www.geologicalservice.eu

Geological Services Knowledge Hub Development & Implementation Plan

Version: v1

Project	101075609 — GSEU — HORIZON-CL5-2021-D3-02	
Deliverable Data		
Deliverable number:	**D9.3 – Geological Services Knowledge Hub Development & Implementation Plan**	
Dissemination level:	**Public**	
Deliverable type:	**Report**	
Work package:	**WP9 – Geological Surveys network and sustainable services for Europe**	
Lead WP/Deliverable beneficiary:	**[EGS / GeoSphere Austria]**	
Deliverable status		
Verified (WP leader):	**Patrick Wall [EGS]**	
Approved (Coordinator):	**Julie Hollis [EGS]**	
Author(s):		Affiliation:
Haya Majid Qureshi		GeoSphere Austria
Martin Schiegl		GeoSphere Austria

Funded by the European Union

EuroGeoSurveys European Geoscience for Society
The Geological Surveys of Europe

Disclaimer

Copyright

Funded by the European Union

Revision History		
Author(s):	**Description:**	**Date:**
Christina Schmutzhard, Robert Supper	1st Draft	26/09/2024
Jørgen Tulstrup, Jasna Šinigoj and Patrick Wall	Revision 1	21/10/2024
Martin Schiegl, Christina Schmutzhard	2nd Draft	25/10/2024
Jørgen Tulstrup, Jasna Šinigoj	Revision 2	08/11/2024
Martin Schiegl, Haya Qureshi, Robert Supper, Christina Schmutzhard	3rd Draft	20/11/2024
Patrick Wall	Revision 3	06/01/2025
Francesco Pizzocolo	Revision 4	06/01/2025
Martin Schiegl, Christina Schmutzhard, Robert Supper	4th Draft	22/01/20205
Patrick Wall	Revision 5	23/01/20205
Francesco Pizzocolo	Final version	24/01/2025

Executive Summary

The GSEU WP9.1 Task aims to establish a sustainable network of European Geological Surveys through the Geological Services Knowledge Hub. This platform will enhance collaboration, standardize research methodologies, and integrate resources across Europe. Recognizing the necessity of a unified approach to geological data management and knowledge sharing, WP9.1 seeks to create a centralized system that fosters cooperation among European Geological Surveys within the GSE.

This document presents the feasibility study and proof of concept (POC) for the Geological Services Knowledge Hub, now referred to as the **Expertise Search Service**. The initiative addresses challenges in managing and utilizing diverse datasets from independent and widely recognized sources, such as project databases and peer-reviewed publications, which are often stored in isolated silos. By leveraging Semantic MediaWiki and related technologies, the project integrates these datasets into a single platform, enabling advanced search, visualization, and collaboration.

Key actions include designing the Expertise Search Service, identifying data sources, and establishing technical frameworks such as knowledge graphs and graph databases. The platform emphasizes robust data governance, incorporating advanced user management systems to ensure data security and controlled access. The POC validated these components, demonstrating their potential for real-world applications.

A SWOT analysis highlighted the advantages of using MediaWiki as a collaborative tool, its adaptability for semantic data management, and its scalability with technological advancements. The initiative also presents opportunities for enhancing pan-European collaboration and addressing challenges related to resource availability and climate change adaptation. The report proposes innovative solutions for assigning persistent Uniform Resource Identifiers (URIs) to facilitate seamless data integration and interoperability.

This deliverable underscores the importance of interdisciplinary cooperation in addressing complex challenges. By enabling both short-term responses and long-term strategic collaborations, the Expertise Search Service positions the GSO network as a pivotal resource for advancing geological research and supporting evidence-based policymaking across Europe.
Moving forward, the findings from this study will guide the implementation of the Knowledge Hub under WP7. The project envisions a scalable platform that evolves with user needs, ensuring that Europe’s geological community remains at the forefront of innovation and sustainability.

Abbreviations	
CERIF	Common European Research Information Format
DOI	Digital Object Identifier
EC	European Commission
EU	European Union
FS	Feasibility Study
GS	Geological Survey
GSO	Geological Survey Organization
HE	Horizon Europe
MW	MediaWiki
POC	Proof of Concept
SMW	Semantic MediaWiki
UML	Unified Modeling Language
URI	Uniform Resource Identifier
WP	Work Package
ZAMG	Zentralanstalt für Meteorologie und Geodynamik

Table of Contents

List of Figures

List of tables

1. Introduction

In the past, cooperation between Geological Survey Organizations (GSOs) was primarily driven by joint project applications, predefined expert group collaborations, and individual bilateral agreements. However, current challenges—including climate change mitigation, energy transition, and critical mineral resource availability—necessitate a pan-European information network of high-quality, standardized geological data for informed decision-making at national and EU levels.

To meet these demands, the European GSOs network, led by EGS, aims to establish a Geological Service for Europe (GSE). This effort requires rethinking the existing cooperation framework to fully leverage expertise across GSOs. The GSE must respond to dynamic and unpredictable policy and research demands, necessitating a system that enables fast, high-quality responses to stakeholder inquiries while fostering long-term collaboration.

Task WP9.1 developed a proof-of-concept for a Geological Services Knowledge Hub, designed as an Expertise Search Service. This tool will facilitate access to GSO expertise and encourage interdisciplinary cooperation. It ensures that assessments of institutional expertise are based on both subjective expert input and automatically harvested data from independent sources such as project databases and peer-reviewed publications. By enabling users to identify organizational strengths objectively, the platform will support the creation of tailored cooperation networks that best fit specific research and policy needs.

In collaboration with GSEU WP7 (European Geological Data Infrastructure), it was determined that this project would focus on a design plan and feasibility study, while WP7 would handle the implementation and technical setup of the prototype. This report details the feasibility study, design plan, and proof of concept for the Geological Services Knowledge Hub.

1.1. Objectives and Scope

Task WP9.1 aims to develop the conceptual design for an Expertise Hub, a centralized platform integrating information on European geological surveys, experts, and projects. Key elements include:

- Links to experts from all European national geological surveys (EGS members).
- Connections between experts, organizations, and projects.
- Discovery and display of relevant information from all EGS members.
- Proof of Concept (POC) demonstration to validate feasibility.

1.2. Rationale and Strategic Importance

European geological surveys generate vast amounts of data on projects, organizations, and experts. However, these datasets often reside in isolated silos, limiting their reuse and analysis. The Expertise Hub will integrate these fragmented data sources into a unified platform, providing users with a comprehensive and accessible knowledge base.

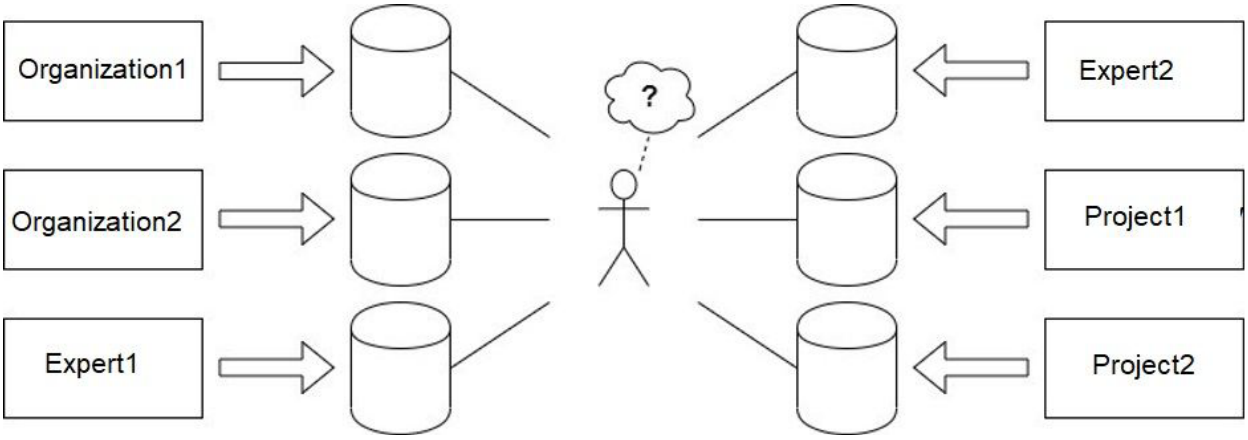

Figure 1: Distributed Data Silos

Objectives: Integrate diverse content silos into a unified platform to consolidate expertise across geological fields, ensuring users have easy access to essential information when needed.

1.3. Target Users, Application and Use Cases

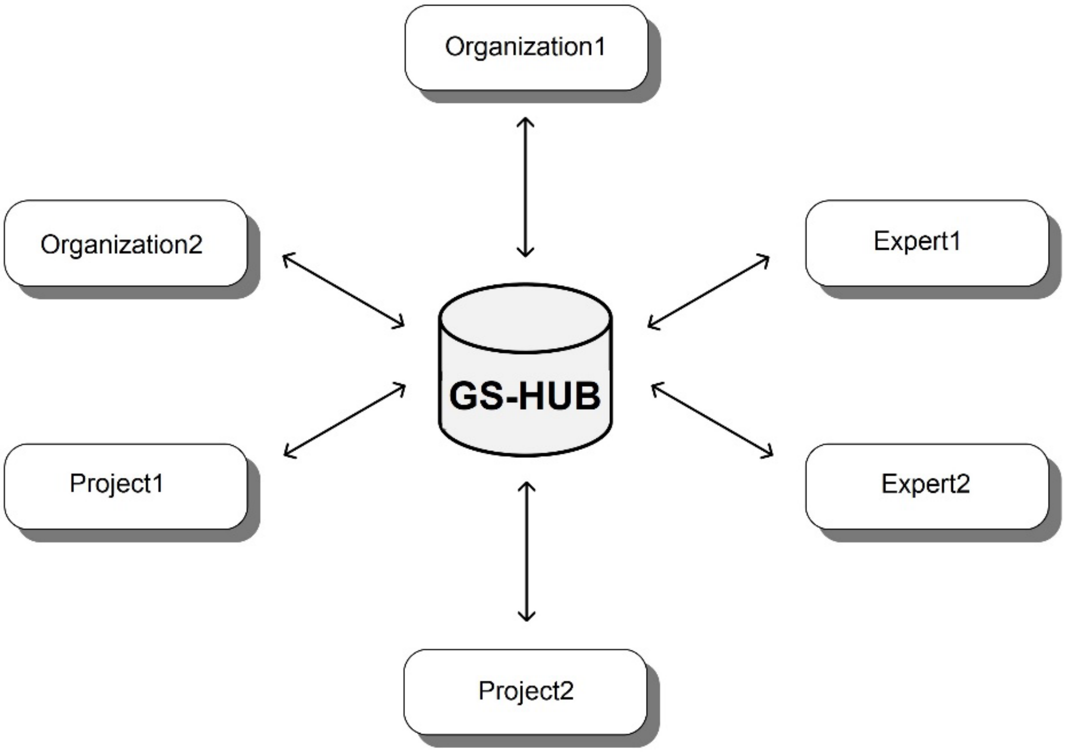

Figure 2: Centralized Expertise HUB of GSO

Target Users: Geological Survey Organizations (GSOs).

A centralized information service providing data on geological organizations in Europe, their projects, and the experts involved can be a valuable resource for all geological organizations:

- **Researchers and Geologists:** Researchers can identify potential collaborators and sources of data, facilitating knowledge sharing and fostering cross-border research initiatives.
- **Networking and Collaboration:** Geologists and organizations can use the platform for networking and forming partnerships.
- **Project Managers:** Professionals managing geological projects can leverage the information for identifying experts and learning from past projects.

1.4. What are the current requirements that this project must meet?

To comprehend the requirements, several use cases were prepared to have a structured way to identify, define, and document system requirements. 36 organizations were asked to vote and 26 organizations participated.

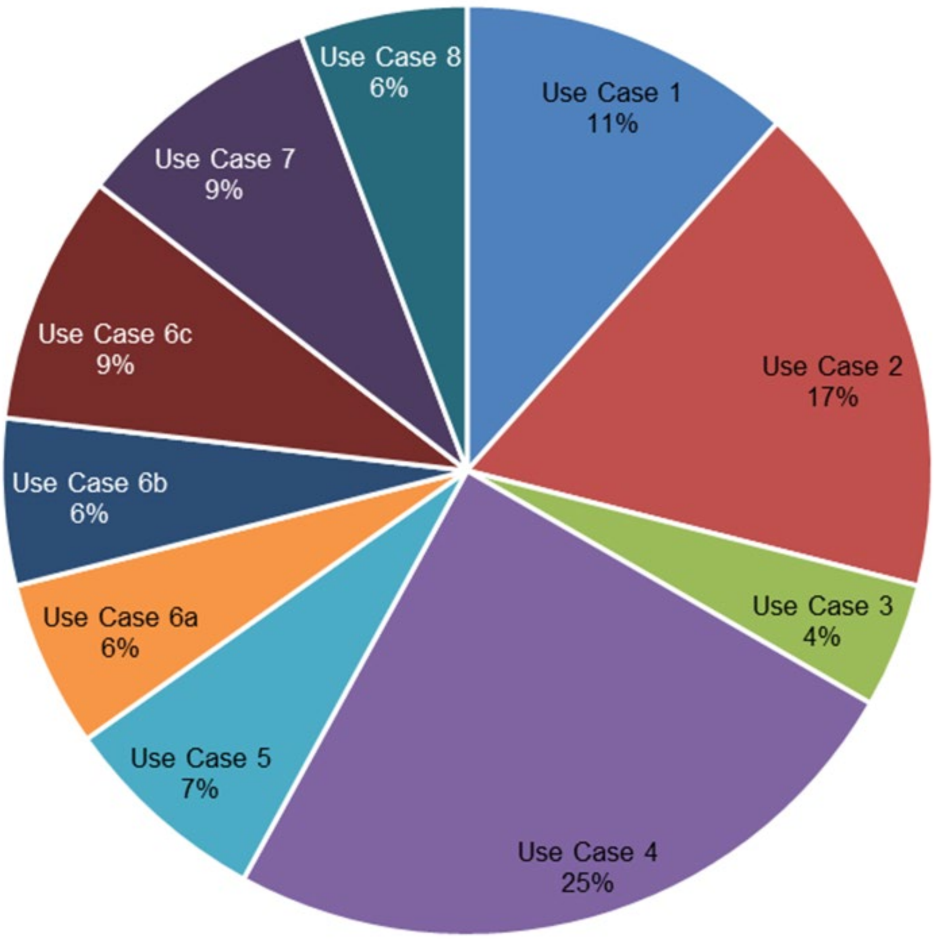

Figure 3: Percentage distribution of use-cases assigned priority 1, 2 or 3

Moreover, use cases support the iterative development process, accommodating changes as the project evolves. 11 use cases were created and it was asked to the stakeholders to vote for the three most important ones. The use cases are as follows:

- **Use Case 1:** Search for Geological Survey Organizations (GSO) and their projects
- **Use Case 2:** Search for project partners
- **Use Case 3:** Search for scientific expertise by "grey literature" and projects
- **Use Case 4:** Search for expertise
- **Use Case 5:** Enhancing/Increasing visibility of existing Working Groups
- **Use Case 6a:** Market Place: Have samples analyzed
- **Use Case 6b:** Market Place: Renting equipment
- **Use Case 6c:** Market Place: Need for a specific service
- **Use Case 6d:** Booking conference rooms
- **Use Case 7:** Search for staff training possibilities
- **Use Case 8:** Provide links to data sets and information on web applications

All the participating organizations showed interest in use case 1, 2 and 4 (see Figure 3). Based on this uses-case study, we break the requirements into the two clusters: user and technical requirements.

1.4.1. User Requirements

- User-Friendly Interface: An intuitive design ensuring seamless navigation and adoption.
- Advanced Search Capabilities: A powerful search system for retrieving relevant projects, experts, and organizations.
- Knowledge Sharing & Access Control: A user management system with controlled access for editing and contributing.

1.4.2. Technical requirements

- Semantic MediaWiki: Chosen for its efficiency in data management and user-friendly structure.
- Knowledge Graphs: Used to structure relationships between experts, projects, and organizations.
- Data Sources: Aggregated from OpenAIRE, Kohesio, and Wikidata.
- Persistent Identifiers (URI System): Implemented to ensure consistency and interoperability across datasets.

Tool
We chose Semantic MediaWiki for efficient data management with user friendly forms. Section 3 explains in detail about the overall MediaWiki system setup and extensions used for it. Its advanced editing supports easy organization, while search features enhance finding organizations, projects, and experts effortlessly.

Data sources
We need data to build a knowledge graph. A knowledge graph is a structured representation of information that captures relationships between entities, such as people, places, and things, in a graph format. It allows for enhanced data connectivity, enabling systems to understand context and make inferences from Linked Data. It can support data entry, store data and visualize data, but its primary function is to organize data from multiple sources and apply context for business purposes: analysis, question answering, search and recommendations. Fundamentally, a knowledge graph is a database.

This data come from a variety of sources, including structured data in databases or spreadsheets, unstructured data in text files or web pages. The Organisations and Projects data is retrieved from OpenAire, KOHESIO and Wikidata. We also take all GSO's employees (experts) data from Wikidata and use it as preload data. We manually evaluated and cross-checked various data sources, selecting these as reliable representations of organizations and projects. Therefore, users of our system do not have to enter data from scratch.

Schema (taxonomy/thesaurus)

Task WP9.1 will reuse classes and properties from various existing specifications such as the described namespaces shown in Table 1. These schemas and namespaces in Table 1 are needed for structuring, describing, and organizing data to enable semantic interoperability and knowledge representation on the web.

Table 1: List of Schemas and their Namespaces

Schema (Ontology/taxonomy/thesaurus)	**Prefix**	**URI**
SKOS (Simple Knowledge Organization System)	skos	http://www.w3.org/2004/02/skos/core#
OWL 2 (Web Ontology Language)	owl	http://www.w3.org/2002/07/owl#
RDF (Resource Description Framework)	rdf	http://www.w3.org/1999/02/22-rdfsyntax-ns#
RDFS (RDF Schema)	rdfs	http://www.w3.org/2000/01/rdfschema#
EURIO (EUropean Research Information Ontology)	eurio	http://data.europa.eu/s66#
DCMI (Dublin Cores Metadata Terms Initiative)	dcterms	http://purl.org/dc/terms/
DINGO (Data Integration for Grants Ontology)	dg	https://w3id.org/dingo#
Schema.org	schema	http://schema.org/
Wikidata Entity	wd	http://www.wikidata.org/entity/
Wikidata Property	wdt	http://www.wikidata.org/prop/

There is also a list of controlled vocabularies used to specify the Expertise, Skill-set and Keywords for Organizations, Projects and Experts. The values belong to the following namespaces.

Table 2: List of Vocabularies

Vocabulary	**Prefix**	**URI**
Keyword Thesaurus	kt	https://data.geoscience.earth/ncl/geoera/keyword/

Note:

CERIF (Common European Research Information Format) modelling approach [4] is a comprehensive and complex framework designed to manage research information, but its intricate structure can be challenging to use in linked open project effectively. Although we are familiar with CERIF, we have opted to use EURIO, which is a published and better suited for our project's needs due to its streamlined approach.

Graph Database

The feasibility in this study was carried out using the GraphDB software. Task WP9.1 plan uses GraphDB for storing and querying the knowledge graph built using Semantic MediaWiki (SMW) an extension of MediaWiki that enables storing, querying, and managing structured, semantic data within wiki pages. Graph databases are optimized for managing graph data and can efficiently perform queries that traverse relationships between nodes.

Query Language

To interact with the knowledge graph, we will use SPARQL query language that allows us to query the data using graph-based patterns and relationships. GraphDB comes with its own SPARQL endpoint.

Data Governance

To ensure the privacy and quality of the knowledge graph, we need to implement data governance processes such as access controls. It allows administrators to regulate user permissions effectively. This feature enables fine-tuned control over who can view, edit, or manage specific content, ensuring data accuracy. We have proposed three user groups with different permissions set for the wiki: EGS, National Delegates and GSOs.

1. **EGS office** enjoys comprehensive access to data in Semantic MediaWiki. Their permission set allows editing any attribute, creating pages, and uploading files and images. This tailored access ensures efficient collaboration and data management within specified parameters (see Figure 13 in Annex 10).
2. Every **national delegate** from European geological surveys enjoys comprehensive access to data in Semantic MediaWiki. Their permission set allows editing any attribute, creating pages (excluding organization and project data from External system), and uploading files and images. This tailored access ensures efficient collaboration and data management within specified parameters (see Figure 14 in Annex 10).
3. Each **European geological survey** gets a single shared login for its employees within their respective organization. Limited access ensures data security, permitting attribute editing only. Employees cannot create pages but may upload files and images in designated expert sections, maintaining control while fostering collaboration within defined parameters (see Figure 15 in Annex 10).

In this way only authorized individuals will have access to sensitive or restricted in- formation, enhancing overall system security.

2. MediaWiki

A wiki is a website designed for collaborative editing of its content, allowing users to easily contribute and modify information. Authors can structure their content using a shorthand (a simplified way of writing or formatting content, using special symbols or syntax) for defining chapters, sections, paragraphs, hyperlinks, and other elements, which the wiki software then renders into web pages.

In order for search engines to provide valuable results, there must be high-quality, well-organized information available online. Wiki technology plays a crucial role in enabling the creation of such repositories, which are continually improved by a large community of contributors and benefit an even larger audience. By facilitating extensive collaboration, wikis help address the challenges posed by the information age.

2.1. MediaWiki Setup

MediaWiki, written in PHP, is compatible with all major operating systems supported by PHP. It relies on a database, which can be MySQL/MariaDB, PostgreSQL, or SQLite. For the WP9.1 Proof of Concept (POC), MySQL is utilized – MediaWiki uses MySQL to manage wiki-based applications, which enable users to collaboratively create, edit, and organize content in the form of interconnected web pages. In contrast, while MySQL stores web page structured data in tables, a Graph database models relationships between data as nodes and edges, providing a more flexible way to represent complex, interconnected information.
Additional information can be found in Annex 10.2. on the technical details of the Media Wiki Setup as well as on Skins which allow users to customize the look and feel of Media Wiki.

2.2. Extensions

A vital component of the MediaWiki system is its extensions, which are designed to enhance the core application.

In this POC, we have utilized approximately 20 extensions, which we will briefly discuss here.

1. **Semantic MediaWiki:** is a full-fledged framework, in conjunction with many spinoff extensions, that can turn a wiki into a powerful and flexible knowledge management system. All data created within Semantic MediaWiki can easily be exported or published via the Semantic Web, allowing other systems to use this data seamlessly
2. **Semantic Result Formats:** is an extension to extension "Semantic MediaWiki" that adds a large number of further result formats, including formats for media, table, export, graph and mathematical functions
3. **Semantic Forms Select:** allows to generate a select field in a form whose values are retrieved from a query
4. **Data Transfer:** allows for importing and exporting data contained in template calls
5. **Page Schemas:** defines the data structure for all pages in a category using XML
6. **Page Forms:** Forms for creating and editing wiki pages
7. **Translate:** translating MediaWiki and beyond
8. **External Redirect:** allows to make redirects to external websites.

9. **Maps:** allows embedding of dynamic maps into wiki pages using Google Maps or Leaflet. Has a visual editor, optionally integrates with Semantic MediaWiki, sup- ports GeoJSON and adds geocoding capabilities
10. **Network:** allows adding interactive network visualizations in your wiki pages
11. **PageOwnership:** Implements page ownership based on users and groups through a user-friendly interface
12. **Widgets:** Allows wiki administrators to add free-form widgets to the wiki by editing pages within the Widget namespace
13. **LinkedWiki:** lets you reuse Linked Data in your wiki. You can get data from Wikidata or another source directly with a SPARQL query
14. **UniversalLanguageSelector:** gives the user several ways to select a language and to adjust language settings
15. **CategoryLockdown:** allows admins to restrict permissions by category and group
16. **DisplayTitle:** uses displaytitle page property in link text, subtitle, and talk page title; provides parser function to query displaytitle
17. **CLDR:** contains local language names for different languages, countries, currencies, and time units extracted from CLDR data
18. **Babel:** refers to the texts on user pages aiding multilingual communication by making it easier to contact someone who speaks a certain language
19. **PagePermissions:** provides per page access based on default / custom roles set by the administrator
20. **Nuke:** gives administrators the ability to mass deletes pages.

3. Proof Of Concept

A proof of concept (POC) is a demonstration aimed at validating the feasibility of a concept or idea. In line with this principle, Task WP9.1 offers a small-scale design study to confirm that the designed conceptual model holds promise for real-world application.

3.1. Conceptual Model

The graphical representation of the conceptual model for WP9.1 is provided in the form of an UML (Unified Modelling Language) class diagram and is depicted in Figure 4. The boxes represent classes while the arrow connections represent properties establishing relations to other classes. The attributes inside boxes represent properties providing either literal data values or relation to other classes that are omitted from the diagram.
The green box related with expert class represent taxonomies or controlled vocabularies that will specify through instances the type of an expertise, knowledge area, and so on. The brown box represents the customs SMW entries that we will establish inside this wiki system.
The legend besides the UML diagram indicates which are the classes that have been reused from other existing ontologies or data sources.

3.2. System Structure

The system structure of our *private* wiki system (POC) is designed to efficiently manage and utilize specified data sources while ensuring secure access and integration of information. At its core, our system leverages a GraphDB triple store to store Wikidata, which can be accessed and queried through a SPARQL endpoint embedded within the wiki interface. Additionally, data from OpenAIRE and Kohesio sources are seamlessly integrated into the wiki environment through direct API connections (see Figure 5). User management is a pivotal aspect of our system, ensuring that access to sensitive information is appropriately controlled. Through the discussed user management system, administrators can assign roles and permissions to users, thus regulating their ability to view, edit, or manage data within the wiki.
The foundation of the system (Private System) tested in the feasibility study consists of three main objects: Projects, Organizations, and Experts, all structured according to the EURIO ontology. This ontology facilitates standardized representation and interlinking of data, enabling seamless navigation and discovery within the wiki environment. Figure 5 shows a possible interaction of publicly accessible components from the existing EGDI infrastructure (knowledge infrastructure, data search, repository and data catalog) - and new components (MediaWiki, GraphDB) for use by selected usergroups (EGS Sectretariate, National Delegates, GSO Employees). In addition, external, already existing resources with machine-readable programming interfaces (APIs) (OpenAIRE, Cordis, Kohesio, Wikidata, Wikibase, ORCID, GeoNames, EGR registry, GeoERA Keyword Thesaurus, EGDI Project Vocabularies, etc.) are presented, from which existing information about organizations, projects and experts can be linked without replicating them.
Furthermore, our system incorporates various ontologies, vocabularies, and thesauri, all residing within the GraphDB. These resources enhance data interoperability and semantic richness, enabling users to perform complex queries and analyses within the wiki interface.
In contrast to our private system, a public system has been developed by WP7, wherein all information is openly accessible to everyone. The key distinction lies in the privacy aspect, as the WP9.1 system is designed to restrict access to authorized users only. Additionally, WP9.1 serves as a conceptual framework for expertise management, which will eventually become a subset of the larger system developed by WP7, showcasing a hierarchical relationship between the two systems.

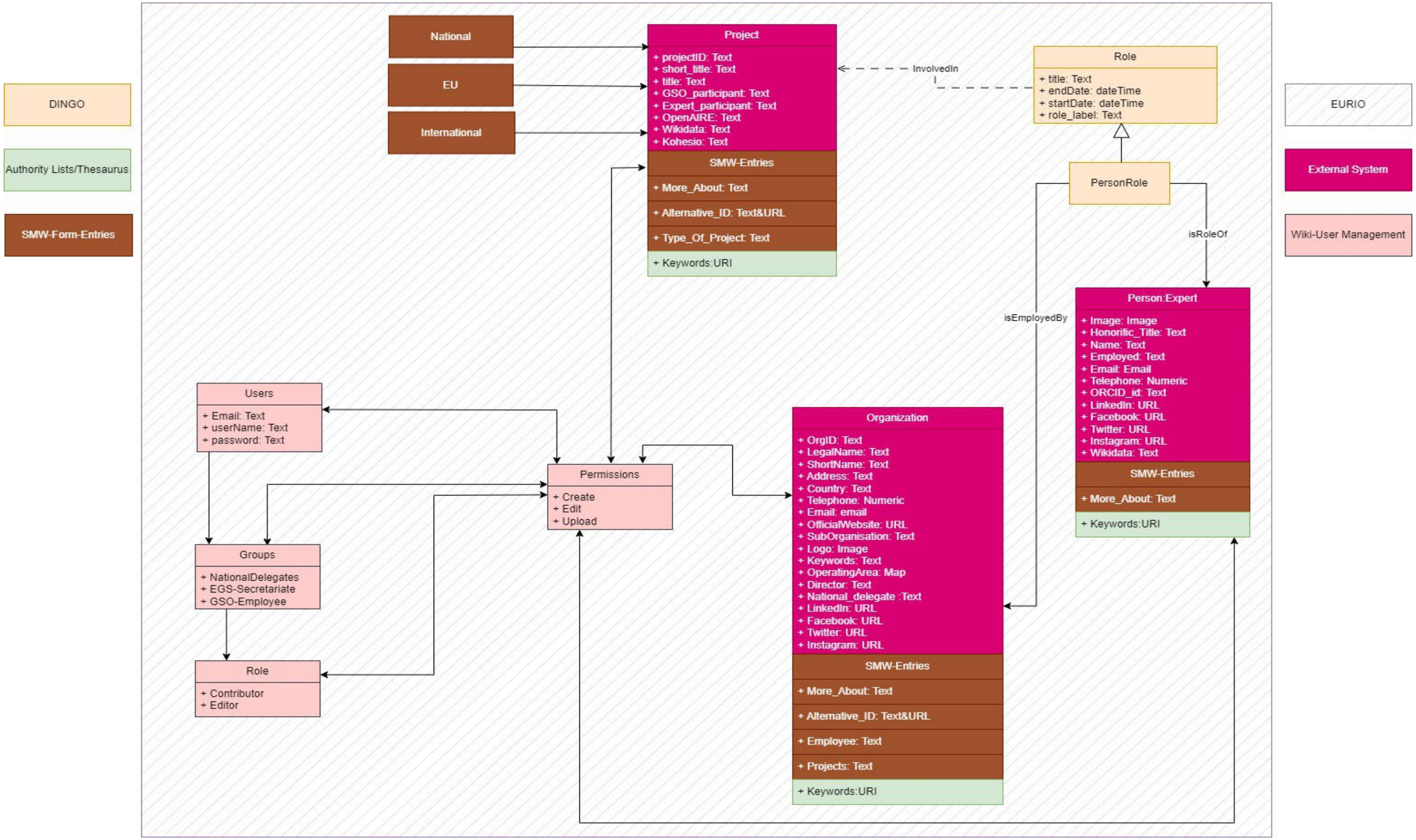

Figure 4: Proposed Data Model for Task 9.1

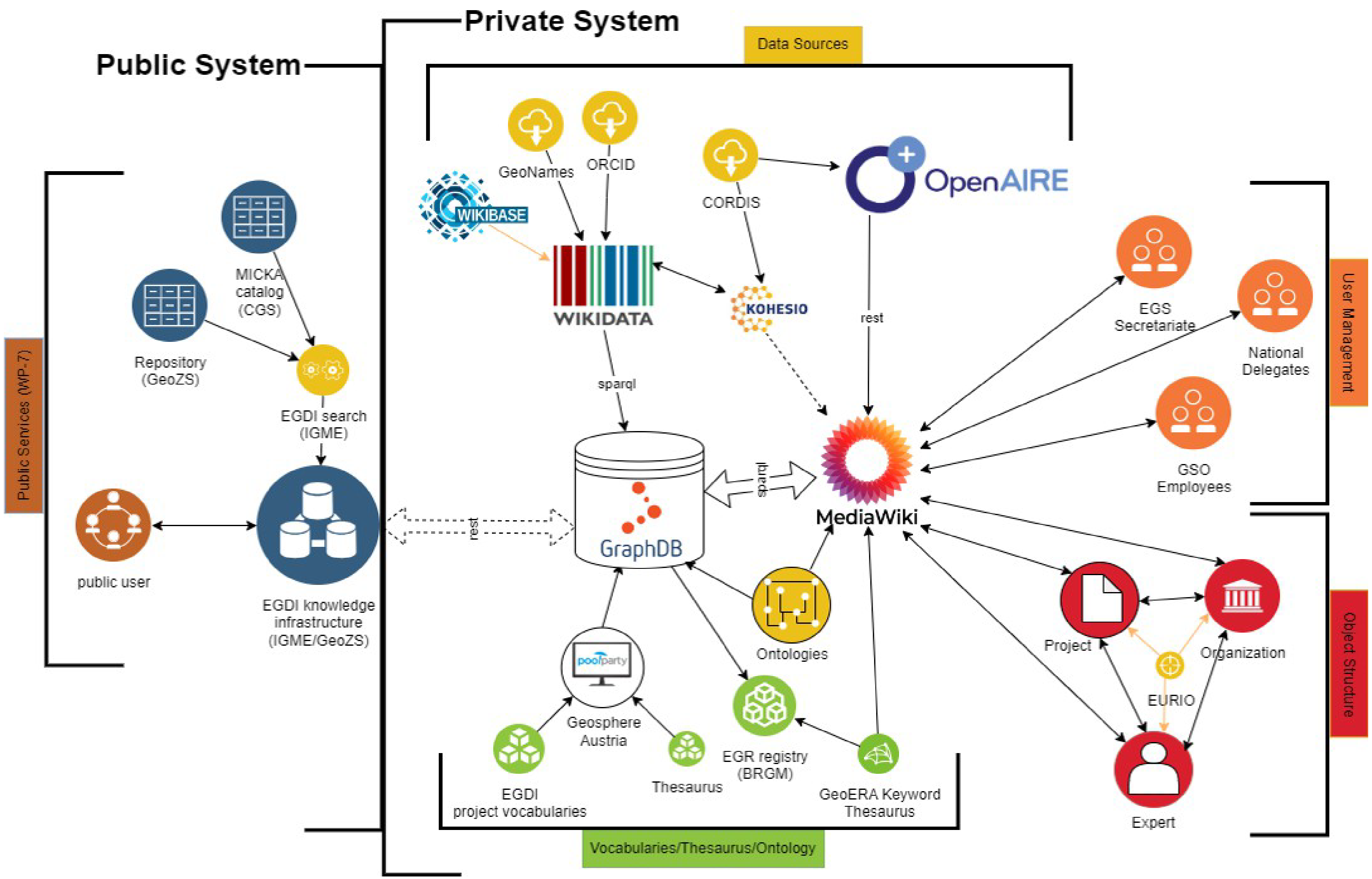

Figure 5: Proposed System Structure for Task 9.1

4. Results

The results of the proof of concept (POC) demonstrate promising feasibility of the conceptual model proposed. It is confirmed that the model effectively addresses the key objectives and showcases the potential for practical application in real-world scenarios.

4.1. Use-Case for GSO – Search for GSO and Projects

Here is an example of a possible question: A GSO wants to find GSOs in a specific European Country. The search result is a list that includes contact information and links to the websites. Also a GSO looks for other GSO's projects on a specific topic. The search in the database results in a list of GSO's projects subdivided into international, national and internal projects. Refer to

Figure 6 for a visual representation of the use-case discussed.

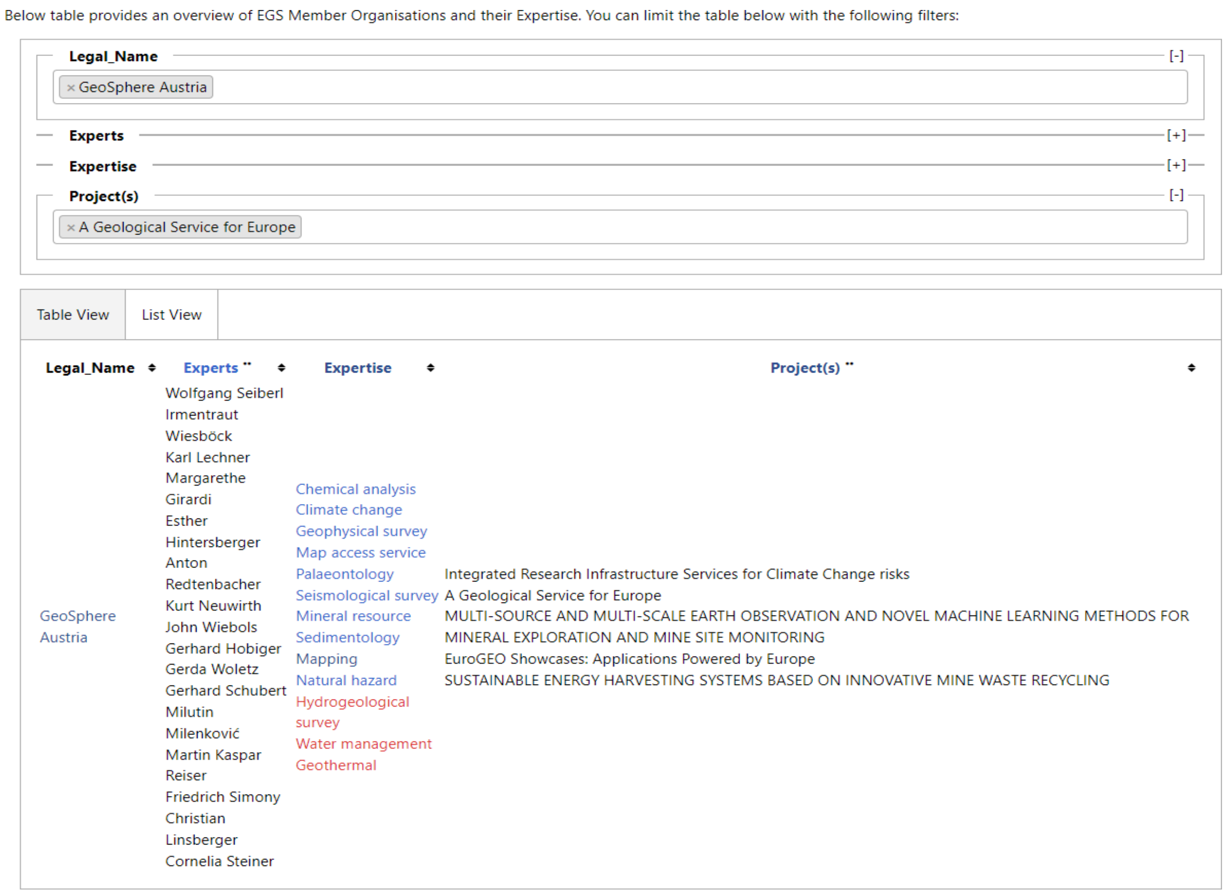

Figure 6: Use-Cases 1 of WP9.1 Task

4.2. Use-Case for Project – Search for Project Partners

The situation of this use case was: a GSO has an idea for a project but experts from other GSO(s) are needed. By doing a search in the database, possible partners can be found and information on specific expertise at a certain GSO is given.

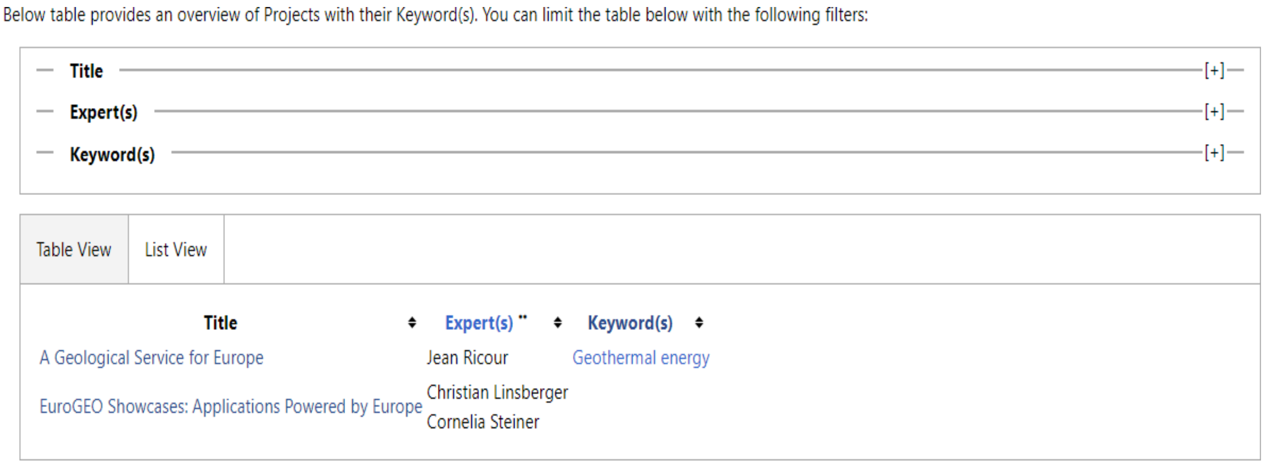

Figure 7: Use-Cases 2 of WP9.1 Task

4.3. Use-Case for Expert – Search for Expertise

Here is an example of a possible question: The European Commission (or a general external stakeholder) asks GSE for an expert on a specific topic for a specific event, review activity, interview, etc. Information on specific expertise at a certain GSO is given. Contact information is provided.

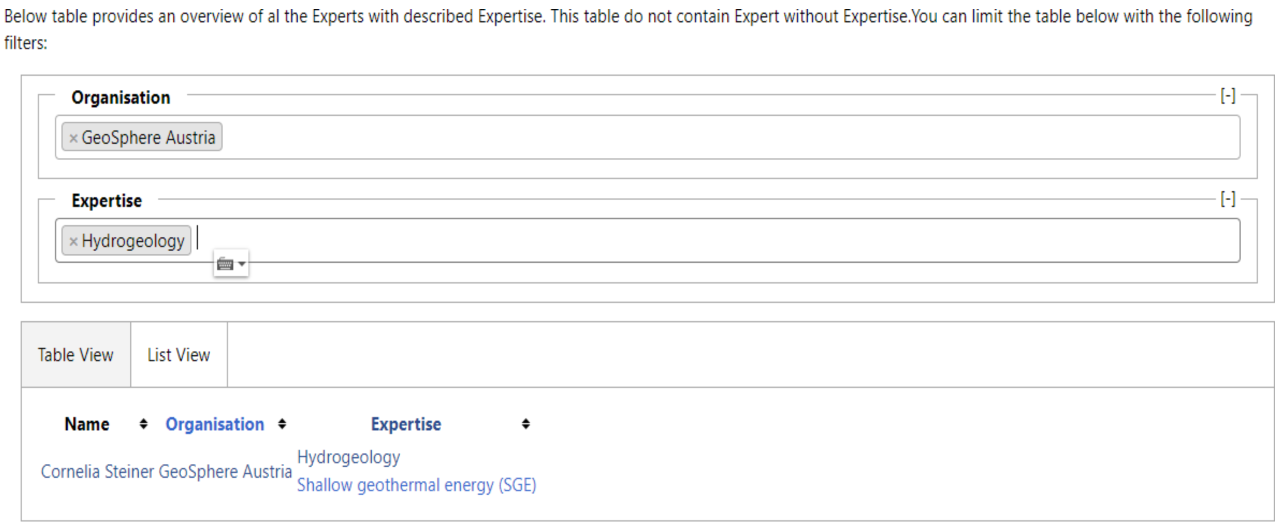

Figure 8: Use-Cases 3 of WP9.1

Another example we suggest is a web application as a map for the spatial search for organizations with references and links to expertise (derived from their completed projects, published data or their experts). This example of a map presentation of organizations and their project partners would also fit in with the implementation of the "Yellow Pages" described in Report D3.5.

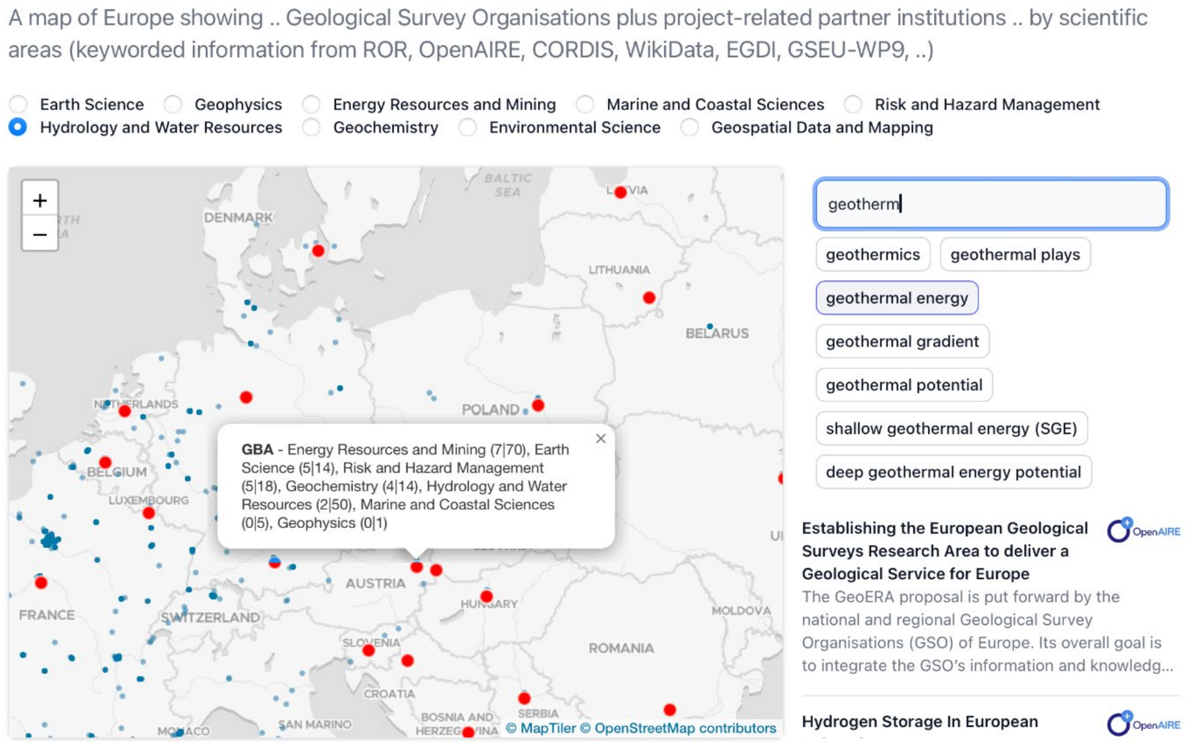

Figure 9: A webmap application to search organization's expertise by scientific area or keywords"

4.4. User Management

EGS and GSE

Members in this group have all the rights over Organization and Project category. However, they have limited rights over Expert category. They are only allowed to add/edit their own experts and not the experts of any geological surveys. Figure 10 shows the claim.

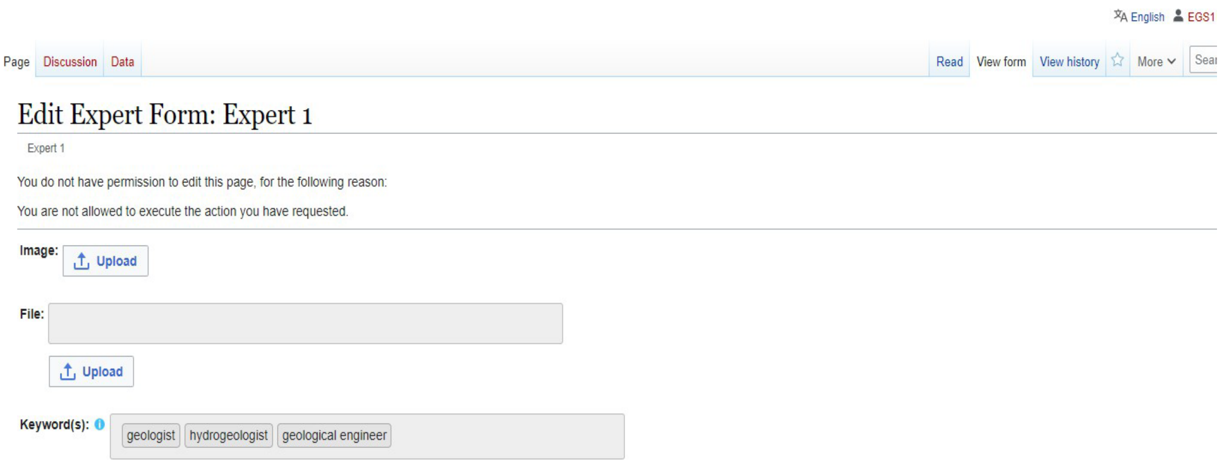

Figure 10: User Management - EGS expert access limitation

User Management – National Delegates

Member in this group have all the rights over Project category. However, they are restricted to edit their organization and expert pages only.

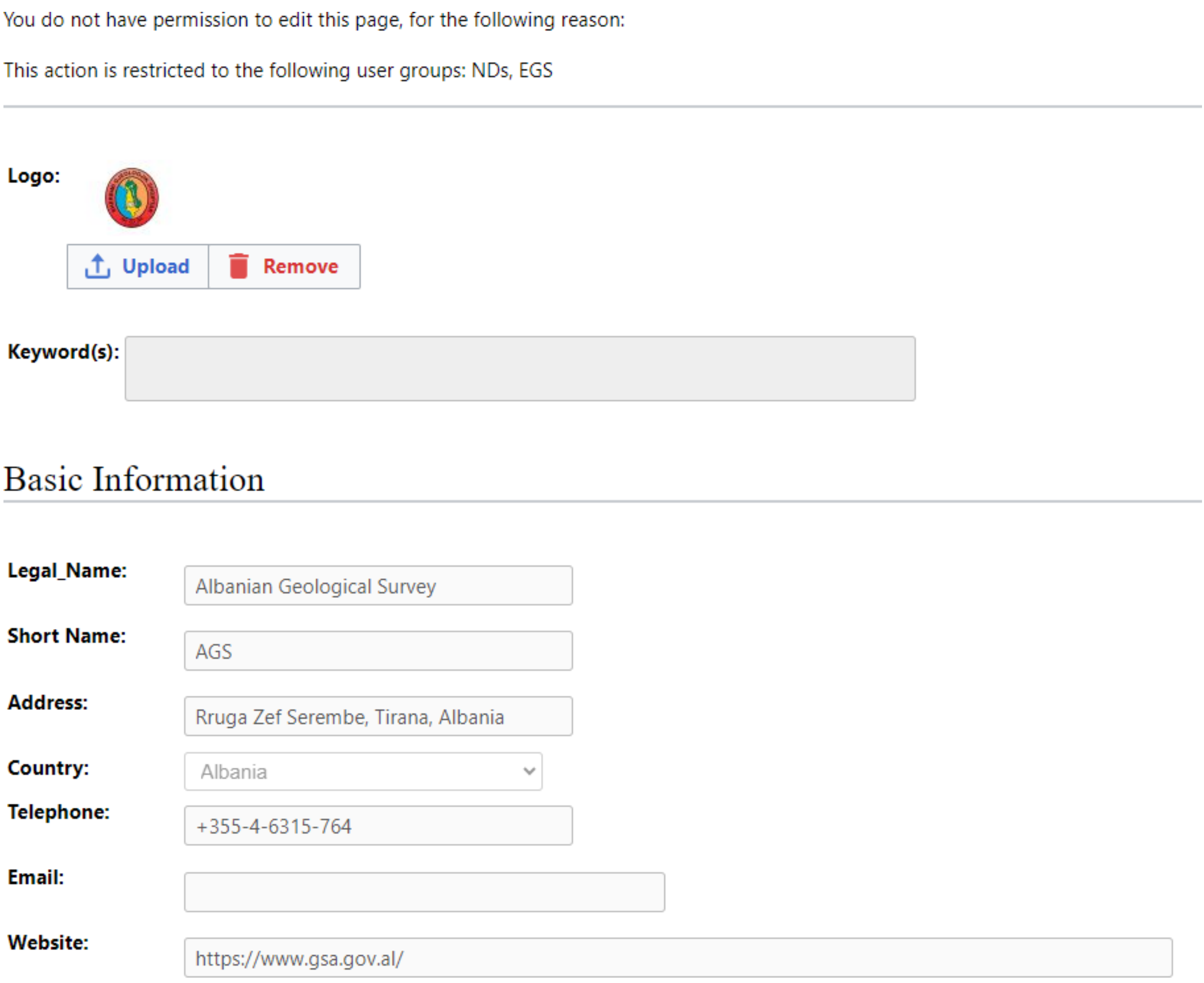

Figure 11: User Management - ND Page Access

User Management – GSO

Member in this group have every right to their own personal pages and do not have edit rights to other categories like their organizations pages or project pages. See further details in the related Annex.

5. Key Considerations for Identifier Selection

When selecting identifiers and URIs, consider their readability, stability, and scalability. Ensure they are easy to understand and remember, remain consistent over time, and can accommodate future growth or changes in your system. This enhances accessibility and reliability for users and systems interacting with your resources. In this section, we will recommend identifiers for experts, organizations, and projects that can link scattered systems more efficiently. Identifiers and URIs are needed to uniquely reference and locate resources on the web, ensuring consistency and enabling precise linking and retrieval of information. In simple words: how can we store knowledge data about organizations, projects or experts, without identifying them unambiguously?

5.1. Uniform Resource Identifier

A URI (Uniform Resource Identifier) is a string of characters used to identify a resource on the internet. Think of it as the web's address system. Just like a home address helps you locate a house, a URI helps you locate a resource, like a webpage or an online file. URIs are crucial because they provide a standard way to access resources, ensuring that different systems can communicate and share information efficiently. They make it easy for users to find, interact with, and share web resources, thus enabling a seamless online experience. Imagine you want to visit a friend's blog. The URI for their blog might look like this: http://www.friendsblog.com.

- **http://** This specifies the protocol, which tells your browser how to retrieve the resource.
- **www.friendsblog.com** This is the address of the blog.

When you type this URI into your web browser, it knows exactly where to go to find your friend's blog. This standard format makes it easy to locate and access online resources.

URIs identify not just Web documents, but also real-world objects like people and cars, and even abstract ideas and non-existing things like a mythical unicorn. We call these real-world objects or things. Two key things about URI:

- **Be on the Web:** Given only a URI, machines and people should be able to retrieve a description about the resource identified by the URI from the Web. Such a look-up mechanism is important to establish shared understanding of what a URI identifies. Machines should get RDF data and humans should get a readable representation, such as HTML. The standard Web transfer protocol, HTTP, should be used.
- **Be unambiguous:** URIs are meant to uniquely identify only single resource, so one URI can't stand for more than one object.

5.1.1. Persistent URI

A Persistent URI is a special type of URI that remains unchanged over time, even if the resource it points to moves to a different location. This ensures long-term access to the resource without the need to update links. Imagine an online research paper with a persistent URI like http://doi.org/10.1234/abcd. Even if the paper is moved to a different server, the URI will still point to the correct resource, ensuring that anyone using the link can always find the paper.

5.2. Challenges in Finding Persistent URIs

Assigning clear and persistent identifiers is crucial for accurately mapping relationships among organizations, projects, and experts. Without them, databases struggle to store structured knowledge, especially when organizations change through mergers or rebranding—as illustrated by GeoSphere Austria's formation from GBA and ZAMG in 2023. Projects face similar challenges due to incomplete metadata and inconsistent naming conventions. In contrast, experts can be reliably identified using ORCID IDs, similar to how DOIs work for scientific publications. The need for ongoing updates and data cleansing, as seen in WikiData and OpenAIRE, highlights the complexity of maintaining these identifiers.

5.3. Proposal for New URI Pattern

In WP9.1 case, a new solution to above issue is proposed. The solution is based on the principles of the Semantic Web and the proper implementation via "Linked Data" in RDF format as a standard.

- **URI for Organizations:** In almost all cases tested, organizations and institutions are represented on the web by an institutional homepage. It is noticeable that their domain name often is only changed if this is accompanied by a change in the organization. One possible solution would therefore be to create a new identifier using the URL domain of the homepage (domain name, top-level domain plus subdomain if not www, to lower case). Protocol and paths are omitted. An example for the new GeoSphere Austria would be <https://org.europe-geology.eu/geosphere-at> as URI. And so to distinguish the predecessor organization GBA by using <https://org.europe-geology.eu/geologie-ac-at>. In a later stage of GSEU the Expertise Hub could be extended with information of organizations regarding laboratory equipment or number of employees by category.

- **URI for Projects:** In a similar way to organizations, this could work with the identification of scientific projects. In scientific text publications, it is common practice to identify projects using their acronyms. Our suggestion is therefore to define URIs using these acronyms (to lower case, replaced special characters). As an example the GSEU project as <https://proj.europe-geology.eu/gseu>. This solution also contains a small residual risk in case that different projects (but usually in different scientific fields) use the same acronyms. An example we discovered here is e.g. the GENESIS project.

In summary, however, we see great benefits of a technically less complicated solution via domain names for organizations and acronyms for projects. One advantage is the better performance of analytical queries when it is no longer necessary to find out the relation between the different identifiers but of the same organization. Another advantage is the easier creation or matching of identifiers when entering and updating the system. Additionally the identifiers (URIs) of the different systems to import (from WikiData, OpenAIRE, metadata catalogue, central database, code list registry, WP9 wiki, etc.) are set in relation to each other and stored separately as a so-called mapping graph (using owl:sameAs).

6. SWOT Analysis

The SWOT analysis is a strategic tool used to assess **Strengths** (e.g., existing resources) and **Weaknesses** (e.g., technical limitations) of the project, as well as external **Opportunities** (e.g., market demand) and **Threats** (e.g., competition or regulatory challenges). This analysis aids in evaluating the project's viability by highlighting key factors that could influence its success or failure.

Figure 12: SWOT Analysis of POC

We conducted a SWOT analysis as part of our feasibility study to assess the use of MediaWiki for defining, editing, and searching information about projects, organizations, and experts. This analysis helped us evaluate the platform's strengths and weaknesses, as well as the opportunities and threats associated with its usage in our POC.
The Proof of Concept (POC) for the GSEU project, WP9 Task1, successfully demonstrated the feasibility of creating a centralized Expertise Hub for geological surveys across Europe. By integrating diverse data silos and utilizing Semantic MediaWiki for efficient data management, the project provided a unified platform that consolidates information about experts, organizations, and projects. The SWOT analyses conducted on both MediaWiki and the POC highlighted the strengths and opportunities of these tools in supporting structured data management while acknowledging the technical challenges and maintenance requirements. The use of a knowledge graph, supported by GraphDB, allowed for efficient querying and relationship traversal, ensuring comprehensive data accessibility. The showcase of data governance processes, including access controls and user group permissions (for better integration into the EGDI, the next step is to use the EGDI Authentication and authorization system), further ensured data privacy and quality. For better integration into the EGDI, the next step is to use the EGDI Authentication and authorization system. Overall, this POC has established a solid foundation for future development, showcasing the potential for improved collaboration and data utilization among geological survey organizations.

In addition, the feasibility study showed various advantages of using a Wiki compared to using a simple database. Wiki facilitates collaborative editing by allowing multiple users to contribute and update content in real time, additionally with version control. With an intuitive interface aimed at non-technical users, it supports the embedding of rich media and includes discussion pages for feedback and collaboration. Wiki also stores a history of changes so that users can revert to previous versions. In contrast, a simple database is characterized by the storage of structured data, but does not have tools for dynamic content creation and collaborative editing. While a database can establish relationships between data points, it does not support the narrative-style linking and navigation that a Wiki provides. Ultimately, a Wiki is more effective in environments where dynamic content sharing and collaboration is essential, while a simple database is best suited for organized data storage and retrieval.

7. Key Findings & Future Outlook

The feasibility study and proof of concept (POC) for the Geological Services Knowledge Hub have demonstrated the viability of a centralized digital platform for expertise discovery, collaboration, and data management across European Geological Survey Organizations (GSOs). The results highlight the transformative potential of this initiative in enhancing accessibility to expertise, fostering interdisciplinary cooperation, and improving strategic decision-making at both national and European levels.

7.1. Key Findings

1. **Successful Validation of Concept**
   - The POC confirmed that Semantic MediaWiki combined with GraphDB provides an effective infrastructure for managing and querying geological expertise and project data.
   - The proposed system integrates structured and unstructured datasets, enhancing searchability and knowledge sharing across various GSOs.
2. **Improved Data Accessibility and Interoperability**
   - The Expertise Search Service enables the identification of key experts, projects, and research activities across Europe, reducing information silos and improving collaboration.
   - The use of Persistent URIs and knowledge graphs enhances interoperability, ensuring data consistency across various platforms such as OpenAIRE, Kohesio, and Wikidata.
3. **Enhanced Governance and Security**
   - The implementation of role-based access control (RBAC) ensures that data governance remains robust, allowing different levels of data visibility and editing rights for GSOs, National Delegates, and European Geological Surveys (EGS).
   - The ability to track revisions and user contributions increases transparency and accountability.
4. **Strategic Alignment with the European Geological Data Infrastructure (EGDI)**
   - The project lays a foundation for better integration with EGDI, supporting a unified approach to European geological data management.
   - The proposed framework can be extended beyond expertise search to include additional services such as laboratory facilities, research collaborations, and data repositories.
5. **SWOT Analysis Insights**
   - The strengths of the platform include its scalability, ease of collaboration, and structured data management.
   - Opportunities lie in extending functionalities, including a marketplace for geological services and a geohazard risk assessment module.
   - Challenges include the need for sustained data governance and regular updates to ensure long-term usability.
   - Potential threats include resource constraints for maintenance and ensuring long-term engagement from all stakeholders.

7.2. Future Outlook & Next Steps

As the project transitions from feasibility to full-scale implementation, several key actions will be critical in ensuring its long-term success:

1. **Integration with EGDI and Other European Geological Initiatives**

- Close collaboration with WP7 will be necessary to align the Knowledge Hub with the European Geological Data Infrastructure.
- Expanding the system to incorporate real-time data updates, automated metadata validation, and enhanced API connectivity will further strengthen its capabilities.

2. **Expanding the Scope of Services**
    - Beyond expertise identification, the platform could evolve into a multi-functional knowledge hub, supporting:
    - A marketplace for geological services (e.g., laboratory equipment sharing, research collaborations).
    - Policy-supporting tools for critical raw materials and geohazard risk assessments.
    - Interactive visualization tools to improve stakeholder engagement.
3. **Data Quality and Sustainability Strategy**
    - A long-term data governance framework will be required to maintain data accuracy, security, and compliance with FAIR (Findable, Accessible, Interoperable, and Reusable) principles.
    - Establishing a dedicated user support and curation team will ensure continued engagement from GSOs and research institutions.
4. **Technical Enhancements & AI-driven Capabilities**
    - Future iterations of the system could leverage AI-powered recommendations for matchmaking experts and projects, streamlining collaboration opportunities.
    - Natural Language Processing (NLP) could be used to extract insights from publications, linking experts to relevant topics dynamically.
5. **Stakeholder Engagement & Funding Opportunities**
    - Sustained engagement with GSOs, policymakers, and the European Commission will be crucial in securing long-term funding.
    - Developing a business model for sustainability, such as a membership-based system or partnerships with industry, could enhance financial viability.

7.3. Conclusion

The Geological Services Knowledge Hub represents a major advancement in European geological collaboration, breaking down information silos and enabling efficient knowledge sharing. By leveraging semantic technologies, structured data frameworks, and robust governance models, the platform is poised to become a cornerstone for geological research, policy-making, and innovation.

The next phase of implementation under WP7 will be crucial in scaling up the system, ensuring its long-term sustainability, and expanding its impact across geosciences, environmental monitoring, and strategic policy development. With continued investment and stakeholder engagement, the Knowledge Hub will significantly contribute to evidence-based decision-making in Europe's geological and environmental sectors.

8. References

[1] Schmid, L.: Evolution of Cooperation via (in)Direct Reciprocity under Imperfect Information. Institute of Science and Technology Austria (2021). https://doi.org/10.15479/at:ista:10293.

[2] Schmid, L., Shati, P., Hilbe, C. et al. The evolution of indirect reciprocity under action and assessment generosity. Sci Rep 11, 17443 (2021). https://doi.org/10.1038/s41598-021-96932-1

[3] Petrillo, M. (12.1.2025). Nine FAQs about Knowledge Graphs for the Enterprise. https://www.linkedin.com/pulse/nine-faqs-knowledge-graphs-enterprise-matthew-petrillo-vu5fe?trk=public_post_main-feed-card_reshare_feed-article-content

[4] Engelman, A., Enkvist, C., Pettersson, K.: A fair archive based on the cerif model. Procedia Computer Science 146, 190–200 (2019)

[5] Leigh, D.: Swot analysis. Handbook of Improving Performance in the Workplace: Volumes 1-3 pp. 115–140 (2009)

9. Annex I – Consortium Partners

	Partner Name	Acronym	Country
1	EuroGeoSurveys	EGS	Belgium
2	Nederlandse Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek	TNO	Netherlands
3	Sherbimi Gjeologjik Shqiptar	AGS	Albania
4	Vlaamse Gewest	VLO	Belgium
5	Bureau de Recherches Géologiques et Minières	BRGM	France
6	Ministry for Finance and Employment	MFE	Malta
7	Hrvatski Geološki Institut	HGI-CGS	Croatia
8	Institut Royal des Sciences Naturelles de Belgique	RBINS-GSB	Belgium
9	Państwowy Instytut Geologiczny – Państwowy Instytut Badawczy	PGI-NRI	Poland
10	Institut Cartogràfic i Geològic de Catalunya	ICGC	Spain
11	Česká Geologická Služba	CGS	Czechia
12	Department of Environment, Climate and Communications - Geological Survey Ireland	GSI	Ireland
13	Agencia Estatal Consejo Superior de Investigaciones Cientificas	CSIC-IGME	Spain
14	Bundesanstalt für Geowissenschaften und Rohstoffe	BGR	Germany
15	Geološki zavod Slovenije	GeoZS	Slovenia
16	Federalni Zavod za Geologiju Sarajevo	FZZG	Bosnia and Herzegovina
17	Istituto Superiore per la Protezione e la Ricerca Ambientale	ISPRA	Italy
18	Regione Umbria	-	Italy
19	State Research and Development Enterprise State Information Geological Fund of Ukraine	GIU	Ukraine
20	Institute of Geological Sciences National Academy of Sciences of Ukraine	IGS	Ukraine
21	M.P. Semenenko Institute of Geochemistry, Mineralogy and Ore Formation of NAS of Ukraine	IGMOF	Ukraine
22	Ukrainian Association of Geologists	UAG	Ukraine

23	Geologian Tutkimuskeskus	GTK	Finland
24	Geological Survey of Serbia	GZS	Serbia
25	Ministry of Agriculture, Rural Development and Environment of Cyprus	GSD	Cyprus
26	Norges Geologiske Undersøkelse	NGU	Norway
27	Latvijas Vides, ģeoloģijas un meteoroloģijas centrs SIA	LVGMC	Latvia
28	Sveriges Geologiska Undersökning	SGU	Sweden
29	Geological Survey of Denmark and Greenland	GEUS	Denmark
30	Institutul Geologic al României	IGR	Romania
31	Szabályozott Tevékenységek Felügyeleti Hatósága	SZTFH	Hungary
32	Eidgenössisches Departement für Verteidigung, Bevölkerungsschutz und Sport	VBS (DDPS)	Switzerland
33	Elliniki Archi Geologikon kai Metalleftikon Erevnon	HSGME	Greece
34	Laboratório Nacional de Energia e Geología I.P.	LNEG	Portugal
35	Lietuvos Geologijos Tarnyba prie Aplinkos Ministerijos	LGT	Lithuania
36	Geosphere Austria - Bundesanstalt für Geologie, Geophysik, Klimatologie und Meteorologie	GeoSphere Austria	Austria
37	Service Géologique de Luxembourg	SGL	Luxembourg
38	Eesti Geoloogiateenistus	EGT	Estonia
39	Štátny Geologický ústav Dionýza Štúra	SGUDS	Slovakia
40	Íslenskar Orkurannsóknir	ISOR	Iceland
41	Instituto Português do Mar e da Atmosfera	IPMA	Portugal
42	Jarðfeingi	Jardfeingi	Faroe Islands
43	Regierungspräsidium Freiburg	LGRB	Germany
44	Geologischer Dienst Nordrhein-Westfalen	GD NRW	Germany
45	Landesamt für Geologie und Bergwesen Sachsen-Anhalt	LfU	Germany
46	Vlaamse Milieumaatschappij	VMM	Belgium
47	Norwegian Petroleum Directorate	NPD	Norway
48	United Kingdom Research and Innovation - British Geological Survey	UKRI-BGS	UK

10. Annex II – Supplementary Technical Details

10.1. User Management

Data governance processes such as access controls allow administrators to regulate user permissions effectively. This feature enables fine-tuned control over who can view, edit, or manage specific content, ensuring data accuracy. We have proposed three user groups with different permissions set for the wiki: EGS, National Delegates and GSOs, see Figure 13 to Figure 15.

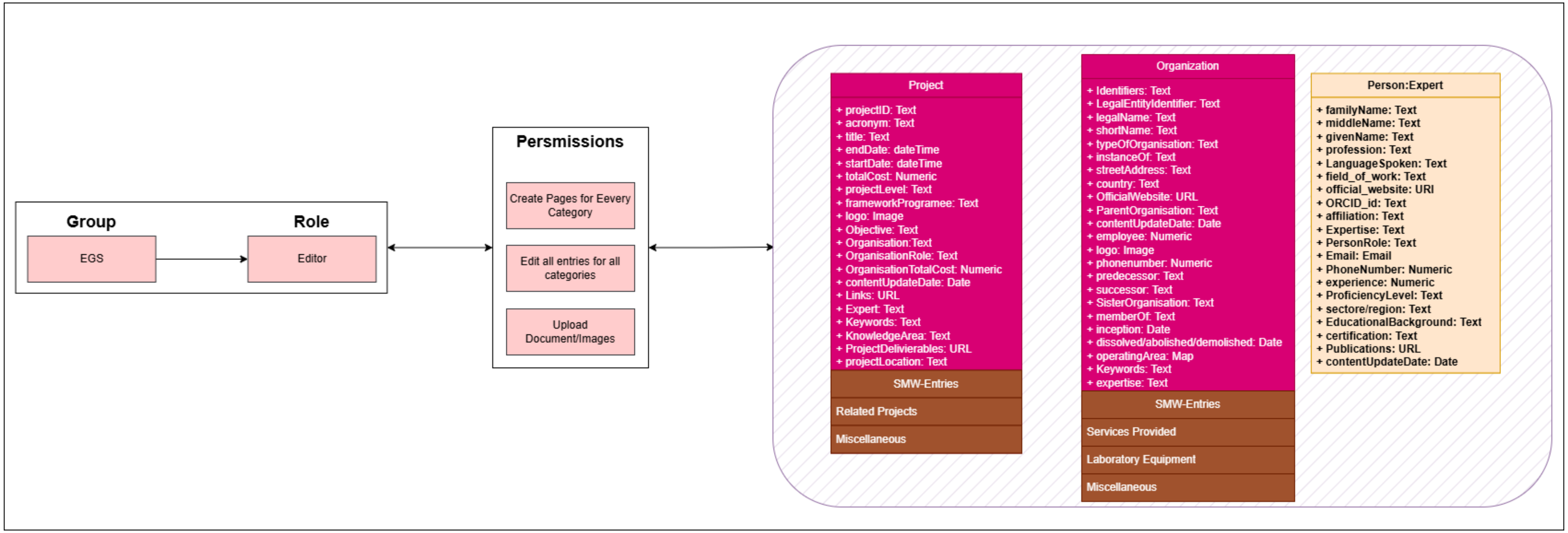

Figure 13: User Management for EGS

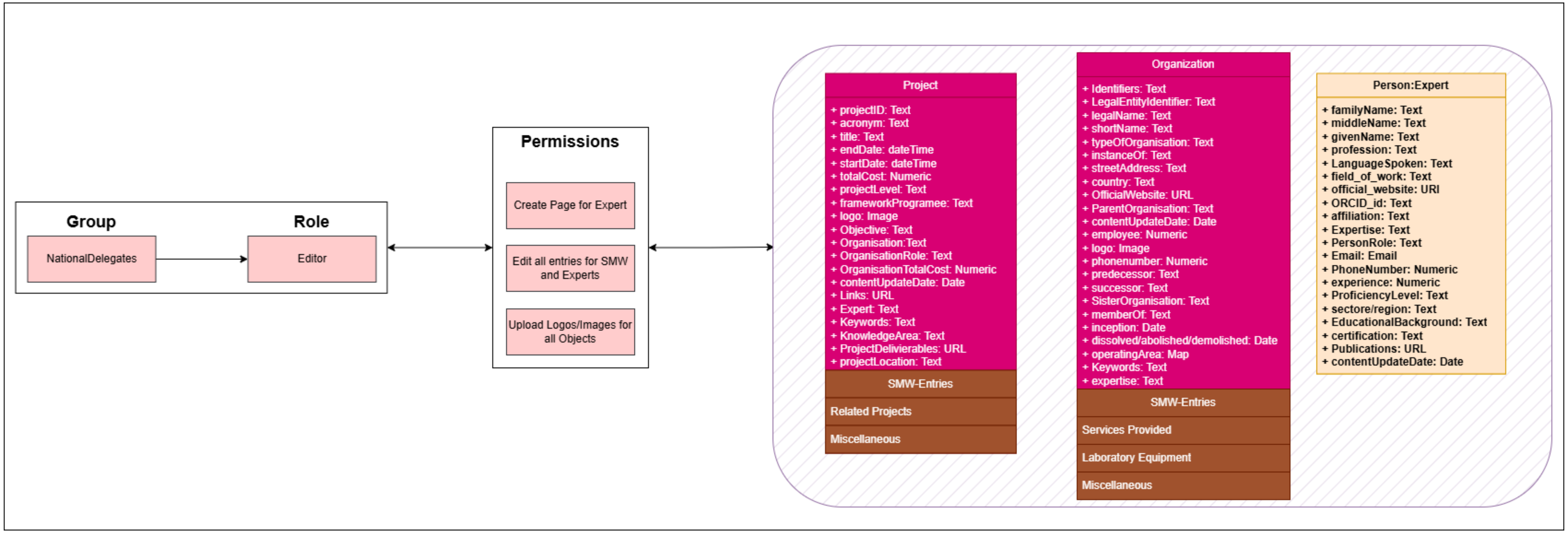

Figure 14: User Management for National Delegates

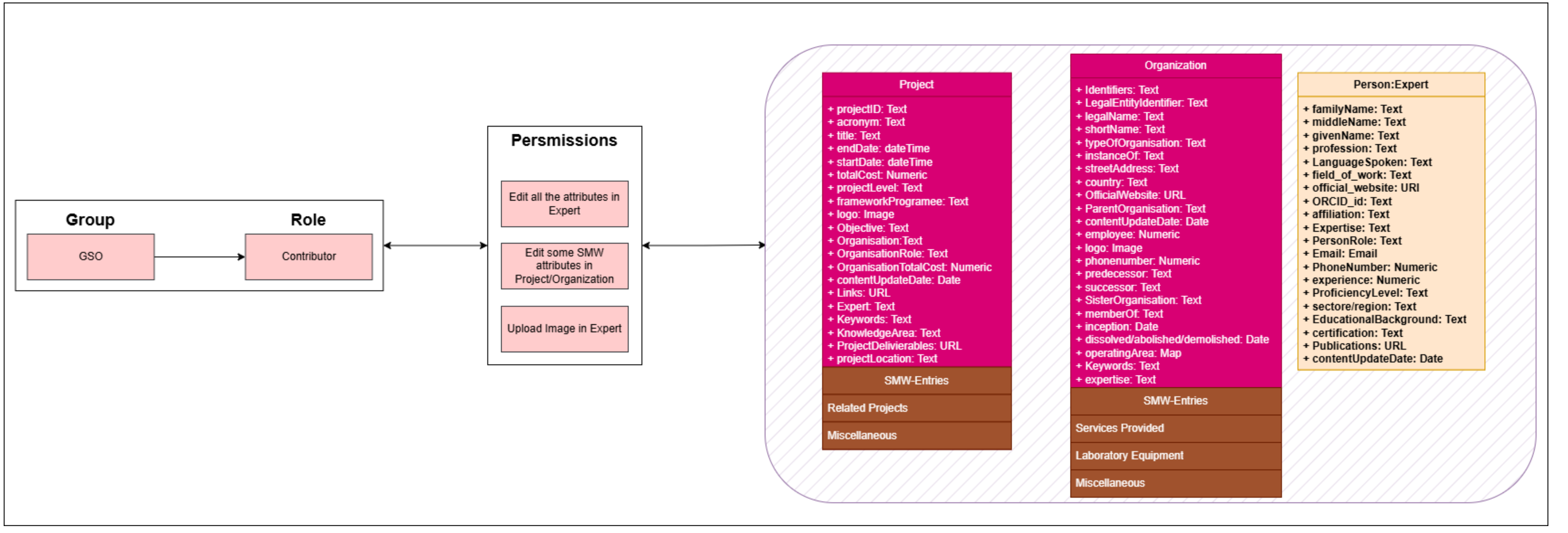

Figure 15: User Management for all Geological Surveys

10.2. MediaWiki Setup

MediaWiki, written in PHP, is compatible with all major operating systems supported by PHP. It relies on a database, which can be MySQL/MariaDB, PostgreSQL, or SQLite. For the WP9.1 Proof of Concept (POC), MySQL is utilized – MediaWiki uses MySQL to manage wiki-based applications, which enable users to collaboratively create, edit, and organize content in the form of interconnected web pages. In contrast, while MySQL stores web page structured data in tables, a Graph database models relationships between data as nodes and edges, providing a more flexible way to represent complex, interconnected information.

To streamline readability, interested readers are directed to the step-by-step download guide available at https://workingwithmediawiki.com/book/chapter2.html.

Product	Version
MediaWiki	1.39.5
PHP	8.1.2-1ubuntu2.17 (apache2handler)
MySQL	8.0.36-0ubuntu0.22.04.1
ICU	70.1
Lua	5.1.5
Pygments	2.11.2

Figure 16: Installed Software for Task 9.1

Technical note for developers: After downloading MediaWiki, simply enter the MW URL into a browser. The MediaWiki code should execute smoothly, prompting the search for a file named **LocalSettings.php**. This file serves as MediaWiki's initialization file, housing all user- modifiable settings for the wiki. Initially, LocalSettings.php is absent, signaling MediaWiki of a new installation. MediaWiki then guides users through a series of steps via the browser, where they specify the wiki's name, database name, and other configurations, including the first user's username and password. Upon completing setup, MediaWiki automatically generates the **LocalSettings.php** file, along with creating a new database in the database system.

Logo Setting

Logo is the customary way to individualize one's wiki. The simplest way to set is is to add something like this to **LocalSettings.php**:

$wgLogos['icon'] = '/path/to/your/logo';

$wgLogos['1px'] = '/path/to/your/logo';

The logo image can be located either within the MediaWiki directory, or at some arbitrary URL. By default, it is located at MW-setup/skins/common/images/wiki. png – you shouldn't replace that file with your logo image, though, because then you run the risk of it being overwritten when you update the MediaWiki code.

URL Structure

By default, MediaWiki URLs appear in a format like:

$*wgServer* = mywiki.com/mediawiki/index.php?title=Main_Page

However, we have preferred something like:

$*wgServer* = https://wp9.geoinformation.dev/

to show its WP9 POC.

Page History

Another notable feature accessible on every content page within MediaWiki is its comprehensive history page. This invaluable tool allows users to view the complete record of all modifications made to a page over time. This feature is important for a system because it ensures transparency, accountability, and traceability, allowing users to track changes, verify the accuracy of content, and resolve disputes by reviewing the complete history of modifications. This feature ensures transparency and accountability by maintaining a detailed log of edits, making it possible to track the evolution of content from its creation to its current state.

Each row on the history page corresponds to a specific edit made to the page, all of which are preserved indefinitely. Each row includes:

"cur" and "prev" links and radio buttons for comparing differences between revisions;

Date and time of the edit, linking to a detailed revision page;

Username of the editor;

An "undo" link for all edits except the earliest one.

In essence, the history page in MediaWiki not only serves as a safeguard against accidental deletions or modifications but also promotes a dynamic editing environment where ideas can evolve through collective effort, ensuring the integrity and continuity of information. It is important for basic information like institutions, projects, and staff, as it ensures accuracy, accountability, and easy tracking of any updates or corrections to this information over time.

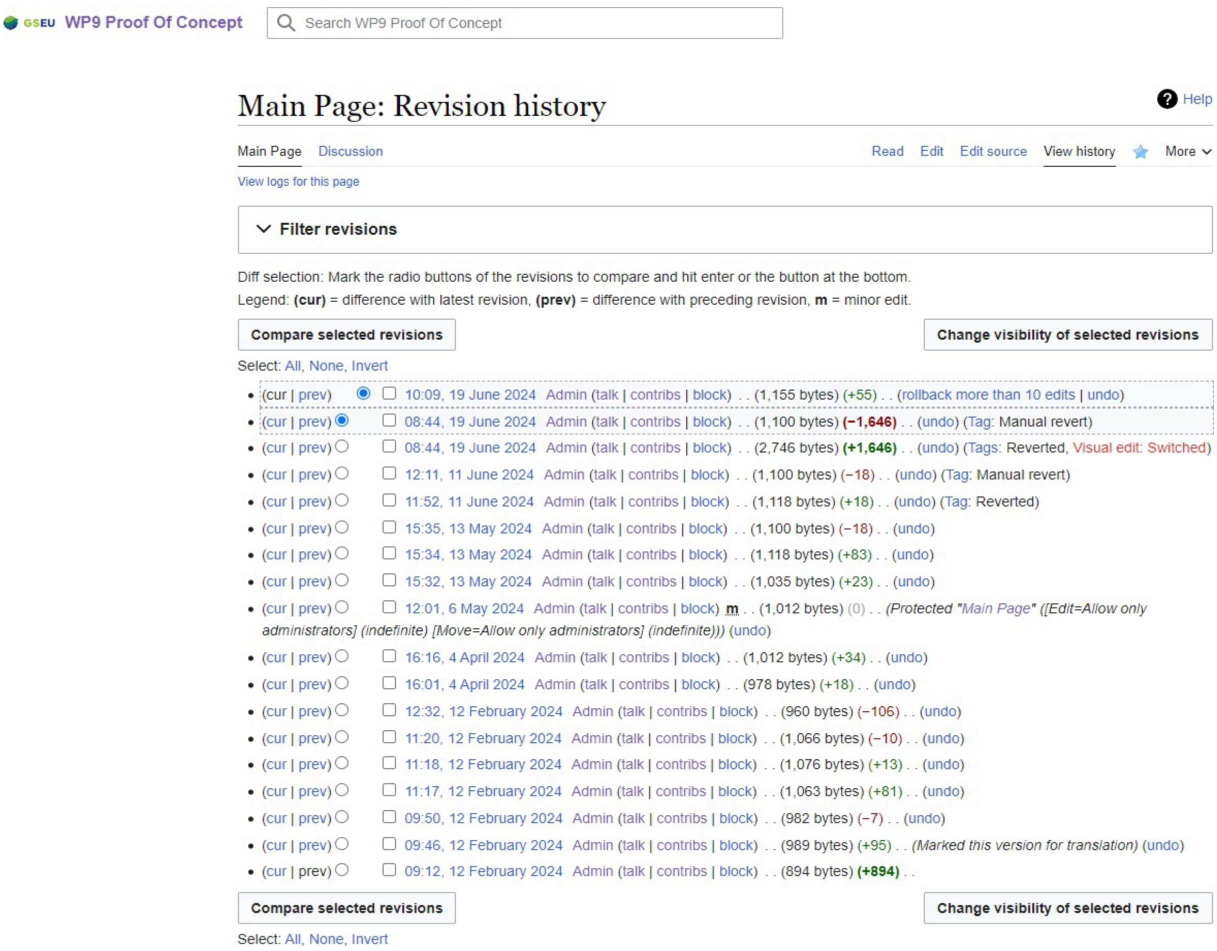

Figure 17: Page History Feature of MediaWiki

Skins

Skins allow users to customize the look and feel of MediaWiki. The default skin for MediaWiki is Vector, which offers a standard layout and design. However, MediaWiki installations typically include several additional skins, allowing users to select an appearance that best suits their preferences or needs. Users can view and choose from the available skins by accessing the preferences page within their MediaWiki account settings.

For example, a user may prefer the "MonoBook" skin, which provides a different layout and styling compared to the default Vector skin. By navigating to the preferences page, the user can select MonoBook or any other available skin to customize their browsing experience, thereby tailoring the interface to their personal or organizational requirements. This feature enhances user engagement by offering flexibility in the presentation of content.

10.3. Results

In the following, additional information on the User Management of GSOs is given:

Figure 18 shows that each expert when login into the system can view the pages that an expert is allowed to edit or manage under the heading name My pages.

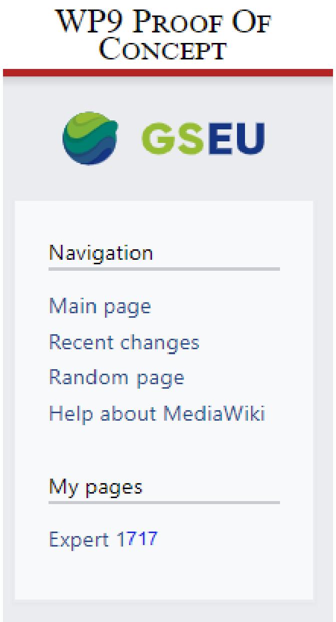

Figure 18: User Management - GSO page access list

Figure 19 shows that expert 1717 is allowed to edit or add any information on the page.

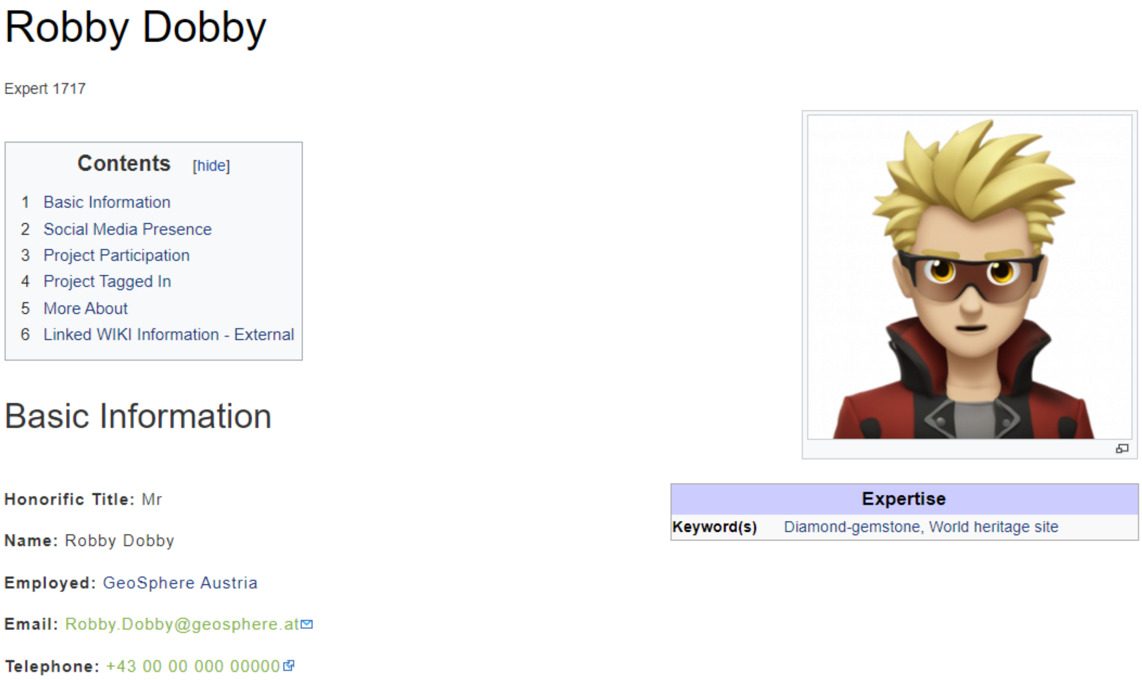

Figure 19: User Management - GSO expert own page access

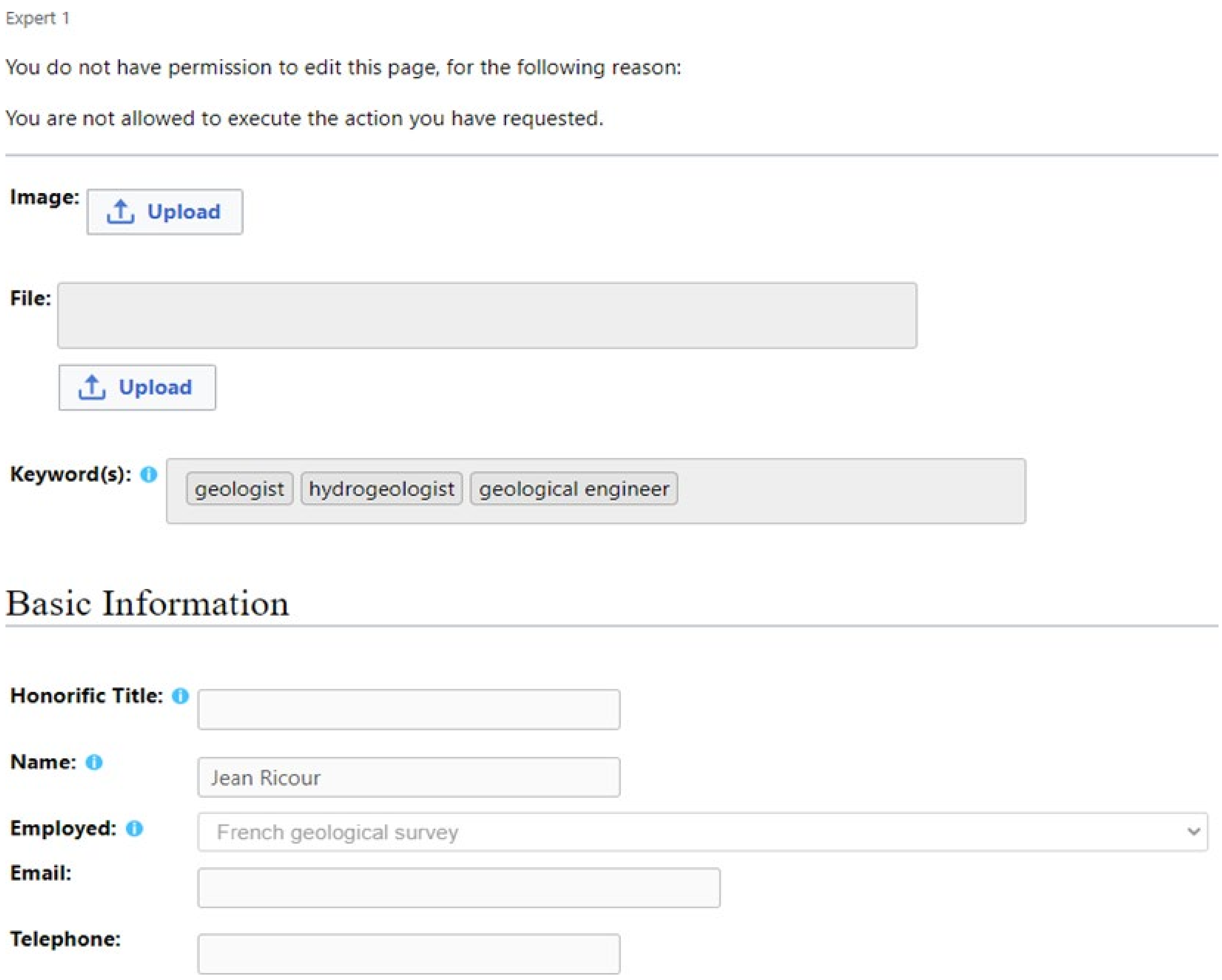

Figure 20 shows that every expert in this system other than expert 1 are not allowed to edit the information on this page.

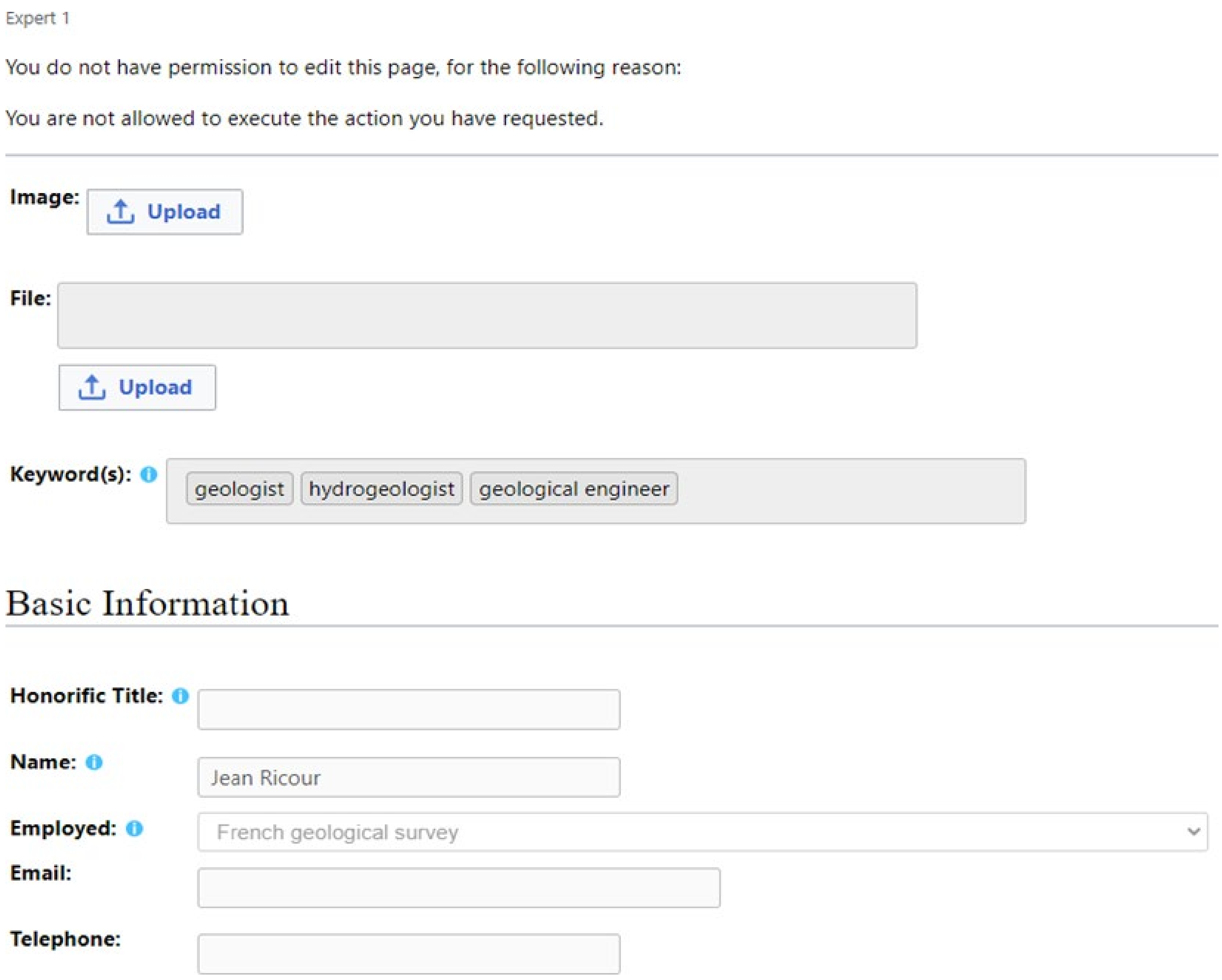

Figure 20: User Management - GSO other expert pages access

Figure 21 shows that experts are not allowed to edit the information on project pages.

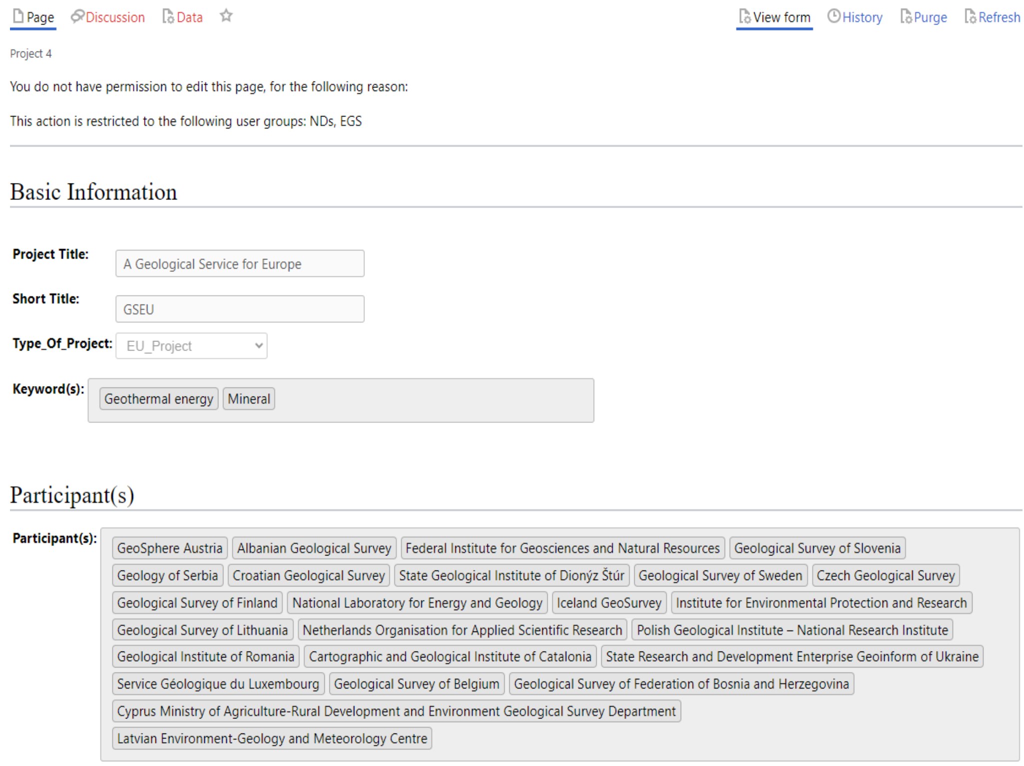

Figure 21: User Management - GSO projects access restriction

Figure 22 shows that experts are not allowed to edit the information on organisation pages.

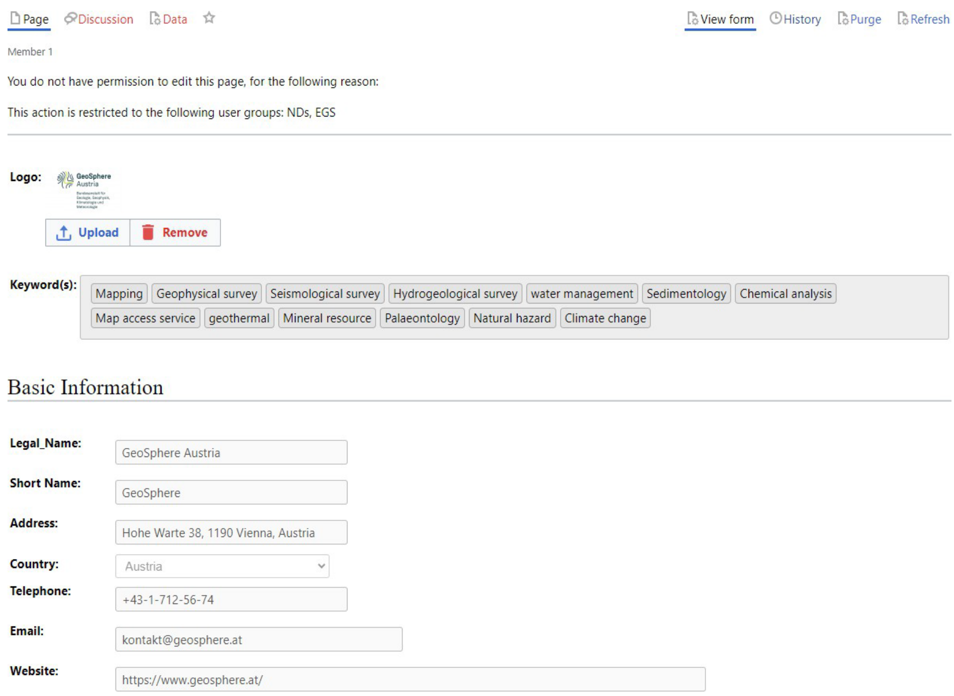

Figure 22: User Management - GSO organisations access restriction