

Recommended citation:

Basili, R., Danciu, L. (2023): Geo-INQUIRE Horizon Europe project. D3.2 WP3 Research infrastructures coordination progress report with definition of KPIs, Potsdam : GFZ German Research Centre for Geosciences, 17 p., Report

DOI: <https://doi.org/10.48440/efvn-4y16>



This work is licensed under a Creative Commons Attribution 4.0 International License. (CC BY 4.0) <https://creativecommons.org/licenses/by/4.0/>



# Geosphere Infrastructures for Questions into Integrated Research



## **D3.2 WP3 Research infrastructures coordination progress report with definition of KPIs**



## Reference information

<b>Deliverable No.</b>	<b>D3.2</b>
<b>Deliverable Name</b>	WP3 Research infrastructures coordination progress report with definition of KPIs
<b>Author(s) (Affiliation)</b>	Roberto Basili (INGV), Laurentiu Danciu (ETHZ)
<b>Lead Beneficiary</b>	INGV
<b>Due Date (month)</b>	M12
<b>Type</b>	R – Document, report
<b>Dissemination level</b>	PU – Public
<b>Reviewer(s)</b>	Helen Crowley (EUCENTRE/GEM), Mateus L Prestes (GFZ)

## Versions

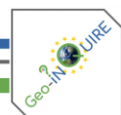
<b>Version No</b>	<b>Date</b>	<b>Changes / Comments / Status</b>
<b>0.1</b>	18.09.2023	Version shared within WP3
<b>0.2</b>	19.09.2023	Version shared with WP7
<b>1.0</b>	29.09.2023	Final version submitted to the EU

*Disclaimer: Geo-INQUIRE is funded by the European Union. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.*



## Table of contents

1	Introduction .....	1
2	Coordination progress report .....	1
2.1	Context and work method to define KPIs .....	3
3	Present status of Installations in WP3 for the portfolio of indicators and their view toward adoption .....	3
3.1	Status of WP3 concerning KPIs.....	3
3.2	WP3 view towards adoption and usage of indicators .....	4
4	Harmonization and uptake strategy of KPIs among infrastructures .....	9
4.1	Service and data-related KPIs and harmonization .....	9
4.1.1	Definition of 'users' and their counting mechanisms .....	9
4.1.2	Tracking country of origin of data or metadata requests .....	10
4.1.3	Definition and counting of datasets.....	10
4.1.4	Harmonization of service and data related KPIs .....	10
4.2	Harmonization of dissemination KPIs .....	10
5	Conclusion.....	11
6	Appendices.....	12
6.1	Appendix 1: Portfolio of KPIs identified by WP7 .....	12
6.1.1	Service status .....	12
6.1.2	Virtual Access (VA) usage.....	12
6.1.3	Accessible datasets .....	13
6.1.4	New datasets.....	13
6.1.5	Transnational Access (TA) usage.....	14
6.1.6	Dissemination.....	14
6.2	Appendix 2: FAIR metrics used to attribute the FAIR score.....	15
6.3	Appendix 3: ESFRI Key Performance Indicators (KPIs) .....	16



## 1 Introduction

Enhancing the sharing and reuse of all data and data products is one of Geo-INQUIRE's strategic priorities, requiring adopting and monitoring the FAIR principles. A suitable set of predefined Key Performance Indicators (KPIs) was proposed to measure the effectiveness of how this goal is pursued. The present deliverable is the first in a set of standardized deliverables across WP2-5 related to the coordination progress of installations concerning such KPIs. The next steps in this respect will be part of deliverables D3.4 (M24), D3.7 (M36), and D3.9 (M48), which will include access and KPI monitoring, as well as service review summaries.

This deliverable includes three main parts (Sections 2-4). In Section 2, the deliverable summarizes the coordination strategy aligned with the transverse work packages WP6, WP7, and WP9. Section 3 describes the status of WP3 installations and their capability toward the adoption of the proposed KPIs. Section 4 presents the long-term view, prepared by WP6 and WP7, on how the European Research Infrastructures can take up the strategy developed in Geo-INQUIRE in a homogenized, robust, and sustainable way. In addition, Section 5 remarks on some critical aspects of adopting the KPIs, and Section 6 is made by a set of Annexes presenting the portfolio of KPIs identified by WP7, the FAIR metrics, and ESFRI KPIs in tabular form with brief descriptions.

The main part of the information behind this deliverable comes from a project-wide survey across all Geo-INQUIRE installations, prepared and carried out in the first six months of the project, combined with a simplified service status assessment carried out by WP6 (preliminary work for D6.1, also delivered at M12). The survey, called the Geo-INQUIRE 1st Installation Survey, was presented in D3.1 (M6), providing an overview of results across all Geo-INQUIRE and WP3 installations.

The present deliverable views the survey specifically in the light of setting a project and a long-term strategy for monitoring progress through well-adapted KPIs, i.e., including a bottom-up approach to feed the KPI strategy with the daily reality of the service operation. It is also based on numerous internal interactions within WP7, which guided the direction taken for deliverables D2.2, D3.2, D4.2, and D5.2. The approach suggested by WP7 was further developed by the Geo-INQUIRE Project Management Board, WP2, WP6, and WP9, and by the Project Management Office.

## 2 Coordination progress report

Work Package 3 (WP3) handles the largest number of project installations, i.e., 72, and directly engages more than 62 experts, including scientists, managers, and developers. Among these installations, there are 49 level-2/3 installations and 23 level-1 installations. All services provided by these installations are intended for virtual access. WP3 is organized around five theme tasks (3.2-3.6), each focusing on a distinct component of service provision. Overarching these is the coordination task (3.1), which is intended to facilitate inter-task coordination as well as liaison



with transversal WPs (6,7,9) and TA (WP8), as well as to stimulate international discipline cooperation.

A coordination group has been established to facilitate the activities of task 3.1. The WP3 lead, co-lead, and contributors from Tasks 3.2-3.6 should form this group. The coordination strategy will be broken down per portfolio, as defined in Tasks 3.2 to 3.6. Task leaders will be responsible for both acting on and monitoring the implementation within their respective portfolios.

In alignment with **Strategic Priority 1**, "Ensure sustainability through a solid integration into ERICs," WP3 has identified 22 installations as 'ready-to-go' services. These include installations spanning various thematic tasks, such as VA2-32-2, VA2-32-3, and VA2-32-9 from Task 3.2, and VA2-36-3 and VA2-36-4 from Task 3.6.

Expanding on **Strategic Priority 2**, "Enhance FAIRness of all data and data products," WP3 is at a crossroads. While its services have grown and been efficiently incorporated into EPOS ERIC, the overarching task of unifying and standardizing metadata and assuring FAIR data remains a significant challenge, particularly for services not part of EPOS yet. Through standardized metadata and improved data management, a task force aims to streamline services and ensure their long-term sustainability and impact, thus effectively fulfilling the project's **Strategic Priorities (1 to 3)**.

An inventory of existing datasets will be performed to determine their current spatial and temporal resolutions. A starting point is the survey conducted in M6 of the project and summarized in Section 3 of this deliverable, as well as in Deliverable D6.1 and Deliverable D6.2.

A major initiative involves the enhancement of datasets to achieve higher spatial and temporal resolutions, opening the door for curiosity-driven research. This aligns closely with Priority 3: "Opening new research opportunities to facilitate an improved understanding of the interface between the solid earth and its fluid envelope". And it addresses, in particular, the Tasks 3.2 and 3.3 portfolios. The cross-coordination between Tasks 3.3, 3.4, 3.5, and 3.6 plays a crucial role in shifting from single-hazard analysis to a multi-risk approach, aligning with Priority 4's objective of supporting integrated studies of extreme geohazards. Task 3.3 lays the foundation by creating a comprehensive geohazard database, which is then elevated by Task 3.4 and 3.5 to include multi-risk modelling by incorporating datasets from other hazard and risk domains. Task 3.6 acts as a feedback mechanism, providing the services for performing testing of multi-hazard and multi-risk datasets and/or models. The installations under each portfolio will be updated to a continuous improvement cycle aligned with the KPI defined in WP7.

WP3 will plan service-specific training and emphasize inclusivity and diversity within its framework. It will collaborate with WP9 to unify and streamline training programs to address this issue. Furthermore, cross-WP liaison has been established to facilitate communication and coordinate with WP6, WP7, and TA (WP8).



## 2.1 Context and work method to define KPIs

Harmonizing KPIs across the Geo-INQUIRE serves multiple objectives. Throughout the project's lifecycle, Geo-INQUIRE aims to monitor and report on the advancements and RI (Research Infrastructure) integration of services (implemented by the Geo-INQUIRE 'Installations'). In parallel, Geo-INQUIRE seeks to assess its milestones and progress on data integration and access, service usage, and training. This leads to the need for a definition at M12 of KPIs to be used for project purposes. For the long term, KPIs must be harmonized inside and across the participating Research infrastructures (Deliverables D6.1, D6.2, D7.1).

To fulfil both the project needs and the needs for RI harmonization, the PMB decided to define a limited set of KPIs, which goes beyond those indicated in the DoA, and use the experience gained to nourish the long-term RI KPI harmonization. For long-term objectives, Geo-INQUIRE needs to consider the ESFRI KPIs (See Annex 3), which, as a consequence, impact the short-term KPIs.

The present report uses the distinction between **KPI**, which means the indicator, while **KPI Value** is the reported value of the indicator at a given time (for example, '*Number of Users*' is a KPI while the KPI Value is the actual number associated, for example, 1000-2000).

The PMB identified 7 groups of KPIs and identified a coordinating group for each of them.

- Service status and RI Integration (WP6, in collaboration with WP2-5)
- VA usage (WP7, in collaboration with WP2-5)
- Accessible datasets (WP7, in collaboration with WP2-5)
- New datasets (WP7, in collaboration with WP2-5)
- FAIR metrics (WP7, in collaboration with WP2-5)
- TA Usage (WP7, in collaboration with WP5 and WP8)
- Dissemination (WP9, in collaboration with WP2-5 and WP8)

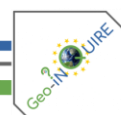
The KPIs will be based on (mainly) raw data from each Geo-INQUIRE installation, data each WP will process depending on specific monitoring and reporting needs, and by Geo-INQUIRE as a whole. Therefore, the KPIs were associated with a method of collecting the data, which is feasible from a practical point of view.

The initial list of KPIs, validated by the PMB, is found in Annex 1.

## 3 Present status of Installations in WP3 for the portfolio of indicators and their view toward adoption

### 3.1 Status of WP3 concerning KPIs

According to deliverable D3.1, WP3 includes 49 level-2/3 installations and 23 level-1 installations, for a grand total of 71. All services distributed by these installations are meant for Virtual Access. In the early phase of the project (February-March 2023), a survey was carried out to capture the



status of each installation through a questionnaire made of six mandatory sections and one optional section. Large parts of the mandatory sections have already been reported in D3.1. Here, we will re-examine several elements of the mandatory and optional parts, focusing on their relevance toward adopting the KPIs.

All 49 level-2/3 installations (100% coverage) responded to the mandatory section of the Questionnaire. The level-1 installations that responded to the Questionnaire are 14 out of 22 (60%). Overall, we could collect a total of 63 responses to the Questionnaire (**Table 1**). The operational status of these 63 installations is summarized in **Table 2**.

The optional part of the Questionnaire was filled in by 17 installations, of which 10 are level 2/3. From now onward, we will report on the status and progress only about the level-2/3 installations because they have full coverage of the mandatory part of the Questionnaire and a significant proportion of the optional part to help devise strategies and plans toward the adoption of KPIs for WP3 within the Geo-Inquire conditions and timeframe.

*Table 1. Number of installations per WP3 task.*

WP3 Task	Level 2-3 Installations	Level 1 Installations
3.2: Harmonized GNSS and Satellite Products for Geohazard Analysis	9	2
3.3: European Catalogues for Geohazard Analysis	17	2
3.4: European Geohazard Products	10	10
3.5: European Multi-risk Products	9	0
3.6: European Geohazard and Risk Testing	4	0
Total	49	14

*Table 2. Operational status of WP3 installations*

Service availability for all installations	Number All	Percent All	Number Level 2/3	Percent Level 2/3
Yes	31	49%	21	43%
Partly	7	11%	6	12%
No	25	40%	22	45%
Total number of responses	63	100%	49	100%

## 3.2 WP3 view towards adoption and usage of indicators

This section reports on the WP3 installation status and plans regarding the portfolio of KPIs identified by WP7 listed in the Annexes (#6 at the end of this document).

### KPIs for Service status and RI Integration (table 6.1.1 Service status, Annex 1)

All WP3 installations belong to the EPOS research infrastructure, except for VA2-33-17 (level 2/3) and VA2-33-20 (level 1), which are currently unavailable and not yet associated with any RI (the





project ARISE funded their development). The metadata standard adopted by most installations is the EPOS-DCAT-AP model. Adoption of KPIs and monitoring of the installation performances are strongly linked to the EPOS strategies already in place.

With all WP32 installations associated with EPOS, the operational status summarized in **Table 2** also reflects a preliminary assessment of their integration with the RI concerning KPIs SSRI-01, SSRI-02, and SSRI-03. Concerning SSRI-04, the survey has not explored the number of services for each installation. WP6 will provide more details on all these KPIs.

### KPIs for VA usage (table 6.1.2 Virtual Access usage, Annex 1)

There are four KPIs for Virtual Access (VA) usage (VA-01 to VA-04). The survey provides some basic information on the capability of the installations to address the KPIs and possibly monitor the KPIs' values. The information about the following KPI IDs was collected using answers to the Questionnaire.

As regards the VA-01, concerning the number (or %) of installations with a logging system in place, the survey is relevant for the 27 operating or partly operating installations. **Table 3** reports the percentage of these installations compared to the total Geo-INQUIRE installations. In WP3, 8 respondents (VA2-33-1, VA2-33-11, VA2-33-7, VA2-33-8, VA2-34-1, VA2-34-7, VA2-34-8, VA2-35-2) have a service usage logging in place. Of these installations, one is classified as partly operating (VA2-35-2). The AAI system regards two operating services, VA2-32-1, GNSS product Gateway, and VA2-32-7, EPOS-GNSS displacements for geohazard and anthropogenic events, for which authentication is mandatory for downloading data.

*Table 3. Percentage of Geo-INQUIRE and level-2/3 WP3 installations with operating or partly operating services with an operating system for usage monitoring and AAI.*

Service usage information	Geo-INQUIRE: 71 installations with operating or partly operating services	WP3: 27 installations with operating (21) or partly operating (6) services, Level 2/3 only
Usage monitoring operating	45%	30%
AAI system in place	20%	7%

VA-02 (Number of data and metadata requests), VA-03 (number of users served), and VA-04 (number of systems capable of geographically locating users) are dependent on a service usage logging system and the compilation of the optional part of the Questionnaire. Since such systems are presently adopted by a minority of installations (30%), and on half of those (10) that responded to the optional part, the values presently obtained are to be considered as rough estimates (**Table 4, 5, 6**).

*Table 4. Heterogeneity of how the presently operating service usage systems collect request numbers. The content of the table is based on survey responses (optional part) of 10 WP3 installations.*

Knowledge of request numbers	Yes	No	N/A	Blank
Number of requests (total)	70%	10%	20%	0%
Number of requests (data)	60%	40%	0%	0%



Number of requests (metadata)	20%	50%	0%	30%
Number of requests (webpage)	80%	0%	0%	20%
Spams filtered?	0%	70%	0%	30%

Table 5. Estimated number of users across Geo-INQUIRE and Level 2/3 WP3 installations with operating or partly operating services. The content of the table is based on survey responses in the mandatory part.

Users		
Number of users	Geo-INQUIRE (71)	WP3 (27)
Less than 100	32%	7%
100 to 1.000	34%	48%
1.000 to 10.000	17%	26%
10.000 to 100.000	13%	11%
More than 100.000	4%	7%

Table 6. Heterogeneity of how the presently operating service usage systems estimate the number of users. The content of the table is based on survey responses (optional part) of 10 WP3 installations.

Knowledge on user numbers	Yes	No	N/A	Blank
Number of users served (total)	60%	10%	20%	10%
Number of users served (data)	50%	40%	0%	10%
Number of users served (metadata)	20%	40%	0%	40%
Number of users served (webpage)	50%	30%	0%	20%
Spams filtered?	0%	60%	0%	40%
Users approximated by IP addresses?	30%	40%	0%	30%
Geographical distribution of users per country	80%	20%	0%	0%
Geographical distribution of users per region	30%	60%	0%	10%

### KPIs for FAIR metrics of Accessible Datasets (table 6.1.3 Accessible Datasets)

Based on the Questionnaire, 20 out of the 27 respondents with level 2/3 operating or partly operating services actively implement the FAIR principles. Still, only 4 of them declare that they are also evaluating the FAIRness level, with two relying on the EPOS strategy to make this evaluation.

Concerning KPI ID AD-01 and AD 02, we find that the number or volume of datasets are not significant measures to qualify the installation/service portfolio in WP3. For example, the dataset number is linked to the data granularity, which is very variable and mainly depends on how the dataset creator has organized the data. The data internal organization may reflect user needs, community long-term practice, and compatibility with older datasets or legacy software. Also, data volume has little to do with quality. A single satellite image file of a small area can be as big as an entire earthquake catalogue covering 1,000 years of all of Europe. However, **Table 6**, based on survey responses (non-obligatory part) of 10 WP3 installations, shows a minority of "blank" responses for data volume and dataset number.

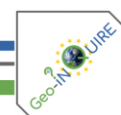


Table 7. Survey feedback concerning knowledge of data holdings. The content of the table is based on survey responses (non-obligatory part) of 10 WP3 installations.

Knowledge of data holdings	Yes	No	N/A	Blank
Data volume	80%	20%	0%	0%
Data volume (open datasets)	70%	20%	0%	10%
Data volume (restricted access)	0%	20%	60%	20%
Data volume (embargoed datasets)	0%	20%	60%	20%
Number of datasets	50%	20%	20%	10%
Number of open datasets	50%	20%	20%	10%
Number of datasets with restricted access	0%	20%	60%	20%
Number of embargoed datasets	0%	20%	60%	20%

The source of user and usage information is reported in **Table 6**, whose content is also based on survey responses (non-obligatory part) of 10 WP3 installations. The two installations that adopt an AAI system did not respond to this part of the Questionnaire. The respondents used alternatively one of three options (No, N/A, Blank). The majority of the respondents rely on the DOI to understand data usage. Although the survey did not ask for more details, usage tracking through the DOI can be done in different ways and for different scopes. For example, distinguishing usage in terms of views and downloads and usage in research outputs, i.e., counts of citations. For views and downloads of DOIs minted through DataCite (<https://datacite.org/>), there is the possibility of using the DataCite Usage Tracker (<https://support.datacite.org/docs/datacite-usage-tracker>) via API. This system, however, is still in an early release phase (beta). For citations, instead, one simple way is to check the dataset citations received by the service on Google Scholar (<https://scholar.google.com/>) or other similar third-party resources. Via API, for the DOIs minted through DataCite (<https://datacite.org/>), there is the possibility of using the DataCite Event Data (<https://support.datacite.org/docs/eventdata-guide>). The actual use of these or other solutions should be further explored.

Table 8. Survey feedback concerning the source of user and usage information.

Source of user and usage information	Yes	No	N/A	Blank
User knowledge through AAI (data only)	0%	50%	20%	30%
User knowledge through AAI (metadata)	0%	50%	20%	30%
Usage knowledge through AAI and user profile	0%	50%	20%	30%
Usage knowledge through DOI	90%	10%	0%	0%

Concerning KPI ID AD-03, we obtained controversial results using the F-UJI tool (<https://www.f-ujl.net/>) in assessing the FAIR score. As an example, **Figure 1** shows the FAIR assessment of the VA2-34-1 (the NEAMTHM18 dataset) service based on four different metadata standards. The score for the same dataset varies between 22-54%. Notice that in this example, the DCAT-AP metadata format, the standard adopted by EPOS and used by 33 out of 63 installations, yields the second-lowest score (41%). The resulting score, therefore, is affected by the combination of

metadata used and the F-UJI tool's capability to read and correctly interpret those metadata. A few strategies can be adopted during the Geo-Inquire project lifespan: 1) adopt the metadata standard that yields the highest score; 2) improve the metadata compilation of that standard; 3) provide continuous feedback to the F-UJI team to enhance the tool capability; 4) explore the capabilities of different tools to assess the FAIRness.

The considerations made for the Accessible Datasets are also valid for the new datasets (table 6.1.4 New Datasets). However, no new datasets have been implemented in WP3 in the first 12 months of the project.

#### Assessment Results:

##### Evaluated Resource:

NEAM Tsunami Hazard Model 2018 (NEAMTHM18): online data of the Probabilistic Tsunami Hazard Model for the NEAM Region from the TSUMAPS-NEAM project

FAIR level: **moderate**

Resource PID/URL: <https://data.ingv.it/metadata/iso19115/309.xml>

DataCite support: enabled

Metric Version: metrics\_v0.5

Metric Specification: <https://doi.org/10.5281/zenodo.4081213>

Software version: 2.2.5

Download assessment results: [\(JSON\)](#)

Save and share assessment results:

##### Summary:



#### Assessment Results:

##### Evaluated Resource:

Home

FAIR level: **moderate**

Resource PID/URL: 10.13127/tsunami/neamthm18

DataCite support: enabled

Metric Version: metrics\_v0.5

Metric Specification: <https://doi.org/10.5281/zenodo.4081213>

Software version: 2.2.5

Download assessment results: [\(JSON\)](#)

Save and share assessment results:

##### Summary:



#### Assessment Results:

##### Evaluated Resource:

NEAM Tsunami Hazard Model 2018 (NEAMTHM18): online data of the Probabilistic Tsunami Hazard Model for the NEAM Region from the TSUMAPS-NEAM project

FAIR level: **moderate**

Resource PID/URL: <https://data.ingv.it/metadata/dcat-ap/309.jsonld>

DataCite support: enabled

Metric Version: metrics\_v0.5

Metric Specification: <https://doi.org/10.5281/zenodo.4081213>

Software version: 2.2.5

Download assessment results: [\(JSON\)](#)

Save and share assessment results:

##### Summary:



#### Assessment Results:

##### Evaluated Resource:

Home

FAIR level: **initial**

Resource PID/URL: <https://data.ingv.it/metadata/dif/309.xml>

DataCite support: enabled

Metric Version: metrics\_v0.5

Metric Specification: <https://doi.org/10.5281/zenodo.4081213>

Software version: 2.2.5

Download assessment results: [\(JSON\)](#)

Save and share assessment results:

##### Summary:



Figure 1. Snapshots of four assessment results for the same dataset (NEAMTHM18) through the F-UJI tool (<https://www.f-uji.net/>) based on four different metadata resources: ISO 19115 (upper left), DataCite (upper right), DCAT-AP (lower left), and NASA DIF (lower right).

### **KPIs for TA usage (table 6.1.5 Transnational Access usage)**

These KPIs do not apply to WP3 because all installations/services are Virtual Access.

### **KPIs for Dissemination (table 6.1.6 Dissemination)**

The survey did not have specific questions to address the five KPIs identified for Dissemination quantitatively, and thus, WP3 relies on the (Seci. However, most installations already have a web page or portal to access the datasets, which include useful information for Dissemination. The installations with a logging system in place (**Table 3**) can track the usage of online documentation (e.g., tutorials) if their logging system has the capability to monitor the views of individual web pages.

Concerning Dissemination in terms of training events, workshops, personalized training, and summer schools, these will be developed in collaboration with WP9. In this respect, WP3 selected 22 installations (VA2-32-2, VA2-32-3, VA2-32-9, VA2-33-1, VA2-33-11, VA2-33-13, VA2-33-14, VA2-33-15, VA2-33-2, VA2-33-3, VA2-33-5, VA2-33-6, VA2-33-7, VA2-34-1, VA2-34-10, VA2-34-3, VA2-34-4, VA2-34-9, VA2-35-1, VA2-35-4, VA2-36-3, VA2-36-4) that could provide detailed information for personalized training to WP9 through another questionnaire whose rationale and results will be given in D9.5.

## **4 Harmonization and uptake strategy of KPIs among infrastructures**

In the Geo-INQUIRE landscape of RIs, several areas of heterogeneity and inconsistency in how KPIs are defined, collected, and interpreted across different installations and infrastructures have been highlighted. Addressing gaps and differences is crucial for successfully implementing and monitoring the Geo-INQUIRE project's objectives. Insights coming from D6.2 further emphasize the need for harmonization along the tracks defined by WP6 and WP7, not only in the context of metadata schemes, mappings, and converters but also in an unambiguous understanding of what exactly is measured and monitored in the end. WP9 addresses harmonization for training in D9.2.

### **4.1 Service and data-related KPIs and harmonization**

#### **4.1.1 Definition of 'users' and their counting mechanisms**

There are varying context-driven understandings of what is meant by 'users' in use across different installations, meanings ranging from users visiting webpages scripts downloading metadata. The acknowledgment of the existence of different types of usage (and users) must be addressed by converging toward common criteria for their identification. In IT Jargon, it is common (read: good practice) to refer to a "consumer" as an agent (both human or machine) that accesses a specific digital resource, and it is exactly this that could be a potential track for discussion in the context of WP7.



Besides developing a common understanding of such a consumer, it is also important to identify possible heterogeneities that merely result from different technical organization principles of these digital resources. As such, the D6.2 report suggests harmonizing data and metadata schemes through conversion to standards. Such harmonization may contribute to a more harmonized counting when it comes to accessing resources that describe the same (or similar) phenomena but are technically organized differently, e.g., a folder of files (which accounts for multiple access to the resources— one for each file) vs a single (huge) file, which is counted as a single access. Ensuring that data access is consistent across datasets and installations also impacts the KPIs.

### 4.1.2 Tracking country of origin of data or metadata requests

Different installations have diverse methods for tracking the country of origin. Again, the D6.2 report emphasizes the need for WP interaction for metadata enhancements, which would trigger relevant track of activities for better defining the metadata and, when needed, for providing a full metadata description by filling in properly the instance of an already existing and agreed metadata scheme. Collaborative efforts with WPs 2-5 can help standardize these tracking methods.

### 4.1.3 Definition and counting of datasets

The varied ways installations define and count datasets, based on disciplinary scientific practice of data collection and analysis, pose challenges in data integration. As mentioned above, the D6.2 report suggests accelerating metadata integration to ensure data is readily accessible and usable. Strategic steps can be taken to integrate outstanding metadata into the RIs, aligning with the project's objectives. The already mentioned harmonization of data and metadata formats would also improve and accelerate the convergence towards harmonized dataset descriptions, at least for the domain where similar data is dealt with.

### 4.1.4 Harmonization of service and data related KPIs

The heterogeneity in KPI definitions and metrics underscores the importance of harmonization. Continuous reviews and iterative enhancements should be planned to optimize the efficiency of metadata, mappings, and converters, ensuring that KPIs are consistent and meaningful across installations.

Most challenges identified in the current deliverable can be addressed by aligning with the strategies and activities proposed in the D6.2, D7.1, and D7.2 (including milestones 7.38 and 7.39) reports. As emphasized in the D6.2 report, regular communication and collaboration among stakeholders will be crucial in addressing these challenges and ensuring the successful implementation of the Geo-INQUIRE project's objectives.

## 4.2 Harmonization of dissemination KPIs

WP9 coordinates the dissemination activities of the project. Larger events (such as workshops, summer schools and hackathons) are anchored directly in WP9 and training events are organised





by the respective work packages. Detailed reporting of these events will, therefore, be part of the WPs training deliverables of M24 and M48, while the overall reporting and impact assessment across WPs will be part of the WP9 deliverables. In order to allow WP9 to measure impact and compare figures across WPs, it is important to measure dissemination KPIs in a uniform way.

WP9 offers general guidelines, together with templates for registration forms and for feedback forms. Both templates can be adapted to the needs of a specific activity. In addition to the KPIs, WP9 will also monitor internal monitoring values (IMV) to ensure that the objectives of the project and the guidelines of the Equality, Diversity and Inclusion Panel (EDIP) are met. The organisers of dissemination activities should contact WP9 a few months before the activity, to ensure that the proper advertisement through the Geo-INQUIRE and other project partners' communication channels, and that the registration and feedback forms collect all necessary information. The organisers will be responsible for collecting feedback forms and reporting KPIs and IMVs to WP9. To ensure that the relevant consortium bodies and work packages have a clear overview of dissemination activities progress, WP9 provides an up-to-date table in the shared filesystem.

## 5 Conclusion

Geohazard analysis and multi-risk assessment are multi-faceted scientific endeavours typically tackled by different specialists and communities working in traditionally separated environments. The hazard and risk data products also have a strong societal impact which requires transparency and repeatability of the adopted data and methods.

The highest priority of WP3 within Geo-Inquire is to exploit the opportunity to consolidate the access to these very different installations and services and enhance their interoperability and FAIRness. In this initial phase of the project, the adoption of KPIs started addressing the performance of the various installations individually.

One of the main challenges for the next phases of the project will be to learn how to use the KPI harmonization to enable us also addressing their collective performance. A critical element is considering the lineage of the data and service production. On the one hand, the good performance in accessing a data product should be reflected into the performance of the data used to produce it. On the other hand, an advanced level of FAIRness of a data product could be of lower value if the FAIRness of the data used to produce it is low. In other words, the performance of many aspects of a RI cannot be limited to the mere summation or average of the performance of its assets.



## 6 Appendices

### 6.1 Appendix 1: Portfolio of KPIs identified by WP7

#### 6.1.1 Service status <sup>1</sup>

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.	
SSRI-02	Service Status and RI Integration	%/number of Installations fully operating	Evaluation of each installation by WP6 in interaction with each installation.	
SSRI-03	Service Status and RI Integration	%/number of Installations under implementation operating	Evaluation by WP6 or input from installations	
SSRI-04	Service Status and RI Integration	Number of services running and reachable publicly	Spreadsheet	

#### 6.1.2 Virtual Access (VA) usage

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: <i>Does your installation have a service usage logging in place?</i>	The survey indicated that almost all operating systems have such a system in place. It is a prerequisite to answer <a href="#">Link to ESFRI KPIs 1,2,17,19</a>
VA-02	Virtual Access	Number data or metadata requests	Question (to be refined) to each installation: <i>If you have a service usage logging in place: How many data and metadata requests do you process each year (spams and robots excluded)?</i>	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 considered in processing of the Installation responses. <a href="#">Link to ESFRI KPIs 1 and 2.</a>
VA-03	Virtual Access	Number of users served	Question (to be refined) to each installation: <i>If you have a service usage logging in place: How many different users does</i>	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

<sup>1</sup> This is a subset of indicators being tracked by WP6 and described in D6.1/6.2





			<i>your installation serve each year?</i>	feasible during the Geo-INQUIRE lifespan. The full usability of this KPI is still to be evaluated. <a href="#">Link to ESFRI KPIs 1 and 2.</a>
VA-04	Virtual Access	Number of systems capable of geographically locate users.	Question (to be refined) to each installation: <i>If you have a service usage logging in place: Do you track the country of origin of data or metadata requests?</i>	Link to ESFRI 9

### 6.1.3 Accessible datasets

KPI ID	KPI Group	KPI definition	Collection Strategy	Comment
AD-01	Existing and accessible data holdings	Number of datasets accessible via the infrastructures in Geo-INQUIRE at the start of the project	Input from table on new datasets (see below)	Definition of a dataset is very different across Geo-INQUIRE installations. The usefulness of this KPI remains closely linked to analysis of the type and geographical distribution of data, and their use science. Efforts of integration of many small datasets will be reflected in this number.
AD-02	Existing and accessible data holdings	Volume of datasets accessible via the infrastructures in Geo-INQUIRE at the start of the project	Input from table on new datasets (see below)	Volume does not indicate quality. However large volume datasets are one of the key targets of Geo-INQUIRE, therefore the low number of high-volume datasets will be reflected through this KPI.
AD-03	Existing and accessible data holdings	FAIR score (0-100%)	Asses the dataset during the project to monitor improvements for the communities working on FAIR profiles with FAIRfuture.	FAIR score assessed with F-UJI <a href="https://www.f-uji.net/">https://www.f-uji.net/</a>

### 6.1.4 New datasets

KPI ID	KPI Group	KPI definition	Collection Strategy	Comment
ND-01	Newly added data holdings	Number of new datasets since the start of Geo-INQUIRE	Input from table on new datasets (see below)	Definition of a dataset is very different across Geo-INQUIRE installations. The usefulness of this KPI remains closely linked to analysis of the type and geographical distribution of data, and their use science. Efforts of integration of many small datasets will be reflected in this



				number.
ND-02	Newly added data holdings	Volume of new datasets since the start of Geo-INQUIRE	Input from table on new datasets (see below)	Volume does not indicate quality. However large volume datasets is one of the key targets of Geo-INQUIRE, therefore the low number of high volume datasets will be reflected through this KPI.
ND-03	Newly added data holdings	FAIR score (0-100%)	Asses the data sate as soon as added and continue periodically to monitor improvements	FAIR score assessed with F-UJI

### 6.1.5 Transnational Access (TA) usage

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
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.
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		

### 6.1.6 Dissemination

KPI ID	KPI Group	KPI definition	Collection Strategy	Comment
DT-01	Dissemination	Number of applications	Logs of registration of participants, unique application	It may necessary to differ between the types of activities. Online training is open to all interested people, while on-site activities (such as summer schools) may have more



				applications then capacity (selection via TEB)
DT-02	Dissemination	Number of actual participants	Logs of online tools used for the activity, signed presence sheet on-site, etc.	
DT-03	Dissemination	Cross-disciplinarity	Data must be collected in the registration forms	Affiliation, field of research, RIs
DT-04	Dissemination	New users	Data must be collected in the registration forms	Questions: "Have you used this infrastructure before? (y/n)"
DT-05	Dissemination	Number/percentage of Early Career Scientists (ECS) <sup>2</sup>	Data collected in the registration forms, percentage of overall attendance	Question about career stage <b>Added based on ESFRI KPIs</b>

## 6.2 Appendix 2: FAIR metrics used to attribute the FAIR score

Definition	F-UJI metric ID	Details of automated test
Data is assigned a globally unique identifier	FsF-F1-01D	The object is identified by a unique identifier (GUID or IRI) that follows a proper syntax. The identifier is web-accessible (not broken)
Data is assigned a persistent identifier	FsF-F1-02D	A data identifier is specified based on a commonly accepted persistent identifier scheme suitable for research data. This id is web-accessible, i.e., it resolves to a landing page with metadata of the data object.
Metadata includes descriptive core elements to support data findability	FsF-F2-01M	Some metadata has been made available via common (web) standards. There, a minimum core citation metadata is specified (creator, title, publication date, publisher, and identifier). A minimum core descriptive metadata is also checked (creator, title, publisher, publication date, summary, keywords, identifier).
Metadata includes the identifier of the data it describes	FsF-F3-01M	Metadata contains a PID or URL which indicates the location of the downloadable data content. A data identifier is included in the metadata and it matches the identifier originally provided.
Metadata is offered in such a way that it can be retrieved by machines	FsF-F4-01M	Metadata of the object is retrievable programmatically through at least one of the following methods: structured data embedded in the landing page; typed links of metadata document or signposting header links; content negotiation with a PID provider service (e.g., DataCite).
Metadata contains access level and access conditions of the data	FsF-A1-01M	Metadata includes the level of data access (e.g., public, embargoed, restricted) and its access conditions using appropriate metadata fields. Access level metadata is machine-readable, and this is verified against controlled vocabularies <sup>345</sup>
Metadata is accessible through a standard communication protocol	FsF-A1-02M	The metadata URI's scheme is based on a common application protocol and is accessible through the identifier provided.

<sup>2</sup> According to the definition of the European Geosciences Union: "An Early Career Scientist (ECS) is a student, a PhD candidate, or a practising scientist who received their highest certificate (e.g. BSc, MSc or PhD) within the past seven years."

<sup>3</sup> [http://vocabularies.coar-repositories.org/documentation/access\\_rights/](http://vocabularies.coar-repositories.org/documentation/access_rights/)

<sup>4</sup> <http://purl.org/eprint/accessRights>

<sup>5</sup> <http://publications.europa.eu/resource/authority/access-right>



Data is accessible through a standardized communication protocol	FsF-A1-03D	The data URI's scheme is based on a shared application protocol. The data is accessible through the identifier provided.
Metadata is represented using a formal knowledge representation language	FsF-I1-01M	The knowledge representation language is one of the following mechanisms: parsable, structured data in the landing page; or parsable, formal metadata (e.g., RDF, JSON-LD) accessible through content negotiation, typed links, or sparql endpoint.
Metadata uses semantic resources	FsF-I1-02M	Namespaces of known semantic resources are present in the metadata; exclude common namespaces (e.g., rdf, rdfs, xsd, owl) from the test.
Metadata includes links between the data and its related entities	FsF-I3-01M	Metadata captures the relation between a data object and its related entities. This relation is expressed using a relation type according to PROV-O or DataCite relation types.
Metadata specifies the content of the data	FsF-R1-01MD	Metadata includes the type of the object and the technical properties of its data file such as format, size, observed or measured variables. The values of the properties comply with the actual data file.
Metadata includes license information under which data can be reused	FsF-R1.1-01M	Metadata contains license information represented using an appropriate metadata element. Preferably, a standard, machine readable license should be specified.
Metadata includes provenance information about data creation or generation	FsF-R1.2-01M	Metadata includes properties representing data creation such as creator, contributors, creation and modification dates and version, source, and relations that indicate data creation activities. Provenance metadata is available in a machine-readable version of PROV-O or PAV.
Metadata follows a standard recommended by the target research community of the data	FsF-R1.3-01M	Metadata is available through at least one of the domain metadata standards listed in RDA Metadata Standards Catalogue
Data is available in a file format recommended by the target research community	FsF-R1.3-02D	Data is available in a long-term file format as defined in ISO/TR 22299. Data is available in an open format <sup>6</sup> . Data is available in a scientific file format <sup>7</sup> .

### 6.3 Appendix 3: ESFRI Key Performance Indicators (KPIs)

Summary of ESFRI KPIs: More information about metrics and other details:

[https://www.esfri.eu/sites/default/files/ESFRI\\_WG\\_Monitoring\\_Report.pdf](https://www.esfri.eu/sites/default/files/ESFRI_WG_Monitoring_Report.pdf)

Objective	KPIs
<b>Enabling scientific excellence</b>	1. Number of user requests for access
	2. Number of users served
	3. Number of publications
	4. Percentage of top (10%) cited publications
<b>Delivery of education and training</b>	5. Number of master and PhD students using the RI
	6. Training of people who are not RI staff

<sup>6</sup> [https://en.wikipedia.org/wiki/List\\_of\\_open\\_formats](https://en.wikipedia.org/wiki/List_of_open_formats)

<sup>7</sup> Library of Congress dataset formats, Wolfram Alpha supported file formats.



<b>Enhancing collaboration in Europe</b>	7. Number of members of the RI from ESFRI countries
	8. Share of users and publications per ESFRI member country
<b>Facilitating economic activities</b>	9. Share of users associated with industry and publications with industry
	10. Income from commercial activities and the number of entities paying for service
<b>Outreach to the public</b>	11. Engagement achieved by direct contact
	12. Outreach through media
	13. Outreach via the RI's own web and social media
<b>Optimising data use</b>	14. Number of publicly available data sets used externally
<b>Provision of scientific advice</b>	15. Participation by RIs in policy related activities
	16. Citations in policy related publications
<b>Facilitating international co-operation</b>	17. Share of users and publications per non-ESFRI member country
	18. International trainees
	19. Number of members of the RI from non-ESFRI countries
<b>Optimising management</b>	20. Revenues
	21. Extent of resources made available

