D4.1 PoliRuralPlus Platform Design

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101136910

Document information

Key information	Data
Project reference number	101136910
Project acronym	PoliRuralPlus
Project title	Fostering Sustainable, Balanced, Equitable, Place-based and Inclusive Development of Rural-Urban Communities' Using Specific Spatial Enhanced Attractiveness Mapping ToolBox
Work package (WP) & deliverable number & name	WP4, D4.1 PoliruralPlus Platform Design
WP leader and co-leader	Stein Runar Bergheim (AV)
Deliverable leader & key author(s)	Stein Runar Bergheim (AV), Karel Charvát (CCSS), John O'Flaherty (MAC), Daniel Molina (SINNO), František Zadražil (CCSS), Otakar Čerba (CCSS)
Contributors and authors	
Peer reviewers	John O'Flaherty (MAC), Daniel Molina (SINNO), David Pešek (CVUT), Raul Palma (PSNC), Aron Rynkiewicz (PSNC), Pavel Kordík (CVUT)
P.R. approval date/version	30.12.2024 | Version V.1
Document type	R — Document, report
For public dissemination	Yes, PU-public
Brief abstract	The PoliRuralPlus Platform Design (D4.1) outlines the development of a modular, scalable digital ecosystem aimed at fostering sustainable, balanced, and inclusive rural-urban development. Building on the foundations of the original PoliRural project, the platform integrates advanced geospatial tools, AI technologies, and data sources to address regional development challenges. Key components include AI-driven models, data management services, and collaborative applications tailored to enhance policy-making, foresight analysis, and stakeholder engagement. The design emphasizes interoperability, real-time data integration, and user-centric development, ensuring adaptability within existing regional infrastructures. The platform supports objectives aligned with the European Green Deal, promoting innovation, resilience, and evidence-based governance strategies across diverse rural and urban regions.

Disclaimer

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Document history

Version	Date	Main measures/changes	Responsible
0.1	15.11.2024	Setting up high-level table of contents, distributing responsibilities	Karel Charvát (CCSS), Stein Runar Bergheim (AV)
0.2	29.11.2024	Initial text in key chapters, initial system architecture diagram	Karel Charvát (CCSS), Stein Runar Bergheim (AV), František Zadražil (CCSS), Otakar Čerba (CCSS), Raul Palma (PSNC), John O'Flaherty (MAC)
0.3	06.12.2024	Enhanced draft	Karel Charvát (CCSS), Stein Runar Bergheim (AV), John O'Flaherty (MAC), Otakar Čerba (CCSS)
0.4	10.12.2024	Enhanced draft	Karel Charvát (CCSS), Stein Runar Bergheim (AV), Aron Rynkiewicz (PSNC)
0.5	13.12.2024	Enhanced draft, reviewed	Karel Charvát (CCSS) , Stein Runar Bergheim (AV), Matúš Botek (CCSS), John O'Flaherty (MAC), Aron Rynkiewicz (PSNC), Daniel Molina (SINNO)
0.6	20.12.2024	Final version	Karel Charvát (CCSS), Stein Runar Bergheim (AV)
V1.0	30.12.2024	D4.1 PoliruralPlus Platform Design, version 1 uploaded to the EU portal/SYGMA	Coordinator

Table of content

Abbreviations

Acronym	Title
AI	Artificial intelligence
API	Application Programming Interface
ChatBot	A software application that behaves as a conversation partner
CLC	Corine land cover
CMS	Content management system
DestinE	Destination Earth
DIH	Digital Innovation Hub
ERA5	Fifth generation ECMWF reanalysis for the global climate and weather
FADN	Farming accountancy data network
GeoAI	Geospatial Artificial Intelligence
GeoJSON	Geographical JavaScript Object Notation, lightweight JSON based format for spatial data representation on the Web
GFS	Global Forecast System
GIS	Geographical Information System
GISCO	EU statistical and foundation GIS data service
GPT	Generative Predictive Text
HILUCS	High-Resolution Layer of Land Use and Cover System
ICT	Information and Communication Technology
IoT	Internet of Things
JSON	JavaScript Object Notation, see also GeoJSON
KML	Keyhole Mark-up Language, XML-based format for spatial data representation on the Web
LLM	Large Language Model
LULC	Land use land cover

Acronym	Title
MAA	Multi-Actor Approach, methodology for project execution
MAE	Mean Absolute Error
MCA	Multi-criteria analysis
MSE	Mean Square Error
NDVI	Normalized difference vegetation index
NER	Named Entity Recognition
OAuth	Authentication standard
OGC	Open Geospatial Consortium
OLU	Open Land Use (map) a data set of land use data in Europe
OSM	Open Street Map, global open GIS data source
PSI	Public Sector Information, as in the PSI Directive
RAE	Rural attractiveness explorer
RAI	Rural attractiveness index
RAP	Regional action plan
SaaS	Software as a Service, term to describe software that the use does not install locally but that relies on web server functionality to provide interactive capabilities through the users web browser
Sentinel	Satellites belonging to EU space programme
UX	User experience, typically used about User experience design
WFS	Web Feature Service
WMS	Web Map Service

Executive Summary

The **PoliRuralPlus Platform Design** (D4.1) outlines a comprehensive framework to foster sustainable, inclusive, and data-driven rural-urban development across Europe. Building on the success of the original PoliRural project, the PoliRuralPlus platform leverages advanced tools, data sources, and AI technologies to address emerging challenges while promoting innovation, collaboration, and regional development strategies.

1. Project Objectives and Platform Contributions

The PoliRuralPlus platform is designed to address seven core objectives:

1. **Foresight-Driven Policy Making**: Tools for governance analysis, scenario planning, and stakeholder engagement support interregional cooperation and integrated strategies.
2. **Economic Growth and Circularity**: Advanced GIS-based tools highlight business opportunities, cultural assets, and regional economic potential.
3. **Improved Service Delivery**: Modules for tracking Regional Action Plans (RAPs) enhance social and economic resilience.
4. **Alignment with EU Green Deal**: Integration of datasets like Copernicus and Eurostat aligns development with biodiversity and sustainability goals.
5. **Cross-Disciplinary Collaboration**: Interfaces for open data exchange foster interdisciplinary decision-making and innovation.
6. **Mission-Oriented Experimentation**: Analytical tools enable scenario testing for circular economies and well-being strategies.
7. **Alignment with NEB**: Facilitates knowledge transfer and project collaboration to meet sustainable development goals.

2. Platform Design and Key Components

The platform is a **modular, scalable ecosystem** that integrates seamlessly with existing systems. Key features include:

- **Data Integration and Management**: Supports real-time data processing through APIs, automated pipelines, and manual fallback storage (e.g., PostgreSQL).
- **AI-Driven Analysis**: Tools like Jackdaw (spatial chat agent), Vulture (productivity sandbox), and Magpie (OGC metadata catalog) enhance decision-making through AI-powered geospatial intelligence and natural language processing.
- **Component Highlights**:
    - **Tourism Suitability Module**: Identifies optimal hotel locations using multi-criteria analysis and gap assessments.
    - **WMS Interpreter**: Converts raster map images into actionable geospatial data.
    - **EO Image Interpretation**: Detects crops, land cover, and changes using Sentinel satellite imagery.
    - **Document Parsing and RAG**: Processes multilingual documents, enabling precise, context-aware answers supported by citations.
- **Technical Architecture**:
    - Built on **FastAPI** (Python backend) and a modern front-end (**React**).

- Utilizes **Qdrant** for vector embeddings and PostgreSQL for structured data storage.
- Ensures security and scalability with OAuth2/OIDC authentication, Docker containerization, and SaaS delivery models.

3. Innovation and Integration

The platform enhances existing components from PoliRural, such as the Digital Innovation Hub (DIH) and the Rural Attractiveness Explorer, while introducing new tools to meet stakeholder needs. It focuses on **AI and GeoAI integration** to streamline workflows, improve user accessibility, and optimize rural-urban linkages. Tools are designed to minimize complexity and ensure broad adaptability for local administrative units across Europe.

Key innovations include:

- **AI-Powered Spatial Analysis**: Jackdaw enriches user queries with geospatial context.
- **Open Data Utilization**: Integration of PSI, INSPIRE, Copernicus, and OpenStreetMap ensures comprehensive data coverage.
- **Interactive Collaboration**: Tools like Map Whiteboard enable real-time, multi-party spatial planning.

4. Deployment and Future Vision

The platform will be implemented incrementally with pilot regions across Europe, incorporating feedback from stakeholders to ensure continuous improvement. By emphasizing integration over standalone development, PoliRuralPlus ensures long-term sustainability and alignment with existing infrastructures.

The vision of the platform is to deliver **scalable, AI-powered tools** that address rural challenges, support evidence-based policy making, and promote sustainable rural-urban development across Europe.

1. Introduction

The PoliRuralPlus project aims to improve the quality of life for people living in urban and rural areas across the EU by promoting coordinated and collaborative development within the regions. It is doing so by fostering a sustainable, balanced, equitable, place-based and inclusive development of rural and urban areas through improved connections, governance arrangements, and integrated territorial policies that prioritise experimentation and innovation in domains that favour bi-directional urban-rural synergies and the development of a well-being economy, driven by foresight, planning and implementation of integrated urban-rural strategies.

1.1 Contribution of the PoliRuralPlus Platform to the Project Objectives

The objectives of PoliRuralPlus are central to shaping the design and functionality of its platform, ensuring its capacity to address complex rural-urban challenges through foresight, innovation, and collaboration.

The first objective (O1) establishes a foresight-based framework for interregional cooperation, which directly informs the platform's core functionalities. This requires tools for effective governance analysis, scenario planning, and stakeholder collaboration, designed to overcome policy barriers and support integrated rural-urban strategies. The inclusion of system dynamic modeling and participatory tools ensures alignment with governance goals and fosters synchronized local and EU policy implementation.

Objective two (O2) emphasizes the enhancement of business opportunities and circular economy initiatives in rural areas. The platform must integrate data visualization tools and mapping services, such as GIS-enabled dashboards, to highlight cultural and natural assets. Features to support connectivity and innovation ecosystem development are crucial, leveraging insights from successful pilots.

The third objective (O3) necessitates tools that strengthen mutual service access and social connectivity. The platform should include modules for tracking progress on regional action plans (RAPs), measuring resilience, and fostering social innovation. Pilot stakeholders can use such tools to build adaptable solutions for service delivery.

Objective four (O4) aligns the platform's development with EU Green Deal priorities, requiring integration with datasets from initiatives like Copernicus and Eurostat to measure progress against biodiversity and sustainability goals. The platform also needs functionalities to simulate impacts of new policies under frameworks like the Farm-to-Fork strategy.

Cross-disciplinary collaboration highlighted in objective five (O5) calls for leveraging European Research Infrastructures and open data spaces to support integrated development. The platform must provide interfaces for data exchange and synthesis, facilitating decision-making across disciplines and regions.

Objective six (O6) drives the need for mission-oriented experimentation capabilities, necessitating advanced analytical tools for system dynamics and collaborative decision-making. These tools are pivotal for stakeholders to explore scenarios promoting circularity and well-being economies.

Finally, objective seven (O7) ties the platform to broader EU initiatives like the New European Bauhaus (NEB). The design must facilitate knowledge transfer and collaborative project development between rural and urban stakeholders, ensuring alignment with aesthetic and sustainable development goals.

This deliverable D4.1 outlines the design of the PoliRuralPlus Platform, which serves as a comprehensive digital ecosystem to support integrated rural-urban development. The platform is designed to facilitate engagement with relevant stakeholders, including farmers, urban consumers, local authorities, businesses, NGOs, and community groups, by establishing effective communication channels and providing tools for collaboration. It forms the foundation for analyzing rural-urban linkages, identifying opportunities, and addressing challenges, including those amplified by the COVID-19 pandemic. The platform supports PoliRuralPlus pilots in designing and implementing integrated urban-rural strategies, which are key components of their Regional Action Plans (RAPs).

The PoliRuralPlus platform is a crucial enabler for showcasing the added value that technical tools can bring to pilot end-users, while its modular, virtual design ensures flexibility for integration within existing tools and organizational contexts in European rural regions. Rather than functioning as an isolated technological solution, the platform acts as a complementary and integrative resource that builds on stakeholders' familiar systems and workflows. Its design minimizes complexity, instead streamlining data integration and collaboration to enhance evidence-based policymaking while reducing effort and increasing efficiency in achieving rural development goals.

Recognizing the challenges posed by an overload of standalone platforms, PoliRuralPlus emphasizes interoperability and modularity. This allows the platform to adapt to diverse local contexts, facilitating its exploitation within existing infrastructures without imposing additional burdens. By embedding itself into the broader policy and operational ecosystem, the platform serves as a catalyst for innovation, tangible outcomes, and collaboration among local authorities, businesses, and community groups.

PoliRuralPlus is a dynamic ecosystem comprising advanced tools such as AI/ML algorithms, GIS-services , and analysis models. These tools empower users with real-time data integration, scenario modeling, and shared analytical capabilities, significantly enhancing the design and implementation of effective Regional Action Plans (RAPs). The platform also incorporates communication and knowledge-sharing mechanisms essential for fostering cooperation, supporting innovation, and strengthening stakeholder engagement.

For rural development professionals, the platform simplifies workflows by automating data processing, generating actionable insights, and offering interactive collaboration interfaces. It reduces manual workloads, facilitates the creation of tailored strategies, and strengthens the monitoring and evaluation of policy impacts. These features collectively ensure that the platform adapts to the operational realities of rural regions, avoiding duplication of efforts and aligning seamlessly with existing practices.

Aligned with the above project objectives, the platform supports foresight-based frameworks, enhances connectivity and governance, and fosters sustainability through innovative, actionable strategies. It contributes to key performance indicators, such as improved interregional coordination, increased business and innovation opportunities, enhanced mutual service development, and significant advancements toward European Green Deal objectives. By addressing these targets, the platform provides a scalable, adaptable, and impactful solution to drive sustainable and integrated rural-urban development across Europe.

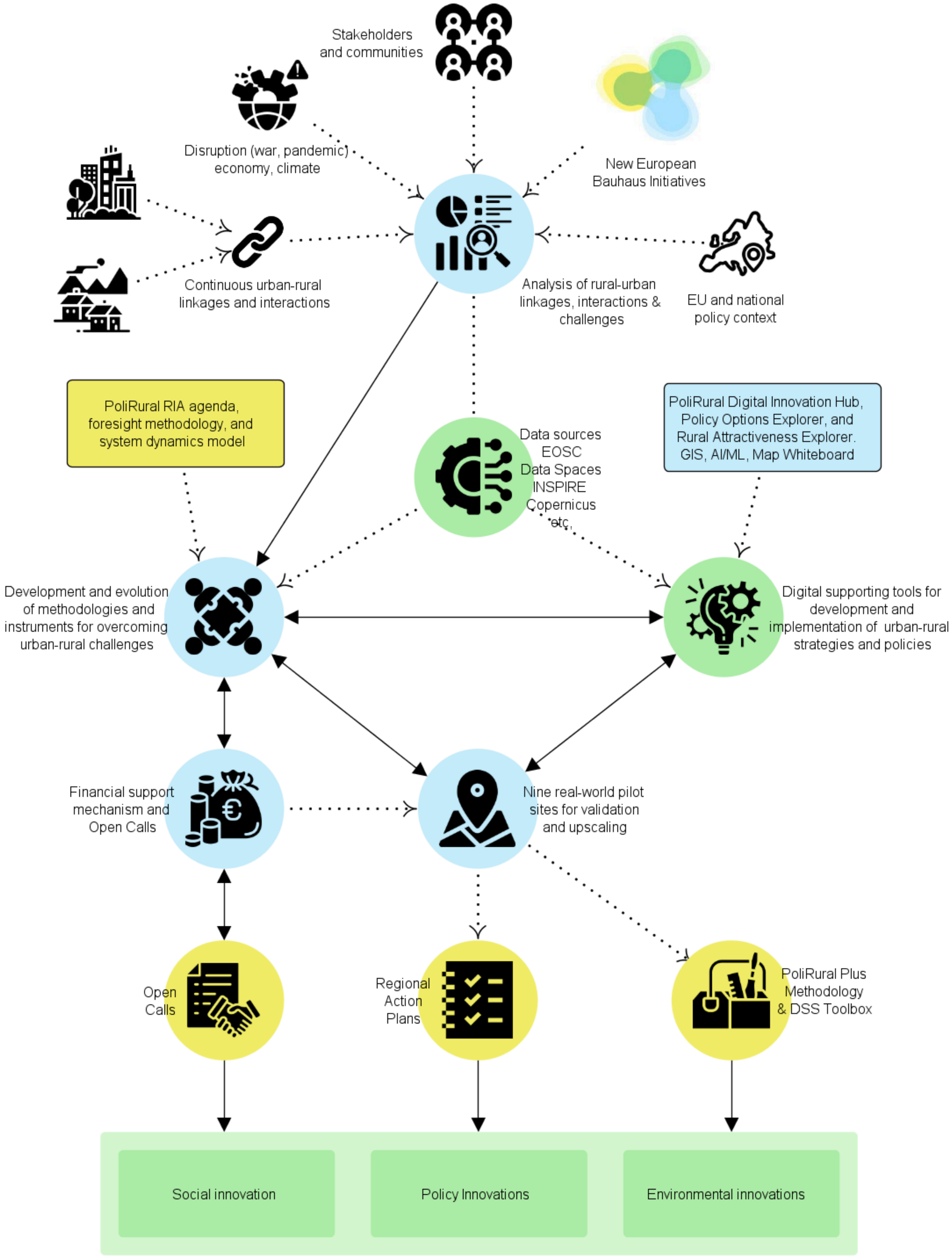


Figure 1: Overall concept of the PoliRuralPlus project

1.2 Building on and Following Up the PoliRural Project

The PoliRural project provided a strong foundation for addressing rural development challenges by creating innovative digital tools, participatory frameworks, and foresight methodologies. Central to its success was the **Digital Innovation Hub (DIH)**, a collaborative platform that integrated data, tools, and methodologies to support policy evaluation, geospatial analysis, and stakeholder engagement. The DIH facilitated co-creation among policymakers, businesses, researchers, and local communities, fostering inclusive and actionable Regional Action Plans (RAPs) aligned with EU priorities such as the Green Deal, CAP reform, and biodiversity strategies.

Key tools developed in PoliRural included the Policy Options Explorer, Rural Attractiveness Explorer, and Semantic Explorer, which supported scenario modeling, impact analysis, and text mining. These tools were instrumental in enabling data-driven decisions and participatory foresight in 12 pilot regions.

PoliRuralPlus builds on and extends the concept of the PoliRural Digital Innovation Hub (DIH) with advanced tools and mechanisms to address emerging challenges and opportunities.

Enhanced Digital Innovation Hub:
The collaborative GIS tools from PoliRural, will have an expanded role in PoliRuralPlus, integrating AI-driven geospatial analytics for real-time co-creation, visualization, and annotation of territorial strategies. Legacy tools like the Semantic Explorer are being replaced by AI-powered solutions that deliver automated data processing, enhanced insights, and user-friendly interfaces. Upgraded tools such as the Rural Attractiveness Explorer and Policy Options Explorer will feature predictive modeling, dynamic scenario testing, and improved regional analysis.

Strengthened Stakeholder Engagement:
PoliRuralPlus emphasizes stakeholder-driven innovation through Open Calls, encouraging the adoption of advanced tools and data services from our platform like the Policy Advisor, Jackdaw, Magpie etc. Tailored training and support from the DIH ensure inclusive participation, empowering stakeholders to leverage new technologies effectively.

Regional Technology Deployment:
The DIH serves as a hub for rolling out AI-enhanced tools and geospatial technologies, fostering innovation and sustainable rural-urban integration. Open Calls promote experimentation and scalability, tailoring solutions to local contexts.

Integrated Modular Ecosystem:
The DIH’s modular design allows users to customize their experience, selecting tools like MapWhiteboard, AI-driven explorers, and system dynamics models based on specific needs. This ensures flexibility across diverse regions and evolving challenges.

By advancing AI integration, enhancing tools, and fostering innovation through stakeholder engagement, PoliRuralPlus creates a dynamic, scalable framework for sustainable and inclusive rural-urban development.

1.3 Technical Foundations and Envisioned Output

When the PoliRuralPlus project was initially conceptualized, its technological baseline was informed by contemporary digital innovation practices, including traditional GIS tools, geospatial data integration, and participatory foresight methodologies. However, since the project's inception, significant advancements in artificial intelligence (AI), particularly the development of Large Language Models (LLMs), have redefined the possibilities for data processing, analysis, and user interaction.

These developments in AI have introduced capabilities for handling complex, unstructured data and enhancing human-computer interactions through natural language processing. The integration of LLMs into geospatial intelligence has enabled unprecedented opportunities for automating and democratizing advanced data analysis processes.

PoliRuralPlus addresses the challenges faced by Europe’s 90,000–100,000 local and regional administrative units. To achieve impact across this broad user base, the project focuses on developing simple, abstract tools that are widely replicable, rather than complex solutions tailored to specific organizations. Tailored tools are costly to develop and maintain, often becoming obsolete due to evolving contexts or generational shifts in staff and management. Realistically, reaching a significant portion of the target audience with a single, unified application is feasible, while the likelihood of success with a range of stand-alone tools is minimal.

To address this, PoliRuralPlus employs an architecture emphasizing integration into existing systems, embodied by the PoliRuralPlus Innovation Hub. This hub, built on the Hub4Everybody platform, mirrors the geoportals and decision support systems commonly used by European administrations, accommodating varying levels of complexity and content. Through these tools, policymakers, decision-makers, and stakeholders can access, visualize, and manage data effectively.

The project’s strategy for tangible sectoral impact is to create modular components that are easy to integrate into existing systems via soft-linking or standardized machine-to-machine protocols, such as the OGC-family standards. Consequently, PoliRuralPlus focuses its technical development on modular experiments that, if successful, can integrate seamlessly with the general-purpose information systems already in place in local administrative units. Based on user requirements, the project is developing the following embeddable modules:

- **Policy Advisor:** A retrieval-augmented generation AI assistant populated with policy and contextual documents relevant to regional development.
- **Jackdaw:** A chat agent that enriches conversations with location-specific information, leveraging PSI and INSPIRE data infrastructures, alongside outputs from EU initiatives such as SPOI and OLU.
- **Vulture:** A productivity sandbox aimed at automating “busy work” by processing information streams associated with individuals or topics using AI.
- **Magpie:** An advanced metadata catalogue enabling natural language queries to identify data resources in PSI portals, INSPIRE geoportals, EU data spaces, EO data, and DestinE. It supports sustainability-focused analyses for specific areas and timelines.
- **Multi-Actor Approach Tool:** A solution for registering, organizing, and managing stakeholders in projects that follow the MAA methodology.
- **Atlas of Best Practices:** A repository of good practices from across Europe, developed and refined over several years.
- **E-Market for Local Food:** A platform concept for managing and marketing local food production for B2C and B2B applications.
- **Map Whiteboard:** A collaborative map-making tool for efficient stakeholder interaction in rural development.

- **Rural Attractiveness and Innovation Potential:** A data analysis suite enabling regional development actors to assess the potential of specific areas.

This modular, integrative approach ensures flexibility, scalability, and alignment with the existing technological and organizational contexts of rural regions across Europe.

1.4 Relationship to the other PoliRuralPlus Work Packages

Work Package 4 (WP4) serves as the technical foundation for the PoliRuralPlus project by developing the **Digital Innovation Hub (DIH)** and its associated tools, such as the **Map Whiteboard** and AI/ML-based GeoAI solutions. WP4 ensures that advanced digital tools are designed, developed, and refined to support regional planning, policy development, and stakeholder engagement. However, the success of WP4 is closely tied to the outputs and contributions of other work packages, ensuring a comprehensive, integrated approach across the project.

WP5 provides crucial input for WP4 by identifying the specific needs of pilot regions through detailed consultations and regional assessments. This information guides the customization of WP4's tools, ensuring that they address the unique challenges faced by each region. Additionally, WP5 has developed the **Multi-Actor Tool**, a collaborative solution that facilitates stakeholder engagement and participatory decision-making. WP4 integrates this tool into the DIH, enhancing its ability to foster inclusive and effective regional planning.

WP2 plays a critical role in managing the testing and validation of WP4 outputs. By coordinating pilot implementations and ensuring rigorous evaluation processes, WP2 guarantees that tools developed under WP4 align with the project's objectives and stakeholder requirements. The testing processes, led by WP2, provide iterative feedback to WP4, enabling continuous improvement of the DIH and its components.

WP3 contributes its Regional Action Plan (RAP) methodology as a foundational element for the DIH's analytical and decision-support tools. WP3’s expertise in foresight and policy planning informs the development of WP4’s tools, such as system dynamics modeling and policy impact assessments. By embedding RAP methodologies into the DIH, WP4 ensures that its tools support evidence-based and participatory policy-making.

WP6 broadens the engagement of WP4 by facilitating Open Calls, which invite stakeholders outside the initial pilot regions to contribute to the development and testing of the DIH. These Open Calls help WP4 refine its tools and methodologies by incorporating diverse perspectives and ensuring their applicability in a wide range of regional contexts. Open Calls also expand the user base of the DIH, fostering wider adoption and impact.

WP7 supports WP4 by disseminating the DIH and its tools to a broader audience and developing business models for their sustainability. By aligning dissemination strategies with WP4’s outputs, WP7 ensures that the tools gain visibility among policymakers, businesses, and researchers. Furthermore, WP7 focuses on creating business models that ensure the long-term viability and scalability of the DIH, extending its impact beyond the project’s timeline.

Finally, WP5’s development of the **LLM Dashboard**, which leverages large language models for intuitive data analysis and decision-making, is integrated into WP4’s DIH. This integration enhances the platform’s functionality, enabling users to analyze and interpret complex data more effectively.

By connecting the analytical, methodological, and participatory efforts of the project, WP4 integrates contributions from WP2, WP3, WP5, WP6, and WP7 to deliver a comprehensive and impactful digital platform. This interconnected approach ensures that the tools developed under WP4 are robust, user-focused, and aligned with the overarching objectives of PoliRuralPlus.

2. Existing Tools, Components, and Data

2.1 Existing Components and their Current State

The **PoliRural** project established a robust technological foundation through its **Digital Innovation Hub (DIH)** and a suite of integrated tools designed for rural policy-making, analysis, and participatory engagement. These tools, validated during **PoliRural**, included advanced geospatial services, foresight methodologies, and collaborative platforms. In **PoliRuralPlus**, this foundation is further enhanced by integrating outcomes and tools from related projects such as **Data4Food**, **THEROS**, and **ALIANCE**. This integration enables more comprehensive data-driven decision-making, fosters improved stakeholder engagement, and supports greater transparency and efficiency in rural development processes.

2.1.1 Key Components and Their Current State

1. **Digital Innovation Hub (DIH):** The DIH acts as a central platform for sharing knowledge, tools, and best practices. It is designed to support:
    - Collaborative decision-making through its social space features, including forums, blogs, wikis, and science shops.
    - Experimentation with data analysis and visualization tools.
    - Hosting and promoting pilot projects and regional policy results.
    - Access to open data and metadata for regional planning and decision-making.
    - The DIH is powered by a cloud-based infrastructure leveraging OpenStack, ensuring scalability and adaptability for diverse regional needs. While functional, some components require modernization to align with current advancements in AI and GeoAI technologies.
2. **Rural Attractiveness Explorer (RAE):** This tool helps evaluate the attractiveness of rural areas by generating the **Rural Attractiveness Index (RAI)**, which aggregates key indicators like social capital, economic opportunities, and environmental assets. The RAE is integrated into the DIH and facilitates comparative regional analyses. (https://sumavaprodukt.regionalnispeciality.cz/)
3. **Atlas of Best practises -** The Atlas of Best Practices is an interactive web tool showcasing transferable solutions for rural development, innovation, and sustainability. Refined under PoliRural and integrated with the PoliRuralPlus Digital Innovation Hub (DIH), it features region-specific best practices from projects like Enabling, Dalia, and Data4Food2030. With interactive mapping and a collaborative framework, stakeholders can explore, contribute, and update content dynamically. The Atlas fosters collaboration, supports knowledge exchange, and offers scalable solutions to common challenges, promoting sustainable regional development.
4. **Multi Actor Approach Tool -** The multi-actor approach tool was initially developed by Social InnoLabs within the Cities2030 project to foster collaboration and co-creation among diverse stakeholders such as policymakers, researchers, civil society, industry representatives, and citizens. During the PoliRuralPlus project, the tool is being redesigned to better fit multi-actor approach practices tailored to a wide range of Labs/Pilots. The new architecture, detailed in the deliverable D5.1, includes the definition of new programs and work entities and new charts and communication engines.
5. **E-market for local food -** The E-Market for Local Food is a digital marketplace developed under the Data4Food2030 and THEROS projects. It supports the direct sale of regionally-produced, organic, and certified foods by connecting local producers, retailers, and consumers. Built on PrestaShop and hosted by

Lesprojekt, it promotes short supply chains and the Farm-to-Fork approach Key features include IoT sensors for real-time monitoring of transport conditions and blockchain technology for traceability and fraud prevention. Integrated data sources like OpenStreetMap and Copernicus optimize logistics and decision-making. The platform enhances transparency, empowers local producers, and promotes sustainability in line with the EU Green Deal

6. **Satellite imagery cloud cover detection, “mosaicking”** - Satellite imagery cloud cover detection is a key focus of the ALIANCE project, which leverages AI to integrate satellite, climate, meteorological, and sensor data. This approach enhances agricultural adaptation to climate change, reduces economic costs, and minimizes environmental impacts. The project developed a validated methodology for estimating missing data in Sentinel-2 imagery, enabling accurate vegetation analysis and yield prediction.
7. **Precise weather forecast -** Precise, localized weather forecasts are the focus of the ALIANCE project, which integrates AI with satellite data, climate datasets, meteorological predictions, and sensor measurements. This approach enhances agricultural adaptation to climate change, reduces economic costs, and minimizes environmental impacts. The project’s validated methodology combines data sources such as Sentinel-2, GFS, HadISD, and ERA5-Land to improve forecast accuracy
8. **Map Whiteboard** - This technology mimics Google Docs for a map based work surface and allows multiple parties to connect to the same map and work together “on the map” much as they would over a printed map copy in a meeting room. Presently, the technology enables multiparty creation and editing of spatial information and building of maps as well as persisting and storing resulting data.
9. **Pilot-Specific Implementations:** Each pilot region utilized the DIH and its tools to address local challenges, such as fostering innovation ecosystems, improving governance, and enhancing rural-urban connectivity. The pilots generated customized action plans and methodologies tailored to regional contexts.

2.1.2 Gaps and Opportunities

Despite the utility of these tools, rapid advancements in AI, particularly in Large Language Models (LLMs) and GeoAI, present an opportunity to modernize and enhance the existing framework. For instance:

- **GeoAI Integration:** The development of JackDaw, a GeoAI solution combining LLM capabilities with geospatial analytics, promises to streamline data processing and visualization.
- **Enhanced Usability:** Simplifying interfaces and enabling natural language interaction through LLMs will make the tools more accessible to non-experts.
- **Broader Data Integration:** Expanding the capacity to integrate real-time and diverse data sources, including IoT and European Data Spaces, can improve the accuracy and relevance of analyses.

These updates, coupled with a stronger emphasis on stakeholder engagement and open innovation through Open Calls, will ensure that PoliRuralPlus builds on the strong foundation of PoliRural while addressing contemporary challenges and opportunities

2.2 Available Data Sources and Models

PoliRuralPlus employs a wide range of data sources to support its objectives, combining **local data** from pilot regions with **global datasets** from Open Data repositories. These sources include Earth observation data, statistical information, socio-economic indicators, geospatial datasets, and participatory insights provided by stakeholders.

Local data captures the specific conditions and challenges of each pilot region, while global data—such as that from Copernicus, INSPIRE, OpenStreetMap, FADN, and national statistics—provides broader contextual and comparative insights. These datasets encompass environmental, agricultural, demographic, and governance dimensions, forming a robust foundation for analysis and planning.

The details of these datasets, their structure, and applications are described in subsequent sections. All data will be integrated into existing and newly developed tools, including the enhanced Digital Innovation Hub (DIH) and advanced GeoAI applications, ensuring seamless analysis and decision support tailored to the needs of regional stakeholders.

2.1.1 Data sources

We will leverage a diverse set of **existing data and services** to support our analysis and decision-making processes. This includes authoritative sources such as **COPERNICUS** for satellite imagery and Earth observation data, **INSPIRE** for standardized spatial data infrastructure, and **Eurostat** for statistical data on regional trends. We will also utilize **SoilGrids** for high-resolution soil property mapping, **Natura 2000** for biodiversity and habitat data, **Open Meteo** for meteorological forecasts, and the **Farm Accountancy Data Network (FADN)** for agricultural economic insights. Additionally, **GISCO** (Geographic Information System of the Commission) will provide geospatial datasets for enhanced spatial analysis.

Incorporating contributions from **voluntary activities** like **OpenStreetMap (OSM)**, which offers detailed user-generated geographic data, will further enrich our spatial analyses.

Complementing these datasets, we will integrate our **own generated data sources**, including **OpenLandUse (OLU)**, a comprehensive dataset on land use patterns, and a **Map of Land Cover and Land Use Changes**, which tracks land use transformations over time.

To achieve a comprehensive and context-specific analysis, we will also integrate **local data sources** such as **local maps** and **local IoT data**. These sources provide granular, real-time information on environmental conditions, infrastructure, and land use, enhancing the accuracy and relevance of our analyses.

By combining standardized datasets, voluntary contributions, custom-generated data, and local data sources, we can achieve a holistic understanding of geographic, environmental, and socio-economic factors. This integration supports advanced spatial analysis, sustainable rural development, and effective policy-making. The synergy among these data sources ensures a robust foundation for accurate, data-driven decision-making and innovation.

Open Land Use (OLU)

The **OLU** database is a multi-scale, open-access database designed to integrate various types of land use and land cover data along with other thematic datasets. The latest version of the OLU database has been developed through an iterative process incorporating user requirements gathered in recent years. This advanced design supports the needs of diverse modeling tasks by offering data in both raw form and through spatio-temporal thematic views.

The OLU database's structure enables it to function as a comprehensive resource, facilitating the integration of multiple data layers to support applications in environmental analysis, land use planning, and sustainable development. By allowing for the combination of land use, land cover, and additional contextual data, the OLU database enhances the accuracy and effectiveness of analytical processes. This capability makes it particularly useful for addressing complex challenges related to **Destination Earth** and other spatial modeling frameworks.

Its open nature ensures that researchers, policymakers, and stakeholders can access and utilize the data for various analytical and decision-making purposes.

Land Use and Land Cover Change Detection

The Land Use and Land Cover Change Detection dataset, generated with the support of the Open Land Use (OLU) database, provides valuable insights into changes in land use and land cover (LULC) across different time periods. Specifically, it tracks growth or loss within various Corine Land Cover (CLC) and High-Resolution Layer of Land Use and Cover System (HILUCS) classifications. This dataset captures land cover dynamics over the periods 1990, 2000, 2005, and 2010, allowing for detailed temporal analysis of spatial changes.

The dataset records both increases (growth) and decreases (losses) in specific CLC and HILUCS classes, such as agricultural land, forests, urban areas, grasslands, and wetlands. By identifying these changes, the dataset supports monitoring of key trends like urban expansion, deforestation, afforestation, and shifts in agricultural practices.

The integration of OLU data enhances the accuracy and comprehensiveness of the dataset by incorporating multi-scale land use and land cover information along with other thematic datasets. This temporal analysis enables stakeholders, including policymakers, urban planners, and environmental managers, to better understand the impacts of land use policies, human activities, and natural processes on the landscape. The information can be used to inform sustainable development practices, land management strategies, and environmental conservation efforts.

By covering multiple reference years, the dataset provides a comprehensive view of land cover transitions, contributing to long-term monitoring, trend analysis, and the assessment of land use changes over several decades.

PoliruralPlus document repository

The PoliRuralPlus Hub (PoliRuralPlus Hub) has evolved into a versatile platform capable of managing not only spatial data but also serving as a document repository. This repository supports various types of documents, including reports, research papers, policy briefs, manuals, and datasets, enabling a centralized resource for comprehensive rural analysis and development.

The PoliRuralPlus Hub's document management capability allows for the storage, categorization, and retrieval of these diverse document types. This feature enhances knowledge sharing, collaboration, and access to critical information among stakeholders such as policymakers, researchers, farmers, and rural development professionals.

Furthermore, the PoliRuralPlus Hub supports the use of these documents for training Large Language Models (LLMs), providing a rich dataset to improve AI-based applications, such as automated content generation, text summarization, and contextual analysis relevant to rural development. This capability ensures that AI models are better aligned with the specific needs and challenges faced by rural communities.

The repository can also facilitate the preparation of Rural Action Plans (RAPs) by offering curated content and reference materials. By leveraging both spatial data and textual resources, the PoliRuralPlus Hub enables evidence-based decision-making and policy formulation, supporting a more integrated and informed approach to rural development.

This integration of spatial data management and document repository functions makes the PoliRuralPlus Hub a comprehensive tool for advancing rural innovation, policy analysis, and sustainable development initiatives.

2.1.2 Models

Cloud removal and index calculations

The **cloud removal and index calculation model** leverages AI-based spatiotemporal techniques to enhance satellite imagery for crop monitoring, addressing the challenge of cloud cover that obstructs optical data. The model integrates **Sentinel-2** optical imagery and **Sentinel-1** radar data, utilizing the **UnCRtainTS neural network architecture** to reconstruct cloud-free images. Sentinel-1 radar data, capable of penetrating clouds, ensures continuous monitoring, while Sentinel-2 provides spectral details necessary for vegetation analysis.

After cloud removal, the model calculates vegetation indices such as the **Normalized Difference Vegetation Index (NDVI)**, which quantifies plant health and density. NDVI is derived from Sentinel-2 bands using the formula:

$$NDVI = \frac{NIR - Red}{NIR + Red}$$

The model is trained and validated using cloud-free target frames from **2022 and 2023** agricultural data, with performance evaluated using **Mean Absolute Error (MAE)** and **Mean Squared Error (MSE)**. Combining Sentinel-1 and Sentinel-2 data significantly improves accuracy compared to using optical data alone. This approach enables reliable, continuous crop monitoring and forecasting, supporting timely interventions in precision agriculture. The result is optimized resource use, reduced economic costs, and minimized environmental impacts, making the model a valuable tool for sustainable agricultural practices.

Rural Attractiveness and Innovation Potential Data

Assessing the attractiveness of rural regions requires a comprehensive approach that combines different data sources. In the Polirural Plus project we use statistical data and data on existing infrastructure. Statistical data provide key indicators on demographic, economic, sociological and other aspects of regions. Their main advantage is their relative availability and standardisation, especially when using data from international organisations such as Eurostat. However, these sources have significant disadvantages, including averaging across statistical regions, which can lead to inaccurate results that do not reflect local differences. Another problem is the frequent obsolescence of data, as statistics take a considerable amount of time to collect and publish. The lack of homogeneity internationally, unless data from global organisations are used, also limits the possibilities of comparing regions between different countries.

An ideal resource for infrastructure assessment is OpenStreetMap (OSM), which offers detailed and up-to-date information on transport networks, public facilities and other types of infrastructure. One of the main advantages of OSM is the homogeneity of the data, which allows its consistent use at an international level. Also important are the frequent updates due to an active community of contributors and the openness of the data, which means free access for different applications and analyses. Nevertheless, OSM also has its drawbacks, such as variable data quality in less covered areas and dependence on contributors.

The combination of statistical and OSM data allows for a more comprehensive view of the attractiveness of rural regions. While statistical data provide a general framework, OSM data offer spatial detail and timeliness. However, for an effective assessment, it is necessary to approach these data critically, with an awareness of their strengths and limitations.

3. Requirements Analysis

Building on the established baseline, the next step was to analyze the evolving requirements and opportunities that have emerged since the PoliRural project, particularly in the context of the policies and challenges addressed by PoliRuralPlus. This analysis focused on identifying gaps, assessing their potential solutions, and designing technical components or solutions tailored to these challenges. It also included experimenting with technological innovations to explore how they can support regional development efforts.

The process will result in the development of tools to assist pilot projects, using direct input from pilot stakeholders. These tools will be abstracted into generic solutions with broad applicability, ensuring their potential for frequent and widespread use across diverse contexts. This approach aimed to bridge identified gaps effectively while fostering innovation and scalability.

3.1 Requirements from PoliruralPlus stakeholders

PolIRuralPlus being a user driven project, the first source of requirements is the project stakeholders. Their requirements and needs are captured on several different levels through activities conducted in WPs2-7. Some requirements have resulted from new technical opportunities as identified by WP4 partners, these are listed in section 3.2 below.

3.1.1 Requirements from WP2: Inputs from D2.1

- **Stakeholder Engagement**: D2.1 outlines the results of stakeholder consultations to identify challenges in rural-urban linkages. Focus areas include governance, resource management, and resilience post-COVID-19.
- **Territorial Insights**: Specific regional needs such as digital divide issues in Vidzeme and energy transitions in Apulia were captured.
- **Policy Gaps**: Rural policy frameworks were benchmarked against EU-level goals (e.g., CAP reform, Green Deal) to identify gaps requiring targeted interventions.

3.1.2 Requirements from WP5: IT Needs from Partner Regions

Each region provided specific technological requirements to support their pilot activities:

1. **County Monaghan, Ireland**:
    - Tools for improving participatory governance, focusing on stakeholder collaboration through digital platforms.
    - Enhanced mapping solutions for infrastructure and service gaps.
2. **Slovakia (Regional Focus)**:
    - Development of decision-support systems for policy-making aligned with rural and agricultural objectives.
    - Analytical tools for evaluating socio-economic impacts of rural development policies.
3. **Central Greece**:
    - Digital solutions for agricultural supply chains, integrating traceability and productivity tools.
    - Tools for climate adaptation and mitigation, focusing on resource management.

4. **Apulia, Italy**:
    - Renewable energy management systems to monitor and optimize local resources.
    - Platforms to engage stakeholders in sustainable agricultural practices.
5. **Mallusjoki, Finland**:
    - Digital innovation hubs to support start-ups and local businesses in bioeconomy sectors.
    - Smart monitoring systems for biodiversity and land use.
6. **Czech-Bavarian Border Region, Czech Republic-Germany**:
    - Tools for cross-border collaboration in regional development and tourism.
    - Data sharing platforms to harmonize economic and environmental strategies.
7. **Spain**:
    - IoT-based tools to modernize agricultural practices and improve water efficiency.
    - Citizen engagement platforms for urban-rural policy coordination.
8. **Vidzeme, Latvia**:
    - Broadband and ICT infrastructure development for rural connectivity.
    - Monitoring frameworks for tracking regional policy outcomes.
9. **Malta**:
    - Platforms for urban-rural coordination in food systems and supply chains.
    - Tools for digital literacy and innovation in agriculture.

3.1.3 Requirements from WP3: Inputs from D3.1

The **Rural Action Plan (RAP)** in the **PoliRuralPlus** project provides a systematic approach to enhancing rural-urban development. The RAP process involves multiple steps, including data collection, analysis, stakeholder engagement, decision-making, and evaluation. Each step is designed to ensure that policies and strategies are evidence-based, collaborative, and responsive to regional needs. The advanced tools and technologies integrated into the **PoliRuralPlus Hub** streamline these steps, making the process more efficient and effective. The following table outlines each step of the RAP, alongside the specific **PoliRuralPlus technologies** that support these steps.

Table 1: PoliRuralPlus Technologies Supporting RAP Steps

RAP Steps	PoliRuralPlus Technologies	Application
1. Data Collection	**Personal Information Stream Processing, Vulture Productivity Tools, Magpie**	Automates the collection of data from emails, URLs, and metadata catalogues, improving productivity and accuracy.
2. Data Integration and Analysis	**Data Extraction and Analysis Tool, Jackdaw (GeoAI Chatbot), Magpie, Rural Attractiveness Explorer**	Combines structured and unstructured data, analyzes spatial and non-spatial data, and identifies key trends.
3. Policy Brief Preparation	**PoliRuralPlus Advisor, Large Language Models (LLMs), Vulture Productivity Tools**	Generates data-driven policy briefs and reports, assisting in informed decision-making.

4. Stakeholder Engagement	**Multi-Actor Approach Tool, MapWhiteboard, Jackdaw (GeoAI Chatbot)**	Facilitates communication, collaboration, and participatory engagement among stakeholders.
5. Decision-Making and Strategy	**PoliRuralPlus Advisor, Jackdaw (GeoAI Chatbot), MapWhiteboard, Vulture Productivity Tools**	Supports decision-making with data-driven insights, AI assistance, and collaborative spatial analysis.
6. Monitoring and Evaluation	**Data Extraction and Analysis Tool, Rural Attractiveness Explorer, Jackdaw (GeoAI Chatbot)**	Monitors policy impacts, evaluates outcomes, and automates feedback collection for continuous improvement.

Explanation of RAP Steps and Technologies

1. **Data Collection**
   The initial step of the RAP process involves gathering relevant data from various sources, such as policy documents, emails, URLs, and metadata catalogues. Tools like **Personal Information Stream Processing**, **Vulture Productivity Tools Sandbox**, and **Magpie** streamline this process by automating data collection and ensuring accuracy.
2. **Data Integration and Analysis**
   In this step, data is processed, integrated, and analyzed to generate meaningful insights. The **Data Extraction and Analysis Tool** and **Jackdaw (GeoAI Chatbot)** support the analysis of both spatial and non-spatial data. **Magpie** helps retrieve and integrate metadata from sources like INSPIRE, and the **Rural Attractiveness Explorer** facilitates the analysis of rural attractiveness indicators.
3. **Policy Brief Preparation**
   Synthesizing data into concise and actionable policy briefs is crucial for decision-making. The **PoliRuralPlus Advisor**, **Large Language Models (LLMs)**, and **Vulture Productivity Tools** automate the generation of these briefs, ensuring they are evidence-based and tailored to stakeholder needs.
4. **Stakeholder Engagement**
   Effective engagement with stakeholders is supported by tools such as the **Multi-Actor Approach Tool** and **MapWhiteboard**, which facilitate collaboration and real-time interaction. **Jackdaw (GeoAI Chatbot)** provides AI-driven insights to enhance communication and engagement.
5. **Decision-Making and Strategy**
   This step focuses on formulating strategies based on data-driven insights. The **PoliRuralPlus Advisor** and **Jackdaw (GeoAI Chatbot)** provide AI-supported decision-making tools, while **MapWhiteboard** enables collaborative spatial planning and analysis. **Vulture Productivity Tools** streamline administrative processes, improving efficiency.
6. **Monitoring and Evaluation**
   Continuous monitoring and evaluation ensure that policies and strategies are effective. The **Data Extraction and Analysis Tool**, **Rural Attractiveness Explorer**, and **Jackdaw (GeoAI Chatbot)** support tracking policy impacts, evaluating outcomes, and collecting stakeholder feedback for ongoing improvement.

By leveraging these advanced **PoliRuralPlus technologies**, the RAP process becomes more streamlined, collaborative, and responsive, supporting sustainable and balanced rural-urban.

3.1.4 Requirements for WP6 and WP7: Communication and Open Call Management

In **Work Packages 6 and 7 (WP6 and WP7)** of the **PoliRuralPlus** project, the **PoliRuralPlus Hub** serves as the central communication platform to support stakeholder engagement, dissemination activities, and the management of open calls. The Hub must facilitate the creation, sharing, and management of information related to blogs, news, training, and project results. Additionally, it must include comprehensive tools for managing the full lifecycle of open calls. The following outlines the specific requirements and functionalities to achieve these objectives.

Main Communication Platform

The **PoliRuralPlus Hub** functions as a primary communication interface, ensuring that project information reaches stakeholders efficiently. It supports the publication and management of various content types:

1. **Blogs and News**:
   The Hub must offer a **Content Management System (CMS)** that allows administrators to create, edit, categorize, and publish blogs and news articles. These posts should be enriched with multimedia content, including images, videos, and hyperlinks, to provide comprehensive updates. The CMS should support scheduled publishing and tagging to facilitate content organization and retrieval.
2. **Interactive Features**:
   To foster engagement, the Hub should include features such as comment sections, feedback forms, and notification systems. These tools ensure that stakeholders can interact with the content and receive alerts about new updates and events.
3. **Event and Training Announcements**:
   The platform must manage information about workshops, webinars, and training sessions. This includes functionalities for event registration, scheduling, and calendar integration to streamline participant management and communication.

Open Call Management

Supporting open calls is a key requirement for WP6. The Hub needs to manage the complete open call process, from submission to result dissemination:

1. **Submission Management**:
   The Hub should provide customizable submission forms that allow applicants to submit proposals and relevant documents securely. It should support file uploads, version control, and data validation.
2. **Review and Evaluation**:
   Tools for assigning reviewers, tracking evaluations, and scoring submissions are necessary. Automated workflows should streamline the review process and notify applicants of the evaluation outcomes.
3. **Status Tracking and Communication**:
   The Hub must include dashboards for administrators to monitor the status of open calls, track application progress, and send automated email notifications to applicants at different stages of the process.

4. **Result Dissemination**:
   A dedicated space for publishing open call outcomes, selected projects, and summaries of funded initiatives should be available. This space should support multimedia content to highlight key results effectively.

Blogs, News, and Training Management

To ensure effective dissemination, the Hub must include:

1. **Content Management System (CMS)**:
   A user-friendly CMS to create, edit, and manage blogs, news, and training content. It should support scheduled publishing and content tagging for easy retrieval.
2. **Search and Filter Options**:
   Advanced search capabilities to allow users to find specific content by keywords, categories, authors, or dates. Filters should streamline access to relevant information such as open calls, news updates, and events.
3. **Content Analytics**:
   Tools for tracking views, engagement, and user feedback on published content. Analytics dashboards help evaluate the effectiveness of communication efforts and guide improvements.

Results Management and Dissemination

The Hub must facilitate the storage and dissemination of project results and deliverables:

1. **Repository for Results and Outputs**:
   A structured repository to store reports, datasets, presentations, and other deliverables. Each item should have associated metadata for easy discovery and citation.
2. **Visualization Tools**:
   Tools for visualizing data, such as interactive charts, graphs, and maps, to effectively present project outcomes and insights.
3. **Public and Private Access**:
   Configurable permissions to manage access rights for different user groups (e.g., public access, partner-only access).

Table 2: Summary of Key Requirements and Functionalities

Functionality	Description
Blogs and News	Creation, categorization, and publishing of blogs and news updates.
Open Call Management	Submission forms, review workflows, status tracking, and result dissemination.

Training Management	Publishing, scheduling, and managing training sessions and events.
Interactive Communication	Commenting, notifications, and feedback forms for stakeholder engagement.
Results Repository	Structured storage and dissemination of project outputs and deliverables.
Content Analytics	Tools for tracking and analyzing engagement with published content.

The **PoliRuralPlus Hub** serves as an integrated platform to meet the requirements of WP6 and WP7 by managing communication, stakeholder engagement, and open call processes. Through functionalities like blogs, news, training management, open call workflows, and result dissemination, the Hub ensures a robust and efficient framework for collaboration and dissemination. These features collectively enhance project visibility, stakeholder participation, and the effective management of rural-urban development initiatives.

3.3 Requirements Derived from Technical Potential (WP4)

While the previous section describes the requirements derived from regional development stakeholders' point of view, this section describes the requirements identified by technical partners as a consequence of changes in technological baseline and the new possibilities these changes enable in improving the efficiency and effectiveness of components and applications already known to be used and be beneficial for rural development purposes.

3.3.1 Harnessing Advanced Technologies

- **State-of-the-Art Integration**: Leverage cutting-edge technologies such as artificial intelligence (AI), semantic data structures, and spatial data analytics to create robust tools and platforms that address the unique challenges of rural-urban linkages.
- **Interoperability Focus**: Emphasize integration over isolated solutions by building on open standards for data, protocols, and APIs. Prioritize shared authentication systems, such as the PoliRuralPlus.eu OAuth provider, for streamlined user access across tools and services.

3.3.2 Utilizing Data from Established Frameworks

- **Open Data Sources**: Tap into established programs such as:
    - **PSI Directive (Public Sector Information)**: Ensure the reuse of public sector data for innovative applications.
    - **INSPIRE**: Leverage geospatial datasets for improved environmental and spatial planning.
    - **Earth Observation (EO)**: Use satellite data for land use monitoring, climate adaptation analysis, and agricultural management.
    - **Destination Earth and Data Spaces**: Capitalize on advanced digital twins and data infrastructures for precise simulations and decision-making support.

- **Historical Data Investments**: Build on the wealth of geospatial and statistical data accumulated over the past two decades to inform predictive models and analytics tools.

3.3.3 Innovative Applications and Models

- **Generative AI and Spatial Data**: Integrate generative AI capabilities with spatial data to develop spatial-enabled conversational agents that allow natural language processing (NLP) queries against GeoJSON and other geospatial formats. This facilitates intuitive exploration of spatial datasets.
- **Area Potential Analysis**: Develop tools to determine the potential of specific areas for applications like agriculture, renewable energy, and urban planning. Utilize predictive analytics to evaluate climate impacts and socio-economic outcomes.

3.3.4 Service-Oriented Architecture

- **Independent Microservices**: Design modular microservices that ensure ease of integration and adaptability for independent exploitation in mid- and long-term scenarios.
- **Abstract APIs**: Provide generalized APIs for calling models, enabling seamless use in diverse applications while abstracting technical complexity for end-users.

3.3.5 Scalable and Accessible Analytics

- **Web-Enabled Analytics**: Deploy analytics services on a web scale to ensure wide accessibility and usability by both technical and non-technical stakeholders.
- **Integration into Commonly Used Applications**: Ensure developed tools are easily integrable with common desktop and web applications, fostering widespread adoption and reducing technical barriers.

3.4 Platform Design Principles

In synthesizing the above requirements into a high-level platform design, technical partners have made a number of design choices that will form the basis for the first version of the platform. These will be revised following implementation, operationalisation and testing of the platform among project stakeholders.

1. **A Unified API End-Point for Access to Microservices:** Developing a single, cohesive entry point, such as a portal or dashboard, is essential to serve as a central hub for all microservices. This unified platform streamlines user access, ensures consistent authentication mechanisms, and acts as the backbone where various services can be seamlessly attached and accessed in an integrated manner.
2. **Pilot-Specific Needs:** The focus must be on delivering services and tools tailored to address the specific challenges of partner regions. For example, decision-support systems are a priority in Monaghan, governance tools are essential in Slovakia, while Malta requires digital agriculture solutions. Customization at the regional level must align with a scalable architecture to facilitate broader adoption across different regions.
3. **Broad Applicability:** Functionalities developed should extend beyond the needs of individual pilots to ensure utility across diverse contexts. This will enhance the overall value of the project for a wider audience. Examples include analytics tools for optimizing rural-urban linkages or improving resource management efficiency.

4. **Integration Over Base Development:** The priority is integrating proven, existing technologies such as Natural Language Processing (NLP) and spatial data analytics rather than building base technologies from the ground up. Leveraging open standards, reusable APIs, and established platforms will accelerate development while ensuring reliability.
5. **Capitalizing on Existing Data:** Utilizing available datasets such as INSPIRE-compliant geospatial data, PSI directive resources, and Earth Observation (EO) repositories will significantly reduce time and costs linked to data acquisition and processing. This strategy ensures faster delivery and optimizes project resources.
6. **Flexible, Parallel Development:** The development process must be designed to accommodate the varied skills and expertise of technical partners. A modular approach, relying on independent microservices, allows parallel progress on multiple components. This ensures seamless integration into the overall project framework without disrupting development timelines.
7. **SaaS Delivery Model:** The emphasis is on adopting a Software-as-a-Service (SaaS) delivery model to maximize accessibility for end-users. SaaS solutions eliminate infrastructure burdens, enabling users to access tools through web-based platforms without requiring complex local installations.
8. **Incremental Development and User-Centric Iteration:** An incremental approach involves breaking development into manageable phases with clear milestones. Early delivery of functional prototypes is critical to iteratively refine solutions based on real-time user feedback. Establishing regular user interaction cycles, including workshops, feedback sessions, and testing activities, will ensure that the solutions address real-world needs and usability requirements effectively.
9. **Long-Term Sustainability:** The project's architecture must be designed for scalability and adaptability to accommodate future enhancements, ensuring long-term relevance and sustainability beyond the project's duration. Solutions must also possess cross-sector applicability, such as spatial decision-support tools or generative AI-enabled chat interfaces. This approach will maximize their impact and ensure continuity across different use cases and sectors.

4. Overall Platform Design

This chapter describes the overall design of the PoliRuralPlus platform, taking into account the data sources, models, apis, applications and how they feed into the overall project objectives.

4.1 Conceptual Design

The conceptual design view of the PoliRuralPlus project is shown in Figure 2, overleaf. It consists of five different elements that are described on a high level below. Each component connects data and/or models via an API and applications to the realisation of one or more project objectives.

Starting off with **data sources and analysis processes** the project relies on a combination of existing well-proven data sources that have an extensive spatial coverage and sufficient spatial and temporal resolution to suit the needs of rural development, policy making and decision taking. The project employs a mix of pre-existing well-known data sources and derived data through custom analysis processes developed or evolved specifically for the PoliRuralPlus project.

Next, the project implements a number of **AI, statistical and spatial models** that provide real-time analysis capabilities as well as modules for preprocessing and denormalization of data for better efficiency in "production" mode. Thes models span a wide range of uses ranging from weather forecast via statistics to document parsing, embedding generation and AI-based information retrieval. In addition to self-built models, the project also relies heavily on mainstream LLM APIs such as OpenAI and Gemini.

All business logic and data are accessed by end-user applications through **an API layer** that is exposed as a self-documenting Open API end-point. Subject to authentication, this API allows users to access high performance processing capabilities from lightweight end-user applications written in client side technologies such as React, Angular etc.

This is where the technological efforts and investments are converted into business value for regional development actors through the **digital innovation hub** as well as the nine embeddable modular applications, each targeting a specific aspect of the PoliRuralPlus objectives with an aim to improve speed, cost or quality of operations and quality data used for decision support.

The embeddable modules include: (1) the **Policy Advisor**, an Open AI RAG that provides access to relevant policy context for regional action plan development; (2) **Jackdaw** - a spatial-enabled chat agent; (3) **Vulture** a sandbox for conducting AI-based productivity tool experiments; (4) **Magpie** - an AI-enabled catalog search and map making application extending the concept of OGC catalog services; (5) **Multi Actor Approach tool** - a software for managing and interacting with multi actors in MAA projects; (6) **Atlas of Best Practises** - a tool that provides access to a growing body of EU-wide good practice in rural development and related disciplines; (7) **E-market for**

local food - a solution for marketing of local production; (8) **Map Whiteboard** - an interactive collaborative map surface allowing multi-party editing and cooperation in map making and data editing and; (9) **Rural Attractiveness Explorer** - a tool to analyse rural attractiveness indicators and determine area potential.

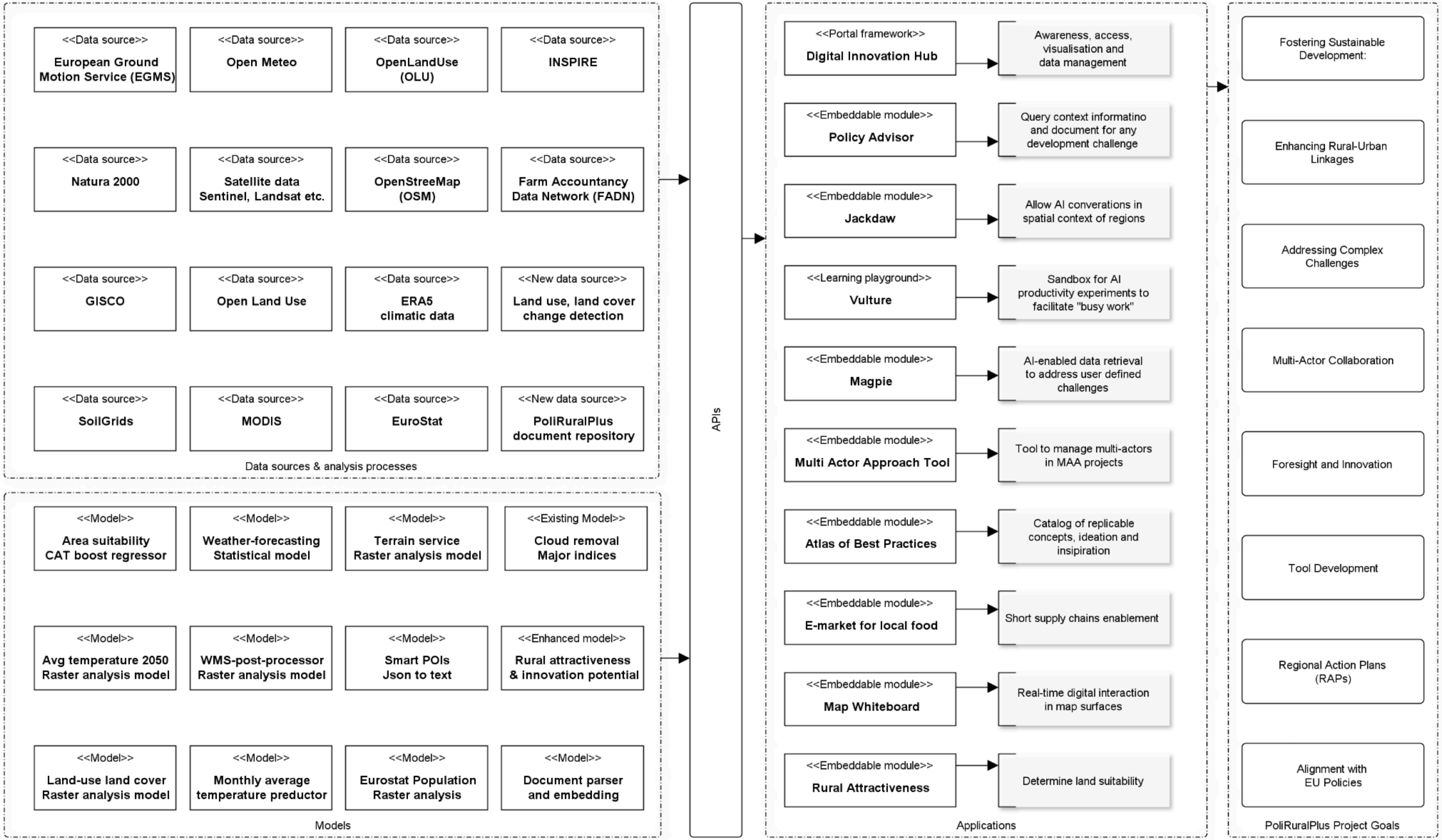

<<Data source>> European Ground Motion Service (EGMS)
<<Data source>> Open Meteo
<<Data source>> OpenLandUse (OLU)
<<Data source>> INSPIRE
<<Data source>> Natura 2000
<<Data source>> Satellite data Sentinel, Landsat etc.
<<Data source>> OpenStreeMap (OSM)
<<Data source>> Farm Accountancy Data Network (FADN)
<<Data source>> GISCO
<<Data source>> Open Land Use
<<Data source>> ERA5 climatic data
<<New data source>> Land use, land cover change detection
<<Data source>> SoilGrids
<<Data source>> MODIS
<<Data source>> EuroStat
<<New data source>> PoliRuralPlus document repository
Data sources & analysis processes
<<Model>> Area suitability CAT boost regressor
<<Model>> Weather-forecasting Statistical model
<<Model>> Terrain service Raster analysis model
<<Existing Model>> Cloud removal Major indices
<<Model>> Avg temperature 2050 Raster analysis model
<<Model>> WMS-post-processor Raster analysis model
<<Model>> Smart POIs Json to text
<<Enhanced model>> Rural attractiveness & innovation potential
<<Model>> Land-use land cover Raster analysis model
<<Model>> Monthly average temperature preductor
<<Model>> Eurostat Population Raster analysis
<<Model>> Document parser and embedding
Models
APIs
<<Portal framework>> Digital Innovation Hub
Awareness, access, visualisation and data management
<<Embeddable module>> Policy Advisor
Query context informatino and document for any development challenge
<<Embeddable module>> Jackdaw
Allow AI converations in spatial context of regions
<<Learning playground>> Vulture
Sandbox for AI productivity experiments to facilitate "busy work"
<<Embeddable module>> Magpie
AI-enabled data retrieval to address user defined challenges
<<Embeddable module>> Multi Actor Approach Tool
Tool to manage multi-actors in MAA projects
<<Embeddable module>> Atlas of Best Practices
Catalog of replicable concepts, ideation and insipiration
<<Embeddable module>> E-market for local food
Short supply chains enablement
<<Embeddable module>> Map Whiteboard
Real-time digital interaction in map surfaces
<<Embeddable module>> Rural Attractiveness
Determine land suitability
Applications
Fostering Sustainable Development:
Enhancing Rural-Urban Linkages
Addressing Complex Challenges
Multi-Actor Collaboration
Foresight and Innovation
Tool Development
Regional Action Plans (RAPs)
Alignment with EU Policies
PoliRuralPlus Project Goals

Figure 2: System architecture (conceptual view)

4.2 Realisation of Design

The conceptual architecture is realised (as shown in Figure 3, overleaf) using well-proven and robust open source components and software built and evolved through prior EU-funded projects. Furthermore, the project seeks to exploit and build on rather than replace or duplicate existing services and infrastructures.

In terms of data storage, the project employs five paradigms, suitable and used for slightly different purposes: RDBMS for structured data, JSON stores for document based data, file storage for file-based data such as PDFs, word documents etc, Semantic Stores for data enrichment and APIs and vector embeddings stores for storage and queries against vector embeddings generated by large language models.

The platform is built using a mix of technologies where Python plays center stage on the server side and HTML5/JavaScript based development frameworks like Angular and React dominate on the client side. The wide array of different technologies is managed through a containerization approach where each service runs in a dedicated container using Docker as virtualisation technology.

The security layer is provided through the Digital innovation Hub, powered by Hub4Everybody with Wagtail as the underlying CMS-technology and OAuth/OIDC provider. This is used to authenticate and authorize requests to the Web APIs and enable a single user account - single sign-on (SSO) experience for all applications in the PoliRuralPlus ecosystem.

While the realisation includes a wide range of embeddable modules that each implement individual user interfaces, the long-term exploitation and deployment strategy is based on integration with existing solutions and data portals that are already in use in local administrative units across Europe. In principle, all containers may be run on a single host machine at any mainstream cloud provider and can be scaled horizontally or vertically to accommodate user growth if offered as a SaaS service.

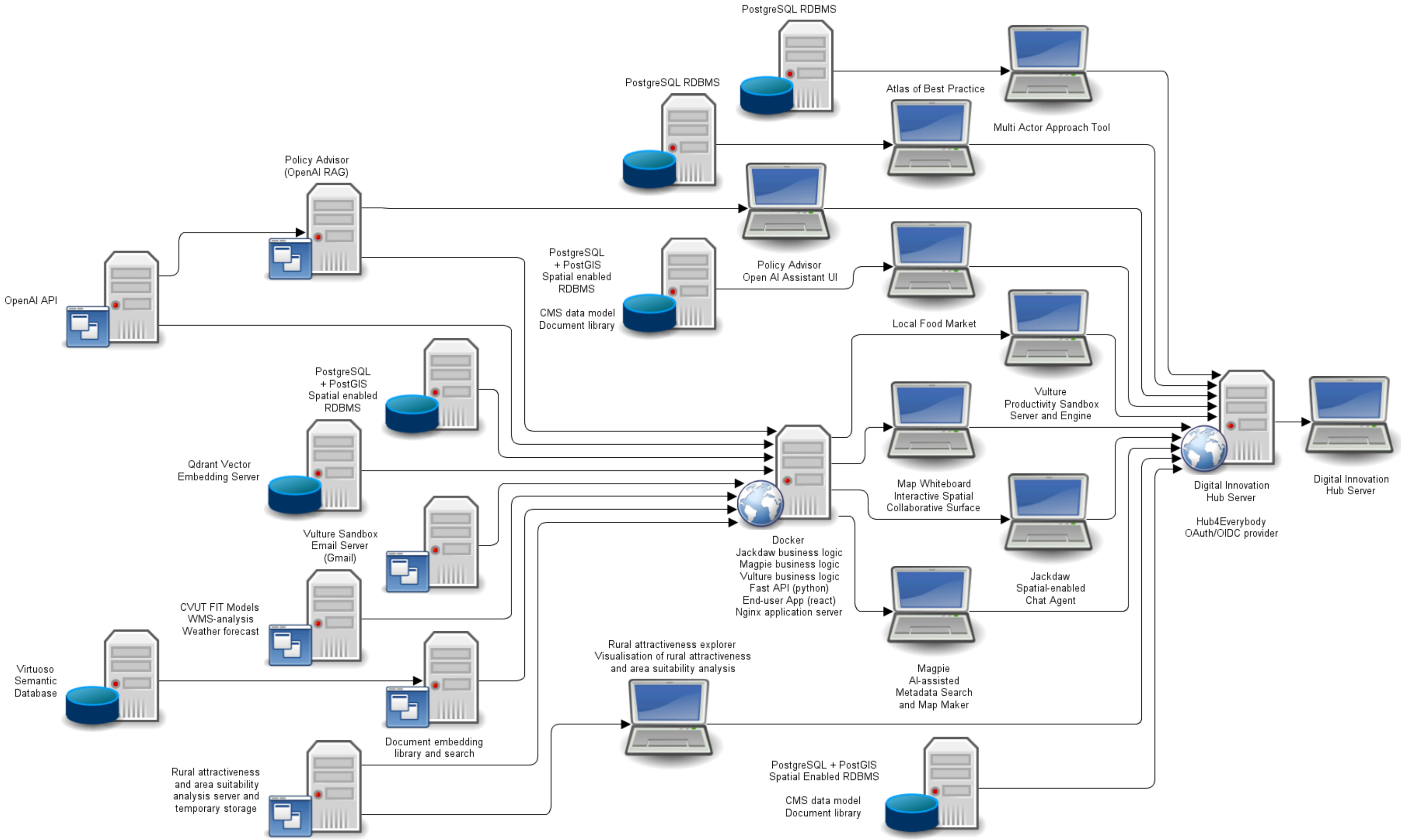


Figure 3: System architecture (realisation view)

5. Component Descriptions

5.1 Core Platform Components

Authentication and Security: The system leverages OAuth2/OpenID Connect (OIDC) protocols, provided by the **Digital Innovation Hub**, to ensure secure user authentication and authorization. This enables a robust, standards-compliant security framework, safeguarding user access and protecting sensitive data while allowing for seamless single sign-on (SSO) capabilities across integrated tools and platforms.

Communication and Protocol Standards: All communication between users, services, and external systems is secured through **HTTPS**, ensuring encrypted data transfer and integrity. This guarantees a high level of security for sensitive information and prevents unauthorized interception of communications.

Web Service Protocols: The application implements **OpenAPI** specifications using **FastAPI**, a modern Python-based framework designed for high-performance web services. OpenAPI ensures that service endpoints are well-documented, enabling interoperability with other systems and tools. FastAPI's asynchronous capabilities allow for efficient handling of concurrent requests, ensuring a responsive and scalable architecture.

Open Data Formats and Messaging: To maximize interoperability and usability, the system supports a variety of open data formats for messaging and data exchange. These include **JSON** and **XML** for structured textual data, **GeoJSON** and **WKT** (Well-Known Text) for geospatial data formats, and **binary file uploads/downloads** for larger datasets. This ensures compatibility with industry standards and a wide range of tools, facilitating smooth integration of geospatial and non-geospatial data.

Website and Data Management: Website content and data management are provided through **Hub4Everybody**, an advanced content management and data-sharing platform powered by **Wagtail CMS**. Built in Python, Wagtail CMS ensures flexibility, scalability, and ease of use for managing websites and spatial content. Additionally, **HSLayers NG**, a comprehensive geospatial tool suite, is integrated into the platform to enable efficient management, visualization, and exploration of geospatial data and metadata. This combination empowers users to interact with spatial datasets seamlessly and derive actionable insights.

Third-Party Services: The system utilizes a **shared API library** to access common back-end infrastructures, streamlining integration with third-party services such as the **OpenAI API**. This shared library reduces development overhead, promotes consistency across tools, and enables AI-driven features like natural language processing, embedding generation, and advanced analytics.

The system combines modern authentication protocols, secure communication standards, flexible open data formats, and a robust geospatial data management platform to deliver a reliable and scalable foundation for AI-powered spatial applications. With seamless integration of third-party APIs, it offers advanced capabilities that support a wide range of use cases.

5.2 Data Services and Data Management

The following sections outline the system's comprehensive strategies for managing, processing, and enriching geospatial and textual data. Through a combination of robust data access pipelines, advanced storage solutions, and AI-driven semantic enrichment, the system ensures scalability, efficiency, and precision in handling complex datasets.

5.2.1 Data Flow, Storage and Processing Solutions

To ensure efficient and up-to-date access to data, the system prioritizes accessing data directly from its **source via provided APIs**. This approach minimizes redundancy, ensures real-time or near real-time data availability, and promotes integration with external systems while maintaining data accuracy.

In cases where APIs prove inefficient—for instance, due to performance issues, bandwidth limitations, or repeated querying of large datasets—or where specific data are required to be stored locally for operational needs, **automatic data pipelines** are employed. These reusable pipelines facilitate the seamless downloading, transformation, and storage of datasets, ensuring they are readily available for processing and analysis without compromising on consistency or scalability.

When neither direct API access nor automated pipelines are feasible—such as when APIs are unavailable, incomplete, or unreliable—**data are manually downloaded** and stored locally. These datasets are organized within structured **files** or managed in a **relational database management system (RDBMS)**, such as **PostgreSQL**. This enables reliable long-term data storage, indexing, and querying while ensuring data integrity and accessibility for downstream processes.

To enable efficient querying, filtering, and transformation of stored datasets, the system implements robust **data services** using **FastAPI** and **Python**. These services provide well-documented, high-performance APIs for interacting with data and include support for advanced geospatial processing. Key libraries such as **GDAL/OGR** and other format-parsing tools are leveraged to handle a broad spectrum of **data formats** and services, including vector, raster, and tabular data. The use of these libraries ensures compatibility with industry-standard formats such as GeoJSON, WKT, Shapefiles, GeoTIFF, and various database-driven geospatial data services.

This multi-layered data access strategy—combining direct API integration, automated pipelines, and manual fallback solutions—ensures that the system can handle diverse data requirements efficiently while maintaining flexibility, scalability, and broad compatibility with geospatial and non-geospatial datasets.

5.2.2 AI-based semantic enrichment

Conversations between users and the chatbot can contain valuable data that may enhance user experience and improve future chatbot interactions. To extract and utilize these insights effectively, an end-to-end entity linking pipeline will be implemented. Entity linking is the process of identifying mentions of entities within text and connecting them to corresponding entries in a knowledge base.

The proposed system will facilitate both entity recognition and entity linking, leveraging general-purpose knowledge bases like Wikidata and domain-specific resources such as Agrovoc, particularly for agriculture-related contexts. The pipeline will begin with Named Entity Recognition (NER), where a fine-tuned Large Language Model (LLM) will identify entities of interest (e.g., Location, Animal, Plant, Climate) from the text.

Following NER, another LLM will generate a detailed profile for each extracted entity, using the surrounding text as context. Each entity profile will include a disambiguated entity name, a description, and its corresponding type. In the final step, these entity profiles will be encoded into high-dimensional vectors. These vectors will then be used to query a vector-based representation of the knowledge base, facilitating accurate and efficient linking of entities to relevant entries.

By integrating NER and EL through advanced LLMs and vector-based representations, this pipeline will enable precise entity recognition and linking, enhancing the knowledge gathering.

Architecture

Semantic enrichment service architecture overview

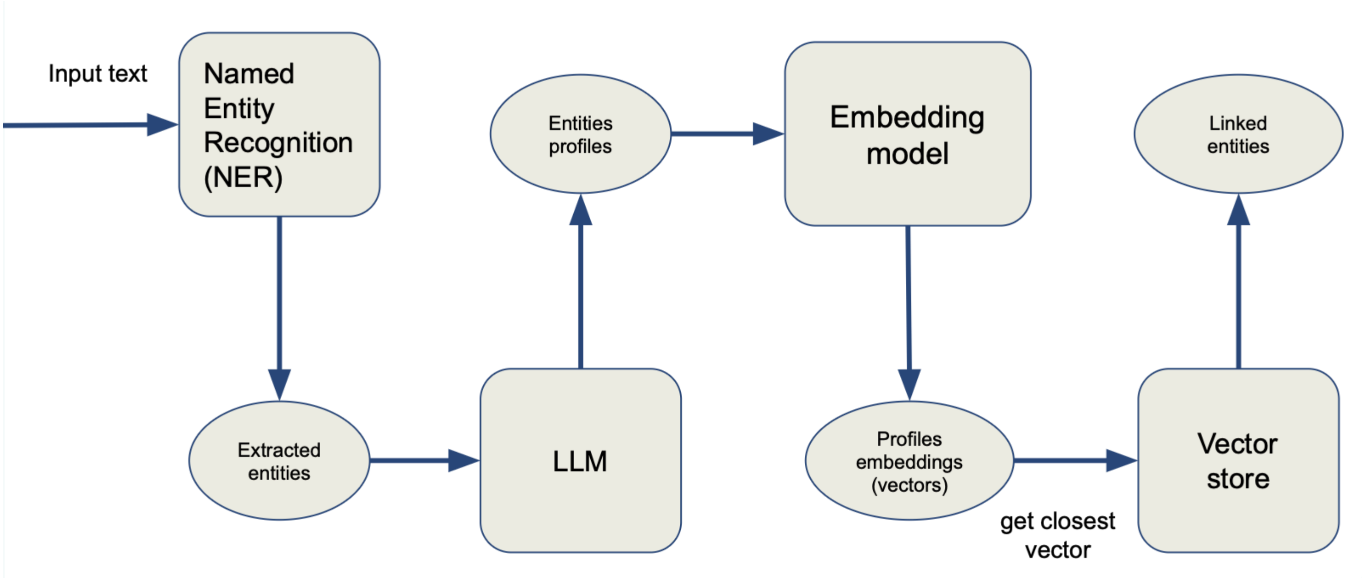


Figure 4: Flow of semantic enrichment service

The semantic enrichment service is designed as a background process, optimized for processing large volumes of text rather than serving as a real-time annotator. This design choice accommodates the service's ability to handle unlimited text length while maintaining efficiency. Although the system prioritizes speed, its primary use case is asynchronous annotation.

Process Overview

1. **Input Text Processing:** The service begins by accepting plain text or text files as input. To accommodate the limited context length of the UniNER LLM used for Named Entity Recognition (NER), the text is divided into smaller, manageable fragments, or "chunks." This ensures that the model processes text effectively without exceeding its token limit.
2. **Named Entity Recognition (NER):** Each chunk is passed through the UniNER LLM, fine-tuned for entity extraction. This model identifies entities of specified types, such as Location, Animal, Plant, or Climate. The output of this step consists of pairs of input chunks and the entities extracted from each chunk.
3. **Entity Profiling:** The extracted entities, along with their corresponding text chunks, are forwarded to another fine-tuned LLM designed for entity profiling. Using the chunk as context, this model generates detailed profiles for each entity. These profiles include a disambiguated entity name, a description, and the entity type, ensuring clarity and context.

4. **Embedding Generation:** The entity profiles are transformed into high-dimensional vector representations using the open-source embedding model mxbai-embed-large-v1. The vectors are encoded with float32 precision to capture the semantic richness of the profiles.
5. **Binary Quantization:** To enhance processing speed and reduce storage requirements, the high-dimensional vectors undergo binary quantization. This step maintains the essential features of the embeddings while optimizing them for rapid querying.
6. **Entity Linking via Vector Search:** The quantized vectors are used to query a Faiss vector database, which stores precomputed vector representations of a knowledge base. This step retrieves relevant profiles from the database, effectively linking the extracted entities to their corresponding knowledge base entries.

By leveraging advanced LLMs, high-dimensional embeddings, and efficient vector-based querying, the semantic enrichment service ensures precise and scalable entity recognition and linking. This architecture supports comprehensive knowledge extraction from unstructured text, enabling downstream applications to utilize enriched and contextually linked information.

5.2 AI Model Development

The following sections outline advanced analytical modules and tools developed to enhance spatial and policy-related decision-making. By combining geospatial analysis, Earth Observation (EO) imagery, and document processing techniques, the system supports a wide range of applications for sustainable development, tourism planning, and knowledge extraction.

5.2.1 Tourism suitability, capacity for establishment of hotel businesses

The **Tourism Suitability and Capacity Analytical Module** is designed to support the strategic planning of new hotel establishments in selected regions. This module assesses tourism potential by integrating various datasets and analytical methods, offering evidence-based recommendations for the number and location of new hotels. The goal is to promote sustainable tourism development by identifying high-potential areas while balancing demand with existing infrastructure. The module will be **integrated into the Jackdaw (GeoAI Chatbot)**, enhancing its functionality with spatial data-driven insights and recommendations for hotel development.

The module evaluates tourism potential through the analysis of key factors such as **land use/land cover (LULC)** data, tourism attractivity indicators, and environmental conditions. By leveraging detailed data from sources like **Open Land Use (OLU)**, **tourism statistics**, **OpenStreetMap (OSM)**, and **climate information**, the module identifies areas with favorable characteristics for tourism development. These include proximity to natural attractions, cultural landmarks, recreational facilities, and accessibility through transport networks.

A crucial part of the module's functionality is the comparison of tourism potential with the distribution and capacity of **existing hotels**. This gap analysis helps pinpoint underserved areas where new hotel infrastructure can meet emerging or unmet demand. The module considers spatial constraints, infrastructure availability, and environmental factors to ensure that recommendations align with sustainable development principles.

The module's workflow begins with the **integration and preprocessing of data** from multiple sources, creating spatial layers that represent land use, climate zones, existing infrastructure, and tourism attractivity. Using a **multi-criteria analysis (MCA)** approach, suitability maps are generated by assigning weights to factors like proximity to attractions, accessibility, and climate conditions. These maps highlight areas with the highest potential for hotel development.

Next, the module performs a **gap analysis** by overlaying the suitability maps with data on current hotel locations. This analysis identifies regions where the tourism potential is high, but existing hotel infrastructure is limited or insufficient. Using **spatial optimization techniques**, the module then suggests optimal locations for new hotels, taking into account factors such as environmental constraints, transport accessibility, and market demand.

The module can also generate different **scenario analyses** based on specific development priorities, such as eco-tourism, cultural tourism, or luxury hotel investments. These scenarios provide tailored recommendations that address various strategic objectives and stakeholder needs.

The key datasets utilized by the module include:

- **Open Land Use (OLU)** for detailed land use and land cover information.
- **Tourism statistics** for insights on tourist arrivals, overnight stays, and seasonal patterns.
- **OpenStreetMap (OSM)** for data on existing hotels, infrastructure, and points of interest.
- **Climate data** to evaluate environmental conditions favorable for tourism activities.
- **Attractivity indicators** such as cultural landmarks, natural attractions, and recreational facilities.

By synthesizing these data sources and applying advanced spatial analysis, the module delivers actionable insights for the planning of new hotel infrastructure. The integration into **Jackdaw (GeoAI Chatbot)** ensures that users can interact with the module through a conversational interface, making it easier to obtain spatially informed recommendations, run analyses, and visualize results dynamically. This approach supports policymakers, investors, and planners in making informed decisions that enhance tourism potential and regional economic growth while ensuring that development is **data-driven**, **sustainable**, and **responsive** to regional tourism dynamics.

5.2.2 Web Map Services (WMS) interpreter for integrated analysis

Most geospatial services, particularly those providing maps and spatial datasets, are primarily available in the form of **Web Map Services (WMS)**. WMS delivers georeferenced map images that are suitable for **visual interpretation** but are typically not very useful for detailed **analytical purposes** due to their raster-based nature. This limitation restricts their direct integration into advanced geospatial analysis workflows.

To address this limitation and support the needs of **Jackdaw (GeoAI Chatbot)**, we will develop a **WMS Interpreter** that can process and analyze WMS images for selected areas. This interpreter will perform **image analysis** on WMS outputs, extracting relevant geospatial information and identifying available objects within the map layers. The extracted information will be converted into a format suitable for further analysis and integration with other datasets.

The WMS Interpreter will enable the following functionalities:

- **Automated Object Detection**: Identify and classify features (e.g., roads, buildings, land cover types) from WMS images.

- **Data Extraction for Analysis**: Convert visual WMS data into structured information that can be analyzed within Jackdaw.
- **Integration with Discovery Services**: Utilize **Magpie** as a discovery service to locate relevant WMS sources and automatically feed the interpreter with WMS layers for processing.
- **Enhanced Decision-Making**: Support Jackdaw in delivering data-driven insights by integrating interpreted WMS data with other spatial datasets.

This development will bridge the gap between **WMS-based visualization** and **geospatial analysis**, making WMS datasets more accessible and valuable for analytical tasks. By transforming raster-based WMS images into actionable information, the interpreter will enhance the capabilities of Jackdaw and provide more comprehensive spatial insights to users.

5.2.3 EO image interpretation

In the **PoliRuralPlus** project, we will develop a comprehensive suite of **classification analyses** for interpreting satellite imagery. These analyses will enhance the capabilities of the platform by enabling detailed and accurate extraction of information from Earth Observation (EO) data. The primary focus areas for these classification tools include:

1. **Crop Detection**:
   - Algorithms for identifying and classifying different types of crops based on spectral signatures and temporal patterns.
   - Detection of crop types at various growth stages to support precision agriculture and resource management.
   - Integration of **Sentinel-2** and **Sentinel-1** data for robust detection under varying weather conditions.
2. **Land Cover Detection**:
   - Tools for classifying land cover types such as forests, grasslands, urban areas, water bodies, and agricultural lands.
   - Use of **multi-spectral and radar imagery** to improve accuracy and reliability of land cover classification.
   - Support for generating and updating **land cover maps** to track current land use.
3. **Change Detection Mapping**:
   - Techniques for identifying and mapping changes in land cover and land use over time.
   - Detection of changes caused by natural processes (e.g., deforestation, urbanization, agricultural expansion) or human activities.
   - Use of multi-temporal EO imagery to assess changes across different time periods (e.g., 1990, 2000, 2005, 2010).

These classification analyses will be integrated with other tools and datasets in the **PoliRuralPlus Hub** and will support functionalities within **Jackdaw (GeoAI Chatbot)**. The interpreted EO data will enhance decision-making by providing accurate and up-to-date information on:

- **Crop health and distribution** for precision agriculture.
- **Land cover dynamics** for environmental monitoring and spatial planning.

- **Land use changes** to inform sustainable development strategies and policy-making.

By leveraging advanced image interpretation techniques, the **EO classification suite** will offer robust capabilities for **crop detection, land cover mapping,** and **change detection**, ensuring that stakeholders have the tools necessary to analyze and respond to spatial and environmental challenges effectively.

5.2.4 Document parser and embedding

The project will develop a dynamic, multilingual repository of policy documents from across Europe, enabling efficient access to relevant information for stakeholders. The repository is structured to provide spatial relevance (e.g., country, county, or municipality), detail relevance (e.g., applicable at national, regional, or field-specific levels), temporal relevance (e.g., specific time periods), and thematic relevance (e.g., environmental legislation). Documents are stored in a file repository implemented within the Digital Innovation Hub, using Wagtail CMS.

The knowledge stored in these documents can be utilized to answer users' questions effectively. Retrieval-Augmented Generation (RAG) is an approach that enables Large Language Models (LLMs) to incorporate pre-existing data into their responses. This method offers three primary advantages:

1. **Reduction of Hallucinations:** LLMs are prone to generating inaccurate or "hallucinated" answers. By incorporating documents during the response generation process, the likelihood of hallucinations is significantly reduced, leading to more reliable outputs.
2. **Enhanced Relevance:** Using already gathered knowledge or organizational policies ensures that the responses are contextually relevant and tailored to the user’s needs. This alignment improves the utility and accuracy of the information provided.
3. **Citable Answers:** Responses are supported by concrete citations from the source documents, allowing users to verify and explore the referenced information. This transparency improves trust and functions as an advanced search engine, where users can access both synthesized answers and source material.

To enhance usability, a chunking and vectorization pipeline generates vector embeddings for documents and their fragments. This allows knowledge within the repository to be effectively utilized using a RAG approach, a technique enabling large language models (LLMs) to incorporate pre-existing data into their responses. A RAG offers three key benefits: it reduces hallucinations by grounding responses in source documents, ensures contextual relevance tailored to user needs, and provides citable answers with concrete references, improving transparency and trust.

Data Preparation Process:

To optimize the repository for the document RAG, a meticulous data preparation process is implemented:

1. **Text Extraction:** Documents, such as PDFs, are processed using advanced techniques to ensure completeness.
2. **Chunking:** Content is divided into manageable fragments using two methods:
   - **Standard Chunking:** Based on word count, ensuring uniform chunk sizes.
   - **Semantic Chunking:** Based on thematic coherence, maintaining logical content groupings. Chunks with insufficient content are filtered out to retain meaningful data.

3. **Embedding:** Each chunk is converted into a high-dimensional vector representation to enable efficient retrieval of semantically similar content.
4. **Vector Storage:** Chunks and their corresponding vectors are stored in a vector database, enabling rapid similarity searches during user queries.

System Architecture:

The RAG system operates as a scalable backend service accessible through a secure REST API. The service processes user queries via a single POST endpoint, ensuring robust and efficient interactions. The workflow includes:

1. **Query Embedding:** User messages are converted into vector representations.
2. **Retrieval:** The vectorized query searches the database for the top k relevant document chunks based on semantic similarity.
3. **Response Generation:** Retrieved chunks and metadata are passed to the chatbot, which generates accurate, evidence-based responses with citations.

This modular and scalable architecture ensures that the chatbot delivers precise, reliable, and transparent information. By leveraging advanced embedding and retrieval techniques, PoliRuralPlus empowers stakeholders to access tailored insights and policy-relevant information efficiently, driving informed decision-making across diverse contexts.

6. Application Development and Evolution

6.1 Hub4Everybody Evolution

The **Hub4Everybody** is set to evolve into a comprehensive and versatile platform capable of supporting a wide range of data management, analysis, and AI integration functions. The most critical aspect of this evolution is the seamless integration of the Hub with key AI tools such as **Jackdaw (GeoAI Chatbot)**, **Magpie**, and advanced textual document analysis capabilities. The primary goal is to establish the Hub as a **universal data repository** that serves as the central resource for training and feeding AI models, particularly **Jackdaw**.

Key Components of the Hub's Evolution

1. **Integration with Jackdaw (GeoAI Chatbot)**:
   - The Hub will provide structured and unstructured data to **Jackdaw**, enhancing its capability to deliver AI-driven geospatial insights and recommendations.
   - Jackdaw will utilize the Hub's data repository for tasks like spatial analysis, WMS interpretation, EO image classification, and decision-making support.
   - Real-time data access and query capabilities will enable Jackdaw to respond dynamically to user requests.
2. **Integration with Magpie**:
   - **Magpie** will serve as a discovery service within the Hub, locating relevant datasets, metadata, and WMS sources for further processing and analysis.
   - The Hub will facilitate the automatic retrieval and integration of data discovered by Magpie, ensuring a streamlined workflow for spatial and textual data management.
3. **Textual Document Analysis**:
   - The Hub will support the ingestion, storage, and analysis of various textual documents, including reports, policy briefs, research papers, and training materials.
   - AI-driven tools will extract named entities, policy changes, events, and trends from these documents, providing insights for decision-making and policy development.
   - This functionality will enhance AI tools like Jackdaw by providing them with up-to-date, context-rich textual information.
4. **Universal Data Repository**:
   - The Hub will function as a **centralized repository** for all types of data, including geospatial datasets, satellite imagery, textual documents, and metadata.
   - It will ensure that data is easily accessible, organized, and ready to be used by AI tools for analysis, training, and decision support.
   - Data formats will be standardized to facilitate interoperability between different components of the Hub ecosystem.
5. **AI Integration in E-Learning**:
   - The Hub will incorporate AI tools into its **e-learning and training modules**, offering interactive and adaptive learning experiences.
   - AI-driven features such as personalized content recommendations, automated assessments, and dynamic learning pathways will enhance user engagement and knowledge retention.

  - The integration of Jackdaw into training modules will provide real-time geospatial analysis and insights to support learning objectives.

Summary of Key Functionalities

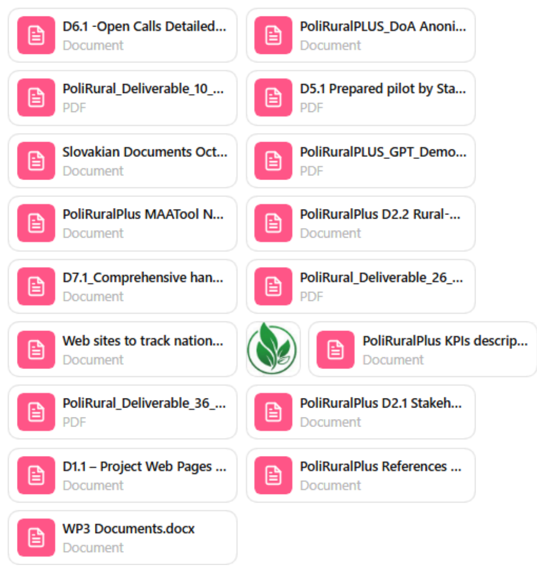
- **Jackdaw Integration**: Provides AI-driven geospatial insights and dynamic analysis.
- **Magpie Integration**: Facilitates data discovery and retrieval for seamless integration with the Hub.
- **Textual Analysis**: Enables extraction and analysis of information from documents for AI-driven decision support.
- **Central Repository**: Serves as a unified source for training and feeding AI models.
- **AI-Enhanced E-Learning**: Integrates AI into training modules for interactive learning experiences.

By integrating **Jackdaw**, **Magpie**, and advanced textual analysis capabilities, the **Hub4Everybody** will evolve into a robust, AI-powered platform. This universal data repository will support comprehensive data analysis, enhance decision-making processes, and provide innovative e-learning solutions, making it a cornerstone of the **PoliRuralPlus** project's digital infrastructure.

6.2 PoliRuralPlus Advisor, Custom GPT

Advisor

The PoliRuralPlus Advisor is an interactive tool and service designed to facilitate and optimize the development and implementation of Regional Action Plans (RAPs) as part of the PoliRuralPlus project, which focuses on fostering sustainable, balanced, and inclusive development between rural and urban communities, leveraging innovative methodologies and digital tools.



Advisor fits into the PoliRuralPlus platform as an user friendly service for all pilot stakeholders to access all of the project's Knowledge Space and References database as it is created and evolved by the ongoing project work[1].

While publicly available to any user with the link, to explore the contents of it Knowledge Space, the PoliRuralPlus Advisor's contents are retained within its Team workspace and are not used by ChatGPT to train its wider models

Advisor's role in the in the PoliRuralPlus ecosystem:

- Support for Regional Action Plans: It aids stakeholders in crafting RAPs that align with EU sustainability goals while considering local needs and capacities.
- Knowledge Dissemination and Collaboration: Centralizes learnings and best practices across pilot projects, encouraging cross-regional learning and scalability.
- Innovation Catalyst: Drives the integration of digital solutions, including low-code/no-code platforms and foresight tools, to modernize rural-urban planning.

[1] See section 1.2 and annexes N and O of PoliRuralPlus D2.2 Rural-urban Linkages & Opportunities post COVID-19 for a discussion of the use of Advisor.

Advisor ensures that the PoliRuralPlus project's objectives of inclusivity, innovation, and sustainability are met by equipping stakeholders with state-of-the-art tools and knowledge systems. It not only enhances the quality of decision-making but also fosters collaboration among diverse stakeholders.

6.2.1 Technical Description

Advisor is built using ChatGPT's Retrieval-Augmented Generation (RAG) custom GPT technology[2] to provide :

Core Functionality:

- Knowledge Hub Integration: Centralizes access to project deliverables, resources, and research findings, making them easily searchable and interactive.
- Interactive Intelligence: Built on advanced language models, the Advisor serves as a virtual assistant for project stakeholders, enabling tailored support for specific regional and thematic tasks.
- Customizability: Tailored to address specific regional contexts and needs, incorporating foresight tools and generative AI capabilities to enhance stakeholder engagement and co-creation processes.

Key Features:

- Stakeholder Engagement: Facilitates the identification and coordination of multi-actor collaborations, central to the project's methodologies.
- Decision Support: Assists policymakers and planners in understanding complex interrelations in rural-urban linkages, providing data-driven insights for RAP development.
- Regional and Sectoral Analysis: Offers tools for analyzing the impacts of variables like COVID-19 on rural-urban linkages and identifying opportunities for economic, social, and environmental improvements.
- Artificial Intelligence and Data Analysis: Supports foresight methodologies and dynamic modelling to anticipate future challenges and identify sustainable solutions.
- Knowledge Automation: Automates documentation and synthesis processes, reducing manual efforts and ensuring consistency across pilot regions.

6.3 Jackdaw: A Spatial Enabled Chat Agent

Jackdaw is an innovative spatial-enabled chat agent that allows users to engage in spatially contextualized conversations with AI agents, addressing challenges often encountered in rural development scenarios. Traditional large language models (LLMs) struggle to provide accurate answers when questions pertain to lesser-known locations or areas with limited internet-based information. Jackdaw overcomes this limitation by proactively enriching user queries with relevant spatial data before presenting them to the AI model.

Users initiate a query by providing a natural language question along with a designated location on a map to spatially limit the scope of the inquiry. Jackdaw parses the user's question to identify key topics, the required spatial and temporal resolutions, and any additional context needed. The system then queries relevant data

[2] Demystifying the Use of RAG in Custom GPT: The Power of Data Retrieval in Answer Generation - RDD10+

catalogs to locate appropriate data sources, services, and models that can inform the response. Concurrent API calls are made to these data sources, retrieving both structured data and textual responses.

The collected spatial context and retrieved data are dynamically integrated into the query's context, allowing the AI to generate human-readable text responses enriched with spatial relevance. In addition to chat-based responses, Jackdaw provides structured outputs that can be visualized through tabular, graphical, or dynamic dashboards. These visualizations include charts, gauges, tables, and maps, enabling users to explore results from multiple perspectives. Users can switch between a standard conversational interface and a dynamic BI dashboard to gain deeper insights.

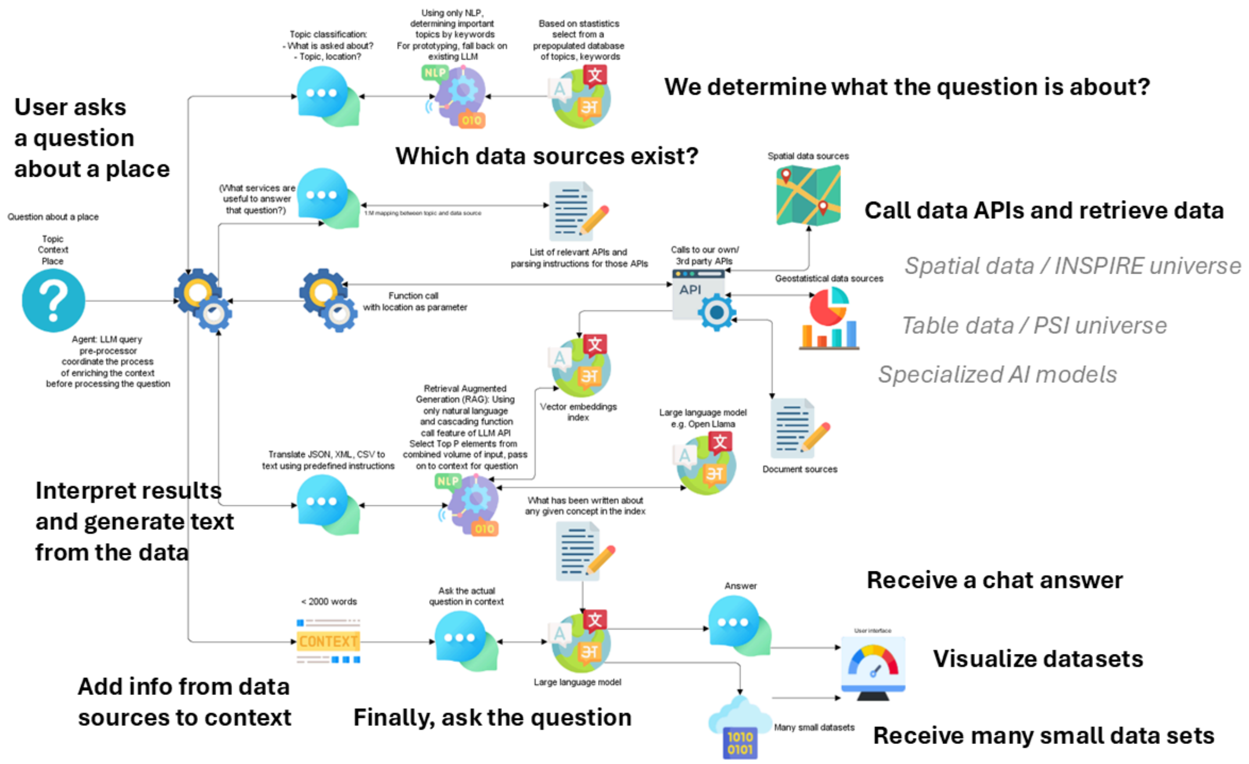


Figure 5: Overview of Jackdaw concept

Jackdaw is purpose-built for rural development practitioners, empowering them to use AI-driven spatial tools for daily management, operations, and evidence-based policymaking. By integrating spatial intelligence with AI capabilities, Jackdaw prototypes the future of geospatial LLMs, preparing decision-makers to harness advanced technology for actionable insights.

Technically, Jackdaw is built on a Python-based FastAPI backend, leveraging a Qdrant vector embedding server and PostgreSQL relational database for robust data storage and retrieval. The API includes business logic and controllers that connect to various tools and data services, enabling seamless integration of geospatial data. Key technologies include OpenAI, LangChain, Pandas, NumPy, and Scikit-learn. Jackdaw operates within a mono-repo alongside its sibling applications, Magpie and Vulture, and is secured via OAuth authentication through the Digital Innovation Hub.

By combining natural language understanding, spatial data integration, and advanced visualization capabilities, Jackdaw sets a new standard for spatially enabled AI applications, fostering smarter, data-driven decisions in rural development.

The application may be accessed here: https://jackdaw.avinet.no

6.4 Vulture: Sandbox for AI-driven Productivity Tool Experiments

Vulture is a flexible and powerful sandbox environment developed to meet user requirements identified through the WP3 buddy exercises. Designed for productivity experimentation, Vulture leverages AI to process information streams efficiently, enabling users to manage, analyze, and retrieve relevant content seamlessly.

Information Ingestion and User Management

To ingest information, Vulture allows users to forward content to a designated email address, either manually or automatically, from their own email accounts. By requiring users to make a deliberate decision to forward information, Vulture ensures compliance with privacy regulations like GDPR and addresses security concerns related to data management. Each forwarded email, including its attachments, is validated and added to a private document collection that is securely tied to the user profile. Access to this collection is strictly limited to the user who forwarded the content, ensuring data confidentiality.

A user profile is created automatically when a user first interacts with the system, and multiple email addresses can be associated with a single profile. This approach offers flexibility by allowing diverse content to be ingested regardless of technologies or platforms. For instance, users can forward WhatsApp messages, PDF documents, web page clippings, or other shareable content from any device or email client, making Vulture highly adaptable to user workflows.

Processing and Storage of Information

Once an email and its attachments are received, Vulture processes the content using an AI-powered workflow. The content is:

1. **Analyzed for Context**: Each email is examined to extract meaningful insights.
2. **Chunked and Vectorized**: The content, along with attachments, is divided into smaller segments (chunks) and converted into vector embeddings. This transformation enables efficient semantic search and retrieval.

The embeddings and associated data are stored in a private user collection within a Qdrant vector database. Users effectively build a self-expanding Retrieval-Augmented Generation (RAG) system, where their personal repository of documents grows dynamically as more content is forwarded to Vulture.

Interactive AI-Driven Search

Vulture provides an intuitive interface where users can query their document collections using natural language. When a question is asked, the system follows a two-step process:

1. **Context Retrieval**: Relevant context documents are extracted from the vector database based on semantic similarity to the user's query.
2. **Answer Generation**: The retrieved context is passed to OpenAI's GPT-4o, which processes the question and generates a comprehensive response.

This functionality supports a wide range of use cases, including but not limited to:

- **Tracking events** and deadlines.
- **Identifying reports** and critical insights.
- **Discovering funding opportunities** or policy updates.

Crawling and URL-Based Interrogation

In addition to email ingestion, Vulture enables users to register URLs for crawling and targeted information extraction. Users can define specific questions or topics of interest, and Vulture will retrieve relevant information from the linked web pages. This feature is particularly useful for policy impact tracking, as demonstrated in the Latvian pilot for the Vidzeme Planning Region, where Vulture is used to monitor and analyze policy developments.

Technical Architecture

Vulture is implemented as a Python FastAPI application, providing robust web service endpoints for seamless backend operations. The platform's architecture includes the following components:

- **Front-End**: Developed using **React**, ensuring an intuitive and responsive user interface.
- **AI Integration**: The system interfaces with **OpenAI APIs** for both natural language processing (GPT-4o) and embedding generation.
- **Vector Database**: Vulture uses **Qdrant** for efficient storage and querying of vector embeddings, enabling fast and accurate semantic searches.
- **Relational Database**: Business logic and metadata are managed in **PostgreSQL** for reliability and scalability.

Vulture is maintained within a mono-repository alongside its sibling applications, Jackdaw and Magpie, fostering a unified and modular ecosystem of tools.

Empowering Productivity

With its email-based ingestion, dynamic RAG system, and AI-powered querying capabilities, Vulture empowers users to organize and process vast amounts of information efficiently. Whether tracking opportunities, analyzing reports, or extracting insights from web content, Vulture provides a flexible and scalable solution for productivity experimentation. By blending intuitive workflows with cutting-edge AI, Vulture addresses real-world needs while ensuring data privacy, adaptability, and ease of use.

The preliminary version of the application may be accessed here: https://jackdaw.avinet.no/vulture (alpha)

6.5 Magpie: AI-driven OGC Metadata Catalog Search

Magpie is an advanced application designed to enhance traditional geospatial metadata catalogs, such as those adhering to INSPIRE and OGC standards, by incorporating AI-driven search capabilities and automated map creation. The system begins by ingesting metadata compliant with established ISO standards like ISO 19115 (geographic metadata) and ISO 19119 (services). This ensures compatibility with recognized protocols, fostering interoperability and maintaining standardization. Metadata records, including Web Map Services (WMS), Web Feature Services (WFS), and APIs, are analyzed to identify actionable and accessible service layers that can be directly integrated into user-facing tools.

To optimize searchability, Magpie consolidates metadata text fields such as abstracts, keywords, and descriptions into a unified text body through normalization processes like deduplication, tokenization, and cleaning. This prepares the metadata for advanced semantic embedding. Using pre-trained or fine-tuned transformer models such as Sentence-BERT or OpenAI's CLIP, the unified text is transformed into vector embeddings that capture the semantic relationships within the metadata. These embeddings, along with essential metadata identifiers and links to original catalog entries, are stored in a scalable vector database, such as Qdrant, ensuring efficient retrieval.

Magpie features a RESTful Web Service API that allows users to interact with the system using natural language or keyword-based queries. User queries undergo parsing and contextual expansion, leveraging NLP techniques to extract intent, integrate synonyms, and enhance relevance. The system then generates vector embeddings for the user query, aligning them with the pre-generated metadata embeddings. By performing nearest-neighbor searches in the vector database, Magpie retrieves metadata records with the highest semantic similarity scores.

The response to user queries includes well-structured metadata records in a JSON format. Each record contains key information such as layer names, descriptions, keywords, and service endpoints (e.g., WMS/WFS). Additionally, Magpie enables users to automatically generate maps by leveraging the identified service layers, seamlessly turning search results into actionable geospatial visualizations.

With its ability to bridge traditional metadata catalog systems and modern AI technologies, Magpie empowers users to explore, discover, and visualize geospatial data in a streamlined and intelligent manner.

The application demonstrator may be accessed here: https://jackdaw.avinet.no/magpie

6.6 Enhancements to Multi Actor Approach tool

The Multi-Actor Approach (MAA) Tool is a flexible, collaborative platform designed to support the effective management of pilots and living labs by engaging diverse stakeholders in rural-urban development initiatives. It integrates configurable workflows, real-time collaboration and advanced data visualization to enhance decision-making and planning. The tool fosters inclusivity and shared ownership, enabling stakeholders to contribute and access knowledge in a secure environment. With customizable dashboards and a modular design, the MAA Tool adapts to evolving project needs while promoting transparency, scalability, and interoperability for long-term, impactful outcomes.

The main technical characteristics of the Multi-Actor Approach (MAA) Tool in the PoliRuralPlus project include the following structural and functional aspects:

1. Work Entities and Data: The tool includes configurable modules for organizing tasks, workflows, and specific project objectives. Additionally, it supports multi-layered data, including geospatial information, and various economic, social, and environmental indicators, offering a comprehensive approach to data management.

2. User Management: The MAA Tool allows for diverse user roles, including administrators, analysts, pilot leaders, external stakeholders, and the general public. Access levels and permissions can be customized to determine who can view, edit, or comment on specific datasets. This flexibility ensures secure yet collaborative engagement among users.

3. Data Visualization and Dashboards: Interactive dashboards are a core feature of the tool, providing users with the ability to visualize key metrics, track indicator progress, and display maps and charts in real time. These dashboards can be tailored to meet specific project needs, such as showing trends or analyzing impact through custom matrices.

4. Scalability and Extensibility: The tool is designed with modularity in mind, allowing it to adapt and expand according to the specific needs of regional pilots. New features can be added as required, making it a flexible solution for evolving project requirements. Furthermore, its interoperability ensures compatibility with external systems, allowing for seamless data import/export and integrated reporting across platforms.

The deliverable D5.1 (Prepared pilot by Stakeholder Multi-Actor Approach v1.0) includes a more detailed description of the characteristics of the MAA Tool.

6.9 MapWhiteboard Evolution

MapWhiteboard is an innovative tool designed to enable collaborative map browsing and spatial data editing, much like how platforms such as Miro and Figma enhance brainstorming and digital whiteboard workflows. It integrates traditional web GIS functionalities with real-time collaboration features, offering a unique solution for shared map-based interactions.

Key features include:

1. **Pre-defined Maps**: Configured with both base map layers and editable layers.
2. **GeoJSON Support**: Uploading, editing, and downloading spatial data.
3. **WMS Layer Integration**: Users can insert and interact with Web Map Service (WMS) layers.
4. **Shared Cursors**: Real-time visibility of all users' cursors on the map.
5. **Collaborative Editing**: Simultaneous editing of features and attributes in editable layers.
6. **Dynamic Layer Creation**: Users can define and create new editable layers.
7. **Third-Party Publishing**: Allows the publication of geometries or features into layers through web services.

The solution is built on the OpenLayers library, using a custom JavaScript extension for event handling, a WebSocket API for real-time communication, and a REST API for managing data and maps. The client application is based on HSLayers-NG, a specialized extension of OpenLayers designed for creating rich web map applications.

While MapWhiteboard remains a powerful tool at TRL 7, further enhancements are necessary to unlock its full potential and expand its usability. The evolution involves both the extensions of its core capabilities and integration with other technologies and workflows. By focusing on user-centric development, interoperability, and AI-powered features, it can become an indispensable tool for collaborative geospatial work across industries and sectors. Below are some proposed development paths:

Integration with AI Tools

Integration with AI tools for MapWhiteboard will focus on intelligent metadata search, enabling users to retrieve relevant information directly applicable to their maps. This functionality will leverage the GeoAI chatbot Jackdaw, which allows users to input text prompts alongside a specified area of interest, streamlining the process of obtaining and applying geospatial insights.

Extended Interoperability

Strengthen integration with the platform of **Hub4Everybody**, to facilitate unified user management and enhance API interoperability.

Enhanced Mobile Support

Refine the mobile user experience for better interaction on smaller screens, including intuitive gestures for navigation, editing, and collaboration.

Usability Focus

Develop step-by-step onboarding tutorials to guide new users in collaborative workflows. Introduce a "light mode" with reduced complexity for non-technical users who require basic map viewing of shared maps from other users.

6.10 Rural Attractiveness Analysis Application

The Rural Attractiveness Analysis Application will be based on an existing application developed in the Polirural project, which offers a wide range of features including data visualisation, filtering of input data, use of map viewer functions (zoom, print, layers) and working with user scenarios. These features provide essential tools for assessing and presenting the attractiveness of rural regions.

Future development of the application will focus on adapting it to the requirements of individual pilots within the Polirural Plus project. Each pilot understands attractiveness differently, from innovation potential to tourism. It will therefore be crucial to have a dialogue with each pilot to ensure that the app reflects their specific needs and expectations.

One of the main improvements will be the combination of existing statistical data with infrastructure data, which will be drawn from OpenStreetMap. This combination will allow a more detailed assessment of attractiveness, as infrastructure plays a key role in deciding how attractive a region is to residents, businesses or tourists.

The next step in the development of the application involves the introduction of AI interpretation of attractiveness assessment models. AI will be able to analyse large amounts of data and provide recommendations or new insights on the attractiveness of regions, which will significantly increase the added value of the app and its usability for a wide range of users.

Overall, the app will continue to expand its potential and bring innovative approaches to assessing the attractiveness of rural regions, while remaining user-friendly and flexible.

7. Implementation Plan

7.1 Roadmap and Timeline for Development

The implementation plan for the project spans three years, structured into distinct phases to ensure systematic development and integration of innovative solutions tailored to the needs of rural and urban regions.

The PoliRuralPlus project commenced in January 2024 and is scheduled to run for 36 months, concluding in June 2026. Building on the successes of its predecessor, PoliRural, the project focuses on advancing digital tools and participatory methodologies to foster sustainable rural-urban integration. Key developments include enhancing the **Digital Innovation Hub (DIH)** with advanced features such as the **Map Whiteboard**, AI/ML-driven analytics, and participatory foresight methods. Stakeholder engagement is central to the project, supported by Open Calls to promote the adoption of innovative technologies and ensure active participation in regional development efforts.

7.1.1 Project Phases and Timeline

Phase 1: Project Setup and Initial Platform Design (January 2024 – December 2024)

The first year of the project is dedicated to designing the platform infrastructure, engaging stakeholders, and laying the groundwork for technology deployment and regional integration.

- **Platform Design**: The **Digital Innovation Hub (DIH)** is conceptualized and designed, focusing on integrating the Map Whiteboard and AI/ML analytics tools. Key functionalities, such as collaborative geospatial analysis and predictive modeling, are defined.
- **Stakeholder Mapping**: Identification and engagement of stakeholders, including farmers, local authorities, businesses, and NGOs, is conducted to establish communication channels and collaboration frameworks.
- **First Platform Release**: By the end of 2024, the first operational version of the DIH, including the Map Whiteboard, is released for initial testing and feedback in the pilot regions.

Deliverables during this period include the preliminary DIH design, stakeholder engagement plans, and the first operational version of the platform.

Phase 2: Tool Refinement and Methodology Development (January 2025 – June 2025)

In this phase, the DIH is refined based on initial feedback, and participatory methodologies are developed to support integrated rural-urban strategies.

- **Platform Refinement**: Enhancements are made to the platform, incorporating feedback from early users to improve usability and functionality.
- **Development of Methodologies**: Advanced system dynamics modeling and foresight techniques are further developed to align with regional needs.
- **Stakeholder Workshops**: Initial workshops and engagement activities are held in pilot regions to validate tools and methodologies.
- **Use Mobilise call** for involved of stakeholders in pilot region to test and used solution

This phase ensures that the platform is robust, user-friendly, and capable of supporting the next phase of implementation.

Phase 3: Full Platform Deployment and Pilot Implementations (July 2025 – December 2025)

The fully operational DIH, along with refined tools and methodologies, is deployed across nine pilot regions. This phase focuses on implementing and validating regional action plans (RAPs).

- **Platform Deployment**: The fully enhanced DIH, including the finalized Map Whiteboard and AI/ML analytics tools, is rolled out across pilot regions.
- **Pilot Implementation**: Stakeholders use the platform to co-create and execute RAPs, addressing key challenges such as governance, connectivity, and socio-economic integration.
- **Open Calls**: Two rounds of Open Calls are launched during this phase to encourage further stakeholder involvement and the rollout of new tools and solutions.

Outputs from this phase include validated RAPs and a fully operational platform demonstrating its applicability in diverse regional contexts.

7.2 Deployment and Testing Phases

The deployment and testing phases of the PoliRuralPlus project are designed to ensure a robust, user-centered validation process that progressively refines the tools and solutions developed. These phases encompass technical testing, stakeholder engagement, feedback assimilation, and iterative improvement across work packages and project partnerships.

Technical Testing Phase

The initial phase focuses on the rigorous technical evaluation of the PoliRuralPlus platform and its components. During this stage:

- Comprehensive **test cases** are developed to cover functional, performance, and integration aspects of the platform.
- Each system component undergoes **isolated testing** to ensure it meets design specifications and operates reliably.
- The entire system is subjected to **end-to-end testing** to validate its operational coherence and readiness for deployment in pilot regions.

Pilot Testing Phase

Once the technical integrity is validated, the platform and tools are deployed in the project's pilot regions for practical evaluation. This phase includes:

- **Use-case testing** within specific pilot contexts, leveraging real-world scenarios to assess the system's effectiveness in addressing regional challenges.
- Active participation of **pilot stakeholders**, including policymakers, planners, and local organizations, to evaluate system usability and relevance.

Wider Stakeholder Testing

To ensure the solutions meet a diverse range of needs, testing is expanded to a broader audience of project stakeholders. This involves:

- Engagement with **external stakeholders**, such as NGOs, SMEs, and educational institutions, to test the platform under varied conditions.
- Testing scenarios designed to reflect the **broader application spectrum** of the tools and technologies, ensuring adaptability and scalability.

Feedback Collection and Analysis

Feedback is a cornerstone of the iterative development process:

- All testers, including pilots and external stakeholders, are encouraged to provide detailed feedback on usability, performance, and functionality.
- Collected feedback is systematically **assessed and abstracted** to identify patterns and prioritize improvements.
- The results of this analysis feed directly into the enhancement process, ensuring the platform evolves to meet user expectations and operational goals.

Cross-Work Package and Partnership Decision-Making

Key decisions related to the deployment and testing phases are made collaboratively across work packages and the partnership. This includes:

- Regular **coordination meetings** to align testing outcomes with project goals and identify cross-WP dependencies.
- Shared responsibility among partners to implement improvements based on testing feedback, ensuring seamless integration and consistency.
- Strategic alignment of testing insights with broader project objectives to inform future development phases.

7.3 Risk Assessment and Mitigation Strategies

The successful execution of Work Package 4 (WP4) within the PoliRuralPlus project requires careful identification, assessment, and mitigation of risks. While the project faces a range of overarching risks, WP4-specific challenges relate to technological development, user engagement, integration, and timeline adherence. This chapter outlines the key risks and corresponding mitigation strategies.

7.3.1 Risk Identification and Assessment

1. Technological Risks

- **Risk**: Development delays due to technical complexities in integrating AI processing capabilities and existing systems.

- **Likelihood**: Medium
- **Impact**: High
- **Mitigation**:
    - Adopt a modular development approach to allow independent progress on individual components.
    - Conduct frequent technical reviews and deploy incremental deliverables to identify issues early.

2. Data-Related Risks

- **Risk**: Insufficient or poor-quality data from pilot regions for model training and validation.
- **Likelihood**: Medium
- **Impact**: High
- **Mitigation**:
    - Collaborate with regional stakeholders to ensure robust data collection processes.
    - Develop fallback mechanisms, such as synthetic data generation, to mitigate data gaps.

3. Stakeholder Engagement Risks

- **Risk**: Limited engagement or misalignment with stakeholder expectations during platform validation.
- **Likelihood**: Medium
- **Impact**: Medium
- **Mitigation**:
    - Facilitate workshops and hands-on demonstrations to align expectations.
    - Integrate user feedback iteratively to ensure alignment with stakeholder needs.

4. Integration Risks

- **Risk**: Integration challenges when combining newly developed AI tools with existing systems in pilot regions.
- **Likelihood**: High
- **Impact**: Medium
- **Mitigation**:
    - Leverage Web APIs and standardized protocols to simplify integration.
    - Conduct pre-integration testing to identify compatibility issues.

5. Timeline Risks

- **Risk**: Delays in platform development affecting deployment and testing schedules.
- **Likelihood**: Medium
- **Impact**: High
- **Mitigation**:
    - Establish a detailed project roadmap with buffer periods for critical tasks.
    - Monitor progress through regular check-ins and milestone evaluations.

6. Adoption and Sustainability Risks

- **Risk**: Difficulty in achieving long-term adoption of the platform due to technical or operational challenges.
- **Likelihood**: Medium
- **Impact**: High
- **Mitigation**:
  - Provide training resources and user-friendly documentation.
  - Engage stakeholders early to co-design features ensuring practical relevance.

7.3.2 Mitigation Framework

To address these risks comprehensively, WP4 employs a layered mitigation strategy:

- **Proactive Monitoring**: Establish key performance indicators (KPIs) for each risk category, enabling early detection and intervention.
- **Iterative Development**: Use agile methodologies to incorporate feedback and make adjustments during each development cycle.
- **Stakeholder Collaboration**: Maintain open communication channels to ensure continuous alignment of technical goals with user requirements.
- **Risk Ownership**: Assign clear responsibilities to partners and WP leaders for managing specific risks, ensuring accountability.

8. Conclusion and Future Outlook

8.1 Summary of Key Design Decisions

The development of the PoliRuralPlus platform has been guided by critical design decisions aimed at ensuring its functionality, adaptability, and user relevance.

- **User-Centric Approach**: The platform was designed to prioritize the needs of diverse stakeholder groups, including policymakers, planners, SMEs, and the education sector. This informed decisions on interface design, functionality, and data accessibility.
- **Integration of Advanced Technologies**: Leveraging AI capabilities and advanced data processing tools, the platform integrates cutting-edge technologies to deliver actionable insights for regional planning and rural-urban linkages.
- **Modular Architecture**: A modular design enables seamless integration with existing systems, adaptability to pilot-specific contexts, and scalability for future applications.
- **Iterative Development**: The platform was built incrementally, incorporating continuous feedback from stakeholders and iterative testing to refine functionalities and enhance user experience.
- **Sustainability and Open Access**: Provisions were made to ensure the platform's long-term sustainability, including potential for open-source deployment and integration with public domain tools.

8.2 Anticipated Challenges and Next Steps

While significant progress has been made, the following challenges are anticipated as the platform moves into broader deployment and operational phases:

- **Complexity of Stakeholder Needs**: Addressing the diverse requirements of stakeholders across pilot regions will necessitate ongoing customization and support.
- **Data Quality and Availability**: Ensuring consistent access to high-quality, region-specific data will remain a critical factor for the platform's effectiveness.
- **Adoption and Engagement**: Encouraging widespread adoption of the platform among stakeholders, particularly those less familiar with digital tools, poses a potential hurdle.
- **Integration with Existing Systems**: Technical challenges related to compatibility and data harmonization with pre-existing solutions may require further refinement.
- **Sustainability and Maintenance**: Developing a robust financial and operational model to sustain the platform beyond the project timeline is imperative.

To address these challenges, next steps include:

- Expanding training and capacity-building initiatives for stakeholders.
- Establishing dedicated support teams for technical integration and customization.
- Strengthening partnerships to enhance data sharing and collaborative use of the platform.
- Implementing feedback mechanisms to refine the platform continuously based on real-world use.

8.3 Long-term Vision for the Platform

The long-term vision for the PoliRuralPlus platform encompasses its evolution as a central tool for fostering sustainable rural-urban linkages. Key objectives include:

- **Scalability Across Regions**: Expanding the platform's reach to additional regions across Europe and beyond, adapting to diverse governance, economic, and environmental contexts.
- **Innovation Hub**: Transforming the platform into a hub for innovation, where new tools, methodologies, and technologies are developed and tested.
- **Integrated Ecosystem**: Enabling seamless interoperability with broader systems and policies, including EU initiatives like the Green Deal and the Long-Term Vision for Rural Areas.
- **Empowerment through Open Access**: Supporting equitable development by ensuring open access to tools and insights, fostering inclusivity among all stakeholder groups.
- **Global Contributions**: Leveraging the platform's capabilities to contribute to global challenges such as climate adaptation, digital transformation, and sustainable development.

This vision ensures that the PoliRuralPlus platform continues to provide meaningful value, driving innovation and inclusivity in regional planning and policy-making for years to come.