hazard assessment: An application to tephra hazard from Campi Flegrei")
Training or Workshop	Training will take place in-person during the stay and through online meetings before and afterwards

Ch-PTHA (TA)

Installation ID	TA2-541-6
Installation Name	PTHA Probabilistic Tsunami Hazard Assessment TA
Hosting Institution	Norwegian Geotechnical Institute (NGI)
Update	The announcement has been finalized (see URL below) and has been advertised actively on outlets such as LinkedIn https://www.linkedin.com/posts/steven-j-gibbons-9319017_transnationalaccess-geoscience-research-activity-7156201154337464320-x3GK?utm_source=share&utm_medium=member_desktop
KPIs	The KPIs applicable to the TA are described in D5.2. The only one applicable prior to the TA visit is TA-03: the number of calls – currently at one for the present call.
Links	
DOI	https://doi.org/10.5281/zenodo.6376607
URL	https://www.geo-inquire.eu/transnational-access-offer/ta2-541-6/ta2-541-7/ta2-541-8
Training & Documentation	
Training Material	Documentation from the private GitLab will be provided prior to and during the visit.
Documentation	A comprehensive overview of the procedure and application to the coast of Eastern Sicily is provided in the Open Access publication https://doi.org/10.3389/feart.2020.591549 ("Probabilistic Tsunami Hazard Analysis: High Performance Computing for Massive Scale Inundation Simulations")
Training or Workshop	Training will take place in-person during the stay and through electronic communication and online meetings before and afterwards.

Ch-Cyber-PSHA (TA)

Installation ID	TA2-541-9, TA2-541-10
Installation Name	Cyber-PSHA - Probabilistic Seismic Hazard Analysis
Hosting Institution	LMU, BSC
Update	First CyberShake applications out of the state of California have been carried out in Southern Iceland (SI) by the BSC scientific group participating in Geo-I. Results from these applications have been summarized in (Rojas et al. 2021; Rojas et al. 2022). In the framework of Geo-I, the BSC scientific group have been progressively improving the execution workflow, originally developed in these former applications, to consider general ERF models than those proposed in SI, where most seismic events are strike slip. A friendly material training has been lately highly progressed to be used as the main educational source before and during the incoming TA's.
KPIs	All TA KPIs will be measured
Links	
DOI	--
URL	https://www.geo-inquire.eu/transnational-access-offer/ta2-541-9 https://www.geo-inquire.eu/transnational-access-offer/ta2-541-10 https://www.scec.org/software/cybershake
Training & Documentation	
Training Material	One-on-one training will be provided during the stay and through online meetings before and afterwards.
Documentation	A user-friendly material training has been lately highly progressed to be used as the main educational source before and during the incoming TA's.
Training or Workshop	-

Ch-MP-PSHA (TA)

Installation ID	TA2-541-11
Installation Name	Ch-MP-PSHA@LMU
Hosting Institution	LMU
Update	SeisSol has been used in several international collaborations towards physics-based seismic hazard assessment which contributed to 11 publications, including in Science and Nature. Specifically, in Li et al., JGR, 2023, we model dynamic rupture, fault interaction, and ground motion for the segmented Húsavík-Flatey Fault Zone, in Northern Iceland. The rupture scenarios obey earthquake scaling relations and predict magnitudes comparable to those of historical events. We show how fault system geometry and segmentation, hypocenter location, and prestress can affect the potential for rupture cascading. The average ground motion attenuation characteristics of dynamic rupture scenarios of comparable magnitudes and mean stress drop are independent

Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call

	of variations in source complexity, magnitude-consistent and in good agreement with the latest regional empirical ground motion models. Our study suggests that an ensemble of physics-based dynamic rupture scenarios can complement classical seismic hazard assessment methods to better characterize the hazard in tectonically and seismically complex regions, especially in data-scarce regions.
KPIs	All TA KPIs will be measured.
Links	
DOI	-
URL	https://www.geo-inquire.eu/transnational-access-offer/ta2-541-11
Training & Documentation	
Training Material	One-on-one training will be provided during the stay and through online meetings before and afterwards.
Documentation	-
Training or Workshop	-

EF-PSHA (TA)

Installation ID	TA2-542-1
Installation Name	EF-PSHA – Probabilistic Seismic Hazard Analysis
Hosting Institution	Eidgenössische Technische Hochschule Zürich (ETHZ), Switzerland
Description	The European Facilities for Earthquake Hazard and Risk (EFEHR) infrastructure provides access to seismic hazard and risk data and products for Euro-Mediterranean region. EFEHR WaaS (Workflow-as-a-Service) infrastructure extends to transnational access with focus on the e-infrastructure as well as the physical infrastructure needed to run hazard and risk computations. Additionally, EFEHR WaaS supports and enables sensitivity analyses, uncertainty quantifications, and comparisons with physics-based PSHA models or alternative hazard models. EF-PSHA relies on sophisticated and computationally demanding workflows using a range of quality-controlled inputs and data products combined into a probabilistic framework. It can deliver time-dependent and time-independent earthquake rate forecasts, as well as access to various seismic hazard models.
Update	The TA call has been opened (until 15th of March).
KPIs	TA KPIs will be measured.
Links	
DOI	n/a
URL	https://www.geo-inquire.eu/transnational-access-offer/ta2-542-1 http://hazard.efehr.org
Training & Documentation	
Training Material	Training materials will be provided on-site, tailored to the specific requirements of the visitor.

Documentation	The EFEHR web platform (http://hazard.efehr.org) provides access to documentation, input computational files and results for the following seismic hazard models: • 2013 European Seismic Hazard Model (https://doi.org/10.1007/s10518-015-9795-1) • 2014 Earthquake Model of the Middle East (https://doi.org/10.1007/s10518-018-0347-3) • 2020 European Seismic Hazard Model (https://doi.org/10.12686/a15) Guidance and support on how to tailor to fit specific model components for complex computations is also part of this call. EFEHR recommends the use of OpenQuake (OQ) open-source engine (https://github.com/gem/oq-engine) and it provides access to various computational servers.
Training or Workshop	At present, there is no scheduled training or workshop. The availability of such sessions will be determined based on the demand for TA.

Appendix 2: KPIs as defined in D5.2

KPIs for Service Status

KPI ID	KPI Group	KPI definition	Collection Strategy	Comment
SSRI-01	Service Status and RI Integration	%/number of Installations with full RI integration	Evaluation of each installation by WP6 in interaction with each installation.	Done by WP6 through implementation level matrix
SSRI-02	Service Status and RI Integration	%/number of Installations fully operating	Evaluation of each installation by WP6 in interaction with each installation.	Done by WP6 through implementation level matrix
SSRI-03	Service Status and RI Integration	%/number of Installations under implementation operating	Evaluation by WP6 or input from installations.	Done by WP6 through implementation level matrix
SSRI-04	Service Status and RI Integration	Number of services running and reachable publicly	Spreadsheet	Done by WP6 through implementation level matrix

VA KPIs

KPI ID	KPI Group	KPI definition	Collection Strategy	Comment
VA-01	Virtual Access	Number (and/or % of installations) of Service Usage Logging systems in place	Question (to be refined) to each installation: *Does your installation have a service usage logging in place?*	The questionnaire indicated that almost all operating systems have such a system in place. It is a prerequisite to answer Link to ESFRI KPIs1,2,17,19

Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call

VA-02	Virtual Access	Number data or metadata requests	Question (to be refined) to each installation: *If you have a service usage logging in place: How many data and metadata requests do you process each year (spams and robots excluded)?*	Note that the input from the installations is somewhat heterogeneous and a homogenization is not fully feasible during the Geo-INQUIRE lifespan. The heterogeneity must be taken into account in processing of the Installation responses. Link to ESFRI KPIs 1 and 2.
VA-03	Virtual Access	Number of users served	Question (to be refined) to each installation: *If you have a service usage logging in place: How many different users does your installation serve each year?*	Note that the input from the installations is very heterogeneous (visits to webpages, all data and metadata download, only data download), a homogenization is not fully feasible during the Geo-INQUIRE lifespan. The full usability of this KPI is still to be evaluated. Link to ESFRI KPIs 1 and 2.
VA-04	Virtual Access	Number of systems capable of geographically locate users.	Question (to be refined) to each installation: *If you have a service usage logging in place: Do you track the country of origin of data or metadata requests?*	Link to ESFRI 9

TA KPIs

KPI ID	KPI Group	KPI definition	Collection Strategy	Comment
TA-01	TA-offer	Number of sites opening calls		A site may concern a testbed, a lab, etc.
TA-02	TA-offer	Aggregated number of unit access		Standard unit of access 1h?
TA-03	TA-offer	Total number of calls		The number of openly published calls per site.
TA-04	TA-support	Amount of time spent on on-site support		Includes training
TA-05	TA-support	Amount of time spent on managing calls		Includes assessment procedures?
TA-06	TA-calls	Number of applicants	General information collected to the TARP decisions	
TA-07	TA-calls	Number of granted applications	Based on TARP decisions	
TA-08	TA-calls	Number of finished programs		
TA-09	TA-calls	Number of countries involved		Spread of applicants over countries.
TA-10	TA-calls	Number of organizations involved		Spread of applicants over organizations.

Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call

TA-11	TA-output	Number of derived data publications		Indicate how much data will be openly shared and how. I.e., are there standard procedures and/or pipelines available for this?
TA-12	TA-output	Number of paper publications		

Additional WP5-VA KPIs

KPI ID	KPI Group	KPI definition	Comment
WP5-VA-06	VA service to Software	No of consultation requests: total and per month	Personalized user support as reported by a service provider
WP5-VA-09	VA service to Software	Number of commits on public repository (e.g. GitHub)	
WP5-VA-10	VA service to Software	Number of bug reports/fixed	e.g. from GitHub issues
WP5-VA-12	VA service to Software	Number of services running and reachable publicly	Stats from various tools and sources.
WP5-VA-13	VA service to Software	Is the software packaged, containerized and downloadable? Ease of deployment.	Through channels like PyPI, Conda or HPC packages (forgot the name)
WP5-VA-15	VA service to Software	No of publications citing software	It might be difficult to monitor

Additional WP5-TA KPIs

KPI ID	KPI Group	KPI definition	Comment
WP5-TA-02	TA service to Software	Amount of computer time (e.g., node-hours) provided by the service	Time used and time granted may differ

Geo-INQUIRE is funded by the European Commission under project number 101058518 within the HORIZON-INFRA-2021-SERV-01 call