# 1. Support for selection and analysis of documents available on Hub for assist to provide thematic analysis

The **PoliruralPlus Hub** serves as a comprehensive repository for documents related to the preparation of regional development plans, including legislation, existing strategies, research papers, and other resources. Its primary objective is to assist users in leveraging this repository to address their challenges effectively. Beyond the analysis of spatial data, users require insights from diverse sources such as textual documents, structured datasets, and multimedia presentations.

The task will consist of two sequential phases:

1. **Document Indexing and Selection**
   Utilize advanced indexing techniques to process and classify documents within the Hub. Employ semantic search and relevance ranking algorithms to identify and prioritize documents based on predefined topics and user-specific queries.
2. **Document Analysis and Query Optimization**
   Analyze the selected documents using state-of-the-art Natural Language Processing (NLP) models to extract actionable insights. Additionally, refine the response generation capabilities of Large Language Models (LLMs) to improve their contextual understanding and accuracy in addressing user queries.

# 2. On line interpretation of OGC WMS services

The task is focused on efficient image processing to enhance the usability of map data provided through Web Map Services (WMS). WMS typically delivers maps as images, where regions are often defined not only by colors but also by various forms of symbology. The objective is to develop capabilities to analyze and interpret the content of such images. This is particularly critical because most map services rely on WMS, which is well-suited for visual interpretation but presents significant challenges for machine-to-machine (M2M) understanding.

An advanced extension of this task involves the semantic interpretation of raster-based WMS outputs, including high-resolution orthoimagery. By enabling the automated extraction and interpretation of meaningful information from these raster images, we can substantially enhance the utility and interoperability of the existing spatial digital infrastructure. This capability would unlock new opportunities for integrating geospatial data into AI-driven workflows and decision-making processes.

# 3. Design and Testing of Metadata API for Integration of New Analyses into JackDaw

The objective is to define a standardized approach for describing new spatial and non-spatial analyses, enabling seamless integration of their APIs into the JackDaw system. This involves the design and testing of a metadata schema to ensure that analyses are appropriately characterized with relevant attributes, including input requirements, processing capabilities, and expected outputs.

The current implementation of JackDaw allows it to select from predefined analyses and datasets, identify the most relevant ones, and incorporate insights from these analyses into the responses generated by Large Language Models (LLMs). Extending this functionality to support dynamic integration of new analyses will enhance the system's adaptability and scalability, ensuring its ability to leverage diverse analytical workflows and datasets effectively.

# 4. Integration of the Atlas of Best Practices with JackDaw

The **Atlas of Best Practices** (available at [https://sumavaprodukt.regionalnispeciality.cz/atlas/](https://sumavaprodukt.regionalnispeciality.cz/atlas/)) provides a comprehensive collection of regional data on food producers, tourism, cultural heritage, and other key regional features. Its integration with JackDaw aims to enrich the system's knowledge base by incorporating localized, context-specific information that can serve a diverse range of stakeholders, including tourists, planners, and regional organizations.

Enhance the capabilities of an AI system (referred to as "JackDaw") by integrating it with a rich geo-based knowledge resource: the **Atlas of Best Practices**. The Atlas is an interactive database offering valuable regional data on food producers, tourism hotspots, cultural heritage, and more. By fusing this information with **Large Language Models (LLMs)**, students will explore innovative ways to provide **personalized recommendations** and insights to users like tourists, urban planners, and regional organizations. This integration will enable JackDaw to dynamically access and utilize data from the Atlas, enhancing its ability to provide tailored insights and recommendations. For instance, users querying JackDaw for tourism or regional planning information will benefit from additional, highly relevant data sourced from the Atlas, thereby improving the quality and relevance of the system's outputs.

**What's in it for you?**

- Work on **real-world AI applications** combining LLMs with geographical and contextual data.
- Learn **popular techniques** such as **Retrieval-Augmented Generation (RAG)** to make AI smarter and more context-aware.

- Build systems capable of **dynamic recommendations** in areas like tourism, urban planning, and cultural heritage promotion.

**Key Tasks**:

1. **Understand the Atlas of Best Practices**:
    - Explore its resources here.
    - Identify key data categories that could enrich JackDaw's AI-driven recommendations.
2. **Integrate Geo Data with AI**:
    - Use LLMs to understand and process user queries.
    - Implement **RAG** by coupling LLMs with a retrieval system to fetch geo-specific data from the Atlas dynamically.
3. **Tailor Outputs to User Needs**:
    - Design scenarios where users benefit from context-enriched recommendations, such as:
        - **Tourists**: Personalized itineraries based on regional highlights.
        - **Planners**: Suggestions for sustainable regional development.
        - **Local Producers**: Marketing insights for regional products.

**Suggested Methods and Tools**:

- **Retrieval-Augmented Generation (RAG)**:
    - Combine LLMs (e.g., GPT models) with a vector database (e.g., Pinecone, Weaviate) to fetch and summarize geo-specific data.
- **Geospatial Data Integration**:
    - Tools like **QGIS**, **PostGIS**, or Python libraries (e.g., **Geopandas**, **Folium**) to visualize and analyze regional data.
- **Information Retrieval**:
    - Implement indexing and search with Elasticsearch or similar systems.
- **Data Enrichment**:
    - Use knowledge graphs (e.g., Neo4j) to model relationships between producers, cultural sites, and tourists.
- **Visualization and Deployment**:
    - Create dynamic dashboards using **Streamlit** or **Dash** for real-time interaction with JackDaw.

**Deliverables**:

1. A brief **proposal** outlining the chosen integration methods and their expected impact.
2. A **prototype or simulation** demonstrating JackDaw's ability to use Atlas data for generating enriched, user-specific outputs.
3. A reflective **report** discussing challenges, solutions, and potential real-world applications.

**Why This Matters**: This project merges **AI innovation** with practical regional development, demonstrating how technologies like LLMs and geospatial data can solve **real-world problems**. The skills you gain here are highly sought-after in industries like tourism, urban planning, and data-driven marketing. Plus, you'll be part of a growing trend in AI that emphasizes **contextual and localized intelligence**.

# 5. Analysis of LPIS Data for Regional Environmental Sustainability and Short Supply Chain Potential

Local and regional environmental sustainability, as well as the potential for developing short supply chains, are influenced by two interrelated factors: the scale of farms and the size of fields. Regions characterized by smaller farms and smaller fields typically exhibit greater potential for short supply chains and demonstrate improved environmental sustainability. Such regions are also better positioned to adapt to climate change due to increased landscape diversity and resilience.

Czech **Land Parcel Identification System (LPIS)** data contains valuable information to support these analyses. The task involves defining appropriate metrics to assess sustainability and short supply chain potential. This includes identifying relevant parameters such as farm scale, field size, and associated environmental indicators. Furthermore, the development of an analytical tool is required to process and evaluate this data effectively. The tool should be integrated with JackDaw to enable automated analyses and provide actionable insights into regional planning, environmental strategies, and the promotion of sustainable short-supply chains.

# 6. Increase Resolution of Climatic Data

The task focuses on increasing the resolution of climatic data, such as ERA5 (spatial resolution in the range of several kilometres), to match the finer spatial resolutions of satellite data, typically ranging from 10 to 100 meters. This requires developing AI models for super-resolution, such as diffusion models or neural operators, capable of translating low-resolution climatic data to higher resolutions by leveraging satellite-derived information, such as the thermal band of Landsat imagery.

The modelling process will involve identifying how these relationships vary across different land cover types, allowing for the construction of more accurate and context-specific models. A critical step will be the integration of climatic data with vegetation and moisture indices derived from satellite data. This will enable an in-depth analysis of the potential for building more precise local-scale models.

By refining these models, the aim is to enhance the spatial accuracy of climatic data for localized environmental assessments, land cover studies, and climate adaptation strategies, ultimately contributing to improved decision-making in regional and local contexts.

# 7. ALIANCE Satellite Images Challenge

The ALIANCE project has developed a methodology and software for reconstructing missing satellite imagery or their obscured parts due to cloud cover. The objective of this challenge is to implement additional analyses on these reconstructed images, enhancing their utility in various domains. The analyses will include, but are not limited to:

- Management zone delineation for precision agriculture.
- Land cover and crop classification to support agricultural planning and monitoring.
- Vegetation condition analysis to assess crop health and growth status.
- Additional spatial and temporal analyses to expand the applicability of reconstructed images in environmental monitoring and sustainable resource management.

This challenge will focus on validating the reconstructed images for these applications and improving the accuracy and efficiency of the analyses. The outcomes aim to increase the value of satellite-derived data, particularly for regions frequently affected by cloud cover, by ensuring they remain viable for advanced geospatial and agricultural workflows.

# 8. ALIANCE Weather Forecast Challenge

The **ALIANCE project** has developed a weather forecast sensor capable of capturing localized weather patterns, such as transitions between rain and sunny conditions. The goal of this challenge is to improve the accuracy and spatial resolution of weather forecasts by integrating climatic data with localized sensor measurements.

Key objectives include:

- **Integration with SensLog and AgroInfo platforms** to facilitate real-time data sharing and advanced analysis.
- Development of additional predictive models based on improved weather forecasts, such as:
  - **Frost prediction** for mitigating crop damage.
  - **Pest risk assessment** to inform pest management strategies.
  - Other weather-dependent analyses to support precision agriculture and environmental monitoring.

This challenge will focus on optimizing data fusion techniques and validating the accuracy of predictions. The ultimate aim is to enable more actionable insights for agricultural and environmental stakeholders by leveraging enhanced weather forecasting capabilities.