FINAL REPORT ON CHALLENGE 5: AI-enhanced geospatial analysis for rural development challenges

Name of mentor(s): Alexander Kovalenko, Karel Charvat
Number of participants: 10

INTRODUCTION

- **Background of the challenge**

Rural areas face persistent challenges related to resource distribution, agricultural productivity, and infrastructure development. Traditional methods often lack the data integration and analytical capabilities necessary for effective decision-making. It is important to note that 80% of the information generated daily has some spatial character, underscoring the importance of geospatial data in addressing these challenges.

By leveraging Artificial Intelligence (AI), particularly Large Language Models (LLMs), combined with geospatial analysis (GeoAI), there is an opportunity to significantly improve rural development strategies. The development of these ideas, and the foundation of the Jackdaw platform, began prior to the hackathon. However, participation in the hackathon accelerated the development process, enabling faster integration, refinement, and deployment of these GeoAI tools.

- **Full explanation of the challenge defining the scope of the effort.**

This challenge focuses on developing **GeoAI solutions** that integrate AI, specifically **Large Language Models (LLMs)**, with geospatial data to enhance decision-making, resource allocation, and rural development planning. The challenge encompasses multiple areas of analysis, including:

- **Open Land Use Analysis**: Assessing land distribution and utilization to optimize planning.
- **Smart Points of Interest (POI) Analysis**: Identifying key locations for infrastructure and development needs.
- **Potential for Tourism Development**: Evaluating suitable areas for building hotels and promoting tourism.
- **Climate Analysis**: Analyzing climate patterns to support agricultural planning and sustainability.
- **Weather Forecasting**: Providing accurate weather forecasts for informed decision-making.
- **Terrain Analysis**: Assessing topographical features to aid in land-use and infrastructure development.

A **connector for INSPIRE metadata** was developed to facilitate streamlined access to geospatial services and data resources from the INSPIRE infrastructure. Additionally, the **PoliRuralPlus API** was implemented to support the integration and processing of these datasets with AI models, enabling efficient and scalable rural development solutions.

An important aspect of the work was the development of a **distributed GeoAI model**. This approach introduced the principle of **AI Spaces**, which could serve as an extension of current **Data Spaces**. AI Spaces enable more dynamic, scalable, and distributed integration of AI capabilities with geospatial datasets, enhancing flexibility and interoperability across different platforms and stakeholders.

Another critical component of the challenge focused on methods to **select the right datasets and analyses** that are most relevant to user queries. This ensured that the solutions provided are not only technically robust but also tailored to real-world needs. By aligning data selection and analytical methods with user questions, the GeoAI solutions delivered actionable insights and supported effective decision-making processes in rural development contexts.

METHODOLOGY

The **methodology** of the report contains the sections that describe the progress and results achieved for the entire challenge.

- **Team description + info about any coordination with other organizations, outside agencies**

The team comprised **10 participants**, forming an **international team** with diverse expertise in **AI, geospatial analysis, and rural development**. The team was composed of both **experienced scientists** and **students**, bringing together a balance of advanced knowledge, innovative thinking, and fresh perspectives. This diversity facilitated the integration of cutting-edge research with creative problem-solving approaches.

Coordination involved collaboration with organizations specializing in **Geographic Information Systems (GIS)**, **AI research**, and **rural policy development**. The combination of seasoned experts and emerging talent allowed for a dynamic and flexible approach to solving the complex challenges associated with rural development. This collaborative environment ensured knowledge transfer, mentorship, and the development of practical, scalable solutions.

- **Technical Background**

The challenge utilized a combination of **AI models**, **GIS platforms**, and **geospatial datasets** to develop effective rural development solutions. Key components included:

- **Large Language Models (LLMs)**: Various LLMs, including **ChatGPT**, **LLaMA (Large Language Model Meta AI)**, and other advanced models, were employed for natural language processing and generating decision-making insights based on geospatial data. These models enabled the interpretation of complex queries and the generation of context-aware responses tailored to rural development challenges.
- **Geospatial Analysis Tools**: Platforms like **QGIS** and cloud-based environments such as **Jupyter Notebook** were used for mapping, spatial analysis, and data visualization.
- **INSPIRE Metadata Connector**: Developed to query and access standardized geospatial data from the **INSPIRE infrastructure**, ensuring compliance with European data standards and improving data interoperability.
- **PoliRuralPlus API**: Facilitated the integration and processing of AI models with geospatial data. The API is accessible via the PoliRuralPlus GitHub repository, offering modular and scalable solutions for rural development challenges.

A critical aspect of the challenge was the **development of methods for selecting the right datasets and analyses** for specific user queries. This process was based on the use of **vector representations** and **knowledge graphs**. By leveraging vectors for semantic data mapping and knowledge graphs for contextual relationships, the system ensured that the datasets and analytical methods chosen were highly relevant to the concrete questions posed by users. This approach enhanced the precision and effectiveness of the GeoAI solutions, aligning the data and insights more closely with real-world rural development needs.

- **Description of the process of solutio**n
- **Identification of Rural Challenges**:
    a. Conducted a comprehensive analysis to define specific challenges related to agriculture, land use, infrastructure, and rural development needs.
    b. Identified key problem areas such as inefficient resource allocation, suboptimal land-use planning, lack of tourism potential assessment, and inadequate infrastructure development.
    c. Engaged with stakeholders, including farmers, local governments, and community organizations, to gather insights and ensure that identified challenges reflected real-world needs and priorities.
- **Data Integration**:
    a. Collected diverse datasets relevant to rural development, including **land use data**, **smart points of interest (POI)**, **climate data**, **weather forecasts**, and **terrain data**.
    b. Integrated these datasets into a unified system, ensuring compatibility with geospatial analysis tools and AI models.
    c. Utilized vector representations and knowledge graphs to facilitate the selection of relevant datasets based on user queries and contextual needs.
    d. Applied preprocessing techniques to clean, standardize, and validate data for seamless analysis and integration with AI models.
- **INSPIRE Metadata Connector**:
    a. Developed and deployed a **connector for INSPIRE metadata** to access geospatial data and services from the INSPIRE infrastructure.
    b. Enabled querying of standardized services such as **Web Map Services (WMS)**, **Web Feature Services (WFS)**, and **Web Coverage Services (WCS)**.
    c. Streamlined the discovery and retrieval of relevant geospatial datasets, ensuring compliance with European geospatial data standards and enhancing interoperability.
    d. Facilitated efficient integration of INSPIRE datasets into AI workflows for analysis and decision support.
- **PoliRuralPlus API Implementation**:
    a. Utilized the **PoliRuralPlus API** to integrate and process various geospatial datasets with AI models.
    b. Enabled the creation of **modular and customizable workflows** for different rural development tasks, including land use optimization, infrastructure planning, and climate analysis.
    c. Supported distributed GeoAI models and introduced the concept of **AI Spaces**, extending current Data Spaces to allow dynamic and flexible AI processing.

   d. Provided tools for efficient data querying, processing, and visualization to support stakeholders in making informed decisions.
- **Model Development**:
   a. Built and trained AI models leveraging **Large Language Models (LLMs)** such as **ChatGPT** and **LLaMA (Large Language Model Meta AI)** to process and analyze geospatial data.
   b. Developed models for tasks such as **land-use analysis**, **POI identification**, **tourism potential assessment**, **climate analysis**, **weather forecasting**, and **terrain evaluation**.
   c. Integrated knowledge graphs and vector-based data selection methods to ensure that analyses were relevant to user-specific queries and contextual needs.
   d. Focused on developing **distributed GeoAI models** to enhance scalability and efficiency in processing large datasets.
- **Validation and Testing**:
   a. Tested the integrated solutions using real-world datasets and practical scenarios reflecting rural challenges.
   b. Conducted iterative validation with feedback from stakeholders, including farmers, planners, and policymakers, to ensure the solutions were practical, relevant, and effective.
   c. Assessed the accuracy and performance of AI models, refining them based on user feedback and validation results.
   d. Evaluated the efficiency of the **INSPIRE metadata connector** and the **PoliRuralPlus API** in providing timely and relevant geospatial insights.

This comprehensive process ensured the development of GeoAI solutions that are practical, scalable, and capable of addressing real-world rural development challenges with precision and effectiveness.

- **Data & Equipment list**

Geospatial Datasets:

- Open land use data
- Climate and weather data
- Terrain and topographic data
- Points of interest (POI) data

Software and Tools:

- GIS Platforms: QGIS, Hub4Everybody
- INSPIRE Metadata Services: WMS, WFS, WCS
- PoliRuralPlus API: For data integration and model deployment

Infrastructure: Cloud services and computational resources for hosting and scaling models

- **Detailed implementation plan**

Phase 1: Data Collection and Preparation

- **Comprehensive Dataset Collection**: Gathered a wide range of datasets relevant to rural development, including:
  - **Land Use Data**: Detailed information on land classification and utilization patterns.
  - **Smart Points of Interest (POI) Data**: Locations relevant to infrastructure, agriculture, and tourism development.
  - **Climate Data**: Historical and current climate patterns to inform sustainability strategies.
  - **Weather Forecast Data**: Real-time and predictive weather information for agricultural and infrastructure planning.
  - **Terrain Data**: Topographic and elevation data to support land-use and infrastructure development planning.
- **Data Preparation**:
  - Preprocessed the data for compatibility with GIS platforms (e.g., QGIS) and AI models (e.g., LLMs like ChatGPT and LLaMA).
  - Standardized data formats and ensured consistency by applying georeferencing and data normalization techniques.
  - Created **vector representations** and **knowledge graphs** to facilitate contextual data selection based on user queries.
  - Validated datasets to ensure accuracy, completeness, and relevance to rural development applications.

Phase 2: Development of INSPIRE Connector

- **INSPIRE Metadata Integration**:
  - Designed and developed a **connector** to access INSPIRE metadata and geospatial services, supporting formats such as **WMS (Web Map Service)**, **WFS (Web Feature Service)**, and **WCS (Web Coverage Service)**.
  - Enabled automated querying, discovery, and retrieval of standardized geospatial datasets from the INSPIRE infrastructure.
- **Connector Functionality**:
  - Ensured seamless integration with existing GIS tools and AI models for efficient data utilization.
  - Facilitated the selection of relevant datasets through user-defined queries, enhancing the precision of geospatial analysis.

- Supported dynamic data loading to accommodate real-time analysis needs in rural development contexts.

Phase 3: Integration with PoliRuralPlus API

- **API Utilization**:
    - Leveraged the **PoliRuralPlus API** to integrate AI models and geospatial data within a unified platform.
    - Implemented functionality for **data ingestion**, **processing**, and **analysis** in modular workflows tailored to rural development tasks.
- **Workflow Development**:
    - Created **customizable and distributed GeoAI workflows** for key challenges such as land-use optimization, infrastructure planning, and tourism potential analysis.
    - Introduced the concept of **AI Spaces**, extending traditional Data Spaces to facilitate distributed AI processing and scalable analysis.
    - Enabled seamless communication between AI models, geospatial tools, and data sources to ensure interoperability and flexibility.

Phase 4: AI Model Deployment

- **Model Implementation**:
    - Deployed AI models leveraging LLMs such as **ChatGPT** and **LLaMA** for tasks including:
        - **Land Use Analysis**: Assessing land distribution and utilization to inform planning.
        - **Smart POI Identification**: Analyzing key locations to identify infrastructure and development needs.
        - **Climate and Weather Forecasting**: Providing predictive insights to optimize agricultural activities and infrastructure resilience.
        - **Terrain Analysis**: Evaluating topographical data to support land-use decisions.
- **Model Integration**:
    - Integrated AI models with the **INSPIRE metadata connector** and **PoliRuralPlus API** to streamline data flow and analysis.
    - Utilized knowledge graphs and vector-based data selection to enhance model relevance and accuracy.

Phase 5: Testing and Validation

- **Real-World Testing**:

  - Applied the developed solutions to real-world datasets and rural development scenarios, focusing on agriculture, infrastructure, and tourism planning.
  - Conducted iterative testing to identify strengths, weaknesses, and areas for improvement.
- **Stakeholder Feedback**:
  - Engaged with end-users, including farmers, local government officials, and community planners, to validate the effectiveness of the solutions.
  - Collected qualitative and quantitative feedback to refine AI models, workflows, and the functionality of the INSPIRE connector.
- **Performance Evaluation**:
  - Assessed the accuracy, efficiency, and scalability of the solutions, ensuring they met user requirements and real-world constraints.

Phase 6: Deployment and Refinement

- **Solution Deployment**:
  - Deployed the finalized GeoAI solutions for practical applications in rural development, ensuring accessibility through the PoliRuralPlus API and INSPIRE connector.
  - Implemented the distributed GeoAI model and AI Spaces framework to support scalable and dynamic analysis.
- **Continuous Refinement**:
  - Established a feedback loop for ongoing refinement based on user experiences and changing requirements.
  - Updated AI models and geospatial workflows to incorporate new data, improve accuracy, and address emerging challenges.
  - Enhanced methods for **dataset selection and analysis** to ensure solutions remained relevant to specific user questions, leveraging vectors and knowledge graphs for context-aware decision support.

This systematic approach ensured the development of robust, scalable, and practical solutions for rural development challenges, integrating cutting-edge AI with geospatial analysis to optimize decision-making and resource allocation.

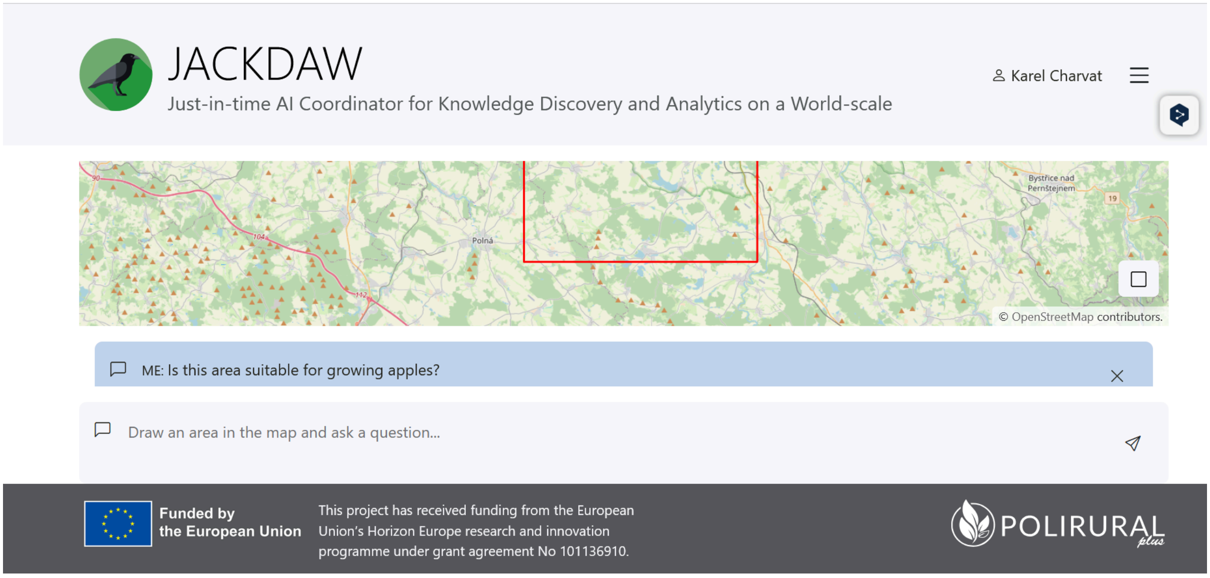
- **Analysis of needs of stakeholder groups**

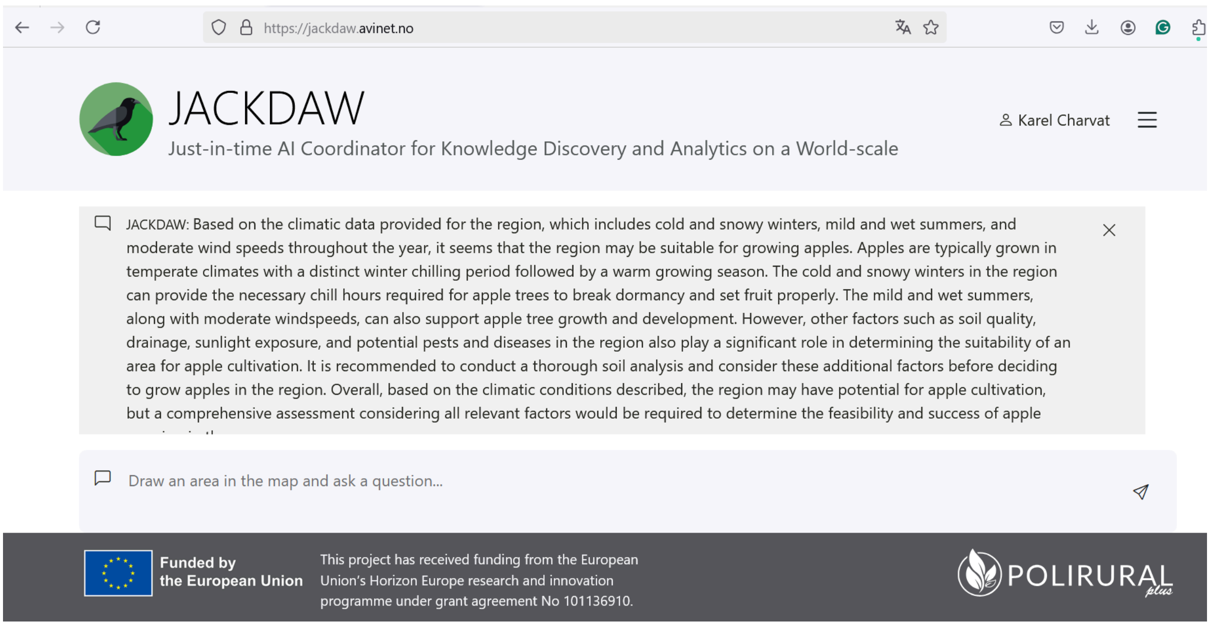
Farmers: Need optimized tools for agricultural planning and resource management.
Local Governments: Require accurate geospatial data for infrastructure and land-use planning.
Community Organizations: Need insights for equitable resource distribution and sustainable development.

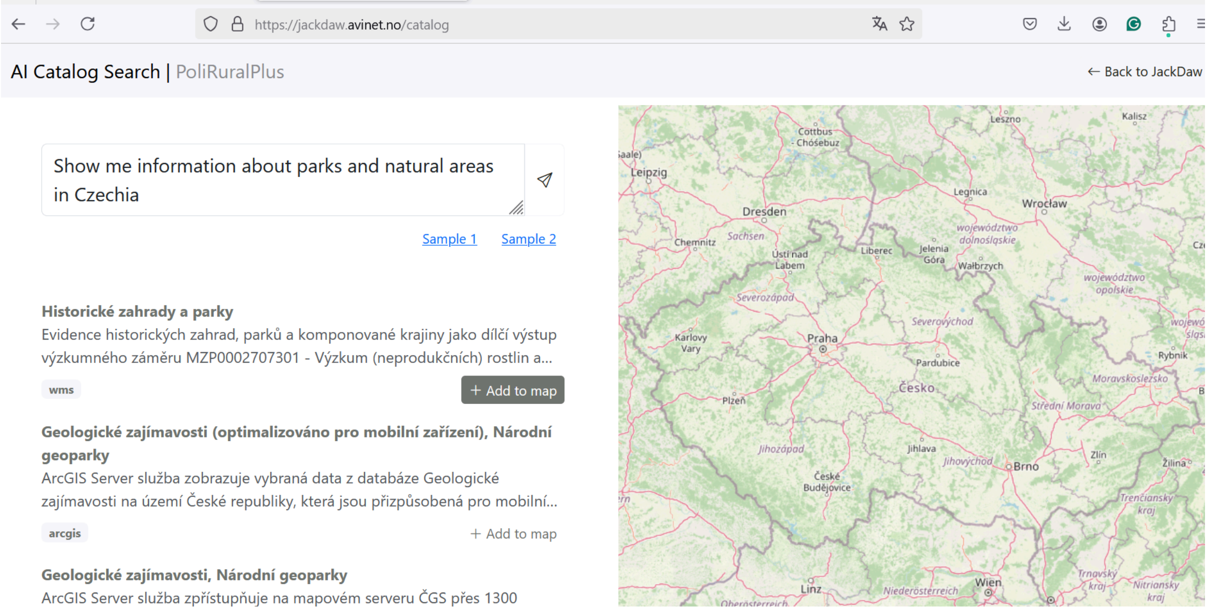
- **Experimental results**

As part of the challenge, we successfully developed and tested **two prototypes** that demonstrate the potential of integrating AI with geospatial analysis for rural development:

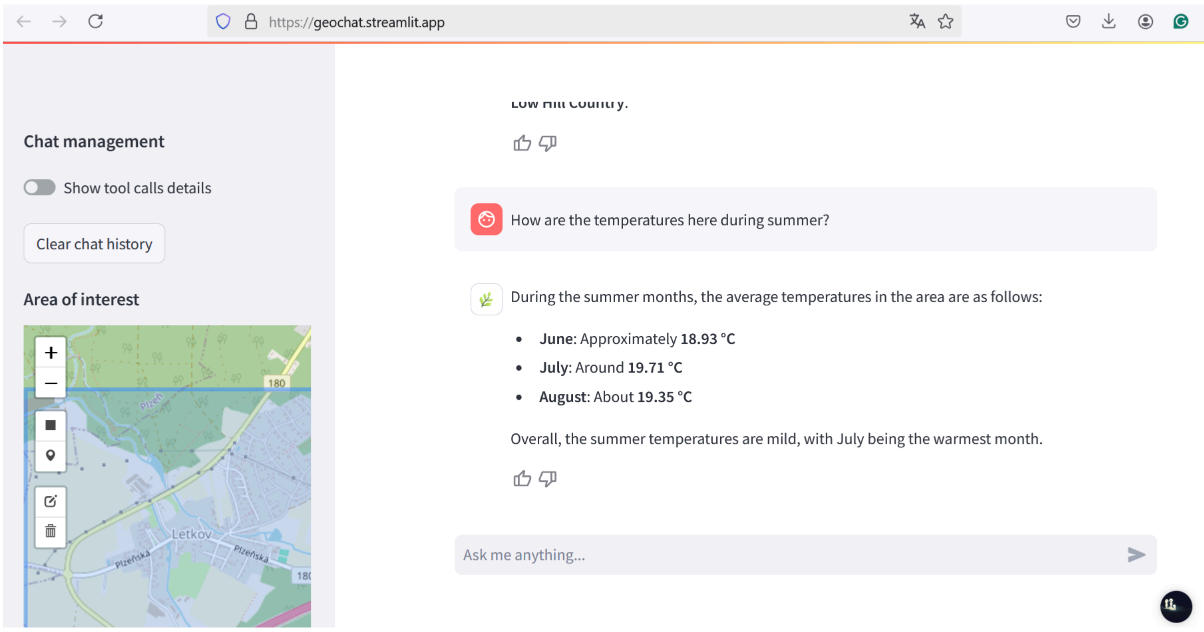
1. **JackDaw** – Accessible at https://jackdaw.avinet.no/, this platform facilitates querying spatial data, analyzing geospatial services, and retrieving insights based on user-defined queries. It supports interaction with INSPIRE metadata, enabling efficient discovery and utilization of geospatial datasets.







2. **GeoChat Chatbot** – Available at https://geochat.streamlit.app/, this chatbot allows users to interact with geospatial data using natural language queries. GeoChat focuses on regional planning tasks such as open land use analysis, smart POI analysis, tourism potential evaluation, climate analysis, weather forecasting, and terrain assessment.



Following the experimental phase, we determined that combining the capabilities of both prototypes into a **single, unified solution** would maximize efficiency and user-friendliness. The merging process will integrate the advanced querying and geospatial service capabilities of **JackDaw** with the natural language processing and conversational interface of **GeoChat**. This unified platform will provide a comprehensive tool for rural development stakeholders, offering intuitive access to geospatial insights, AI-powered analysis, and streamlined decision support.

This integrated solution will deliver the following benefits:

- **Enhanced User Experience**: Seamless interaction with geospatial data through a conversational interface.
- **Comprehensive Analysis**: Combining spatial queries, map services, and AI-driven insights in one platform.
- **Scalability**: Supporting distributed GeoAI models and AI Spaces for large-scale applications.
- **Efficiency**: Faster access to relevant datasets and analyses based on user-specific queries.

This approach ensures a practical and scalable tool that addresses the complex challenges of rural development through advanced AI and geospatial integration.

FINDINGS & CONCLUSION

- **Discussion of the results and findings**

The integration of **AI models** with **geospatial analysis** successfully addressed key rural development challenges. The combination of the **PoliRuralPlus API** and the **INSPIRE metadata connector** enabled seamless access, integration, and analysis of diverse geospatial datasets. This integration facilitated improvements in the following areas:

- **Open Land Use Planning**: Optimized land utilization strategies through accurate and context-aware analysis.
- **Infrastructure Development**: Data-driven planning and prioritization of infrastructure projects based on geospatial insights.
- **Tourism Potential**: Identification of suitable locations for tourism-related investments, such as hotel development, based on POI, terrain, and climate data.
- **Climate and Weather Analysis**: Enhanced forecasting capabilities to support sustainable agricultural practices and resilient infrastructure planning.
- **Terrain Analysis**: Informed land-use decisions by incorporating topographical and elevation data into planning processes.

This solution demonstrates the significant **potential for the development of GeoAI** in Europe. By integrating AI and geospatial analysis within a distributed framework, the system offers a **basic concept for solutions that can generate platforms with lower energy consumption**. This is achieved through optimized data processing, modular workflows, and distributed computational resources, reducing the energy footprint compared to centralized models.

Additionally, the distributed model architecture allows the integration and use of **various Large Language Models (LLMs)**, such as **ChatGPT**, **LLaMA**, and other models. These LLMs can be **trained for specific purposes**, enabling targeted applications for different rural development challenges. For instance:

- LLMs trained for **agricultural planning** can provide specialized insights on crop optimization and resource management.
- LLMs tailored for **infrastructure development** can assist with spatial planning and logistics.
- Models focused on **climate and weather forecasting** can deliver context-specific predictions for rural areas.

This flexible, distributed approach enhances the adaptability and scalability of GeoAI solutions, making them more efficient, sustainable, and applicable across various rural contexts in Europ

- **Further improvements**
- **Connector Enhancement**: Expand the **INSPIRE connector** to support additional services, datasets, and geospatial data formats, ensuring broader interoperability and enhanced data coverage for rural development needs.
- **API Development**: Enhance the **PoliRuralPlus API** with new functionalities, such as advanced data querying, visualization tools, and integration with additional data sources and AI services, to enable a wider range of applications.
- **Model Refinement**: Continuously improve **AI models** through training with localized datasets to increase accuracy and relevance for specific rural development challenges. Implement domain-specific tuning of LLMs for tasks like agricultural planning, infrastructure development, and climate analysis.
- **Testing of Different LLMs**: Evaluate and integrate various **Large Language Models (LLMs)**, such as **ChatGPT**, **LLaMA**, and others, to identify the most effective models for specific geospatial analysis tasks. This approach ensures flexibility and optimization based on use-case requirements.
- **User Training**: Develop comprehensive **training programs** and materials for stakeholders, including farmers, policymakers, and community organizations, to effectively adopt and utilize GeoAI tools in their workflows.
- **Scalability**: Optimize the computational infrastructure to support **large-scale deployments** of GeoAI solutions across different regions. Implement a **distributed GeoAI model** to enable efficient processing, reduce energy consumption, and facilitate scalability.
- **Interface Development**: Create **user-friendly interfaces** that simplify the process of adding new types of analyses and datasets. These interfaces will allow stakeholders

and developers to expand the system's capabilities without extensive technical knowledge, fostering flexibility and adaptability.

These enhancements will ensure that the GeoAI solutions remain **versatile, sustainable, and effective**, addressing the evolving needs of rural development and maximizing the potential of AI-driven geospatial analysis.