

# Requirement Analysis and Specification Document

HAZARD MAPS PROJECT SE4GEO IN ITALY  
POLITECNICO DI MILANO

## AUTHORS

Giovanni Montefoschi

Zhanbin Wu

Xiao Tan

Firoozeh Rahimian

## INSTRUCTOR

Giovanni Quattrocchi

## DATE PREPARED

26/04/2024

## VERSION

0.0.0



## Table of content

Introduction .....	2
Context and motivations .....	2
Solution overview .....	2
Scope and limitations .....	3
Stakeholders .....	3
User characteristics.....	4
Use cases .....	5
Use case diagram .....	5
Use cases .....	6
User stories.....	10
Requirements and domain assumptions .....	12
Functional requirements.....	12
Technical requirements .....	13
Domain assumptions .....	13

# Introduction

## Context and motivations

Italian public authorities collect, process, and analyze hazard-related datasets to produce landslide hazard maps. By combining the hazard maps with population and building censuses on the regional level of Italy, one can assess the exposure to a relevant hazard. Therefore, decision makers require effective means to access, visualize and analyze this information.

## Solution overview

Component	Description	Technologies/Tools
Database	Stores pre-ingested data from the IdroGEO API, specifically designed to handle geographical data.	PostgreSQL, PostGIS
Web Server (Backend)	Manages API requests, data processing, and interacts with the database to fetch and send data. Handles data cleaning and preprocessing.	Flask (Python), REST API
Dashboard (Frontend)	Provides a user interface for data interaction and visualization. Enables users to perform queries and view results through maps, charts, and other graphical displays.	Jupyter Notebooks, Matplotlib, Plotly, Folium, Jupyter Widgets
Data Management	Processes and manipulates data for analysis and visualization, handling complex data operations efficiently.	Pandas, GeoPandas, Dask, Xarray
Integration Strategy	Ensures seamless data flow between the backend and the frontend and outlines how different system components communicate.	API connections, CI/CD practices
Development Approach	Utilizes Agile methodologies, focusing on iterative development with regular reviews and user feedback integration.	Agile methodologies, GitHub for version control, Visual Studio Code
Scalability and Security	Addresses how the system will handle increased data loads and security measures to protect data integrity and user privacy.	Scalability solutions, data security protocols

## Scope and limitations

The system will include a database, a RESTful web server, and a Jupyter Notebook-based dashboard that interact to provide real-time data processing and visualization capabilities.

## Stakeholders

For our project, the user groups we identify as stakeholders are the following:

### 1) Civil Protection Analysts

Professionals involved in disaster management, focusing on prevention and preparedness.

Role	Civil Protection Analysts
Technical Proficiency	High
Needs	Access to real-time data, ability to perform complex queries and receive quick responses
System Interaction	Frequent use of dashboard for visualization and data analysis features

### 2) Insurance Company Analysts

Users who evaluate risks related to natural disasters to adjust insurance policies and coverage.

Role	Insurance Company Analysts
Technical Proficiency	Moderate to high; need to use analytical tools
Needs	Detailed reports, historical data analysis, and risk forecasting
System Interaction	Regular use for accessing reports and performing risk assessments

### 3) Government Officials

Decision-makers who need to understand hazard exposures to better plan and implement safety measures.

Role	Government Officials
Technical Proficiency	Low to moderate; primarily interested in summary information and conclusions
Needs	High-level overviews, summaries, and the ability to drill down into data as needed

System Interaction	Occasional use with focus on dashboard summarizations and key metrics
--------------------	---

#### 4) Technical Support Staff

Individuals who maintain and troubleshoot the system, ensuring it operates smoothly.

Role	Technical Support Staff
Technical Proficiency	Very high; in-depth understanding of both the front-end and back-end systems
Needs	Detailed system logs, alerting mechanisms, and efficient ways to update or troubleshoot the system
System Interaction	Daily interaction with more technical and administrative components of the system

## User characteristics

The users of our application will be of two kinds. One for accessing and analyzing the data, which is the bulk of the application. The other for system administrators, who will need the capability to maintain the application's database to assure its data is of quality and availability.

### 1) Main User

This user seeks access to data regarding PIR Hazard and Risk indicators. Therefore, his actions can be divided in browsing the available data, performing queries in a database, performing analysis on the data, customizing the data view and finally downloading the desired data.

Specifically, while browsing the data, the user must be able to view information about different datasets such as risk areas for landslides.

Once a dataset is selected by the user, he can further perform queries on it using either spatial attributes such as the administrative level, or other non-spatial attributes such as population counts. These non-spatial attributes are also browsable and can be displayed in custom views on the application's map.

Once a query is performed the application offers the user the possibility to perform some analysis on the data, such as extrapolating statistics based on risk levels, population, and other information.

Finally, the user can customize the data view on the application's map to have a better picture of the analysis performed, which can be saved and utilized outside of this application's domain.

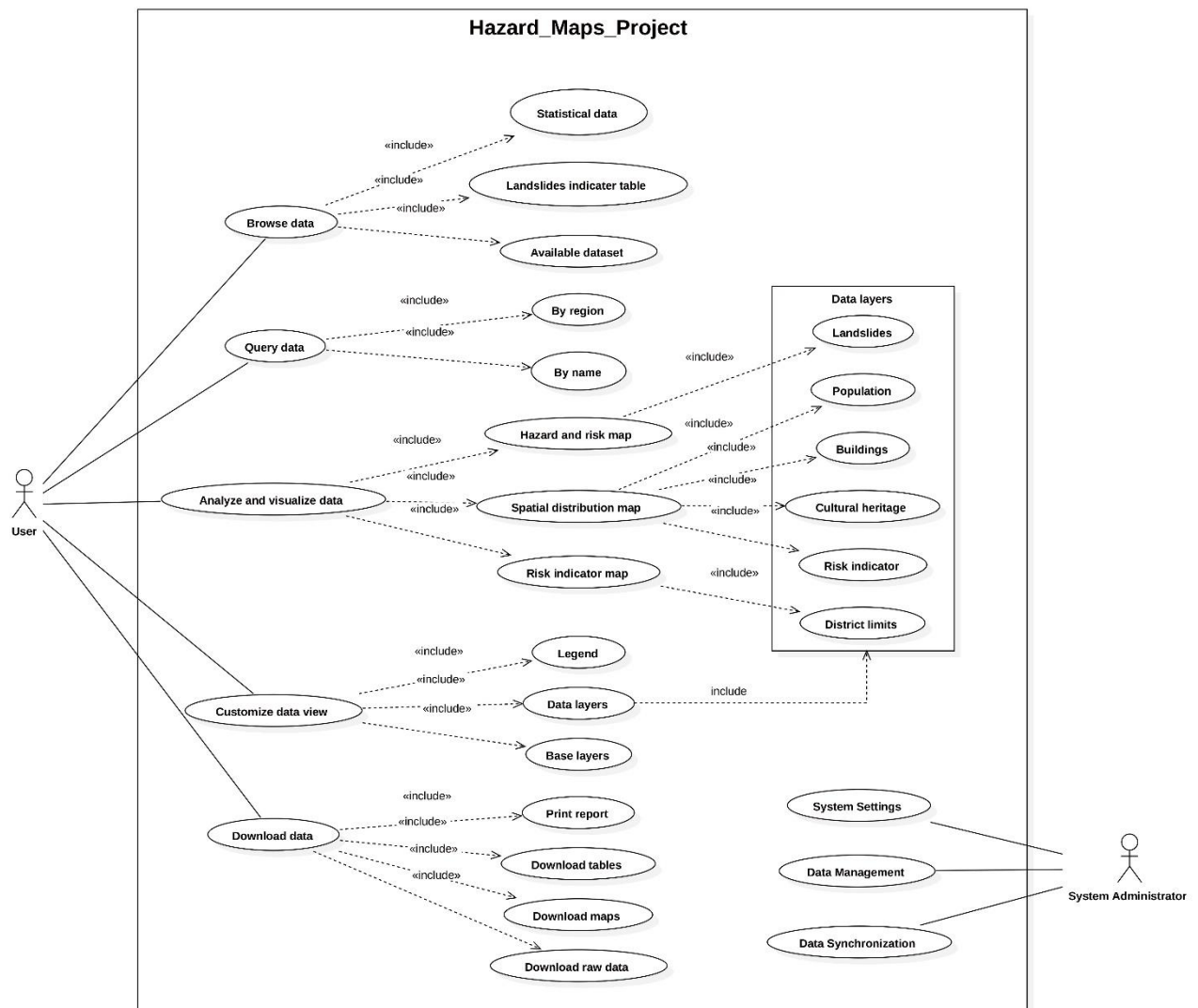
The user may want to download the results of the queries he performed while using the application and save them in different data formats such as csv, xml or json.

## 2) System administrator

As a system administrator a user can perform actions that are related to monitoring the status of the application and of its database. Update and synchronize the data with that provided by authorized public authorities and access the database for its proper management.

## Use cases

### Use case diagram



## Use cases

### 1) Business Use Cases

Use Case 1	Browse data
Participant	User
Goal	Allow users to browse available disaster data in the system.
Entry Condition	Relevant data from the IdroGEO API PIR is present in the database.
Flow of Events	<p>The user accesses the system's data browsing page.</p> <p>The system displays a list of available disaster datasets, which include statistical landslide indicator table and available datasets for user to download.</p> <p>The user selects a desired disaster dataset to view detailed information.</p>
Exceptions	If no disaster datasets are available in the system, an appropriate prompt message is displayed to the user.
Exit Condition	The user successfully browses available disaster datasets.

Use Case 2	Query dataset
Participant	User
Goal	Allow users to query landslide disaster data at region levels in Italy, and return the analyzed map to the user.
Entry Condition	Relevant data from the IdroGEO API PIR is present in the database.
Flow of Events	<p>The user accesses the system's data query page.</p> <p>The user selects the regional level or the name of location to query.</p> <p>The user selects the census group to query (population, buildings, or both).</p> <p>The system executes the query based on the user's selections and retrieves relevant map from the database.</p> <p>The system returns the map to the user.</p>
Exceptions	If the system fails to retrieve relevant data, an appropriate error message is displayed to the user. If the selected administrative level, disaster type, or census group is unavailable, the system displays corresponding prompt messages to the user.
Exit Condition	The user successfully queries disaster data.

<b>Use Case 3</b>	<b>Analyze and visualize data</b>
Participant	User
Goal	Allows users to analyze and visualize the disaster data they query for deeper insights, and to view and further process the analysis results on the Jupiter Notebooks dashboard.
Entry Condition	The user successfully queried the disaster data. Procedure
Flow of Events	Users navigate to data analysis and visualization tools. The user selects the dataset to be analyzed and visualized. Users perform the required analytical map, such as hazard and risk map, spatial distribution map and risk indicator map. Users view the analysis results on the Jupiter Notebooks dashboard for further processing and visualization.
Exceptions	If the data set selected by the user cannot be analyzed or visualized, the system displays the appropriate error message to the user.
Exit Condition	Users successfully analyzed and visualized disaster data.

<b>Use Case 4</b>	<b>Customize data view</b>
Participant	User
Goal	Allows users to customize the display of disaster data according to their needs.
Entry Condition	The database contains relevant data from the IdroGEO API PIR.
Flow of Events	The user navigates to the data view customization page. The user selects the different data layers to be customized. The user selects the data fields to be displayed. The user selects which data filters and sorting methods to apply. The user sets map styles and other visualization options. Users save customized data views.
Exceptions	If the user did not successfully query the disaster dataset, the appropriate error message is displayed to the user. If the user does not select any data fields or chart types, all available data fields and the default chart types are displayed by default.
Exit Condition	Users have successfully customized the display mode of disaster data and can view and use the customized view in the system.



<b>Use Case 5</b>	<b>Download data</b>
Participant	User
Goal	Allows users to download the queried data to the local as a file.
Entry Condition	The user has executed the data query and obtained the results.
Flow of Events	<p>Users perform data query operations on the data query page.</p> <p>After you obtain the query results, the system provides export options on the page.</p> <p>The user clicks the export option and selects the format to export the data (e.g. PDF, CSV, Excel, etc.).</p> <p>The system provides generates the query result and download links or downloads files directly to the user's device.</p>
Exceptions	<p>If the system cannot generate the exported file, an error message is displayed to the user.</p> <p>If you cancel the export operation, the process is complete, and you can continue to view data or perform other operations.</p>
Exit Condition	The queried data is successfully exported to the local PC and can be further processed or analyzed in the local environment.

## 2) Technical Use Cases

<b>Use Case 6</b>	<b>System Settings</b>
Participant	Administrator
Goal	Allow the administrator to configure system parameters and settings.
Entry Condition	The administrator is logged into the system.
Flow of Events	<p>The administrator logs into the system.</p> <p>The administrator accesses the system settings page.</p> <p>The administrator browses through the configurable parameters and settings options.</p> <p>The administrator modifies the necessary parameters and settings.</p> <p>The administrator saves the modifications and applies them to the system.</p>
Exceptions	<p>If the system encounters an error while saving the modifications, it displays the corresponding error message.</p> <p>If the administrator cancels the modification operation, the flow ends, and the administrator's changes are not applied to the system.</p>
Exit Condition	The system successfully applies the new parameters and settings made by the administrator.

Use Case 7	Data Management
Participant	Administrator
Goal	Allow the administrator to manage datasets, data sources, and data cleaning processes within the system.
Entry Condition	The administrator is logged into the system.
Flow of Events	<p>The administrator logs into the system.</p> <p>The administrator navigates to the data management page.</p> <p>The administrator browses the list of datasets and data sources and selects the item to manage.</p> <p>The administrator performs the necessary operations, such as adding a new dataset, editing data source information, or deleting a dataset.</p> <p>The administrator saves the modifications and applies them to the system.</p>
Exceptions	<p>If the system encounters an error while saving the modifications, it displays the corresponding error message.</p> <p>If the administrator cancels the modification operation, the flow ends, and the administrator's changes are not applied to the system.</p>
Exit Condition	The data management operation is successfully completed.

Use Case 8	Data Synchronization
Participant	Administrator
Goal	To regularly synchronize and update data within the system.
Entry Condition	The administrator is logged into the system.
Flow of Events	<p>The administrator logs into the system.</p> <p>The administrator opens the data synchronization settings page.</p> <p>The administrator configures the frequency and time for data synchronization.</p> <p>The system periodically checks external data sources and updates data within the system based on the configured settings.</p> <p>The administrator saves the modifications and applies them to the system.</p>
Exceptions	<p>If the system encounters an error during data synchronization, it displays the corresponding error message.</p> <p>If the administrator cancels the data synchronization operation, the flow ends, and the system's data remains unchanged.</p>
Exit Condition	The system successfully completes data synchronization and updates.

## User stories

### 1) User stories for Main User

- a) **As a Civil Protection Analyst**, I want to browse through the latest statistical data and access the Landslide Indicator Table so that I can assess and prepare for potential landslide risks in various regions.

Acceptance Criteria:

- (1) The system presents me with the latest statistical data and allows me to access the Landslide Indicator Table.
- (2) I can view and compare different datasets to understand the risk areas for landslides.

- b) **As an Insurance Company Analyst**, I need to query data by region and by name to generate detailed risk assessments for our insurance policy adjustments.

Acceptance Criteria:

- (1) From the main dashboard, I use the search functionality to input specific regions or hazard names.
- (2) The system retrieves and displays the requested data, which I can filter further based on region levels and disaster types.
- (3) I analyze the results, which are crucial for adjusting our insurance coverage strategies.

- c) **As a Government Official**, I want to analyze hazard maps and visualize spatial distributions of population and buildings to make informed decisions on urban planning and disaster management.

Acceptance Criteria:

- (1) On the platform, I navigate to the 'Analyze and Visualize Data' section after identifying a potential risk area.
- (2) I use the tools provided to overlay demographic data on hazard maps to understand the potential impact on populations and infrastructure.
- (3) The insights gained aid in planning safety measures and resource allocation for affected regions.

- d) **As a Civil Protection Analyst**, I need to customize how disaster data is displayed by adjusting legends and data layers, enabling me to present the information in an accessible way during stakeholder briefings.

Acceptance Criteria:

- (1) Within the dashboard, I select 'Customize Data View' to tailor the visualization of risk data.
- (2) I modify data layers for clarity and choose legends that help in interpreting the maps easily.
- (3) The customized views allow me to present complex data in a format that stakeholders can easily understand and act upon.

- e) **As Technical Support Staff**, I want to assist users by downloading various data formats like CSV, Excel, and PDF, ensuring they have offline access to essential information for detailed analysis and reporting.

Acceptance Criteria:

- (1) I perform data queries as requested by our analysis team and visit the 'Download Data' section.
- (2) I select the preferred data formats and initiate the download, ensuring the team can continue their work with local copies.
- (3) I verify the integrity of the downloaded files and distribute them to the relevant departments for further action.

## 2) User stories for System Administrator

- a) **As a System Administrator**, I need to access and configure system settings to ensure the application runs smoothly and meets user needs.

Acceptance Criteria:

- (1) The administrator logs into the system management interface.
- (2) They navigate to 'System Settings' to review and modify configurations as needed.
- (3) The changes are saved and applied to the system without disrupting user activities.

- b) **As a System Administrator**, I want to manage and curate datasets to maintain the integrity and relevance of the data provided to users.

Acceptance Criteria:

- (1) The administrator enters the 'Data Management' section.
- (2) They review the datasets, add new ones or update existing entries, ensuring all data sources are current and accurate.
- (3) Any changes are committed, and the system updates its databases accordingly.

- c) **As a System Administrator**, I must synchronize the system data with external data sources regularly to keep the system updated with the latest hazard information.

Acceptance Criteria:

- (1) The administrator sets up data synchronization intervals via the 'Data Synchronization' settings.
- (2) They configure the system to automatically check for and integrate new data from trusted external sources at the set intervals.
- (3) The synchronization process runs at the specified times, and the administrator receives notifications upon completion or in the event of errors.

## Requirements and domain assumptions

### Functional requirements

#### 1) Browse Data:

- a) The application should allow users to browse region(s) for its/their statistical data of the following groups (population, buildings, and cultural heritages).
- b) The application should allow users to have access to the Landslide hazard indicator table and the hazard map.
- c) The application should allow users to have access to available dataset.

#### 2) Query data:

- a) The application should allow users to query data from database through a search bar in dashboard by region's name.
- b) The application should allow users to query hazard map through a search bar in dashboard by name of location.

#### 3) Analyze and Visualize Data:

- a) **Hazard and Risk Map:** The dashboard should allow users to analyze and visualize the landslide hazard map classified by very high p4, high p3, medium p2, moderate p1 and attention zones AA.
- b) **Spatial Distribution Maps:** The dashboard should enable users to visualize the region through Spatial Distribution Maps, depicting the distribution of various features including population, buildings, and cultural heritage sites.
- c) **District Limits:** The dashboard should allow users to visualize district limits with boundary maps.
- d) **Risk Indicator:** The dashboard should allow users to represent the intensity or concentration of risk factors across the region by Heatmap.

#### 4) Customize Data View:

- a) **Legend:** The dashboard should allow users to have access to the different Legend of the map.
- b) **Data layers:** The dashboard should allow users to change the transparency and the color of different data layers.
- c) **Base Layers:** The dashboard should allow users to choose their preferred base layer from Open Street Map, Bing Map, and ESRI Topo Map.

#### 5) Download Data:

Users should be able to export query results in PDF, CSV, and XLS formats in their own devices through:

- a) Print Reports

- b) Download Tables
- c) Download Maps
- d) Download Raw Data

## Technical requirements

- 1) The back-end web server must interact with the database through a REST API.
- 2) Data is returned to the front end in JSON format.
- 3) The dashboard must be built using Jupyter Notebooks.

## Domain assumptions

- 1) Assuming the data retrieved from the IdroGEO API PIR is accurate and has been verified and validated.
- 2) Assuming the data retrieved from the IdroGEO API PIR is complete and covers all key information required for the project.
- 3) Assuming the stability and availability of the IdroGEO API are reliable, and the system can regularly access and retrieve data from it.
- 4) Assuming the system has sufficient performance and capacity to handle user requests and data processing operations, completing them within a reasonable time frame.
- 5) Assuming the end users of the system have basic computer and internet skills and can correctly utilize the functionalities provided by the system.
- 6) Assuming the accuracy of geographic information data (such as map data) is limited and may have a certain degree of error and uncertainty.