

# Design Document

HAZARD MAPS PROJECT SE4GEO IN ITALY  
POLITECNICO DI MILANO

## AUTHORS

Giovanni Montefoschi

Zhanbin Wu

Xiao Tan

## INSTRUCTOR

Giovanni Quattrocchi

## DATE PREPARED

31/05/2024

## VERSION

0.0.0



# Table of content

Table of content .....	1
1. Introduction .....	2
1.1 Context and motivations.....	2
1.2 Definitions and Acronyms.....	2
1.3 Purpose .....	3
1.4 Scope and Limitations .....	4
2. Architectural Design.....	5
2.1 Overview .....	5
2.2 Component Diagrams.....	5
2.3 Technology Stack .....	7
3. Software Design .....	8
3.1 Database Design .....	8
3.2 REST APIs.....	10
3.3 Dashboard Design .....	11
4. Implementation and Test Plan .....	14
4.1. Implementation Plan.....	14
4.2 Test Plan .....	16
5. Bibliography .....	16

# 1. Introduction

## 1.1 Context and motivations

The Italian public authorities are responsible for collecting, processing, and analyzing datasets related to natural disasters to generate landslide hazard maps. By combining these hazard maps with data on population, households, buildings, industries and services, cultural heritage, and territories across various regions of Italy, a comprehensive assessment of regional exposure to landslides can be conducted. This multidimensional data analysis aids in evaluating the potential impact of landslides on different areas, providing critical information for decision-makers.

The analysis specifically covers the following data aspects:

- **Territory:** Geographical and physical characteristics of the region.
- **Population:** Distribution and density of the population in each region.
- **Families:** Structure and distribution of households.
- **Buildings:** Number and types of buildings.
- **Industries and Services:** Distribution and economic significance of industries and services.
- **Cultural Heritage:** Distribution and historical and cultural value of heritage sites.

The scope of this project includes both national and regional levels. By integrating these multidimensional datasets, decision-makers can more effectively access, visualize, and analyze relevant information, thereby developing scientific disaster management and prevention measures, enhancing disaster response capabilities and public safety levels.

## 1.2 Definitions and Acronyms

- **ISPRA:** Istituto Superiore per la Protezione e la Ricerca Ambientale (Italian Institute for Environmental Protection and Research). ISPRA is responsible for environmental protection and research, providing data related to natural disasters.
- **PIR:** Hazards and Risk Indicators. These indicators are used to assess and analyze the risk of natural disasters such as landslides.
- **REST API:** Representational State Transfer Application Programming Interface. An interface design style for web services allowing interaction between clients and servers.
- **GeoJSON:** Geospatial JavaScript Object Notation. An open standard format used for encoding geographic data structures, commonly used in web mapping applications.
- **WMS:** Web Map Service. A standard protocol used to request geographic information services over the internet to generate maps.

- **PostgreSQL:** An open-source object-relational database management system emphasizing extensibility and standards compliance, widely used for complex data analysis tasks.
- **PostGIS:** An extension to PostgreSQL adding support for geographic objects, enabling the handling of geographic information system (GIS) data.
- **Flask:** A lightweight web application framework written in Python, designed for simplicity and flexibility, suitable for rapid development of web services and REST APIs.
- **GeoServer:** An open-source server software allowing users to share and edit geospatial data. It is designed to interact with many different data sources and output data to various clients, including web maps.
- **Pandas:** A Python data analysis library providing high-performance, easy-to-use data structures, and data analysis tools.
- **GeoPandas:** An extension to Pandas supporting geospatial data types, enabling the handling and analysis of geospatial data.

### 1.3 Purpose

The aim of this project is to develop an interactive web application to support disaster management activities in Italy. The application will integrate PIR data from ISPRA, presenting landslide risk and exposure information related to population, buildings, households, industries and services, cultural heritage, and territories through a user-friendly interface, assisting decision-makers in effective disaster management.

The main interactive features of the application include:

- **Region Selection:** Users can select different regions from a dropdown menu to view detailed statistical information and exposure data for that region, including population, households, buildings, industries and services, cultural heritage, and territories.
- **Map View:** The application provides landslide risk maps showing the distribution of landslide risks across Italian regions. The maps support zooming and panning, allowing users to view detailed risk information for specific areas.
- **Data Display:** Statistical information and landslide risk data for the selected region are presented through data cards and stacked bar charts. Users can clearly see the distribution of different categories of data, such as population, households, buildings, industries and services, and cultural heritage across different risk levels.
- **Data Download:** The application allows users to download statistical information and landslide risk data for specific regions, supporting further analysis and use.
- **Legend and Styles:** The application includes a detailed legend to help users understand various risk symbols and data representations on the map.

- **Map Layer Selection:** Users can select different map layers to view more detailed information. These layers include:
  - OpenStreetMap: Provides a basic map view.
  - Carto Positron: Offers a clear and simple map style suitable for background maps.
  - Esri Topographic: Provides a detailed topographic map view.
  - Esri World Imagery: Offers satellite imagery views.
  - Landslide Hazard Map: Displays the landslide risk map layer, indicating areas at risk of landslides.
  - Population Density: Displays the population density distribution layer.
  - Risk Indicator Buildings: Shows the building risk indicator layer.
  - Risk Indicator Cultural Heritage: Displays the cultural heritage risk indicator layer.
  - Risk Indicator Families: Shows the family risk indicator layer.
  - Risk Indicator Industries and Services: Displays the industries and services risk indicator layer.
  - Risk Indicator Population: Shows the population risk indicator layer.
  - Risk Indicator Territory: Displays the territory risk indicator layer.
  - Regions: Displays region boundary layers.
  - Selected Region: Highlights the selected region.

These interactive features enable users to easily access and analyze relevant information, allowing them to develop scientific disaster management and prevention measures, enhancing disaster response capabilities and public safety levels.

## 1.4 Scope and Limitations

### Scope:

Develop a client-server application comprising a database, REST API, and interactive dashboard. Data will be sourced from the ISPRA IdroGEO API, including landslide risk and census data on population, buildings, households, industries and services, cultural heritage, and territories. The application will support national and regional-level data presentation and analysis.

### Limitations:

The current project does not support real-time data integration, though this can be a direction for future development. The design and functionality of the dashboard are limited by the technical capabilities of Dash and related graphical libraries. The analysis primarily focuses on regional levels, which may not sufficiently meet the needs of smaller geographic units (e.g., neighborhoods or communities).

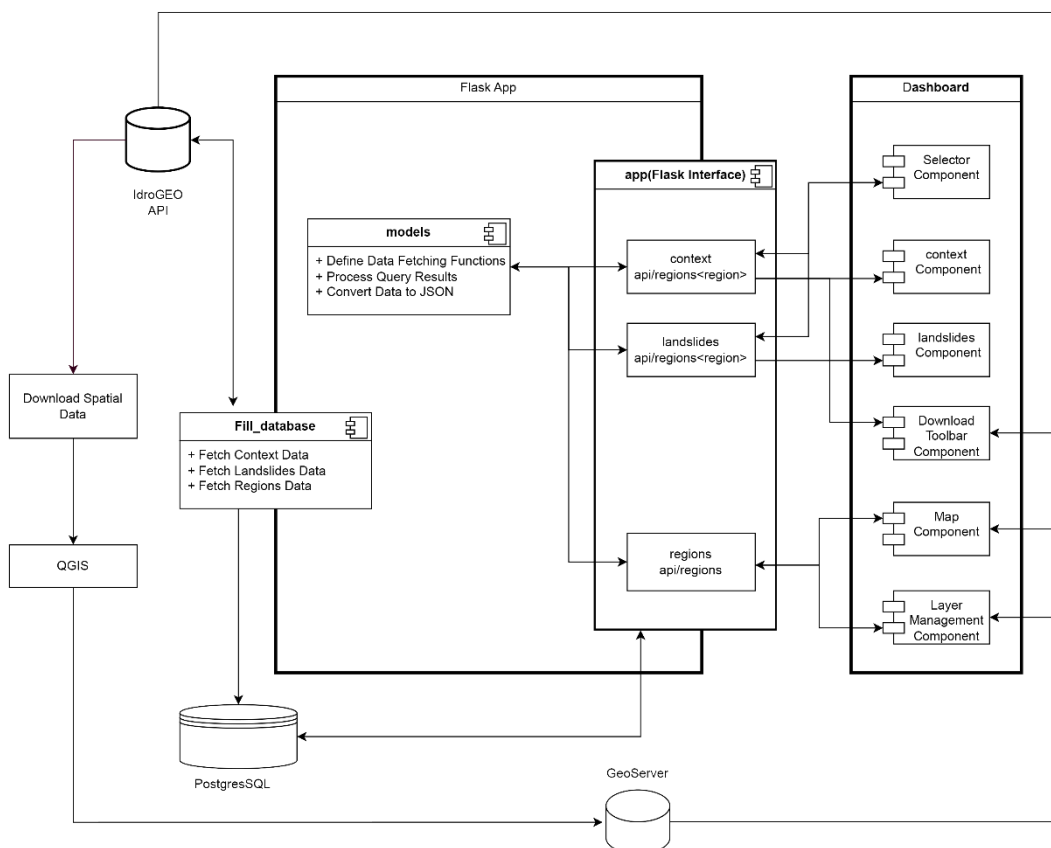
## 2. Architectural Design

### 2.1 Overview

The system is composed of three main components: the database, the web server (REST API), and the interactive dashboard (frontend). The database is responsible for storing and managing landslide risk data and related statistical information obtained from the ISPRA IdroGEO API. The web server provides query and retrieval services to the database through REST API interfaces. The interactive dashboard is used for visual data presentation, supporting users in selecting different regions, viewing detailed statistical information and risk analysis, thereby achieving interactive data analysis and decision support.

### 2.2 Component Diagrams

The component diagram below illustrates the relationship between the database, web server, and dashboard, as well as the interaction and data flow between each component.



#### Component Description:

- **IdroGEO API:** Provides API interfaces for landslide risk and related statistical data.

- ♦ **Function:** Acquire landslide risk data and related statistical information through HTTP requests from the ISPRA IdroGEO API.
- **Download Spatial Data:** Provides spatial data download.
  - ♦ **Function:** Download spatial data from external sources for further geospatial analysis.
- **QGIS:** Processes and converts spatial data.
  - ♦ **Function:** Process downloaded spatial data using QGIS and publish these data through GeoServer.
- **Fill\_database:** Acquires data from the API and stores it in the database.
  - ♦ **Function:** Use the fill\_database.py script to acquire statistical data, landslide data, and regional data from the IdroGEO API and store them in the PostgreSQL database.
- **PostgreSQL with PostGIS:** Stores and manages geospatial data.
  - ♦ **Function:** Use the PostgreSQL database extended with PostGIS to store and manage geospatial data.
- **GeoServer:** Publishes geospatial data.
  - ♦ **Function:** Publish spatial data processed from QGIS through GeoServer, providing map layer services.
- **Flask App:** Manages data acquisition, processing, and API interface provision.
  - ♦ **Function:** Acts as the core of the system, responsible for data acquisition, processing, and provision of data services through REST APIs.
  - ♦ **Models:**
    - Define Data Fetching Functions: Define functions to fetch data from the database.
    - Process Query Results: Process query results.
    - Convert Data to JSON: Convert data to JSON format.
  - ♦ **API Interfaces:**
    - Context API:
      - context api/regions<region>: Acquires statistical data for a specific region.
    - Landslides API:
      - landslides api/regions<region>: Acquires landslide data for a specific region.
    - Regions API:
      - regions api/regions: Acquires data for all regions.
- **Dashboard:** Visualizes and interacts with data.
  - ♦ **Function:** Provides a user interface for selecting different regions and viewing detailed statistical information and risk analysis.
  - ♦ **Components:**
    - Selector Component: Allows users to select different regions.
    - Context Component: Displays statistical data for the selected region.

- **Landslides Component:** Displays landslide risk data for the selected region.
- **Download Toolbar Component:** Provides data download functionality.
- **Map Component:** Displays landslide risk maps and other related layers.
- **Layer Management Component:** Allows users to manage and switch between different map layers.

## 2.3 Technology Stack

- **Database:**
  - **PostgreSQL:** An open-source object-relational database management system with powerful extensibility and standards compliance, suitable for complex data management tasks.
  - **PostGIS:** An extension to PostgreSQL adding support for geospatial objects, enabling the handling and storage of geographic information system (GIS) data.
- **Web Server:**
  - **Flask:** A lightweight web application framework written in Python, designed for simplicity and flexibility, suitable for rapid development of web services and REST APIs.
  - **GeoServer:** An open-source server software for publishing geospatial data. GeoServer is designed to interact with many different data sources and output data to various clients, including web maps.
- **Data Processing:**
  - **Pandas:** A powerful Python data analysis library providing high-performance, easy-to-use data structures and data analysis tools, suitable for various data processing tasks.
  - **GeoPandas:** An extension to Pandas supporting geospatial data types, enabling the handling and analysis of geospatial data.
  - **Requests:** A simple and easy-to-use Python library for sending HTTP requests, suitable for interacting with APIs and acquiring data.
  - **QGIS:** An open-source geographic information system application for processing and analyzing spatial data and publishing data through GeoServer.
- **Dashboard:**
  - **Dash:** A Python framework for building analytical web applications based on Flask and Plotly, suitable for rapid development of interactive data visualization applications.
  - **Plotly:** A Python library for creating interactive charts, supporting various chart types, including scatter plots, line charts, bar charts, and more.



- Building interactive maps, spatial data.
- Leaflet into Dash, for applications.
- Making code changes and providing version development and code
- Development environment assistance, debugging, support.

- metric information  
utes.

EA13000060FB59D0455FB1A1645DDC72E7A0534180F3FD94  
EAE400002007372D7D1741E4B8C896046C534100E2F1927;  
40000000001915CFE81A2104176E8CB2899045341200DFD2E  
720E0000E0E034D41D3AE2641984C15BC6CE9334100098A9F1  
EA1300001074366B087127415E40CB5CF0FE5341686C74931  
E8000000A083E575F99244A21C98B8834654543410172653F  
1100000028A91380475721412EB290B7E896524128A913703  
200000000072980BCFD241EE54A26A48AB5241283A926B  
AC00000000605457F952441782D2137F6D0A514168D5E7A1  
E80000003BE838452CB6274184AC0C75F46E52410E2406291  
E3090000F80FE9777D20362410E30630BB9852410E959F24  
AA00000008060E00EBBB2A412063E2242A2541138E3F38A5  
E8E900039C1AC09B7D85284198B8964062D95241084B8FD01  
38080000004A3921A1A05E2A1FEB2783261D514118D09C2D  
9900000078832FE24A82741FAS6C62029P950414850FC58D  
00000000E09CE931C7C8A334124586F0FE65A01EC9E39C8D  
60000007424970E28630417E4290861F550417424970F2C  
C00000004C5986D08B7E3041EAB6622FD7D850414C5986D8  
200000090707FBB8EA1C29411C38675235064E41D0463B80  
17000000E05179986641B41EA4E04B0F0689041F0045F9D1

- region, including  
stries and services,  
dependent table, with the

	Adults (15-64) % double precision	Elderly (65+) % double precision
009	65.153	20.838

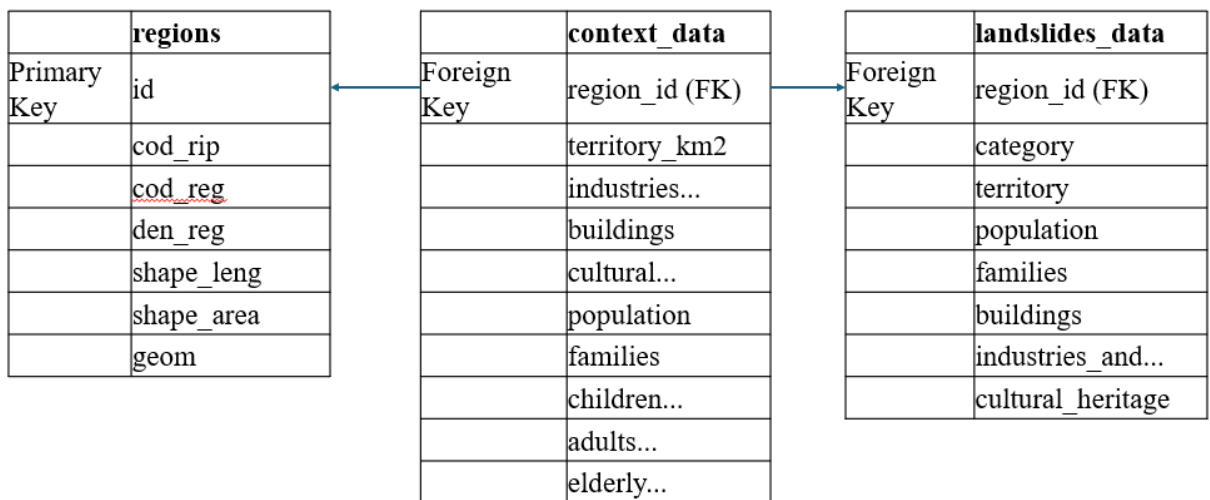
- specific region,  
region has an  
identifier of the region.

	Category text	Territory text	Population text	Families text	Buildings text	Industries and services text	Cultural heritage text
1	Very high P4	9494 (3.143%)	499749 (0.841%)	206968 (0.841%)	223065 (1.537%)	31244 (0.65%)	5351 (2.508%)
2	High P3	16890 (5.592%)	803917 (1.353%)	340926 (1.385%)	342483 (2.359%)	53197 (1.107%)	7182 (3.366%)
3	Medium P2	14551 (4.817%)	1720208 (2.894%)	727315 (2.955%)	562800 (3.877%)	127356 (2.65%)	10728 (5.028%)
4	Moderate P1	12555 (4.157%)	2006643 (3.376%)	844536 (3.431%)	522206 (3.598%)	147766 (3.075%)	12390 (5.807%)
5	Attention zones AA	6987 (2.313%)	676948 (1.139%)	271208 (1.102%)	216540 (1.492%)	45677 (0.95%)	2502 (1.173%)
6	P4 + P3	26385 (8.735%)	0 (0%)	0 (0%)	565548 (3.896%)	0 (0%)	0 (0%)
7	None zones	241576 (79.978%)	53726722 (90.397%)	22222522 (90.286%)	12648136 (87.137%)	4400737 (91.568%)	175207 (82.118%)

## Table Descriptions:

- **regions Table:**
  - ◆ **Fields:**
    - id: Primary key, identifying the unique ID of the region.
    - cod\_rip: Regional code.
    - cod\_reg: Area code.
    - den\_reg: Region name.
    - shape\_leng: Region boundary length.
    - shape\_area: Region area.
    - geom: Geometric information, used to store the geographical shape of the region.
- **context\_data\_{id} Table:**
  - ◆ **Fields:**
    - Territory (km<sup>2</sup>): Territory area.
    - Industries and services: Number of industries and services.
    - Buildings: Number of buildings.
    - Cultural heritage: Number of cultural heritage sites.
    - Population: Number of people.
    - Families: Number of families.
    - Children (0-14) %: Percentage of children aged 0-14.
    - Adults (15-64) %: Percentage of adults aged 15-64.
    - Elderly (65+) %: Percentage of elderly aged 65 and above.
- **landslides\_data\_{id} Table:**
  - ◆ **Fields:**
    - Category: Risk category (e.g., very high risk P4, high risk P3, moderate risk P2, low risk P1, attention area AA).
    - Territory: Affected territory area (and percentage of the total area).
    - Population: Affected population (and percentage of the total population).
    - Families: Affected families (and percentage of the total number of families).
    - Buildings: Affected buildings (and percentage of the total number of buildings).
    - Industries and services: Affected industries and services (and percentage of the total number).
    - Cultural heritage: Affected cultural heritage sites (and percentage of the total number).

- **Database ER diagram:**



## 3.2 REST APIs

The REST API provides the following endpoints for client applications to communicate with the server, retrieve, and process data from the database:

### 3.2.1 GET /context/ string:id

- **Description:** Retrieve statistical data for a specific region.
- **Request Parameters:**

**id:** Unique identifier of the region (string format).

- **Response:** Returns statistical data for the specified region, including information on population, households, buildings, industries and services, cultural heritage, and territories.
- **Code:**

```
@app.route("/context/string:id", methods=["GET"])
def get_context_endpoint(id):
    table_name = f'context_data_{id}'
    return get_context(conn=conn, table_name=table_name)
```

```
# models.py function
def get_context(conn, table_name):
    df = pd.read_sql_table(table_name=table_name, con=conn)
    df_json = df.to_json(orient="records")
    return df_json
```

### 3.2.2 GET /landslides/ string:id

- **Description:** Retrieve landslide risk data for a specific region.
- **Request Parameters:**

**id:** Unique identifier of the region (string format).

- **Response:** Returns landslide risk data for the specified region, including data for different risk categories such as very high risk (P4), high risk (P3), moderate risk (P2), low risk (P1), and attention area (AA).
- **Code:**

```
@app.route("/context/string:id", methods=["GET"])
def get_context_endpoint(id):
    table_name = f'context_data_{id}'
    return get_context(conn=conn, table_name=table_name)

# models.py function
def get_context(conn, table_name):
    df = pd.read_sql_table(table_name=table_name, con=conn)
    df_json = df.to_json(orient="records")
    return df_json
```

### 3.2.3 GET /regions

- **Description:** Retrieve data for all regions.
- **Request Parameters:** None.
- **Response:** Returns basic information and geometric data for all regions.
- **Code:**

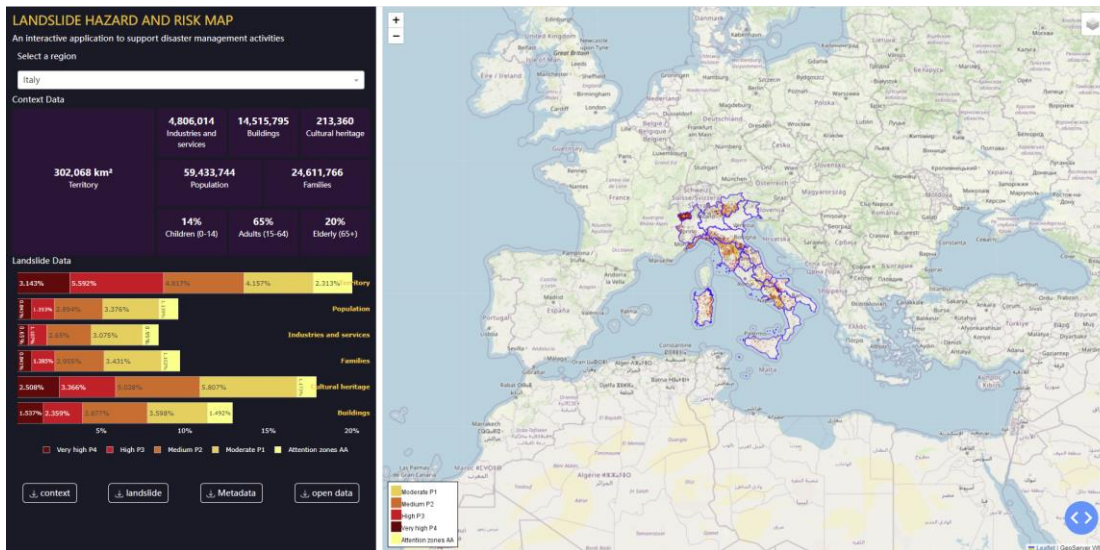
```
@app.route("/regions", methods=["GET"])
def get_regions_endpoint():
    return get_regions(conn=conn)

# models.py function
def get_regions(conn):
    df = pd.read_sql_table(table_name='regions', con=conn)
    df_json = df.to_json(orient="records")
    return df_json
```

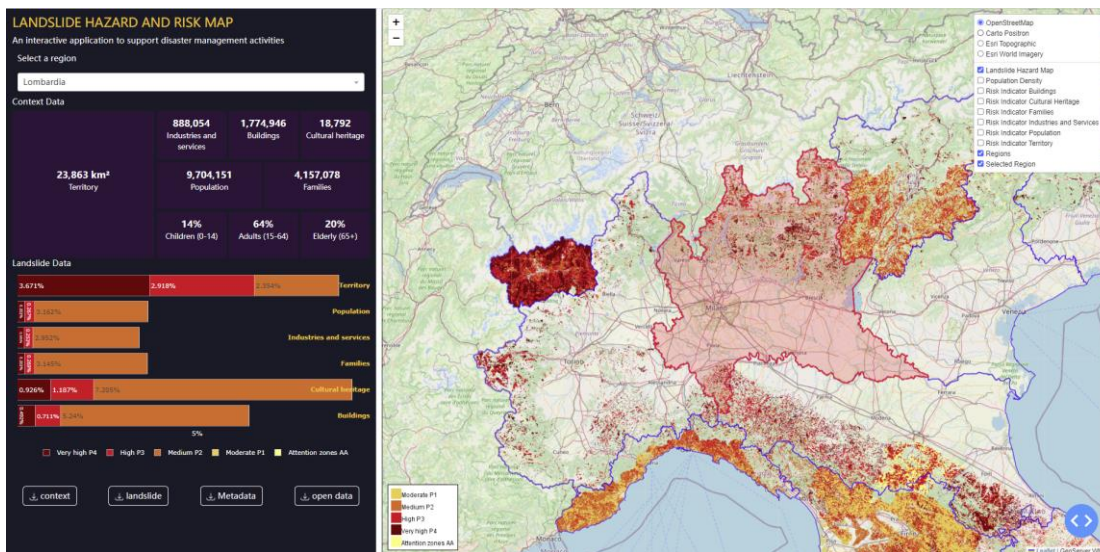
## 3.3 Dashboard Design

The dashboard plays a crucial role in disaster management and decision support. The design of this system's dashboard aims to enhance the efficiency of landslide risk analysis and management through interactive data visualization and an intuitive user

interface. The main features of the dashboard and their detailed descriptions are as follows:



Initial interface.



The interface for the user to select an area.

## Main Features:

- **Map View:**
  - **Description:** Displays landslide risk maps, allowing users to select different regions to view detailed information. The map view is built on Leaflet and Dash Leaflet, providing highly interactive geospatial data presentation.
  - **Functionality:**



- **Region Selection and Highlighting:** Users can click on regions on the map to highlight the boundaries and view detailed information.
- **Layer Switching:** Provides multiple map layers, allowing users to switch between different types of data, including landslide risk layers, population density layers, industrial and services distribution layers, etc.
- **Map Zooming and Panning:** Supports free zooming and panning of the map, allowing detailed viewing of specific area's geographic information.
- **Technical Implementation:**
  - Uses the Leaflet library to build the basic map view, integrated seamlessly with the Dash framework using Dash Leaflet.
  - GeoJSON formatted data is used for drawing and highlighting region boundaries.
- **Data Display:**
  - **Description:** Displays statistical data and landslide data for the selected region in the form of charts and data cards. The data display part is built on Dash and Plotly, offering diverse charts and intuitive data cards.
  - **Functionality:**
    - **Statistical Data Display:** Shows detailed statistical data for the selected region, including territory area, population, number of households, number of buildings, number of industries and services, number of cultural heritage sites, percentage of children, percentage of adults, and percentage of elderly.
    - **Landslide Risk Data Display:** Displays landslide risk data for the selected region, covering different risk levels such as very high risk (P4), high risk (P3), moderate risk (P2), low risk (P1), and attention area (AA), with affected area, population, number of households, number of buildings, number of industries and services, and number of cultural heritage sites.
    - **Data Charts:** Provides stacked bar charts to intuitively display the detailed classification and proportion of landslide risk data.
  - **Technical Implementation:**
    - Uses the Dash framework to build an interactive data display interface.
    - Employs the Plotly library to create various types of charts for data visualization.
    - Data is dynamically rendered on the front end, fetched from the backend through REST APIs.
- **Download Functionality:**
  - **Description:** Allows users to download data for specific regions, facilitating offline analysis and report preparation. The download

functionality is implemented using Dash's download components, ensuring convenience and flexibility in data acquisition.

- **Functionality:**
  - Data Selection and Download: Users can select and download statistical data and landslide risk data for the selected region, with the downloaded data files formatted in CSV, suitable for further processing and analysis.
  - Download Button: Provides a download button on the data display interface, allowing users to easily obtain relevant data files with a single click.
- **Technical Implementation:**
  - Uses Dash's dcc.Download component to implement the download functionality.
  - The backend provides data interfaces that generate CSV files and respond to frontend requests.

With these features and layout design, the dashboard will provide users with an intuitive, user-friendly interface supporting landslide risk analysis and decision-making. Users can gain a comprehensive understanding and analysis of landslide risks and related statistical data through the map view, data display, and download functionality.

## 4. Implementation and Test Plan

### 4.1. Implementation Plan

#### Hazard Maps Project

This project aims to provide hazard maps for disaster management using Python, Flask, GeoServer, and PostgreSQL with PostGIS extension. The following instructions will guide you through setting up and running the project.

#### Prerequisites

- Python 3.8+
- PostgreSQL 12+
- Node.js and npm (for frontend dependencies)
- pgAdmin (for database management)
- GeoServer (for map layer publishing)

#### Setup Instructions:

##### 1. Clone the repository

First, clone the repository to your local machine:

```
git clone https://github.com/tanxiaoo/Hazard_Maps_Project-SE4GEO.git
cd Hazard_Maps_Project-SE4GEO/APP
```

## 2. Download large files

Due to GitHub's file size limitations, the spatial\_datafolders contain large files that need to be downloaded separately. Please download the necessary files from the following links and place them in the backend/resources/geoserver/spatial\_datafolders:

- [Download spatial\\_data folder](#)

## 3. Create the database and enable the PostGIS extension

In pgAdmin:

- i. Open pgAdmin and connect to your PostgreSQL server.
- ii. Create a database named se4g24.
- iii. Select the newly created database, open the query tool, and run the following command to enable the PostGIS extension:  
CREATE EXTENSION postgis;
- iv. Run the fill\_database.py file to populate the database:  
cd backend  
python fill\_database.py

## 4. Configure GeoServer

- i. Download and install GeoServer:
  - Visit the [GeoServer download page](#) and download the latest version of GeoServer.
  - Follow the installation instructions.
- ii. Start GeoServer and access the GeoServer Web Interface at:
- iii. `http://localhost:8080/geoserver`
- iv. Publish layers and styles:
  - In the left menu, select Data -> Layers.
  - Click Add new Layer, select the data store, and choose the data from backend/resources/geoserver/spatial\_data folder, then follow the prompts to complete the layer publishing.
  - In the left menu, select Style -> SLD.
  - Click Add a new style, upload the style files from backend/resources/geoserver/SLD directory, and apply them to the corresponding layers.

## 5. Create a .env file

Create a .env file in the APP directory and add the following content:

```
DATABASE_URL=postgresql://postgres:yourpassword@localhost:5432/se4g24
GEOSERVER_URL=http://localhost:8080/geoserver
```

Ensure to replace yourpassword with your PostgreSQL password.

## 6. Install Python dependencies

Ensure you are in the APP directory, then run the following commands to create and activate a virtual environment and install the required Python packages:

```
python -m venv venv
```



```
source venv/bin/activate # For Windows, use `venv\Scripts\activate`
pip install -r requirements.txt
```

### 7. Install Node.js dependencies

In the APP directory, install the Node.js dependencies:

```
npm install
```

### 8. Run the Flask application

In the APP directory, run the Flask application:

```
cd backend
```

```
python app.py
```

### 9. Run the frontend dashboard

In the APP directory, run the frontend dashboard:

```
cd frontend
```

```
python dashboard.py
```

Follow the above steps and you will be able to get the project up and running properly. If you encounter any problems during the process, please refer to the project documentation or contact the project development team for assistance.

## 4.2 Test Plan

### Unit Testing:

- Write unit tests to cover each module, ensuring the correctness of each function.
- Test database connections, API endpoints, data processing functions, etc.

### Integration Testing:

- Test the integration between modules, ensuring correct data flow between different components.
- Test the database population script, API queries, and dashboard data display.

### User Testing:

- Invite users to trial the dashboard, collect feedback, and make improvements.
- Test user interactions, data visualization effects, and download functionality.

## 5. Bibliography

- ISPRA IdroGEO API Documentation: [ISPRA IdroGEO API Documentation](#)
- GeoPandas Documentation: [GeoPandas Documentation](#)
- Flask Documentation: [Flask Documentation](#)
- Dash Documentation: [Dash Documentation](#)
- GeoServer Documentation: [GeoServer Documentation](#)
- PostgreSQL Documentation: [PostgreSQL Documentation](#)