

OGC® DOCUMENT: 25-010

External identifier of this OGC® document: <http://www.opengis.net/doc/PER/CDRP-2024.2/D001>



Open
Geospatial
Consortium

OGC CLIMATE AND DISASTER RESILIENCE PILOT 2024.2 ENGINEERING REPORT

ENGINEERING REPORT

PUBLISHED

Submission Date: 2025-03-31

Approval Date: 2025-03-31

Publication Date: 2025-07-03

Editor: Stelios Contarinis, Vyron Antoniou, Micah Brachman

Notice: This document is not an OGC Standard. This document is an OGC Public Engineering Report created as a deliverable in an OGC Interoperability Initiative and is *not an official position* of the OGC membership. It is distributed for review and comment. It is subject to change without notice and may not be referred to as an OGC Standard.

Further, any OGC Engineering Report should not be referenced as required or mandatory technology in procurements. However, the discussions in this document could very well lead to the definition of an OGC Standard.

License Agreement

Use of this document is subject to the license agreement at <https://www.ogc.org/license>

Copyright notice

Copyright © 2025 Open Geospatial Consortium

To obtain additional rights of use, visit <https://www.ogc.org/legal>

Note

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. The Open Geospatial Consortium shall not be held responsible for identifying any or all such patent rights.

Recipients of this document are requested to submit, with their comments, notification of any relevant patent claims or other intellectual property rights of which they may be aware that might be infringed by any implementation of the standard set forth in this document, and to provide supporting documentation.

CONTENTS

I.	EXECUTIVE SUMMARY	vi
I.A.	Purpose of the Pilot	vi
I.B.	Key Objectives and Scope	vi
I.C.	Summary of Findings and Recommendations	viii
II.	KEYWORDS	x
III.	CONTRIBUTORS	x
1.	INTRODUCTION	2
1.1.	Pilot's Background and Motivation	2
1.2.	Importance of GenAI in Climate Resilience and Disaster Management	3
1.3.	Alignment with UN SDGs and Climate Resilience Goals	3
1.4.	OGC's Role in Advancing AI Integration	4
2.	CHALLENGES IN EXPLOITING GENAI WITHIN THE CLIMATE RESILIENCE DOMAIN	6
2.1.	Geospatial Awareness in Generative AI Systems	6
2.2.	Hallucinations in the Geospatial Domain	7
2.3.	Aligning GenAI with Geospatial Standards	9
2.4.	Integration with Application Programming Interfaces (APIs)	10
3.	ANALYSIS READY DATA (ARD) MATURITY REPORT (D010)	13
3.1.	Data Maturity Framework	13
3.2.	Tools for Self-Service Assessments	13
4.	GENERATIVE AI FOR WILDFIRE REPORT (D030)	16
4.1.	Use Cases and Functionalities	16
4.2.	Data Sources and FAIR Evaluation	16
4.3.	OGC Compliance and Interoperability	16
4.4.	Findings and Recommendations	17
5.	INFORMATION INTEROPERABILITY REPORT (D040)	19
5.1.	Objective and Approach	19
5.2.	Mapping Core Concepts	19
5.3.	Use Cases and Practical Applications	19
6.	GENERATIVE AI VIRTUAL ASSISTANTS: DESIGN AND IMPLEMENTATION	22
6.1.	Objectives and Use Cases for Virtual Assistants	22
6.2.	Design and Architecture of Demonstrators	22

6.3. Training Data: Types, Sources, and Preprocessing	23
6.4. Validation Practices Against Hallucinations	23
7. CASE STUDIES AND PROTOTYPES	27
7.1. Demonstrator for Virtual AI Assistants (D100 – GeoLabs)	27
7.2. Demonstrator for Virtual AI Assistants (D110 – CCMEQ):	30
7.3. Demonstrator for Virtual AI Assistants (D110 – Danti)	32
7.4. Demonstrator for Virtual AI Assistants (D120 – CRIM):	34
7.5. Demonstrator for Virtual AI Assistants (D120 – GIS.FCU):	47
7.6. Demonstrator for Virtual AI Assistants (D120 – TerraFrame):	48
7.7. Demonstrator for Sectoral Applications (D101 – Hartis):	51
7.8. Demonstrator for Sectoral Applications (D101 – Pixalytics):	54
8. PILOT FINDINGS & RECOMMENDATIONS	59
8.1. Findings	59
8.2. Recommendations	60
9. STAKEHOLDER PERSPECTIVES	62
9.1. Stakeholder Roles and Contributions	62
9.2. Feedback and Requirements from Community and Industry Stakeholders	62
9.3. Opportunities for Collaboration and Future Partnerships	63
10. FUTURE OUTLOOK	65
10.1. Recommendations for Enhancing GenAI Integration for Climate Resilience	65
10.2. Strategic directions for Stakeholders and Developers	67
10.3. Areas for Further Research and Development	68
10.4. Security, Privacy and Ethical Considerations in GenAI for Geospatial Applications	69
BIBLIOGRAPHY	72
ANNEX A (NORMATIVE) ABBREVIATIONS/ACRONYMS	75
ANNEX B (NORMATIVE) DEMO ILLUSTRATIONS FOR VIRTUAL AI ASSISTANTS	
(D100 – GEOLABS)	78
B.1. Demo Illustration 1: Searching data source	78
B.2. Demo Illustration 2: General climate-related query	78
B.3. Demo Illustration 3: Generate code example for climate data access and analysis	79

LIST OF FIGURES

Figure 1 – CDRP2024.2 Pilot - Application Domains & Platforms Ecosystem	2
Figure 2 – D100 (GeoLabs) - High Level Workflow	29
Figure 3 – D110 (Danti) - Logical Architecture Diagram	33
Figure 4 – D120 (CRIM) - Street Map	35

Figure 5 – D120 (CRIM) - Flooding Map	36
Figure 6 – D120 (CRIM) - Combined Map	36
Figure 7 – D120 (CRIM) - Architecture Diagram	37
Figure 8 – D120 (CRIM) - Agent Screenshot	38
Figure 9 – D120 (CRIM) - Lonqueil	40
Figure 10 – D120 (TerraFrame) - Use of SKGs for Cascading Effects	49
Figure 11 – D120 (TerraFrame) - Insights in a human-readable format	50
Figure 12 – D101 (Hartis) - AI-Assisted Coastal Vulnerability Analysis	51
Figure 13 – D101 (Hartis) - AI-Assisted Coastal Vulnerability Analysis Architecture	52
Figure 14 – D101 (Hartis) - AI-Assisted Coastal Vulnerability Analysis Results	53
Figure 15 – D101 (Pixalytics) - Example plot of SPI, UTCI and Health Index for the South Eastern UK from 2022 to 2024:	56
Figure B.1 – D100 (GeoLabs) - Demo 1	78
Figure B.2 – D100 (GeoLabs) - Demo 2	79
Figure B.3 – D100 (GeoLabs) - Demo 3	80
Figure B.4 – D100 (GeoLabs) - Demo 4	80



EXECUTIVE SUMMARY

I.A. Purpose of the Pilot

The **Climate and Disaster Resilience Pilot (CDRP) Phase 2** aimed to explore the integration of **geospatial data** and **generative AI** for enhancing climate resilience and disaster management. The pilot demonstrated various **AI-enabled tools**, **engineering reports**, and **sectoral applications** designed to improve data analysis, communication, and decision-making.

I.B. Key Objectives and Scope

The pilot undertook a comprehensive evaluation of the readiness of generative AI (GenAI) technologies to address pressing climate-related challenges, including **floods** and **wildfires**. It also emphasized the need to enhance data standards and compliance, foster collaboration among diverse stakeholders, and develop actionable tools and workflows. These efforts are aligned with FAIR principles (Findable, Accessible, Interoperable, Reusable) to advance global resilience strategies and support informed decision-making in climate resilience and disaster management.

I.B.1. Objective 1: Integration of Generative AI Virtual Assistants for Climate Resilience

The first objective focused on integrating **generative AI virtual assistants** into existing geospatial data frameworks to enhance climate resilience. By leveraging AI, the goal was to improve data accessibility, usability, and decision-making by bridging the gap between complex geospatial datasets and actionable insights. The pilot assessed platforms such as Copernicus Climate Change Service (C3S) and WEkEO to determine their interoperability with GenAI tools and ensure alignment with FAIR principles. Key outcomes included the development of prototype virtual assistants capable of:

- **Improving Data Discoverability** – Enabling users to efficiently find relevant datasets and services.
- **Providing Actionable Insights** – Transforming raw geospatial data into comprehensible, decision-ready information.
- **Enhancing User Engagement** – Offering plain-language responses and contextual guidance tailored to various stakeholder needs.

I.B.2. Objective 2: Development of GenAI Prototypes for Data and Service Environments

The second objective focused on developing GenAI prototypes tailored for diverse data and service environments. These prototypes demonstrated the practical capabilities of GenAI tools in enhancing data usability, accessibility, and stakeholder engagement across multiple domains, including climate resilience, health, energy, and insurance. The prototypes were designed to:

- **Improve Findability** – Enhance the discoverability of relevant data and services, particularly from key platforms like Copernicus Climate Change Service (C3S) and WEkEO.
- **Facilitate Informed Decision-Making** – Offer stakeholders actionable insights derived from structured geospatial data, contextual knowledge, and domain-specific expertise.
- **Deliver Plain-Language Responses** – Provide users with clear and comprehensible answers to complex queries, including references to trusted sources, visualizations, and associated links.

I.B.3. Objective 3: Assessment of Data Maturity and Interoperability for GenAI Integration

The third objective assessed the maturity and interoperability of existing data and service platforms to support GenAI integration. This evaluation pinpointed that data ecosystems are robust, accessible, and capable of supporting the GenAI's advanced capabilities in climate resilience and disaster management workflows. This objective involved:

- **Evaluating Data Maturity** – Assessing the readiness of platforms such as Copernicus Climate Change Service (C3S), WEkEO, and NOAA datasets based on criteria like FAIR principles, AI-readiness, and cloud optimization.
- **Enhancing Interoperability** – Identifying gaps and barriers in dataset interoperability, ontologies, and APIs across different platforms, including OGC-compliant services.
- **Addressing Challenges** – Overcoming issues related to inconsistent metadata standards, varying data formats, and limited cross-platform compatibility to facilitate seamless GenAI integration.
- **Crosswalks for Ontology Alignment** – Developing mappings between geospatial ontologies and data models to ensure consistent and interoperable data usage.

I.C. Summary of Findings and Recommendations

I.C.1. Key Findings

One of the key findings was the effectiveness of generative AI in climate data analysis. AI-powered virtual assistants were developed to help users explore data requirements, analyze climate impacts, and support decision-making. These assistants leveraged large language models (LLMs), such as Llama 3, to efficiently process structured and unstructured climate data.

A major focus was on data integration and interoperability. Participants assessed Copernicus Climate Services, advanced analysis-ready data (ARD), and integrated OGC geospatial standards. Notably, GeoLabs, IIT Bombay, and the University of Alabama collaborated to develop a shared ontology that enhances semantic interoperability between OGC standards and environmental datasets.

Several AI demonstrators were developed for specific sectoral applications:

- **Coastal Resilience** – Hartis built an AI-powered demonstrator to assess the Coastal Vulnerability Index (CVI) using geospatial and environmental datasets.
- **Drought and Heat-Related Health Risks** – Pixalytics integrated Copernicus API data to calculate drought and heat indices.
- **Flood Hazards and Mitigation** – GIS FCU developed a virtual assistant providing insights on flood risks, mitigation strategies, and response plans for Canada.

The pilot also explored AI applications in wildfire risk assessment and insurance. Xentity and NRCan conducted a state-of-the-art review on AI in wildfire risk modeling, identifying tools to assist insurers, policymakers, and affected communities in risk assessment.

Beyond data analysis, AI-powered knowledge discovery and decision support systems were also demonstrated. Dante's Knowledge Engine showcased multimodal AI search, integrating government, commercial, and social media data. TerraFrame introduced a graph-based AI model to analyze cumulative climate disaster impacts, offering insights into the indirect effects of climate hazards on infrastructure, education, and healthcare.

Several challenges were identified, including:

- **Lack of geospatial awareness in AI models**, requiring additional geocoding and ontology mapping.
- **AI hallucinations**, where generative models produced incorrect or misleading responses, especially in disaster scenarios.
- **Data interoperability issues**, requiring improvements in metadata standards and cross-platform compatibility.

To address these challenges, participants recommended utilizing **Retrieval-Augmented Generation (RAG)** and **Chain-of-Thought (CoT)** reasoning — two complementary techniques designed to enhance the quality and reliability of generated responses.

RAG models improve language models by retrieving relevant information from external sources before generating a response. This retrieval process allows the model to augment its internal training data with verified, up-to-date knowledge, resulting in responses that are more accurate, contextually appropriate, and factually reliable. Rather than relying exclusively on internalized knowledge, RAG actively integrates real-world information throughout the generation process.

In contrast, CoT reasoning is a method where the model is encouraged to break down complex problems into a sequence of coherent, logical steps, enabling more structured and interpretable reasoning.

I.C.2. Recommendations

To enhance AI applications in climate resilience and disaster management, participants proposed several key strategies aimed at improving geospatial awareness, aligning AI systems with industry standards, integrating AI into early warning frameworks, fostering collaboration, and enhancing financial risk modeling.

A critical recommendation was to **improve AI's geospatial awareness** by developing domain-specific AI training and employing graph-based knowledge representations. This approach would strengthen spatial reasoning capabilities in AI models, allowing them to better understand and process geospatial data.

Another essential aspect involved **aligning AI systems with geospatial standards**. Implementing OGC-compliant APIs and expanding semantic ontologies would enhance AI-driven geospatial analysis, ensuring consistency, interoperability, and accuracy when dealing with climate-related data.

The **integration of AI into early warning systems** was also emphasized. By operationalizing AI-driven climate risk assessments within emergency response frameworks, AI could significantly enhance preparedness and response to events such as wildfires, flooding, and droughts.

Participants further highlighted the importance of **fostering collaboration** between AI developers, climate scientists, and insurers. Strengthening these partnerships could improve predictive modeling for various climate risks and support the development of more accurate and reliable forecasting tools.

Additionally, recommendations focused on **enhancing financial risk modeling** through AI, providing valuable tools for climate-related financial risk assessment. This would be particularly beneficial for insurance companies and government agencies seeking to quantify and mitigate potential economic impacts.

The findings from this pilot are publicly accessible through the CDRP Phase 2 engineering reports and project website, along with demonstrators illustrating AI's potential in climate resilience. Future initiatives will concentrate on enhancing AI demonstrators, refining user interfaces, improving predictive modeling, and conducting thorough uncertainty analysis. Moreover, ongoing engagement with policymakers, industry leaders, and community

organizations will be essential to facilitate real-world deployment of AI solutions in climate resilience.



KEYWORDS

The following are keywords to be used by search engines and document catalogues.

Generative AI, Virtual Assistant, Geospatial Data, Climate Data, Retrieval-Augmented Generation (RAG), Natural Language Processing (NLP), FAIR Principles (Findable, Accessible, Interoperable, Reusable), OGC Standards, OGC API – Processes, Machine Learning (ML), Climate Trend Analysis, Copernicus Data Sources, ECMWF Climate Data Store (CDS), Green Deal Data Space, WEkEO Environmental Data Hub, GIS Tools, Data Search and Discovery, Multi-Modal AI, Satellite Data Processing, Prompt Engineering, Geospatial Intelligence, GPU-Powered AI Inferencing, Data Cleaning and Indexing, Training Data Markup Language (TDML), Machine Learning as a Service (MLaaS), Code Generation for Climate Data Analysis, Chain of Thought (CoT) Reasoning, AI Hallucination Mitigation, Data Visualization, Multi-LLM Model Selection, Spatial Knowledge Graphs, Environmental Monitoring



CONTRIBUTORS

All questions regarding this document should be directed either to the editor or to the contributors.

NAME	ORGANIZATION	ROLE	ORCID
Stelios Contarinis	HARTIS	Editor	https://orcid.org/0000-0002-5789-4098
Vyron Antoniou	HARTIS	Editor	https://orcid.org/0000-0002-7365-9995

NAME	ORGANIZATION	ROLE	ORCID
Loukas Katikas	HARTIS	Contributor	https:// orcid. org/ 0000- 0003- 1886- 4125
Samantha Lavender	Pixalytics	Contributor	http:// orcid. org/ 0000- 0002- 5181- 9425
Gérald Fenoy	GeoLabs	Contributor	https:// orcid. org/ 0000- 0002- 9617- 8641
Chetan Mahajan	Indian Institute of Technology Bombay	Contributor	https:// orcid. org/ 0009- 0001- 9632- 184X
Surya Durbha	Indian Institute of Technology Bombay	Contributor	https:// orcid. org/ 0000- 0003- 1022- 8378
Rajat Shinde	University of Alabama in Huntsville	Contributor	https:// orcid. org/ 0000- 0002- 9505- 6204
Nathan McEachen	TerraFrame	Contributor	https:// orcid.

NAME	ORGANIZATION	ROLE	ORCID
			org/ 0009- 0009- 7419- 4905
Matt Tricomi	Xentity	Contributor	
Joost van Ulden	Natural Resources Canada, Canada Centre for Mapping and Earth Observation	Contributor	
Micah Brachman	OGC	Editor	https:// orcid. org/ 0009- 0008- 6198- 5145
Ingo Simonis	OGC	Editor	https:// orcid. org/ 0000- 0001- 5304- 5868



1

INTRODUCTION

1.1. Pilot’s Background and Motivation

The Climate and Disaster Resilience Pilot 2024.2 emerges as a response to the escalating impacts of climate change and natural disasters, driving the need for advanced solutions that leverage cutting-edge technologies and collaborative approaches. This initiative capitalizes on the rapid advancements in **generative AI**, **Earth Observation (EO) technologies**, and geospatial platforms to address gaps in disaster management and climate resilience workflows.

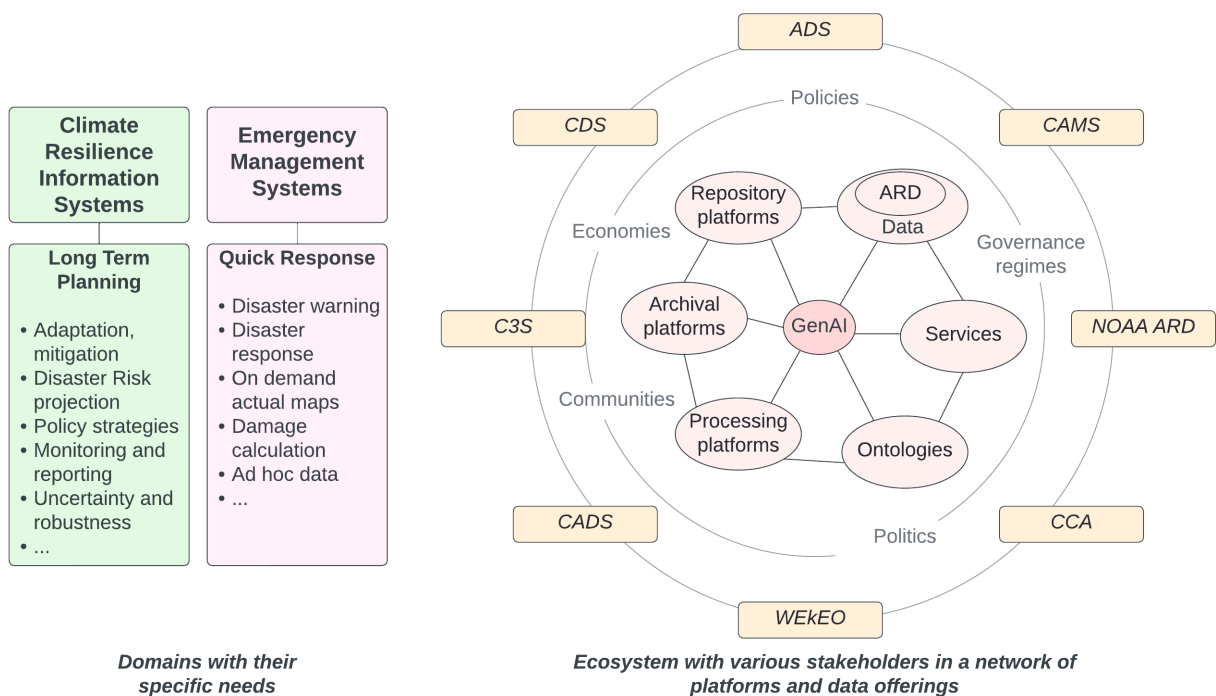


Figure 1 – CDRP2024.2 Pilot - Application Domains & Platforms Ecosystem

Building on established frameworks like the **FAIR principles** and evolving standards such as those developed by the Open Geospatial Consortium (OGC), the pilot prioritizes interoperability and accessibility across diverse datasets and services. The integration of EO technologies with generative AI offers transformative potential, enabling actionable insights derived from vast and complex geospatial data. Platforms like **Copernicus** and **WEKEO** serve as foundational resources, providing the critical data infrastructure required for meaningful analysis and innovation.

1.2. Importance of GenAI in Climate Resilience and Disaster Management

Generative Artificial Intelligence (GenAI) serves as a transformative tool in addressing the challenges posed by climate change and disaster management. Its ability to process, analyze, and generate insights from vast volumes of geospatial and Earth Observation (EO) data provides significant advantages in enhancing resilience and response capabilities.

One of the critical strengths of GenAI is its capacity to **bridge data complexity and usability**. The integration of spatial and textual information enables decision-makers to derive actionable insights from otherwise fragmented and complex datasets. This approach proves invaluable for predicting the impacts of climate-related hazards, including coastal flooding and wildfires, and for optimizing response strategies.

GenAI fosters **stakeholder engagement and accessibility**, translating complex data into plain-language insights, tailored recommendations, and intuitive visualizations. These capabilities ensure that vital information reaches not only technical experts but also policymakers, community leaders, and the general public.

Efficiency in data workflows also improves through GenAI, automating processes such as data discovery, analysis, and reporting. Integration with FAIR-compliant platforms such as Copernicus and WEkEO amplifies the potential of existing technologies, ensuring that climate and disaster resilience initiatives remain adaptive and forward-looking.

1.3. Alignment with UN SDGs and Climate Resilience Goals

The Climate and Disaster Resilience Pilot 2024.2 aligns closely with the United Nations Sustainable Development Goals (UN SDGs) and broader climate resilience objectives, reinforcing global efforts to combat the impacts of climate change and promote sustainable development.

The pilot directly supports **SDG 13: Climate Action**, focusing on improving preparedness and response to climate-related disasters. Emphasis on generative AI and geospatial data integration enhances capabilities for monitoring, predicting, and mitigating climate impacts, contributing to informed policy-making and community resilience.

Efforts also advance **SDG 11: Sustainable Cities and Communities**, fostering solutions that strengthen urban resilience against disasters. AI-driven tools empower local governments and urban planners to identify vulnerabilities, optimize resource allocation, and improve disaster management strategies.

Support for **SDG 17: Partnerships for the Goals** reflects the collaborative approach involving governments, private sector entities, academic institutions, and NGOs. This partnership-driven

strategy ensures the development of scalable, standards-compliant solutions that address the needs of multiple sectors and regions.

1.4. OGC's Role in Advancing AI Integration

The **Open Geospatial Consortium (OGC)** contributes to the integration of artificial intelligence (AI) into geospatial systems and technologies. As a global organization responsible for developing open standards for geospatial data, OGC facilitates interoperability and accessibility, establishing a framework that supports the adoption of AI-driven geospatial technologies in a wide range of applications.

OGC's work focuses on creating standards that enable AI systems to interact with geospatial platforms effectively, addressing data formats, processing workflows, and service interoperability. These efforts support AI models in utilizing geospatial data for applications such as **environmental monitoring, disaster prediction, and urban planning.**



2

CHALLENGES IN EXPLOITING GENAI WITHIN THE CLIMATE RESILIENCE DOMAIN

CHALLENGES IN EXPLOITING GENAI WITHIN THE CLIMATE RESILIENCE DOMAIN

Generative AI has immense potential for transforming geospatial applications, particularly in areas like climate resilience and disaster management. However, leveraging this potential comes with unique challenges, as GenAI systems often struggle to handle the complexities of geospatial data and domain-specific requirements. The following sections explore critical issues such as geospatial awareness, hallucinations, climate-specific contexts, and integration with geospatial standards and APIs.

2.1. Geospatial Awareness in Generative AI Systems

Generative AI systems face inherent challenges in understanding and referencing geographic locations. Unlike humans, who perceive locations through lived experiences, context, and intuitive spatial reasoning, GenAI's grasp of geospatial information is limited to the data it is trained on or linked to. This limitation raises several issues critical to the effective application of GenAI in domains requiring geospatial awareness, such as climate resilience or disaster management.

2.1.1. Text-Based Knowledge

The static nature of GenAI's text-based training data constrains its ability to interpret geospatial context dynamically. For instance, while GenAI might “know” that Athens is the capital of Greece, this knowledge is abstract, disconnected from physical coordinates, and often devoid of real-time or topographical details. The lack of temporal and spatial contextuality means that while GenAI can generate plausible text about a location, it cannot independently verify spatial relationships, boundaries, or proximity without additional data.

2.1.2. Ambiguity in Place Names

GenAI models also struggle with the inherent ambiguity in place names. Many locations share identical names (e.g., Springfield in the United States), and disambiguating them requires more than linguistic understanding—it demands access to hierarchical geospatial data or contextual user input. Without this, responses may be inconsistent or outright incorrect, particularly in areas with lesser-documented geographic features.

2.1.3. Fragmented Geospatial Data

Another challenge lies in the fragmented nature of geospatial data sources. While authoritative platforms such as Copernicus or OpenStreetMap exist, their integration with GenAI systems is

not seamless, often requiring manual effort to reconcile differing data standards and formats. This fragmentation exacerbates inconsistencies in GenAI's understanding of spatial relationships, reducing its utility for complex geospatial tasks.

2.1.4. Knowledge Gaps in Undocumented Areas

GenAI models exhibit clear limitations in understanding remote or less-documented regions. For example, a well-documented city like London may yield detailed and accurate responses, while a lesser-known location such as Eresos, Greece, may only elicit generic or partial information. This disparity can be attributed to uneven representation in training datasets, further complicating the equitable application of GenAI across diverse geographies.

2.1.5. Temporal Changes and Updates

Geospatial knowledge is not static; changes in urban development, climate impacts, or administrative boundaries can render training data obsolete. GenAI models cannot account for such temporal variations unless they are continually retrained with updated datasets or connected to live geospatial information systems. This presents a challenge for applications requiring real-time situational awareness, such as disaster response.

2.2. Hallucinations in the Geospatial Domain

Generative AI systems are prone to hallucinations—instances where they generate inaccurate, inconsistent, or entirely fabricated information. In the geospatial domain, such hallucinations can have particularly significant consequences, given the importance of precision and reliability in geographic and spatial data. These hallucinations often arise due to inherent limitations in the training data, lack of real-world validation mechanisms, and the complexity of spatial relationships.

2.2.1. Causes of Hallucinations

1. **Incomplete or Biased Training Data:** GenAI models trained on incomplete datasets may lack comprehensive information about certain locations, leading to fabricated details when queried about these areas. For instance, when asked about a remote village, the system might create plausible but inaccurate descriptions or relationships based on similar but unrelated data.
2. **Ambiguity and Overgeneralization:** The ambiguity of geographic names or features can cause hallucinations. For example, querying a GenAI model about “Springfield” could result in a conflation of characteristics from multiple locations with the same name. Also, overgeneralization occurs when the system extrapolates patterns from well-known locations to poorly documented ones, often producing erroneous outputs.

3. **Disconnected Context:** GenAI lacks inherent spatial reasoning, which can lead to inconsistencies in understanding geographic hierarchies. For example, it might erroneously describe a city as being within the boundaries of another city or region due to a misunderstanding of administrative divisions.
4. **Temporal Mismatches:** The static nature of training data means GenAI may hallucinate when describing locations that have undergone recent changes. For instance, a region affected by natural disasters or urban development might be described based on outdated data.
5. **Fabrication of Nonexistent Features:** In the absence of specific information, GenAI may invent geographic features, such as creating fictional landmarks, rivers, or infrastructure, to provide what it perceives as a complete answer.

2.2.2. Implications of Hallucinations

Hallucinated geospatial information can have serious consequences, particularly in critical areas like disaster management, urban planning, and climate resilience. Misleading data can result in ineffective or even harmful strategies, undermining the effectiveness of decision-making. Persistent inaccuracies further erode trust in AI tools, especially in high-stakes applications where precision is essential. Moreover, such errors can propagate across systems if flawed GenAI outputs are used to train other models or update databases, compounding the problem and spreading inaccuracies further.

2.2.3. Understanding the Dynamics of Climate Systems

Climate systems are complex and challenging for Generative AI (GenAI) to understand. They involve feedback loops, non-linear changes, and variations across different locations and time periods. These factors require specialized data and modeling to make accurate predictions, which general AI systems often cannot handle without significant adjustments.

1. **Feedback Loops:** GenAI systems trained on static or simplistic datasets may fail to capture the cascading effects of such feedback loops, leading to incomplete or misleading predictions. Climate systems are replete with feedback mechanisms, where outputs of a process loop back to influence the same process. For example:
 - The melting of polar ice reduces surface albedo (reflectivity), leading to greater heat absorption and accelerated warming.
 - Vegetation loss increases carbon dioxide levels, exacerbating warming and further vegetation degradation.
2. **Non-Linear Changes:** Unlike many domains where relationships are linear, climate systems exhibit non-linear changes. Small variations in one variable, such as temperature, can lead to disproportionate impacts, such as sudden shifts in weather patterns or ecosystem collapses. Capturing and predicting these non-

linear dynamics requires specialized modeling approaches and datasets that account for thresholds, tipping points, and chaotic behaviors.

3. **Spatial and Temporal Variability:** Climate impacts vary significantly across spatial (local to global) and temporal (short-term to long-term) scales. General-purpose AI systems may not be equipped to address this multi-scale variability effectively without customization.

Translating general AI capabilities to work effectively in climate resilience is not straightforward and comes with several challenges. AI models need significant customization to handle complex climate datasets like satellite imagery and weather models, often requiring specialized preprocessing. Collaboration with climate experts is essential to incorporate critical insights about unique variables, such as atmospheric or oceanic dynamics. Additionally, integrating AI with established climate models, like General Circulation Models (GCMs), is necessary but challenging due to their specialized assumptions. Lastly, the vast scale and complexity of climate data demand high computational resources, making effective adaptation both technically and resource-intensive.

1. **Customization for Climate Data:** General AI models must be extensively retrained or fine-tuned on climate-specific datasets, such as satellite imagery, weather models, and hydrological simulations. These datasets are often complex, multidimensional, and require domain-specific preprocessing steps, including reformatting, resolution alignment, and noise reduction.
2. **Domain Expertise Requirements:** Successful application of AI in climate resilience necessitates collaboration with climate scientists, hydrologists, ecologists, and other domain experts. These experts provide critical insights to ensure AI systems account for the unique variables and relationships in climate data, such as atmospheric circulation patterns or ocean heat dynamics.
3. **Integration with Existing Models:** Climate-specific AI applications often need to interface with established climate models, such as General Circulation Models (GCMs) or Regional Climate Models (RCMs). These models operate on specialized assumptions and parameters that general-purpose AI may not natively understand.
4. **Data and Computational Complexity:** Climate data is often vast in volume and requires intensive computational resources for analysis and simulation. Translating general AI capabilities to climate-specific contexts involves addressing these scale and resource challenges.

2.3. Aligning GenAI with Geospatial Standards

The integration of Generative AI (GenAI) with geospatial standards represents both an opportunity and a challenge in the climate resilience domain. Geospatial standards are essential for ensuring interoperability, data quality, and consistency across platforms and datasets, while

GenAI relies on these structured inputs to generate meaningful outputs. Misalignment between GenAI capabilities and geospatial standards can limit the potential of AI-driven solutions in geospatial applications.

1. **Complexity of Geospatial Data:** Geospatial data often involves multiple layers, formats, and coordinate systems. GenAI systems may struggle to interpret or align these data layers without adherence to standards. For example, integrating raster data from satellite imagery with vector data from cadastral maps requires strict compliance with projection and resolution standards.
2. **Lack of AI-Specific Standards:** Existing geospatial standards are not explicitly designed for AI workflows, leading to gaps in how data is prepared, annotated, or served to GenAI systems. This lack of alignment can reduce the efficiency of AI applications in geospatial analysis and decision-making.
3. **Variability Across Platforms:** Different geospatial platforms implement standards in varying ways, creating interoperability challenges for GenAI. For example, the way WMS is implemented may differ slightly between providers, complicating its integration with AI workflows.
4. **Metadata Inconsistencies:** While standards like ISO 19115 exist, metadata practices are not always consistent, especially for legacy datasets. GenAI relies heavily on well-structured metadata to contextualize geospatial information.
5. **Dynamic Data Requirements:** Real-time or near-real-time geospatial data, such as weather updates or sensor networks, often operate outside traditional standards, creating integration challenges for GenAI systems that require immediate and standardized inputs.

2.4. Integration with Application Programming Interfaces (APIs)

Connecting Generative AI systems with Application Programming Interfaces (APIs) makes them much more powerful. APIs allow GenAI to get real-time data, perform analyses, and provide accurate and relevant answers. However, working with APIs also brings some challenges that need to be carefully managed.

1. **Rate Limits and Access Restrictions:** Many APIs limit how often they can be used or how much data can be requested at a time. This can slow down GenAI in situations where fast responses are needed, like tracking disasters or weather changes. Some APIs charge fees or require subscriptions, which can be too expensive for smaller teams or organizations.
2. **Latency and Performance:** Getting data from an API takes time, especially if the data is large or if there are multiple requests. This delay can affect how quickly GenAI can provide answers, which is a problem for real-time tasks like navigation.

or emergency planning. Improving speed requires careful system design, but it can be tricky to balance performance and complexity.

3. **Data Privacy and Security:** APIs often handle sensitive data, so it's important to keep this information safe. If security measures like encryption or user authentication aren't properly set up, there's a risk of data leaks or unauthorized access. Setting up and maintaining these protections adds extra work and complexity to the system.
4. **Dependency Management:** Relying on APIs means depending on third-party services. If an API changes, limits access, or stops working, it can disrupt how GenAI functions. To avoid problems, systems need to be flexible enough to handle changes or switch to alternative APIs when needed.



3

ANALYSIS READY DATA (ARD) MATURITY REPORT (D010)

ANALYSIS READY DATA (ARD) MATURITY REPORT (D010)

The Analysis Ready Data (ARD) Maturity Report ([link](#)) evaluates the maturity of crucial ARD sources, focusing on NOAA datasets for disaster risk response and climate assessments. The report reviews three major data maturity models and develops an updated **Data Maturity Matrix** to enhance data quality, accessibility, and interoperability. It also identifies tools for automating data maturity evaluation and advancing ARD readiness.

3.1. Data Maturity Framework

The report integrates three existing data maturity assessment models:

- **Data Stewardship Maturity Matrix (DSMM)** – Focuses on data stewardship best practices.
- **CEOS Analysis Ready Data (CEOS-ARD)** – Ensures satellite data is pre-processed and analysis-ready.
- **WGISS Data Management and Stewardship Maturity Matrix (DMSMM)** – Provides a structured approach to Earth Observation (EO) data stewardship.

An updated Data Maturity Matrix is proposed, categorizing datasets into four levels:

- **L0 (Not Managed)** – No formal management practices.
- **L1 (Partially Managed)** – Basic metadata and accessibility with limited quality assurance.
- **L2 (Managed)** – Standardized, well-documented, and quality-controlled data.
- **L3 (Fully Managed)** – Optimized, validated, and FAIR-compliant data.

3.2. Tools for Self-Service Assessments

To support automated and self-service ARD maturity evaluations, the report highlights key tools, including:

- **Data Maturity Assessment Templates** (e.g., DSMM Model Template, CEOS ARD Self-Assessment Guide)

- **Compliance Test Tools** (e.g., OGC Compliance Test Suites, CF-Checker, Geospatial Metadata Validation Service)

The report emphasizes the need for a **comprehensive suite of automated tools** to streamline ARD maturity assessments. Future developments should focus on:

- **Enhancing FAIRness and AI readiness** in data.
- **Improving metadata indexing and interoperability** for disaster response applications.
- **Strengthening integration with cloud-native storage and high-performance computing frameworks.**



4

GENERATIVE AI FOR WILDFIRE REPORT (D030)

GENERATIVE AI FOR WILDFIRE REPORT (D030)

This engineering report D030 ([link](#)) builds upon the findings of Phase 1 (D-123) from **OGC Disaster and Climate Resilience Pilot III**. The primary focus is on advancing GenAI applications for wildfire risk analysis, social impact, emergency response as related to wildland fire insurance workflows, specifically in the Canadian context. The Wildland Fire (WF) community depends on robust data insights and advanced tools to bolster planning and operational decisions—augmented rather than replaced by the experiential knowledge of stakeholders.

4.1. Use Cases and Functionalities

This report outlines **key GenAI-driven use cases** relevant to wildfire resilience, response, and risk assessment. This report centers on leveraging GenAI to strengthen wildfire insurance and preparedness efforts in Canada, addressing social impact, operational efficiency, and business resilience. Specifically, the use case focus, and needed data focuses on Helping People and Business Management as it relates to Wildland Fire Insurance Stakeholders.

4.2. Data Sources and FAIR Evaluation

Phase 2 includes an inventory of over **200 Canadian wildfire-related data sources** categorized in data subject areas of Wildland Fire National Strategy & Management, National Base Data Layer Information, and Risk Indicators, Analysis, and Assessment which would be needed for GenAI Training data.

4.3. OGC Compliance and Interoperability

This report relies on the Phase 1 report basis which aligns with OGC best practices, ensuring cross-agency data integration and AI model transparency including references to OGC APIs and Data Standards, Metadata and Traceability, and AI Model Governance.

4.4. Findings and Recommendations

This report focuses on By focusing on Phase 2 priorities, combined with Phase 1 inputs, which provides a forward-looking roadmap for GenAI adoption in wildfire resilience and risk management including consideration of the following:

- **Key Wildland Fire Business Objectives for Canadian Insurance Sector:** Insurers lack granular, AI-driven wildfire risk assessment tools . Current models would benefit to leverage high-resolution geospatial and ecosystems datasets for social impact and business management. Develop GenAI-powered wildfire risk models that integrate geospatial, fuels, topography, weather, historical fire data, and predictive analytics for improved underwriting and risk-based insurance pricing.
- **Data Needs:** Generative AI requires domain-specific, structured, and unstructured wildfire datasets to enhance predictive accuracy. Over 200 Canadian wildfire-related datasets were identified, categorized, and assessed for AI readiness. Establish continuous training and labeled data improvement lifecycle to refine AI models, ensuring real-time API integrations where necessary.
- **Mapping Use Cases to Dataset Readiness and Priority:** Data gaps in Canadian wildfire analytics exist , particularly in structure materials/fuels, fuel moisture levels, and community vulnerability metrics . Prioritize high-impact AI use cases (e.g., Community Risk & Resilience Assessment, Grant & Funding Strategy Development, and Asset Risk Reduction & Loss Prevention) by expanding integration with national datasets and real-time wildfire data sources.
- **Gen AI WF Capabilities to Support Use Cases:** AI language models (LLMs) struggle with contextual wildfire decision-making without enhanced domain adaptation, retrieval-augmented generation (RAG), and multi-modal AI . Implement RAG and knowledge graph-based AI architectures to improve wildfire intelligence extraction, risk communication, and operational decision-making .
- **GenAI Roadmap Recommendations for Wildfire Insurance:** Regulatory frameworks and AI governance policies are not well established for Generative AI in wildfire insurance and risk assessment. Align AI model development with OGC interoperability standards to ensure data provenance, auditability, and regulatory compliance . Adopt OGC Training Data Markup Language (TDML-AI) to ensure traceability, validation, and ethical AI deployment in wildfire analytics.
- **Findings on Stakeholder Engagement:** AI adoption in wildfire management is hindered by organizational awareness, data silos and in cases cultural resistance . Establish cross-sector collaboration between wildfire agencies, insurance companies, and AI developers to accelerate GenAI adoption. AI pilot projects are essential for proving the effectiveness of wildfire AI applications. Conduct targeted AI prototype testing (e.g., Wildland Fire Customer Awareness tool, Predictive Risk Dashboard, and Claims Automation System) with measurable success metrics.



5

INFORMATION INTEROPERABILITY REPORT (D040)

INFORMATION INTEROPERABILITY REPORT (D040)

The **Information Interoperability Engineering Report** ([link](#)) explores methods for enhancing interoperability between different geospatial and environmental data retrieval systems. The report focuses on aligning the **OGC Environmental Data Retrieval (EDR) API** with the **Common Core Ontology (CCO)** to improve information exchange across various domains such as emergency response, environmental monitoring, and defense applications.

5.1. Objective and Approach

A key objective of the report is to resolve **semantic and syntactic heterogeneities** in geospatial data. While the OGC EDR API provides structured environmental data queries, CCO offers a formal ontology that standardizes domain-agnostic concepts, including entities, events, and roles. The report proposes a **shared ontology framework** to establish a common vocabulary that facilitates seamless data integration. This framework is developed using **OWL/RDF**, ensuring that equivalent, broader, or narrower concepts from both representations are mapped correctly.

5.2. Mapping Core Concepts

The report outlines a methodology for achieving interoperability by identifying and mapping **core concepts** between OGC EDR and CCO. For example, `edr:Time` is mapped to `cco:TimeInterval`, while `edr:Geometry` corresponds to `cco:SpatialRegion`. These mappings allow geospatial and environmental data from different sources to be queried and interpreted in a unified manner. Additionally, bridge classes are introduced where direct mappings do not exist, ensuring a flexible and scalable approach.

5.3. Use Cases and Practical Applications

Several **use cases** are presented to demonstrate the real-world benefits of this interoperability framework.

- **Semantic Query for Flood-Prone Areas:** Integrates flood-prone area data from OGC EDR API with hydrographic features in CCO.

- **Retrieval of Weather Data:** Queries temperature at a specific location and time using mapped ontology concepts.
- **Average Rainfall Over a Region:** Retrieves average rainfall for a given spatial extent over a specific period.

The report suggests further refinement of ontology mappings, enhanced **SHACL-based validation rules**, and broader adoption of this interoperability framework in various geospatial applications. By addressing interoperability challenges, the proposed approach aims to **improve data discoverability, query efficiency, and cross-domain collaboration** in geospatial information systems.



6

GENERATIVE AI VIRTUAL ASSISTANTS: DESIGN AND IMPLEMENTATION

GENERATIVE AI VIRTUAL ASSISTANTS: DESIGN AND IMPLEMENTATION

GenAI assistants for the geospatial domain can leverage Retrieval-Augmented Generation (RAG) to combine real-time data retrieval with generative capabilities, ensuring that insights are both accurate and contextually relevant. By combining RAG architecture with training and data strategies, the GenAI Virtual Assistants can deliver high-value, actionable insights across domains in climate resilience and disaster management. The section outlines the objectives, use cases, design architecture, and data strategies that underpin the pilot implementations.

6.1. Objectives and Use Cases for Virtual Assistants

The primary objective of GenAI Virtual Assistants is to bridge the gap between complex geospatial datasets and actionable insights for diverse stakeholders, including urban planners, disaster response teams, and policymakers. The use cases focus on enabling intuitive interactions with data and delivering decision-ready insights through plain-language explanations.

- **Data Discovery and Accessibility:** Assisting users in locating relevant geospatial datasets across platforms like Copernicus Climate Change Service (C3S) and WEkEO.
- **Real-Time Disaster Monitoring:** Providing up-to-date insights on hazards such as floods, wildfires, and hurricanes by integrating live data streams.
- **Policy Support:** Offering contextual recommendations for urban planning, risk mitigation, and resilience strategies based on current geospatial data.
- **Community Engagement:** Enhancing public awareness by providing easily understandable summaries and visualizations of complex environmental data.

6.2. Design and Architecture of Demonstrators

The demonstrators can utilize a RAG-based architecture to enhance performance, reliability, and context-awareness in AI outputs. Key components of the architecture include:

- **Data Retrieval Layer:** Integrates APIs, live data streams, and geospatial repositories for real-time information gathering.
- **Generative Model Core:** Utilizes Large Language Models (LLMs) fine-tuned on domain-specific datasets to generate insights and explanations.

- **Knowledge Graph Integration:** Enriches generative outputs with structured geospatial ontologies to ensure spatial and contextual accuracy.
- **Validation Pipeline:** Implements feedback loops, confidence scoring, and expert validation to reduce hallucinations and improve output reliability.
- **User Interface:** Provides an interactive, user-friendly platform that supports text-based queries, visualizations, and contextual recommendations.

6.3. Training Data: Types, Sources, and Preprocessing

The effectiveness of Virtual Assistants depends on the quality and diversity of training data. Data preprocessing and curation ensure that inputs are both representative and FAIR-compliant. Key considerations include:

- **Data Types:** Incorporate geospatial datasets (raster and vector data), climate records, sensor network outputs, and socio-economic indicators.
- **Data Sources:** Draw from authoritative platforms such as C3S, WEkEO, OpenStreetMap, and NOAA, along with community-contributed datasets.
- **Preprocessing Techniques**, such as:
 - **Data Cleaning:** Remove inconsistencies, outliers, and noise to improve data quality.
 - **Normalization:** Align data formats, projections, and coordinate systems for interoperability.
 - **Data Augmentation:** Generate synthetic datasets to simulate disaster scenarios and address data gaps in underrepresented regions.
 - **Metadata Enhancement:** Ensure datasets include sufficient metadata for contextualization and adherence to FAIR principles.
 - **Model Fine-Tuning:** Tailor pre-trained LLMs with domain-specific geospatial and climate data to improve relevance and accuracy.

6.4. Validation Practices Against Hallucinations

As discussed, GenAI systems can sometimes produce hallucinations—outputs that are inaccurate, inconsistent, or entirely fabricated. This is especially critical in geospatial and climate-related applications, where precision and reliability are paramount. The following practices can help validate AI outputs and minimize the risk of hallucinations.

1. **Cross-Referencing with Authoritative Data:** Compare AI-generated outputs with verified datasets or authoritative sources, such as government geospatial

databases, OpenStreetMap (OSM), or satellite imagery. For example, you can compare a bounding box generated by GenAI against trusted geographic platforms like Google Maps or Copernicus.

2. **Ground-Truth Datasets:** Use curated and validated datasets as training and testing benchmarks to minimize inaccuracies during model development. Train models with datasets from established organizations like NOAA or ESA for accurate climate predictions.
3. **Spatial Consistency Checks:** Verify spatial relationships in AI outputs to ensure they align with known geographic rules and structures. For instance, check that cities are not placed in oceans and that bounding boxes do not overlap invalid regions.
4. **Confidence Scoring:** Require AI models to assign confidence scores to their outputs, indicating the certainty of their predictions. Low-confidence results can then be reviewed more carefully before use.
5. **Human-in-the-Loop Validation:** Incorporate domain experts or users into the validation process to review and correct AI outputs. For example, emergency planners can verify AI-generated evacuation routes against real-world conditions during disaster management.
6. **Feedback Loops:** Implement mechanisms for users to provide feedback on errors or inaccuracies in outputs, enabling iterative improvement. Allow corrections to AI-generated maps or bounding boxes, and retrain the model with updated data.
7. **Multimodal Validation:** Use multiple data types (e.g., textual descriptions, satellite images, GIS layers) to cross-validate AI outputs. For example, flood predictions can be verified using elevation data and past flood records.
8. **Consistency with Domain Knowledge:** Ensure that outputs align with known domain-specific principles and constraints. Validate climate-related predictions against meteorological principles, such as seasonal trends or geographic climate zones.
9. **Synthetic Data Testing:** Test the model with synthetic or simulated scenarios to assess its performance in edge cases. For instance, create hypothetical locations and evaluate whether the AI can distinguish plausible outputs from impossible ones.
10. **Regular Model Updates:** Periodically retrain models with updated data to account for temporal changes, such as new developments or natural disasters. Reflect recent urban expansions or environmental changes by incorporating updated satellite imagery.
11. **Chain-of-Thought Reasoning:** Integrate Chain-of-Thought (CoT) Reasoning to guide AI through step-by-step logical processes, improving its ability to reason accurately in geospatial and climate-related queries. This method involves breaking down complex problems into a sequence of intermediate reasoning

steps, allowing the AI to explicitly justify each decision before producing an output.

The background features a dark blue field with several thin, light yellow lines intersecting at various points. Three of these intersection points are marked with small yellow dots. One dot is located in the upper right quadrant, another is further to the right and slightly lower, and the third is in the lower left quadrant. The lines create a network of triangles and other geometric shapes across the page.

7

CASE STUDIES AND PROTOTYPES

7.1. Demonstrator for Virtual AI Assistants (D100 – GeoLabs)

The D100 demonstrator chatbot can be accessed online at <https://cdrp.geolabs.fr:8503> and the presentation portal is available at <https://geolabs.github.io/CDRP/D100/>. We present three demo illustrations in the appendix Annex B.

7.1.1. Use Cases and Functionalities

The D100 virtual assistant enhances the accessibility and usability of various geospatial and climate-related data sources. It serves as an interactive tool that enables users to efficiently search, retrieve, and explore relevant datasets. The assistant's core functionalities include:

- **Data Discovery & Search:** Users can query the assistant to find relevant datasets from multiple data sources, improving data findability.
- **Retrieval-Augmented Generation (RAG):** The chatbot provides responses by fetching and indexing relevant documents, ensuring context-aware answers.
- **Web Search-Based Responses:** The assistant is capable of conducting a web search to fetch information from online sources.
- **LLM Model Selection:** Users can choose from various large language models to generate responses, optimizing results based on specific use cases.
- **FAIR Data Principles Compliance:** Developed based on aligning with findable, accessible, interoperable, and reusable (FAIR) principles.
- **OGC Standard Integration:** Designed with an aim to ensure compliance and interoperability with geospatial data processing workflows and existing OGC API implementations.

7.1.2. Data Sources and FAIR Evaluation

The virtual assistant is developed using multiple data sources and evaluated based on the FAIR principles—ensuring that data is Findable, Accessible, Interoperable, and Reproducible.

7.1.2.1. Data Sources

The assistant leverages the following key data sources:

- Copernicus Data Sources
- ECMWF Climate Data Store (CDS)
- Green Deal Data Space
- Copernicus Atmosphere Monitoring Service (CAMS)
- WEkEO Environmental Data Hub
- [GitHub Repository of WEkEO](#)
- [GitHub Repository of ECMWF](#)

7.1.2.2. Evaluation

The assistant is developed using publicly available data sources indexed in a structured database. It is built using free and open-source tools and powered by Large Language Models (LLMs). The system prioritizes fairness, accessibility, interoperability, and reproducibility to ensure a robust and ethical AI framework.

7.1.3. OGC Compliance and Interoperability

The assistant builds upon previous efforts in OGC initiatives, including:

- OGC CDRP.1 (Landslide Demonstrator)
- OGC Testbed 19 – Machine Learning Activity (Inferencing-as-a-Service)

These efforts explored integrating machine learning (ML) inferencing as OGC API Processes, particularly for landslide detection. The workflows were developed using the OGC Training Data Markup Language (TrainingDML) and OGC API Processes (Parts 1 and 2). The virtual assistant acts as a user-interaction layer for these implementations, providing an end-to-end Generative AI stack for scientific applications while ensuring seamless interoperability.

7.1.4. Findings and Recommendations

The D100 virtual assistant improves data search and provides an interactive way to explore the datasets mentioned above.

The figure below illustrates the high-level workflow of the assistant:

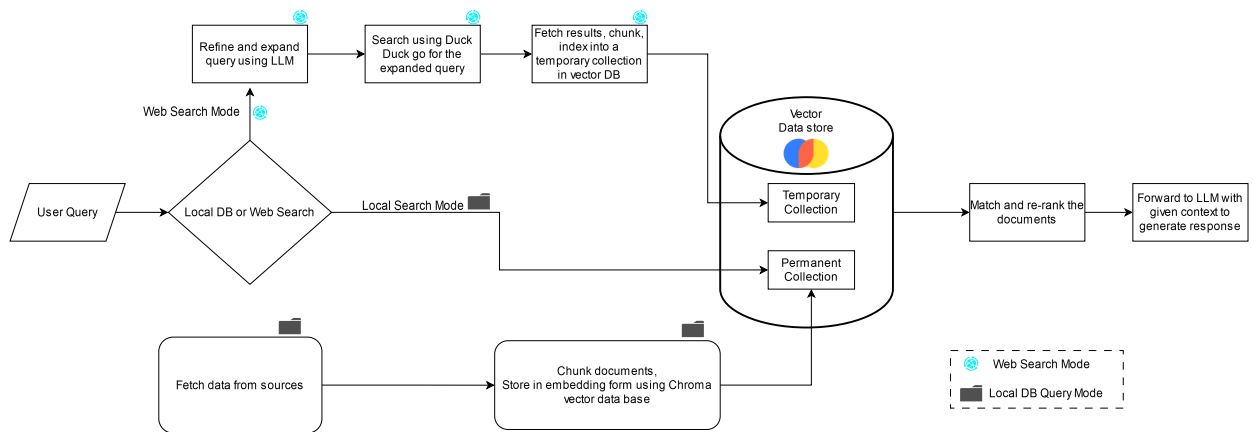


Figure 2 – D100 (GeoLabs) - High Level Workflow

Users interact with the assistant through a chatbot interface, where queries are processed using one of the following approaches:

Approach 1: Retrieval-Augmented Generation (RAG) using a Local Database **Fetch data from various sources** Index the data in a Chroma vector database **Match stored documents with the user query** Retrieve the top 5 matching documents ** Generate a response using an LLM based on the retrieved data

Approach 2: Web Search-Based Responses **Conduct a web search for the user's query** Fetch relevant data and index it in a temporary vector database **Match the query with indexed documents** Retrieve the top 5 documents ** Generate a response using an LLM based on the retrieved data

Users can select from the following LLM models for response generation:

- [phi4:latest](#)
- [llama 3.2:latest](#)
- [llama 3.1:8b](#)
- [llama3:latest](#)
- [deepseek-r1:1.5b](#)

Key Learnings:

- Data indexed in markdown format improves LLM processing and comprehension. In future, standardization for LLM friendly input and output data formats can be researched.
- High-quality data cleaning and indexing enhance search efficiency and response accuracy.
- The BM25 algorithm efficiently re-ranks search results for improved document retrieval.
- GPU-powered LLM inferencing significantly reduces response generation time.

- Effective prompt engineering on user queries enhances query relevance and response robustness.

7.1.5. Future Directions

To further enhance the assistant, we plan to:

- Evaluate Chain of Thought (CoT) reasoning for improved query understanding.
- Assess and mitigate hallucinations in generated responses.
- Implement dynamic validation mechanisms to ensure response reliability.
- Refine guardrails for LLM responses for safe and reliable output generation.
- Extend support for additional OGC standards, including:
 - Training Data Markup Language (TDML)
 - OGC API Processes Part 1: Core
 - OGC API – Processes – Part 2: Deploy, Replace, Undeploy
- Investigate the feasibility of integrating multi-modal capabilities, such as image and satellite data processing.
- Enhance documentation and user guidance to improve adoption and usability of the virtual assistant.

7.2. Demonstrator for Virtual AI Assistants (D110 – CCME0):

The proposed use case for embedding a generative AI chatbot on GEO.ca focuses on enhancing the search and discovery of FAIR and Open geospatial data through conversational AI. Although CCME0 was assigned to support this deliverable, we committed only to providing in-kind contributions rather than developing a full solution. Since then, however, prototype development has been initiated to explore the feasibility of this solution. It aims to support diverse user groups such as researchers, policymakers, educators, and citizens by providing intuitive natural language query capabilities and personalized data discovery features. The plan includes functionalities like contextual search, integration with GEO.ca's APIs, and enhanced user interaction through visualization tools. Since the proposal, development of a prototype has begun, including efforts to experiment with OGC GeoPackage visualization on a map client. These activities align with Canada's commitments to accessibility, reconciliation, and compliance with geospatial standards.

7.2.1. Use Cases and Functionalities

The generative AI chatbot use case on GEO.ca is designed to enhance the search and discovery of FAIR and Open geospatial data through a conversational interface. It supports various user scenarios, including researchers seeking climate change impact data for analysis, policymakers needing land use data to inform urban planning, educators accessing open data for classroom demonstrations, and citizens viewing historical temperature trends to understand and advocate for climate action.

The chatbot's primary functionalities include natural language query processing, contextual, geospatial data-driven search, personalized responses based on user needs, and integration with existing GEO.ca APIs for seamless access to data.

Additionally, outside of the provided use case, we explored data visualization capabilities in a subsequent ongoing proof-of-concept development. This involved investigating how the LLM could be augmented with tools to support visualizing data resources (e.g., GeoPackage) associated with the generated results.

7.2.2. Data Sources and FAIR Evaluation

The AI chatbot leverages data repositories from GEO.ca, adhering to FAIR principles. The data is made findable through clear metadata and search functionalities that enhance discoverability. Accessibility is ensured by providing open access to data via CCMEQ's APIs and services. The data is interoperable, formatted to integrate with various GIS tools following OGC standards. Additionally, datasets are reusable as they comply with licensing and metadata standards such as ISO-19115.

The chatbot intends to further promote these principles by offering guided assistance on accessing and interpreting data. When search results are returned within the chatbot, more information about each dataset is displayed, including links to associated data and metadata resources. Contextual record searching ensures that when users request more details about a specific result, the LLM focuses solely on the metadata for that record. This focused approach helps to streamline user interactions by presenting only relevant information for a specific record, reducing confusion caused by unrelated search results and reinforcing the authoritativeness of the information and data returned.

7.2.3. OGC Compliance and Interoperability

The planned integration of the chatbot aims to align with OGC (Open Geospatial Consortium) standards to support key objectives. First, interoperability is prioritized by ensuring that the chatbot interfaces with GEO.ca's systems using OGC-compliant APIs, thereby adhering to the Canadian Geospatial Data Infrastructure (CGDI) and Canada's Standard on Geospatial. Additionally, the chatbot is designed to leverage OGC standards, including OGC API – Features and OGC API – Records, to handle geospatial queries effectively. As part of the development, we have also begun experimenting with the GeoPackage standard to improve data handling and visualization capabilities. Although the solution remains in the prototyping stage, ongoing

efforts are focused on refining these functionalities. Finally, accessibility and inclusivity are central considerations, with efforts to train the chatbot's NLP on multilingual and region-specific geospatial terms, respecting Canada's official languages, commitments to reconciliation with Indigenous Peoples, and the goals outlined in the Accessibility Action Plan.

7.2.4. Findings and Recommendations

Findings: The chatbot was not fully implemented, so there are no direct findings from its deployment on GEO.ca. However, experimentation with a generative AI chatbot currently under development revealed several key challenges. These include hallucinations, where responses contain information not based on authoritative sources, and low confidence in results, particularly when data is not retrieved from trusted APIs like those provided by CCMEQ. Additionally, infrastructure costs for training, fine-tuning, and operating such models pose a significant barrier, particularly during initial deployment and whenever new data is added to the catalogue or data repositories.

Recommendations: To address these challenges, responses generated by the chatbot prototype should clearly indicate when information is not sourced from authoritative APIs, particularly those not provided by CCMEQ. Ongoing research and experimentation will focus on mitigation of hallucinations and improvements to model reliability. Optimizing the infrastructure for scalability and cost-efficiency is necessary to ensure the sustainability of the system as usage grows, especially with periodic updates to data catalogs and repositories. Additionally, pilot testing will help validate the prototype's performance and provide opportunities to gather user feedback before moving toward full implementation.

7.3. Demonstrator for Virtual AI Assistants (D110 – Danti)

Danti's demonstrator (<https://gov.danti.ai/log-in> contact gov-suppot@danti.ai for account approval) focuses on leveraging generative AI and spatial intelligence to improve data discovery, accessibility, and analysis. The objective is to develop a knowledge engine that enables users of all skill levels to interact with multimodal geospatial data. Dante.ai integrates large language models (LLMs), geospatial analytics, and real-time data sources to provide actionable insights for decision-makers. The system is designed to support both government and commercial sectors, offering enhanced capabilities for **geospatial intelligence, disaster response, and climate risk assessment**.



Figure 3 – D110 (Danti) - Logical Architecture Diagram

7.3.1. Use Cases and Functionalities

The demonstrator showcases its capabilities through real-world geospatial analysis scenarios. Key functionalities include:

- **Intelligent Data Discovery:** Users can search for locations (e.g., wildfire-prone areas in California) and receive AI-generated summaries, relevant datasets, and geospatial insights.
- **Multi-Source Integration:** The system aggregates data from government, commercial, and social media sources, including satellite imagery, climate data, and news feeds.
- **Geospatial Awareness & Querying:** AI models enable context-aware retrieval of geospatial data, refining **search results based on location, time, and data relevance.
- **Decision-Support Analytics:** Provides visual representations, mapping insights, and predictive analytics to support policy-making and emergency response.
- **Automated Alerts & Monitoring:** Users can save searches and receive notifications when new data is available for their areas of interest.

7.3.2. Data Sources and FAIR Evaluation

The demonstrator integrates various data sources, including:

- **Government Imagery:** Optical, SAR, and fire event data (e.g., NOAA, NASA, NGA).
- **Commercial Earth Observation Data:** Providers such as Planet Labs.

- **Social Media & News Data:** Real-time event tracking from open sources.

7.3.3. OGC Compliance and Interoperability

The Danti demonstrator currently indexes multiple OGC and NSG compliant data formats including GeoTIFF and NTF. Future plans include investigating the implementation of OGC API-Records to streamline and standardize metadata management, as well as OGC-API Processes to enable cross-platform analytical capabilities.

7.3.4. Findings and Recommendations

The demonstrator evaluated different methods for enhancing AI-driven geospatial awareness in LLMs, including:

- **Retrieval-Augmented Generation (RAG):** AI-enhanced information retrieval from multiple sources.
- **Knowledge Graph Integration:** AI-driven connections between related datasets and geospatial queries.
- **AI-Powered Query Generation:** Allowing LLMs to generate structured queries for geospatial data retrieval.

Key findings indicate that integrating AI-driven search with knowledge graphs significantly improves geospatial data interpretation and usability.

7.4. Demonstrator for Virtual AI Assistants (D120 – CRIM):

CRIM's demonstrator (<https://ogc-demo.crim.ca/>) leverages generative AI to enable a virtual assistant capable of interacting with geospatial data, using maps to drive conversations and provide insights, moving away from structured APIs to more intuitive map-based queries. A prototype was developed to test this concept, focusing on flood risk maps as a use case.

7.4.1. Use Cases and Functionalities

The demonstrator explores scenarios where AI assists users in interpreting flood risk maps and other geospatial overlays. Users can interact with the system by entering prompts such as specific addresses to determine flood risks. Functionalities include:

- Querying geospatial data overlays to answer location-specific risk questions.

- Adjusting map opacity to enhance model explanations.
- Intercepting queries for custom tools, such as geocoding locations and retrieving related map images.

Challenges include improving model reliability and minimizing hallucinations, such as misinterpreting map features or geographic contexts.

7.4.2. Data collection tool

The Canadian flooding imagery data collector allows extraction of data by coordinates and zoom level. These data are extracted from the Flood Susceptibility Index 2015 map layer hosted on geo.ca then overlaid on the street map layer from [ArcGISOnline](https://arcgis.com/arcgisonline) with a specified opacity level.



Figure 4 – D120 (CRIM) - Street Map

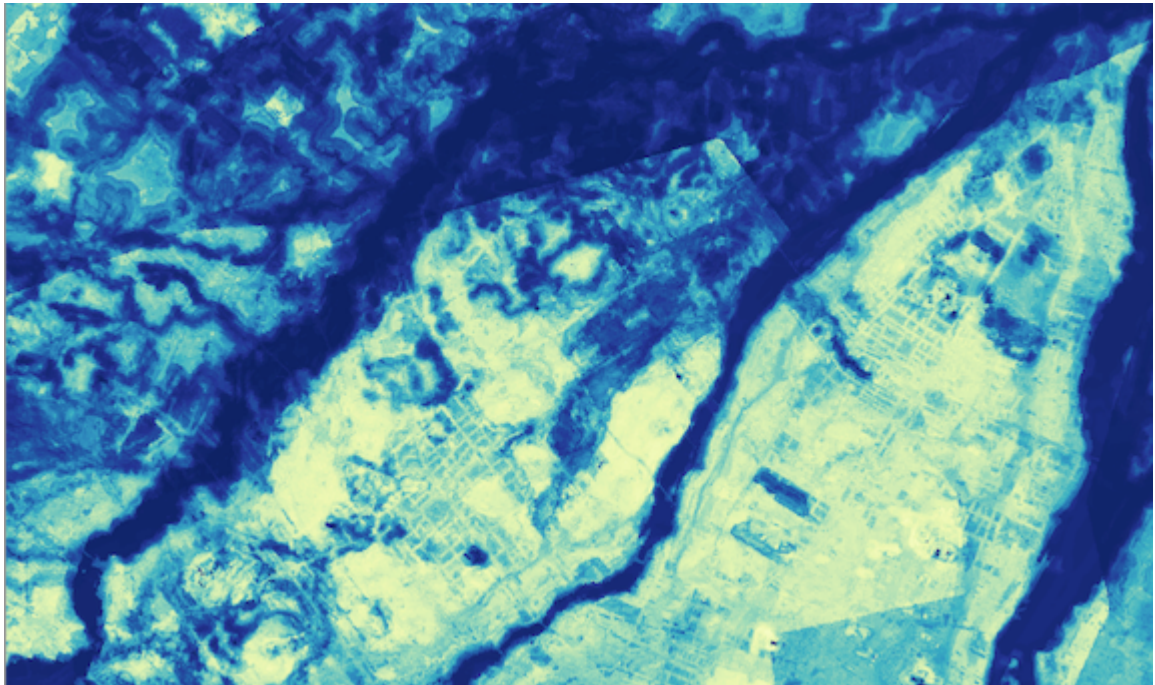


Figure 5 – D120 (CRIM) - Flooding Map

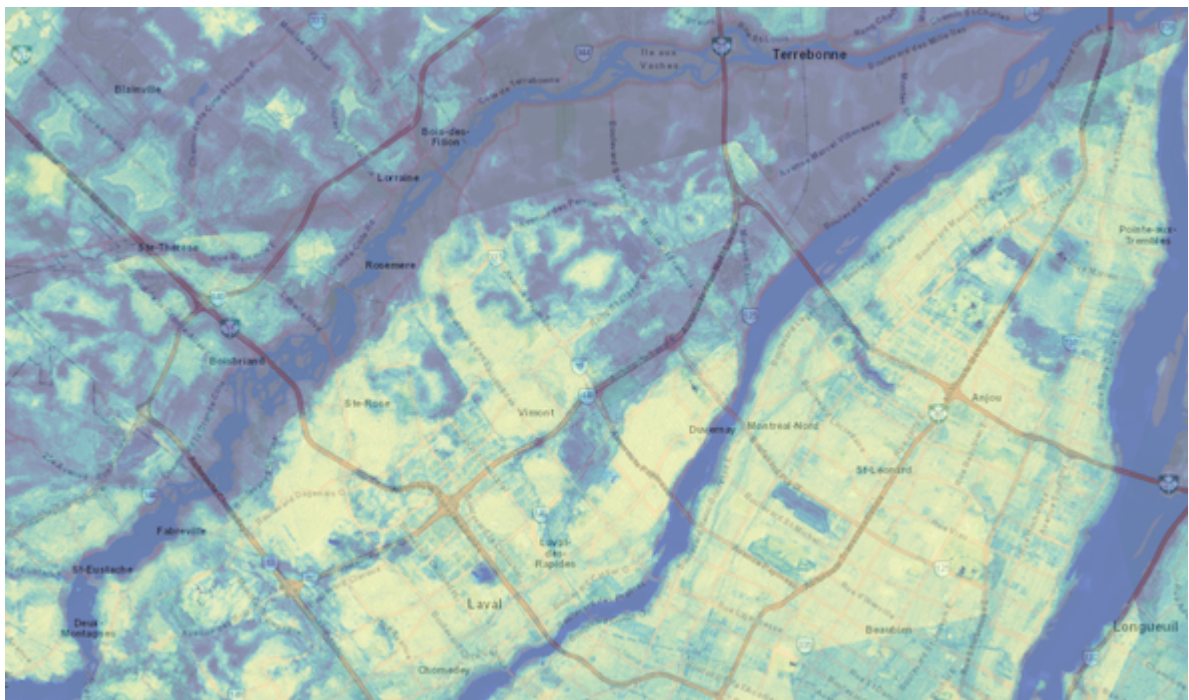


Figure 6 – D120 (CRIM) - Combined Map

1. Street map: This is the base image accessed from the coordinates provided. Because the street map is made from indexed tiles, the tool must determine which tile is needed to construct a landscape user-friendly view. Based on the tile from which the coordinate is found in, it will form a 2:1 ratio image. Since each tile is 256×256 pixels it forms a 1024×512 pixels image, this format is

important for the Flood image. Finally a pre-determined zoom is set to have enough information in the image.

2. Flooding susceptibility image: This image is extracted from Geo.ca flooding-susceptibility map based on a coordinate, zoom and the image width and height in pixels, which is the values of the previous street map.
3. Combine overlapped image: This is the last step in which the flooding susceptibility image is overlapped on the street map with a specify opacity, a balance between losing saturation of the image and the readability of street and neighborhood names.

Central to this tool is the Processing Pipeline. This pipeline consists of a series of operations that transform user input into actionable data. Within this pipeline, four key tasks are performed:

1. Identifying the location from the user message.
2. Geocoding the identified location.
3. Fetching flood imagery corresponding to the location.
4. Analyzing the flood risk based on the gathered data.

These steps are illustrated in the following architecture diagram.

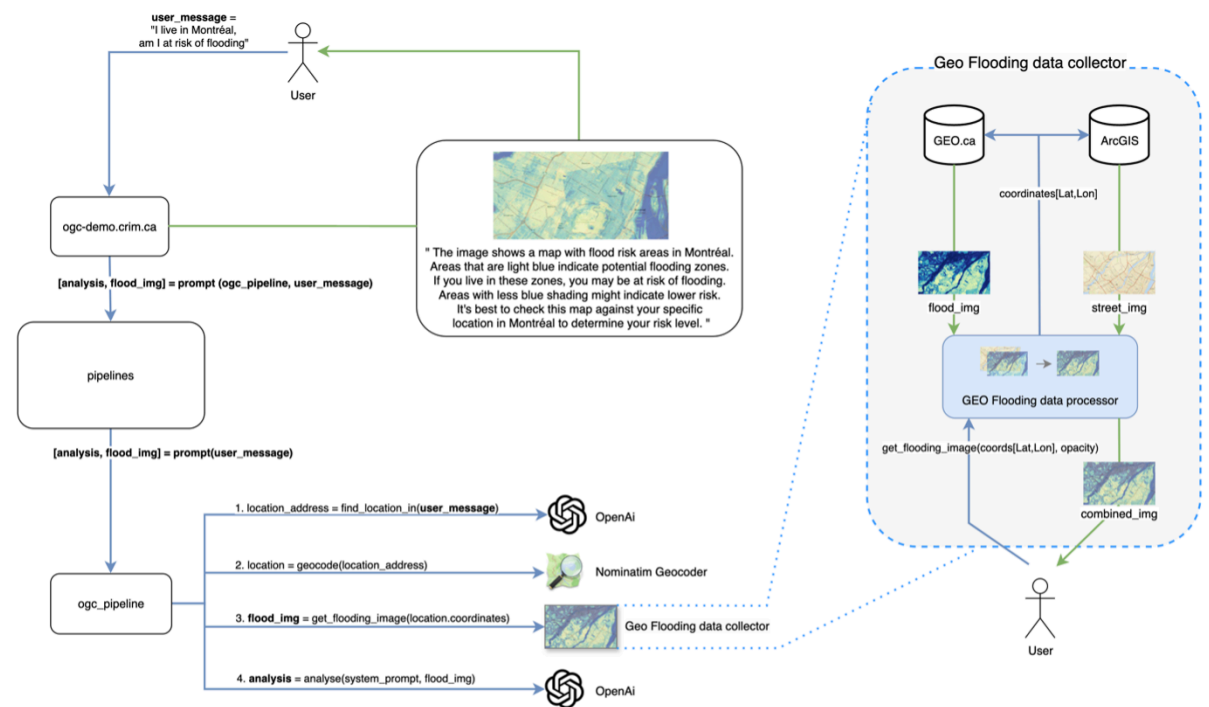


Figure 7 – D120 (CRIM) - Architecture Diagram

The Location Processing begins by finding the location within the user message using a first LLM call. Once identified, the address is fed into the Geocoding Service using the Nominatim Geocoder, which converts the address into precise geographic coordinates. Following geocoding,

the Flood Image Retrieval step leverages a Geo Flooding Data Processor. This component accesses flood susceptibility imagery based on the coordinates obtained, providing visual data that is crucial for the analysis. The retrieved flood image is then added as context to the original user prompt. Grounding the response generation on this image helps limit hallucinations and allows the user to verify the claims made. This analysis yields insights into flood risks, considering the specific conditions of the user's location. The result is communicated back to the user, enhancing their understanding of potential flood threats in their area.

An example of the use of this agent is shown in the figure below.



Figure 8 – D120 (CRIM) - Agent Screenshot

Canadian GEO flooding susceptibility agent

Using the previous tool, we can combine it with an AI model to build an agent which will be able, from a user natural language format input, extract location coordinates information to collect the flooding image which will be sent to the AI model for flooding analysis. In our case, we built this system using gpt4o and gpt4o-mini. The User Input Module is where the user provides a message indicating their location, such as “I live in Montréal, am I at risk of flooding”. This message serves as the input for the subsequent processing steps. The System Interface, represented by “ogc-demo.crim.ca” acts as an intermediary, taking the user message and preparing it for analysis. This step involves creating a prompt that combines the user message with the chosen processing pipeline, initiating the data flow through the system.

7.4.3. Experiments and Findings

As part of the project, we conducted several experiments to assess the reliability of responses provided by such an assistant. We decided to evaluate the performance of the system in a Visual Question Answering (VQA) setting. This task involves asking a model to answer questions based on an image. While this task is relatively classic and often associated with vision tasks with a limited set of answers, the recent arrival of Large Language Models (LLMs) and, more importantly, Large Vision Models (LVMs)—which integrate a visual component—has renewed its approach.

Existing publicly available datasets VQA datasets generally do not include maps, even less ones relevant to flooding. For reference, MAPQA (<https://arxiv.org/abs/2211.08545>) only contained very general maps, and Floodnet (<https://arxiv.org/abs/2012.02951>) focused on satellite images showing the presence of floods but without precise localization. We decided to build a dataset that would more closely align with the interests of OGC and the sponsors for this pilot. We think that map-understanding or explanation through VLM is an interesting research avenue that has a lot of potential of impact in the geospatial domain due to the universality of images as a format. More specifically, allowing location-specific queries grounded in trusted sources of data on flooding and other related events seems to be of great interest for both governments and the insurance industry.

Dataset Creation

The dataset was built in a semi-automatic way using GPT-4, followed by human validation. It includes two types of questions:

- **Categorical questions** expecting answers such as “yes”, “no”, or “partially”, accompanied by an explanation.
- **Open-ended questions** created manually, requiring an answer that specifies an area on a map.

The process involved using flood risk maps with a color scale ranging from light yellow to dark blue, where dark blue indicates high flood risk areas. The Flood Susceptibility Index 2015 (<https://geo.ca/flood-mapping/flood-map-gallery/flood-susceptibility-index/>) was used as a data source for this project. The index exposed as a map layer in geo.ca showcases the risk of flooding across Canada. We took screenshots of different areas of Québec (9 different images) as the following:

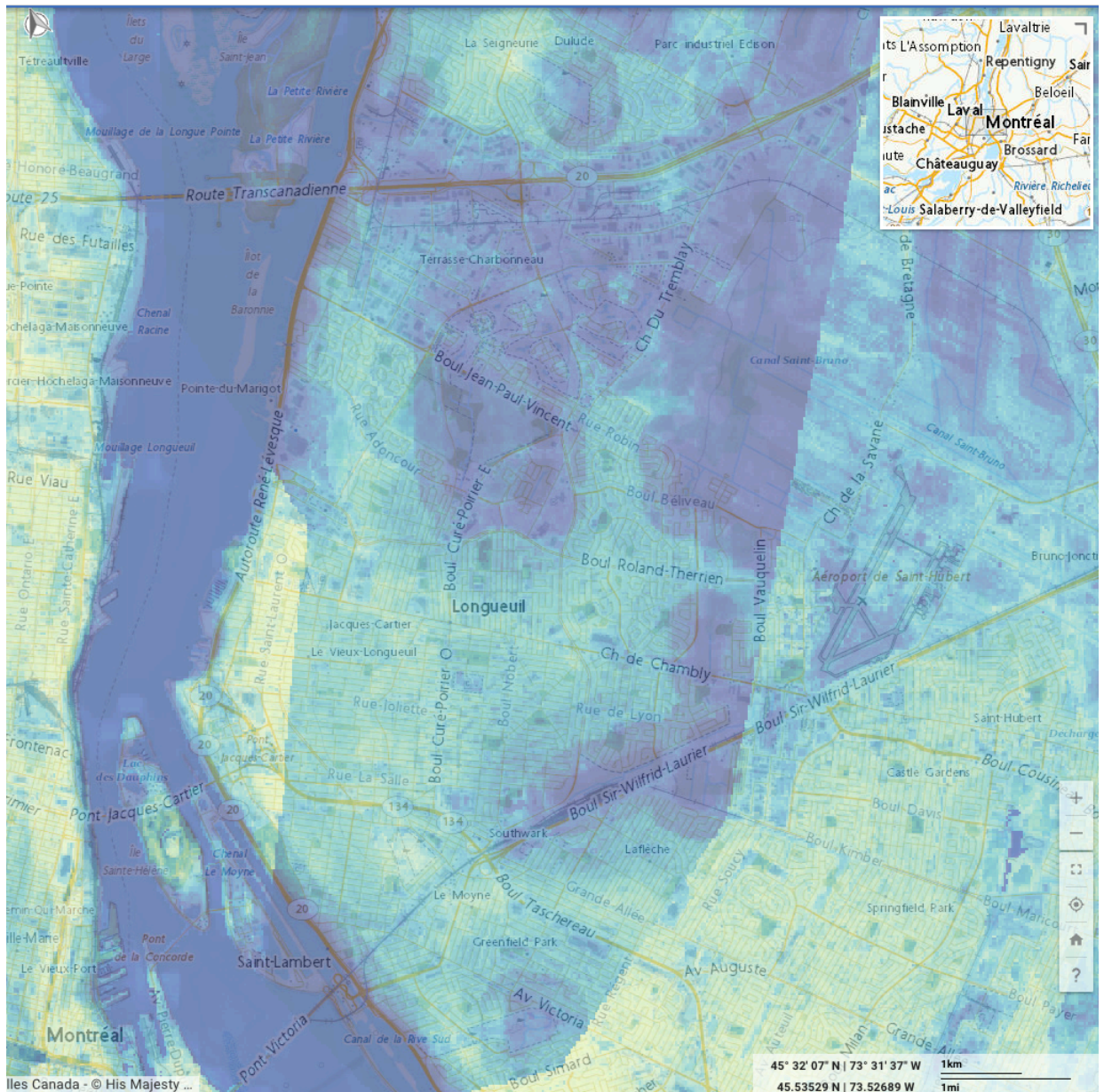


Figure 9 – D120 (CRIM) - Longueuil

The goal was to generate precise questions about each image using targeted prompts, forcing the model to identify specific street names, intersections, or neighborhoods with unambiguous zone selections (to avoid referencing areas that include both flood-prone and non-flood-prone sections).

We then asked the model to generate answers based on the questions and corresponding images. These answers were manually validated by two annotators.

Dataset Overview

In this modest-sized dataset (104 occurrences), we can distinguish two types of question-answer pairs:

- **Categorical Questions** (95 Q&A pairs)

Expected answers: “yes”, “no”, or “partially”, accompanied by an explanation.

Question : *Est-ce que les zones près de la rivière des Outaouais sont plus sujettes aux inondations que celles éloignées de la rivière ?*

Answer: *Oui.*

Explanation: *les zones proches de la rivière des Outaouais présentent des teintes bleues plus sombres, ce qui indique un risque plus élevé d'inondation par rapport aux zones éloignées*

Question : *Est-ce que les zones proches du lac Leamy sont indiquées comme étant à risque élevé d'inondation sur cette carte ?*

Answer:** *Oui.*

Explanation: *Lower Town montre un niveau de risque plus élevé (couleur bleue plus intense) par rapport à Rockcliffe Park, qui est situé en hauteur et semble moins vulnérable.*

- **Open-Ended Questions**(9 Q&A pairs)

Expected answers: precise locations on the map.

Question : *Quelles sont les zones principales de Montréal les plus à risque d'inondation (en bleu foncé) ?*

Answer: *Collège Reine-Marie entre la 125 et Boulevard Saint-Michel, entre Parc Extension et Outremont, pied du Mont Royal au nord-ouest*

Challenges: Model Hallucination

One of the primary challenges in creating the dataset was model hallucination, both in question generation and answers. In this context, hallucination is defined as:

- Mentioning a location that does not exist on the provided map (e.g., referencing an island that isn't there).
- Incorrectly localizing a place, such as stating a zone is in the southeast when it is actually in the north.
- Using vague references requiring implicit knowledge, such as “the industrial area” or “the largest infrastructures”.

For categorical questions, we identified:

- 5 cases of hallucinations in answers
- 19 cases of hallucinations in question generation

Additionally, inconsistencies were noted between the explanations and short answers (yes, no, partially). In these cases, the explanation should have led to a different response than the one provided, with this phenomenon occurring 5 times in closed questions.

Thanks to manual validation performed after this first production, questions containing exclusively geographic hallucinations (places not present on the map) were removed from the experiments, leaving 82 valid occurrences. The 5 cases of hallucinations in the answers were not an issue since we used manual validation, considered the ground truth, to evaluate the results.

Experiments Conducted

We conducted three series of experiments to assess the assistant's answer quality. We used exclusively OpenAI's GPT-4-o model for this purpose.

1. Evaluation of Generation Techniques for Categorical Questions

We applied different generation techniques to answer the dataset's questions and evaluated these responses against the manually validated ground truth. The image associated with the question was always provided to the assistant.

The generation techniques used are as follows:

- **Direct Generation:** Directly predicting the response using the question and the associated image as input.
- **CoT (Chain of Thought):** Generating a response step by step based on the instruction "Let's think step by step."
- **Structured Generation:** Generating responses while ensuring adherence to a predefined response format described by a JSON schema.
- **Structured Generation with CoT:** Combining structured generation with reasoning; an initial reasoning step allows the model to process the problem within the token space, potentially improving classification.
- **Structured Generation with Fallback:** If direct generation does not follow the expected format, an attempt is made to extract the response from the generated text while enforcing the structure.
- **Structured Generation with Fallback and CoT:** Combination of the three previous techniques. The evaluation metrics used are precision, recall, and F1-score.

The results are as follows:

7.4.4. Prediction report for categorical questions

Table 1

PREDICTIC MODE	PRECISION (YES)	PRECISION (NO)	PRECISION (IN PART)	RECALL (YES)	RECALL (NO)	RECALL (IN PART)	F1 (YES)	F1 (NO)	F1 (IN PART)
Direct Generation	NA	NA	NA	NA	NA	NA	NA	NA	NA
Chain-of- Thought	NA	NA	NA	NA	NA	NA	NA	NA	NA
Structured Generation	63	57	29	73	16	54	67	25	38
Structured Generation 56 with CoT	56	71	17	73	20	23	63	31	19
Structured Generation 64 Fallback	64	75	33	80	36	38	71	49	36
Structured Generation Fallback + CoT	69	73	44	91	44	31	78	55	36

Note : NA signifie “Non Applicable”.

We observed that the combination of the three techniques yields the best results. Although the results are encouraging, they vary significantly depending on the type of response: the “yes” response category scores significantly higher than the others.

We also identified cases where the model refuses to generate a response, negatively affecting the scores. Further analysis would be needed to understand the conditions leading to these refusals.

2. LLM as a Judge for Open-Ended Questions

We tested the use of a language model (LLM) as a judge to evaluate open-ended questions. Although our dataset is limited and not statistically significant, this approach was considered for future experiments with an expanded dataset. The LLM was used to verify: — The factual accuracy of a statement based only on the facts present in the provided map. — Whether the response contains knowledge explicitly present in the provided map or relies on the model’s implicit knowledge.

This method compares the LLM’s evaluation as a judge for automatically generated and manually generated responses. If the method is effective, the LLM should judge manually generated responses as 100% correct.

However, in our experiment, the LLM judged its own generated responses as factually correct 74% of the time, compared to 62% for manually generated responses. For explicit/implicit knowledge classification, it evaluated its own responses as explicit 80% of the time, compared to 77% for manual responses.

These results are not yet conclusive, given the 38% gap for manual responses (compared to the expected 100%), indicating that further experiments are needed, and probably an optimization of the LLM as a judge query.

3. Exploration of Implicit Knowledge

In the same spirit, we conducted a small exploratory experiment to assess the actual use of the flood layer in the image by the model. Specifically, this involved randomly adding an artificial blue layer to one part of the map to observe whether the model produced different responses based on this addition.

The model should theoretically adjust its responses based on the blue layer. If, however, the responses remain identical, this would indicate that the model relies more on its implicit knowledge of the location rather than the actual flood layer present in the image.

Similarly, one could ask whether the presence of a river on the map automatically leads the model to identify flood-prone areas near the watercourse, even in high-altitude areas where flooding would be unlikely.

This experiment was conducted exploratorily, without definitive conclusions. However, we observed variations in responses depending on the location of the overlay, which seems promising but requires further large-scale investigation.

Final Insights on the Integration of Generative AI for Flood Risk Map Analysis

At the end of the project, the creation of the assistant, the dataset, and the execution of various experiments allowed us to highlight key points regarding the integration of generative AI to support flood risk map analysis.

- **Targeted Prompting for Specific Applications**

When the assistant is applied to a narrow domain (e.g., flood map analysis in specific areas), using structured prompting techniques (Chain of Thought, JSON) results in more reliable and easier-to-validate responses. Additionally, fallback mechanisms prove essential: when the model's output does not conform to the expected format, it is possible to re-query the model or post-process its response to enforce a predefined schema (e.g., a JSON format). This avoids many inconsistencies and forces the model to clarify its information.

Conversely, for more general applications, the model's performance decreases if its reasoning is not carefully structured, as it becomes more likely to drift towards approximate associations or irrelevant generalizations.

- **Access to Geospatial Tools**

Integrating specialized APIs (geocoding services, on-the-fly map retrieval, detailed cartographic databases) significantly improves the quality and accuracy of responses. By having access to updated and localized information, the model is less likely to fabricate non-existent locations or overlook geographic specifics. These tools provide a factual grounding that can be cross-referenced with the model's internal knowledge, reducing hallucination risks.

- **Difficulty in Building a Generalist System**

Each map has its own characteristics (legend, scale, symbolization style, colors, etc.). In this context, relying solely on general knowledge (e.g., that a river implies a flood risk) can lead to inaccurate generalizations. Furthermore, a model trained on specific maps may struggle to extrapolate to a different type of map with a very different style (e.g., a completely different color palette or the use of non-standard icons). This highlights the need for additional research and adaptation efforts, so that the model truly takes into account the legend and visual specifics of each cartographic source.

- **Guardrails and Verification Strategies**

To reduce hallucinations, several strategies can be implemented:

- **Grounding Verification:** Systematically comparing the model's output against the map or a certified database.
- **Self-Consistency:** Generating multiple responses and keeping the most frequent or the most well-argued one.
- **API Utilization:** Calling an external service (e.g., a geocoder) to verify that a mentioned location actually exists and is in the claimed area.
- **Automatic Disclaimers:** Clearly indicating when the required information is not found on the map, or when responses rely on unverified assumptions.

These guardrails would be essential complementary mechanisms, as they help mitigate the model's automatic biases and misinterpretations.

7.4.5. Future Work and Conclusion

Overall, this experience highlights the value of integrating generative AI in flood risk map analysis. The results obtained with our prototype are promising, although this integration requires particular attention in prompt design, response validation, and hallucination management. The experiments conducted open the door to further research aimed at improving the reliability and generalizability of such a system.

We therefore propose the following research directions to continue this work:

Expanding the Dataset

- **Increasing map diversity:** Including different styles, legends, and geographic areas.
- **Strengthening open-ended questions:** Generating more scenarios where the answer must specify precise areas, to test the model's ability to pinpoint the correct map segments.
- **Reducing geographic hallucinations:** Refining prompts and generation techniques to avoid fabricated or vague locations.

Impact of Transparency and Layer Styles

- **Flood zone opacity:** Studying whether visibility (more or less opaque shades) affects the model's ability to detect high-risk areas.
- **Other graphical variations:** Testing how symbolization (icons, non-standard colors) influences response accuracy.

Automatic Legend Retrieval and Analysis

- **Automated extraction:** Enabling the model to identify and read the legend directly from the image, rather than relying solely on general knowledge.
- **Dynamic adaptation:** Adjusting explanations based on color codes detected in the legend.

Detection of “No Answer” Cases

- **Handling insufficient data:** Investigating how to encourage the model to respond “I don't know” when the map does not allow for a conclusion.
- **Managing uncertainties:** Proposing a confidence level in the response, or even a mention of missing information.

Fine-Tuning and Specialized Training

- **Training on cartographic images:** Fine-tuning a visual or multimodal model (LVM) on a broader and more varied set of maps, with precise annotations (e.g., flood zone masks).
- **Integrating elevation data:** Adding altimetry data to improve the model's understanding of flood vulnerability.

Disclaimers and Contextualization

- **Automatic warnings:** Generating a message when the requested information cannot be inferred solely from the map (e.g., presence of local flood protection infrastructures).
- **Dynamic contextualization:** Linking the assistant with weather data or other external sources to provide a more precise risk assessment.

Improving LLM “as a Judge”

- **Optimizing evaluation prompts:** Better calibrating queries to the LLM to reduce self-evaluation biases.
- **Using multiple judges:** Comparing evaluations from multiple models to minimize the risk of incorrect or overly subjective judgments.
- **Validating through human consensus:** Keeping an independent manual validation to assess the reliability of automatic evaluations.

The research avenues outlined above provide concrete levers to better counter hallucinations, improve reasoning transparency, and make the tool more robust and adaptable to various cartographic environments.

7.5. Demonstrator for Virtual AI Assistants (D120 – GIS.FCU):

The GIS.FCU D120 demonstrator (<https://ryan19981229.github.io/ogc-demo-web/>) showcases a generative AI-enabled virtual assistant designed to enhance data-driven decision-making in geospatial contexts, particularly for flooding-related scenarios. Leveraging advanced AI techniques such as semantic embeddings and large language models (LLMs), the demonstrator provides efficient data retrieval and synthesis, ensuring users receive accurate and contextually relevant information. This tool exemplifies the potential of AI to streamline workflows in GIS and disaster management.

7.5.1. Use Cases and Functionalities

The demonstrator showcases a virtual assistant enabled by the large language models (LLMs), capable of answering users' queries in plain text about the risks, hazards, and impacts of coastal flooding in Canada. Functionalities include:

- Querying relevant data on flooding in Canada to generate an appropriate answer.
- Adjusting the knowledge database to enhance the accuracy of responses.
- Leveraging Retrieval Augmented Generation (RAG) by integrating retrieval accuracy of language models.

7.5.2. Data Sources and FAIR Evaluation

The project integrates datasets including the cause of flooding, financial impacts, insurance solutions, risk impacts, risk to building/ housing/real estate, economy, residential lending, etc. The data collection methodology of this project emphasizes identifying content with clear topics and focuses on organizing key points and relevant information within each topic.

The data sources are designed to align with the FAIR principles and are stored on platforms that are at least compliant with HTTPS standards, ensuring secure access. All data is open, free, and accessible, provided in standardized formats in English with detailed source information to ensure reusability.

7.5.3. Findings and Recommendations

- **Effectiveness of RAG Implementation:** Integrating Retrieval-Augmented Generation (RAG) significantly enhanced the accuracy and contextual relevance of the virtual assistant's responses.
- **Data Quality and FAIR Principles:** Adhering to FAIR Principles ensured high-quality, reusable data was critical to the demonstrator's success. The process highlighted the importance of comprehensive metadata and standardized formats to facilitate interoperability across datasets.
- **Challenges in Data Integration:** Integrating diverse datasets from multiple sources presented challenges, including discrepancies in data formats. These issues underscored the necessity of robust preprocessing steps to harmonize data for AI workflows.
- **Scalability and Performance:** The demonstrator handled a wide range of queries, but scaling the system to accommodate larger datasets and more complex scenarios will require optimized computational resources and more sophisticated indexing strategies.
- **Future Enhancements:** The project revealed opportunities for future development, including the incorporation of real-time data streams, support for multilingual queries, and advanced visualization tools to complement textual outputs.

7.6. Demonstrator for Virtual AI Assistants (D120 – TerraFrame):

The D120 demonstrator by TerraFrame (<https://genai.usace.geoprism.net/> contact info@terraframe.com for login credentials) leverages generative AI and spatial knowledge graphs (SKGs) to address complex challenges in climate disaster impact analysis. The primary objective is to develop a methodology for understanding the immediate affected areas and the cascading effects on critical infrastructure and services such as transportation, public health, and education. This project focuses on building repeatable, sustainable, and scalable processes that can be applied beyond climate domains. Integrating large language models (LLMs) with SKGs enables advanced querying and relationship traversal to provide insights and traceability.

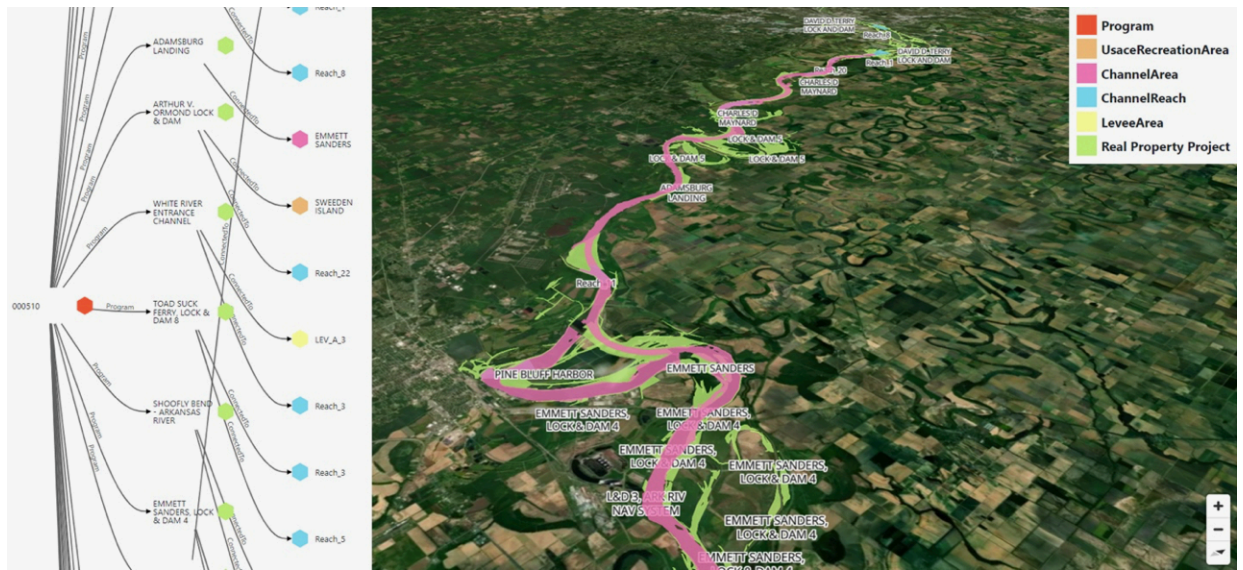


Figure 10 – D120 (TerraFrame) - Use of SKGs for Cascading Effects

7.6.1. Use Cases and Functionalities

The demonstrator showcases its functionality through real-world scenarios. The pilot used an initial geo-ontology to explore methods for bringing geospatial awareness into LLMs and identified school zones impacted by flooding due to levee breaches. The use cases emphasize:

- **Impact Analysis:** Mapping and tracing the downstream effects of disasters across interconnected systems.
- **Query Automation:** Using natural language to generate complex graph queries and return insights in a human-readable format.
- **Multi-Domain Application:** Adapting the methodology for various domains by modifying graph schemas and expanding node definitions.
- **Dynamic Insights:** Providing summarized results for large datasets and maintaining traceability for decision-making.

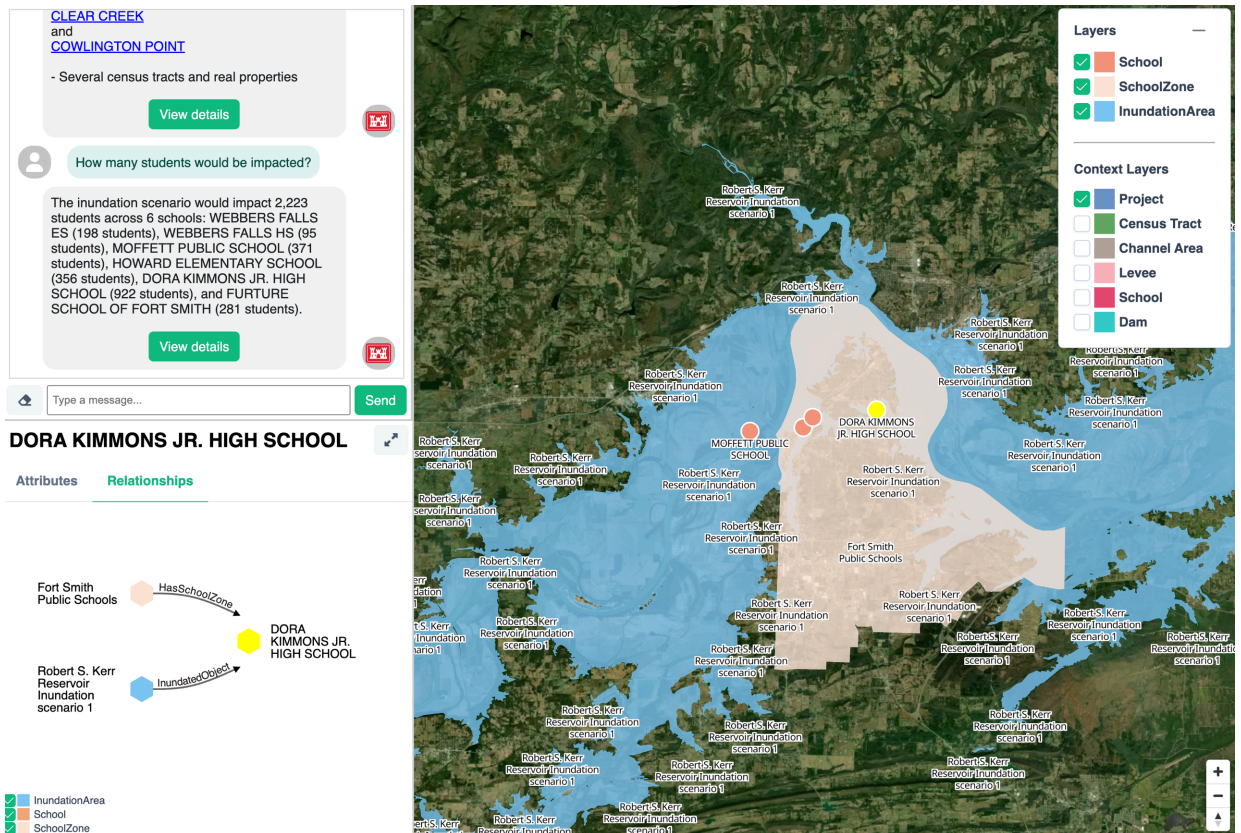


Figure 11 — D120 (TerraFrame) - Insights in a human-readable format

7.6.2. Data Sources and FAIR Evaluation

The project integrates diverse datasets, including spatial, infrastructure, and administrative data sourced from organizations like the US Army Corps of Engineers. The next phase will also include civil infrastructure data from the Internet of Water Coalition and the U.S. Census.

7.6.3. OGC Compliance and Interoperability

One of the effort's objectives is to evaluate OGC standards for making independently developed SKGs interoperable using machine-to-machine readable interfaces, maintaining referential integrity, and propagating changes over time. To that end, the following required metadata properties were identified: defining geo-ontology (i.e., schema for node types and edges), source, temporal validity, and version. Since SKGs are semantic structures that are curated in a decentralized manner, a fixed approach to a modal problem is needed. There does not appear to be an OGC standard to facilitate this.

7.6.4. Findings and Recommendations

Different methods for integrating geospatial awareness modeled in SKGs into LLMs were explored, including natural language vector-embedded Retrieval-Augmented Generation

(RAG) built from RDF triples, creating a RAG from structured Cypher queries with a Planner and ReAct agent, and using LLMs to generate Cypher queries. The latter was the most straightforward and dependable method for identifying cascading effects with multiple degrees of separation modeled in an SKG. It is recommended that the OGC continue to explore an SKG interoperability standard to enable the sustainable creation and maintenance of cross-domain SKGs.

7.7. Demonstrator for Sectoral Applications (D101 – Hartis):

The D101 Coastal Vulnerability Demonstrator (<https://cdrp.hartis.org>) integrates Generative AI to enhance the analysis and visualization of coastal resilience metrics. The tool utilizes the **LLaMA 3 model** to interpret user queries, retrieve relevant **geospatial data from Copernicus services**, and compute the **Coastal Vulnerability Index (CVI)**. This innovative approach combines AI-driven workflows, dynamic mapping, and plain-language explanations, making complex environmental data accessible and actionable, for a diverse range of users, including policymakers, researchers, and the public.

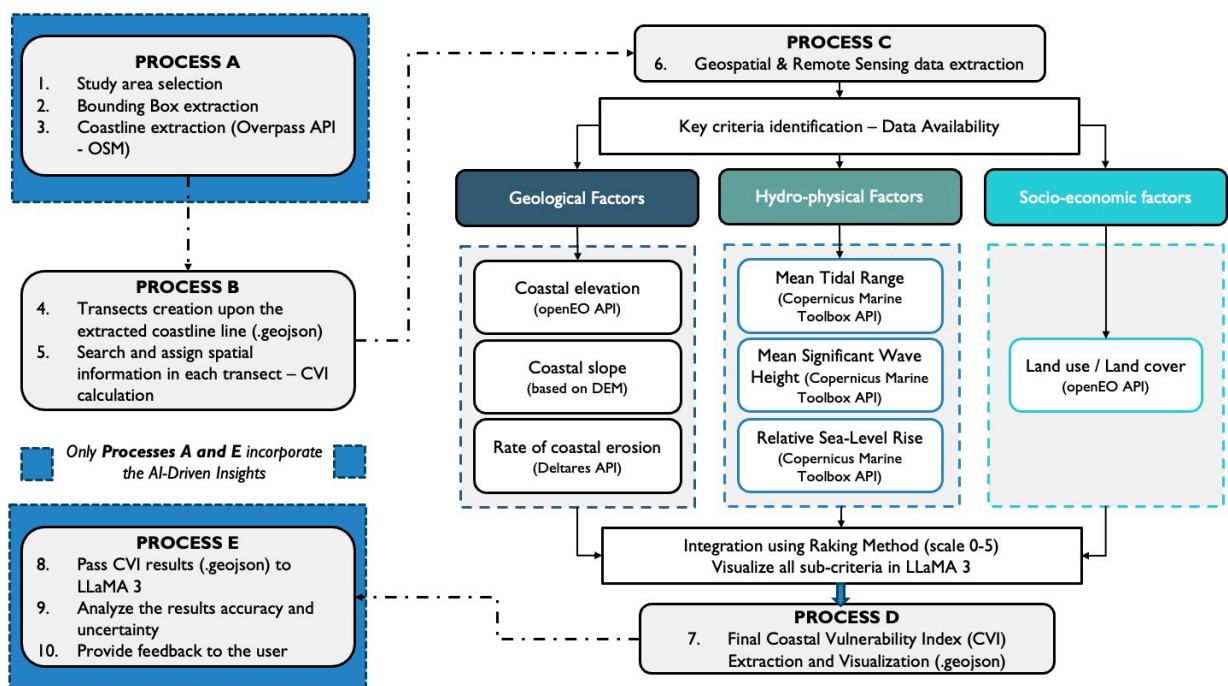


Figure 12 – D101 (Hartis) - AI-Assisted Coastal Vulnerability Analysis

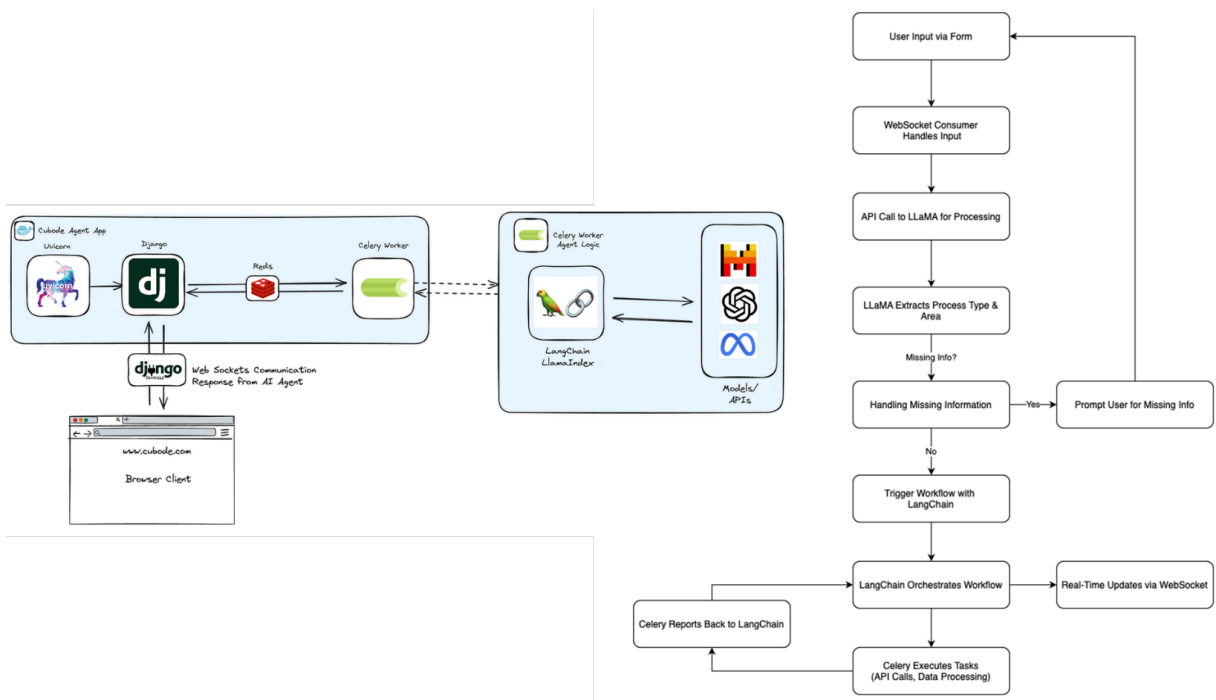
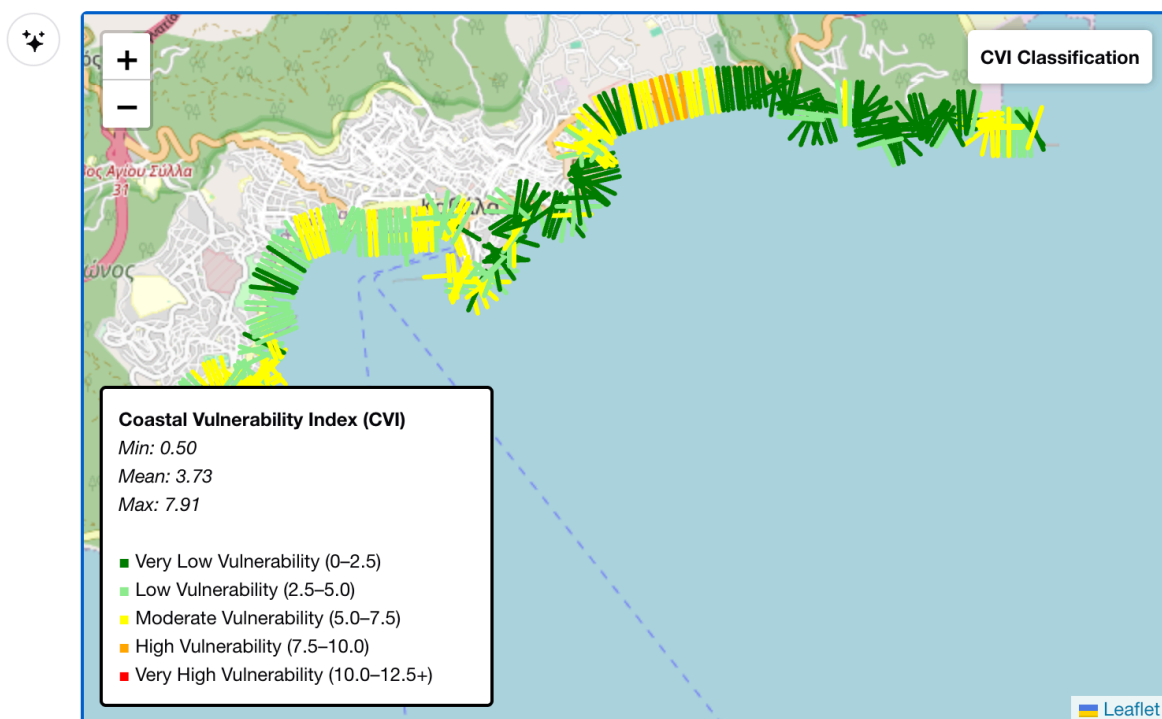


Figure 13 – D101 (Hartis) - AI-Assisted Coastal Vulnerability Analysis Architecture

7.7.1. Sectoral Use Case – Coastal Resilience

The integration of GenAI extends the tool's usage enabling advanced query interpretation and contextual responses for any coastal area in the world. The demonstrator assists in identifying vulnerable coastal zones, providing actionable insights for engaging communities through intuitive, understandable outputs. Applications in test areas such as Eresos Village (eastern Greece) and Kayafa Lake (western Greece) showcase the tool's ability to dynamically adapt to different geographic and environmental conditions. The system computes and visualizes coastal vulnerability based on key geospatial indicators—such as land cover, coastal elevation, slope, erosion rates, and tidal influence—offering interactive visualizations and clear, actionable insights. This approach supports both expert users (e.g., urban planners, disaster response teams) and non-experts (e.g., local communities, decision-makers) in understanding and addressing coastal risks.

Calculating CVI..



CVI calculation completed.

Figure 14 — D101 (Hartis) - AI-Assisted Coastal Vulnerability Analysis Results

7.7.2. Data Sources and FAIR Evaluation

The demonstrator leverages multiple data sources, including Copernicus services (e.g., C3S, CAMS, and EUMETSAT), along with external geospatial platforms compliant with OGC standards. The integration of GenAI enhances the data retrieval and processing pipeline, ensuring seamless interaction with users.

To align with FAIR principles, the system prioritizes:

- Interoperability – Ensuring compatibility with standard geospatial ontologies and APIs.
- Reusability – Allowing scalable and replicable application of AI-driven coastal analysis across different regions.

7.7.3. Evaluation and Lessons Learned

The pilot process highlighted several key takeaways:

- The **AI-driven analysis of the CVI index** provided by the LLaMA 3 model, translated the geospatial data into actionable insights, providing real-time responses related to coastal hazards, sea-level rise, and climate adaptation strategies.
- Its **scalability across different coastal regions** is a major strength, as the demonstrator operates dynamically without requiring pre-defined study areas. Users can analyze any coastal location, making it adaptable for global applications.
- However, **computational challenges and optimization** remain important considerations. Large-scale calculations in computing the coastal vulnerability index (CVI) across extensive coastal regions—require optimization. Strategies like limiting the analysis zones are necessary to ensure faster processing times.
- The system's focus on **user interaction and interpretability** also proved valuable. The ability to generate user-friendly reports and interactive visualizations make technical geospatial assessments more accessible.
- Another important aspect is **uncertainty analysis and weighting mechanisms**. Further work is planned to improve uncertainty estimation in CVI calculations. Integrating weighted data processing techniques, the influence of different environmental factors can be adjusted, enhancing the accuracy and interpretability of the vulnerability assessment.
- Looking ahead, plans are in place to broaden user testing and incorporate feedback from policymakers, urban planners, and local communities. These insights will guide future refinements, including enhanced visualization, interactive AI guidance, and more flexible scenario modeling.

7.8. Demonstrator for Sectoral Applications (D101 – Pixalytics):

The D101 Health Impact Demonstrator (<https://github.com/pixalytics-ltd/Climate-drought>) leverages data-driven workflows to address the **health impacts of heat and drought**. This pilot integrates advanced climate data processing with analytical tools to support informed decision-making in the health sector. Using Python-based libraries and ERA5 datasets, the workflow processes precipitation and temperature-related indices, including the Standardized Precipitation Index (SPI) and the Universal Thermal Climate Index (UTCI). The aim is to provide **analyst-ready data** that can be queried, visualized, and combined to derive actionable insights. The workflow has been successfully deployed to the ECMWF supported, **Copernicus WEkEO platform**, ensuring accessibility and scalability for future developments.

7.8.1. Sectoral Use Cases (Health)

The primary use case for the D101 demonstrator is in public health, focusing on understanding and mitigating the impacts of heatwaves and droughts. Specific applications include:

- Assessing heat stress risks using UTCI data to **inform public health interventions**.
- Correlating SPI and UTCI indices to analyze the **combined effects of drought and temperature extremes on vulnerable populations**.
- Supporting urban planning and emergency preparedness by **identifying regions at higher risk** of heat and drought impacts.

Future extensions will explore integrated heat and drought indices to provide more robust insights into climate-driven health challenges.

During 2022, the UK endured record high temperatures of 40 degC and there were around 3 000 more deaths in the over-65s than usual in [England and Wales](#). So, the indices have been tested over the UK (location in East Anglia, which is in the South East) for 2022.

7.8.2. Data Sources and FAIR Evaluation

Pixalytics [previously developed](#) an OGC-compliant API service, combining meteorology, hydrology, and remote sensing open data/datasets to produce ARD data based on a composite of different indicators. The original API access was been set up following the [Building Blocks for Climate Services](#) approach, and within this pilot the workflow is accessed through a Jupyter Notebook that will be deployed on the WEkEO platform.

The Standardized Precipitation Index (SPI) and Universal Thermal Climate Index (UTCI) are being calculated using ERA5/ERA5-Land reanalysis data from the Copernicus Climate Change Service (C3S) [Climate Data Store](#). It is also possible to download both parameters, pre-calculated [ERA5-HEAT dataset](#), from C3S but there were issues with the retrieval failing.

SPI is used to characterize meteorological drought, quantifying the observed precipitation standardized departure from a selected probability distribution function that models the input precipitation data. UTCI is the air temperature of a reference outdoor environment that would elicit the same physiological response as the actual environment ([Di Napoli et al. 2020](#)). Its inputs include mean radiant temperature (MRT), which representing how human beings experience radiation, and the temperatures are classified into categories as grades of thermal stress.

Heat stress follows a diurnal pattern with UTCI values at 06:00 or 18:00 generally lower than UTCI values at 12:00 or 15:00, and it's also more dominant in southern Europe ([Di Napoli et al. 2017](#)). To follow the approach of the pre-calculated ERA-HEAT dataset, [hourly ERA-5 data](#) was downloaded. Initially, the UTCI was calculated for just summer months (June to August) as these months have the greater likelihood of showing occurrences of heat stress. The calculated health index can also be used to model the impact of cold weather on health so the full year was calculated as well. Values of UTCI were calculated hourly using the [xclim](#) Python module, then the maximum value from the day was chosen and the average was calculated across each calendar month.

The SPI needed to be calculated over a long period of time of time and so daily data from 1985 to the present data was used, which took a long time to download as it needed to be downloaded a month at a time to comply with C3S's policy. So, the data is cached for future analyses run at the same location and monthly data was also tested that would allow for a

quicker download. Negative values indicate the situation is drier than normal while positive values indicate it is wetter.

The figure shows an example **plot of SPI, UTCI and the Health Index** for the UK in 2022 with, for example, low SPI and high UTCI values in the summer. The Health Index is calculated from the combination of SPI & UTCI, so has the high/positive values in the summer indicating heat/drought impact, with low/negative values indicating the health impact of cold/wet weather.

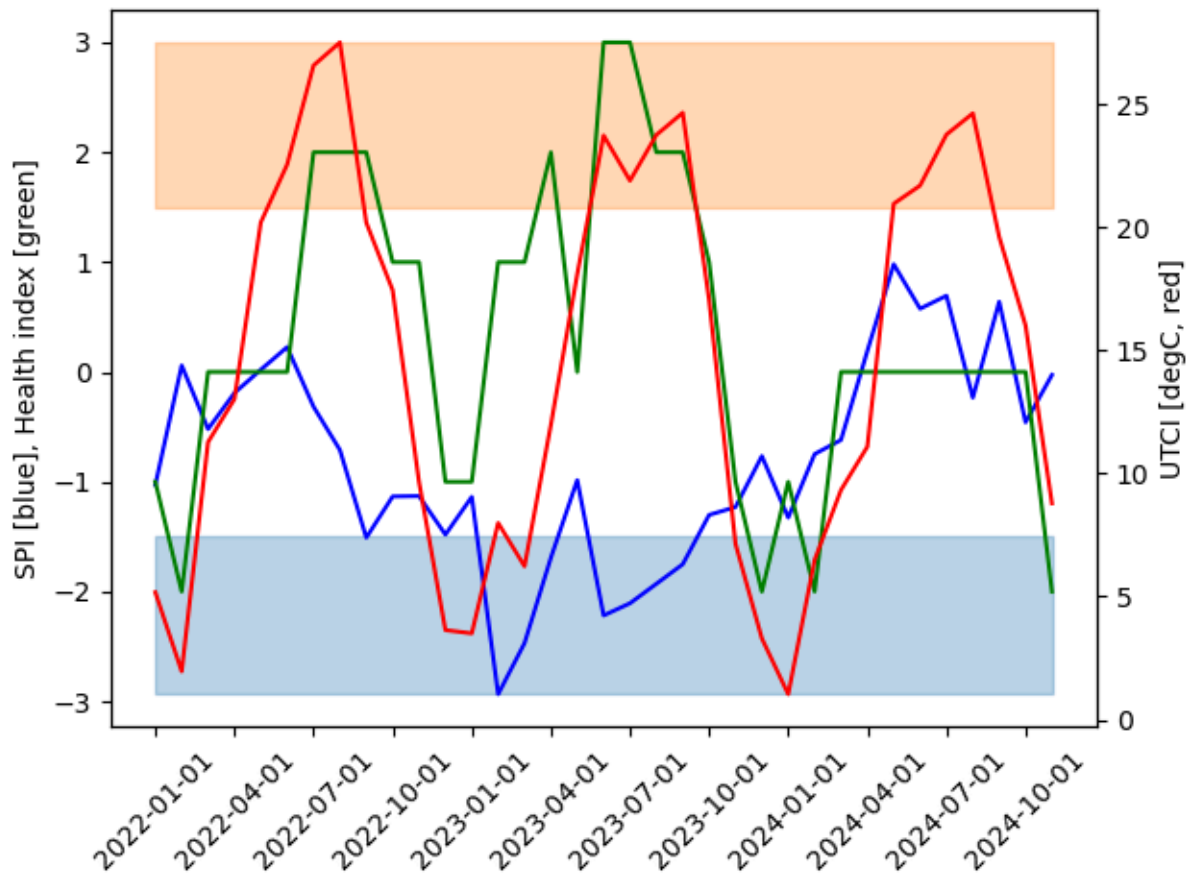


Figure 15 — D101 (Pixalytics) - Example plot of SPI, UTCI and Health Index for the South Eastern UK from 2022 to 2024:

7.8.3. Evaluation and Lessons Learned

As we used daily maximum values for this analysis, so the highest value during the day was captured. Also, different temporal compositing was investigated (monthly versus daily) to speed up the CDS download time, and it was decided monthly values could be used for SPI while hourly values were needed for the UTCI calculation.

There is a delay while the data is accessed from the CDS and downloaded. The delay is reasonable for the SPI calculation, but more significant for UTCI and so users need to be patient. A month at a time is downloaded, so if the connection fails it can be restarted and previously downloaded data used, providing a level of resiliency.

The developed code has been deployed and demonstrated using the WEkEO platform, which allows users to run applications within a pre-existing cloud infrastructure with harmonized data access. The developed code has been made available within a public GitHub repository, with archiving using Zenodo providing a DOI for released versions. The aim will be to build on this code base further within future pilot activities, supporting the testing of standards and interoperability.



8

PILOT FINDINGS & RECOMMENDATIONS

The **CDRP 2024.2 Pilot** has identified key challenges and opportunities in the integration of GenAI with geospatial applications for climate resilience and disaster management. These findings emphasize the need for structured data handling, enhanced AI geospatial awareness, and improved interoperability between AI and geospatial systems.

8.1. Findings

- **Geospatial Awareness in AI Models**
 - AI models exhibit **limited understanding of geographic references**, requiring additional geocoding and ontology mapping for accurate results.
 - AI-generated responses often lack spatial context, necessitating **validation mechanisms** to enhance reliability.
- **AI Hallucinations and Result Confidence**
 - Generative AI models can produce **misleading or incorrect geospatial insights** when relying on incomplete or ambiguous data.
 - Low confidence in AI-generated results stems from datasets that lack **clear metadata or structured retrieval mechanisms**.
 - **Structured AI prompting** (e.g., JSON formatting, Chain-of-Thought reasoning) improves AI's ability to generate reliable responses.
- **Data Interoperability and Metadata Challenges**
 - Inconsistent metadata standards and **cross-platform compatibility issues** hinder seamless AI-geospatial data integration.
 - A lack of **OGC-compliant APIs** and **semantic ontologies** limits AI's ability to process and understand geospatial data effectively.
- **Infrastructure and Scalability Barriers**
 - **High costs for training, fine-tuning, and scaling AI models** create barriers to full implementation.
 - AI models struggle with processing large-scale geospatial datasets in real time due to **computational and infrastructure constraints**.

8.2. Recommendations

To improve AI applications in climate resilience and disaster management, the following actions are proposed:

- **Improve AI's Geospatial Reasoning and Retrieval Capabilities**
 - Implement **retrieval-augmented generation (RAG)** models to ensure AI retrieves verified knowledge from structured geospatial sources.
 - Expand the use of **geospatial APIs** (e.g., geocoding services) to **link AI-generated outputs with authoritative spatial data**.
- **Enhance Metadata Standards and Data Interoperability**
 - Adopt comprehensive metadata schemas such as **ISO 19115** to ensure structured and discoverable datasets.
 - Develop **ontology crosswalks** to align **OGC standards** with AI-driven geospatial reasoning models.
- **Strengthen AI Validation Mechanisms**
 - Implement **automated geographic validation** in AI workflows to detect and mitigate hallucinations in location-based responses.
 - Utilize **knowledge graphs and graph-based representations** to structure AI's spatial reasoning processes.
- **Optimize AI for Emergency and Climate Risk Applications**
 - Deploy **AI-driven risk assessments** within emergency response systems to support **real-time decision-making**.
 - Leverage AI for **financial risk modeling**, helping insurers and government agencies assess the impact of climate-related disasters.
- **Develop Scalable Infrastructure for AI-Geospatial Integration**
 - Invest in **cloud-native architectures** and **edge computing solutions** to enhance AI's ability to process real-time geospatial data.
 - Strengthen collaboration between **AI developers, climate scientists, and emergency management experts** to co-develop predictive models for wildfires, floods, and droughts.

9

STAKEHOLDER PERSPECTIVES

The integration of GenAI and geospatial data for climate resilience and disaster management requires active involvement from a diverse set of stakeholders. Governments, private organizations, academic institutions, and community groups each play a critical role in ensuring the success of these initiatives. This section explores the unique contributions of stakeholders, gathers feedback on current implementations, and identifies opportunities for fostering collaboration and future partnerships.

9.1. Stakeholder Roles and Contributions

Each stakeholder group contributes uniquely to the advancement of GenAI and geospatial applications:

- **Governments and Policy Makers:** Provide regulatory frameworks, funding for research and infrastructure, and enforce data-sharing policies to ensure accessibility and inclusivity.
- **Private Sector:** Drives innovation by developing AI tools, investing in cloud infrastructure, and commercializing generative AI solutions tailored to geospatial use cases.
- **Academia and Research Institutions:** Focus on advancing knowledge through interdisciplinary studies, producing high-quality datasets, and improving AI models tailored for geospatial applications.
- **Community Organizations:** Offer ground-level insights on local needs, validate the usability of AI solutions in real-world scenarios, and promote community-driven data initiatives.
- **NGOs and Advocacy Groups:** Bridge gaps between technical teams and vulnerable communities, ensuring that GenAI applications prioritize equity and ethical use.

9.2. Feedback and Requirements from Community and Industry Stakeholders

Feedback from stakeholders emphasizes several critical areas for improvement:

- **Data Quality and Accessibility:** Communities and industry stakeholders stress the importance of standardized, high-quality, and easy-to-access datasets for GenAI workflows.

- **User-Centric Design:** End-users, including disaster managers and urban planners, require tools that are intuitive, customizable, and adaptable to diverse operational needs.
- **Fairness and Equity:** NGOs and advocacy groups highlight the need for inclusive solutions that address underrepresented regions and marginalized communities.
- **Interoperability:** Industry stakeholders seek seamless integration of datasets across platforms to avoid redundancies and inefficiencies in disaster response planning.
- **Training and Capacity Building:** There is a clear demand for training programs to help stakeholders understand and effectively use GenAI-powered tools in decision-making processes.

9.3. Opportunities for Collaboration and Future Partnerships

Collaboration among stakeholders offers transformative potential in advancing the role of GenAI in climate resilience and disaster management. Key opportunities include:

- **Public-Private Partnerships:** Governments can team up with private companies to leverage funding and technical expertise for large-scale implementations, such as developing real-time disaster monitoring systems.
- **Academic-Industry Collaborations:** Universities and research institutions can work with tech companies to refine AI algorithms and improve the FAIRness of geospatial datasets.
- **Community-Led Data Initiatives:** Encourage participatory data collection by involving local communities in the creation and validation of geospatial datasets, ensuring relevance and inclusivity.
- **Intergovernmental Cooperation:** Foster cross-border data sharing and joint research efforts to address global climate challenges more effectively.
- **Innovation Hubs and Hackathons:** Establish collaborative platforms where stakeholders can co-develop prototypes, share best practices, and accelerate innovation in GenAI applications for geospatial data.



10

FUTURE OUTLOOK

10.1. Recommendations for Enhancing GenAI Integration for Climate Resilience

GenAI has the potential to revolutionize climate and disaster resilience efforts by enabling **predictive analysis**, **early warning systems**, and **adaptive planning**. The integration of GenAI with geospatial data can support decision-making for disaster response, climate adaptation, and risk mitigation. GenAI offers transformative potential in geospatial applications, but its **effective integration** depends on **strategic actions**, such as the **prioritization** of data quality, availability, and interoperability.

10.1.1. Data Quality and Availability

High-quality, climate-specific geospatial data is essential, as GenAI models depend on accurate datasets, including **historical weather records**, **floodplain maps**, **hazard zones**, and **critical infrastructure data**. Real-time data from Earth Observation (EO) systems, IoT climate sensors, and social media can enhance early warning systems and disaster impact assessments. Investing in **data cleaning**, **validation**, and **quality control** processes is paramount.

10.1.2. Bias and Data Gaps

Addressing data gaps and biases is vital, as geospatial data is often **unevenly distributed**, leading to **biased models**. Filling **data gaps** in climate-vulnerable regions, such as **small island states**, **coastal zones**, and **low-income urban areas**, is essential. Techniques like **data augmentation** and **synthetic data generation** can simulate disaster scenarios to improve model robustness. Promoting data sharing and interoperability is equally important. Siloed data hinders progress, so encouraging **open data initiatives**, establishing standards like OGC ones, and developing tools for data transformation and interoperability are critical. The effort should include documented APIs and open data formats for easier access and integration.

10.1.3. Interdisciplinary Collaboration

Collaboration between **climate scientists**, **urban planners**, **disaster management agencies**, and **GenAI developers** is paramount. This could include developing flood prediction models, drought resilience strategies, and wildfire monitoring frameworks.

10.1.4. Metadata and Standards

Metadata specific to disaster resilience should include attributes such as **evacuation routes**, **response times**, and **climate vulnerability indices**. Standards like OGC should extend to disaster-relevant features, covering aspects such as provenance, accuracy, spatial and temporal resolution, and known biases.

Fostering interdisciplinary collaboration is essential to bridge expertise gaps. Effective integration of GenAI and geospatial data requires collaboration across fields such as geospatial science, AI/ML, software engineering, and domain-specific disciplines. Establishing collaborative platforms, including joint research projects, hackathons, workshops, and online communities, can facilitate knowledge sharing and accelerate innovation.

10.1.5. Scalability and Robustness

The development of robust, scalable, and explainable AI models tailored to geospatial data is another key area of focus. Generic GenAI models may not be optimal for geospatial applications, necessitating research into architectures and training techniques designed to handle unique geospatial characteristics such as spatial autocorrelation, scale dependency, or complex spatial relationships. Scalability is also important to process massive geospatial datasets in real-time during disasters, leveraging cloud and distributed computing frameworks, requiring models and infrastructure that can handle large volumes of data efficiently.

10.1.6. Explainability and Ethical AI

Explainability is equally important in life-critical applications like disaster risk prediction. For example, communities must trust flood forecasts to act promptly. Ethical AI must prioritize inclusion of **indigenous knowledge** and equity in **resilience planning**, ensuring outputs do not marginalize vulnerable populations. Addressing bias, ethical considerations, and trustworthiness is paramount. Methods for detecting and mitigating bias in geospatial data and GenAI models should be based on inclusive and fairness-aware datasets and careful evaluation of outputs.

Additionally, establishing trust and accountability is essential, requiring benchmarks and inspection rules to evaluate the accuracy and reliability of autonomous GIS solutions, including reviewing workflows, code, and outputs for potential biases. Promoting open-source development and standardization is another important step. Encouraging the development and sharing of open-source tools, libraries, and platforms can foster community involvement, accelerate innovation, and promote wider adoption.

10.1.7. Functionality Enhancements

Focusing on enhanced functionality will unlock new possibilities. **Autonomous systems** should enable **real-time disaster detection** (e.g., floods, wildfires) by filtering and analyzing live geospatial data catalogs and web services, enabling LLMs to understand metadata, quality, and relevance for specific tasks. LLMs trained on disaster-specific geospatial datasets should assist

in identifying safe evacuation routes, prioritizing infrastructure repairs, and simulating disaster scenarios.

Efforts should enable GenAI to answer and **provide reasoning**, requiring deeper geospatial knowledge and hypothesis generation. Investment in training Large Spatial Models (LSMs) on extensive geospatial data, akin to LLM training on text corpora, can significantly enhance spatial awareness and reasoning.

10.2. Strategic directions for Stakeholders and Developers

10.2.1. Directions for Stakeholders

Governments and funding agencies should prioritize GenAI applications for climate resilience, such as adaptive planning, early warning systems, and post-disaster recovery. Investments should target data acquisition for high-risk regions and tools for climate impact attribution. To effectively guide the integration of GenAI with geospatial data, distinct yet interconnected steps are necessary for both stakeholders and developers. For stakeholders, including governments, research institutions, businesses, and funding agencies, a primary focus should be on building a robust data ecosystem. This involves:

- Significant investment in acquiring, curating, and maintaining high-quality, representative geospatial datasets.
- Establishing and enforcing data standards and metadata schemas to ensure interoperability.
- Actively promoting open data initiatives and collaborative data-sharing platforms.

Stakeholders should also foster interdisciplinary collaboration by:

- Creating funding opportunities.
- Establishing joint research programs.
- Organizing workshops and conferences that bring together geospatial experts, AI researchers, and domain-specific specialists.

Furthermore, stakeholders must ensure that AI solutions are ethical and equitable, focusing on inclusivity in disaster resilience planning by:

- Establishing clear guidelines and regulations for the use of GenAI in geospatial applications.

- Addressing concerns related to data privacy, algorithmic bias, and potential societal impacts (see also Section 9).

Finally, stakeholders should invest in education and training programs to develop a skilled workforce capable of working with GenAI and geospatial data, ensuring a smooth transition and maximizing the benefits of this technology.

10.2.2. Directions for Developers

Developers should focus on building robust, scalable, and user-friendly GenAI tools and platforms. This involves:

- Developing novel AI architectures and training methodologies specifically designed for the unique characteristics of the geospatial domain.
- Prioritizing the scalability of their solutions to handle large volumes of geospatial data, leveraging cloud computing and distributed processing frameworks.
- Enhancing the explainability and interpretability of GenAI models through techniques like workflow visualization and explainable AI (XAI) methods to build trust and facilitate user understanding.

Developers should also actively engage with stakeholders to understand their needs and priorities, ensuring that GenAI solutions are relevant, practical, and address real-world challenges.

10.3. Areas for Further Research and Development

10.3.1. Research Directions

Several key areas require further research and development to fully realize the potential of GenAI in geospatial applications. Firstly, to develop GenAI models for climate resilience that predict **cascading impacts**, such as floods following hurricanes or power outages due to extreme heat. Focus on training Large Spatial Models (LSMs) for disaster scenarios, akin to LLMs for text, to enhance geospatial reasoning and response efficiency.

Research should also focus on developing methods for **handling uncertainty** and **imprecision** inherent in real-world geospatial datasets, including techniques for **data imputation**, **error propagation**, and **uncertainty quantification**. Secondly, enhancing explainability and interpretability of GenAI models in the geospatial domain for building trust and ensuring responsible use, given that geospatial data is often used in life-critical applications. This involves developing techniques that provide insights into the **internal workings** and potential biases of

these models, allowing users to understand why a particular prediction or classification was made.

Developing multimodal GenAI models that combine EO, LiDAR, sensor networks, and crowd-sourced disaster data for comprehensive analysis is a major research direction. This involves developing techniques for **data fusion**, **multimodal representation learning**, and **cross-modal reasoning**. Research is also needed on developing efficient and scalable infrastructure for deploying and managing GenAI models in real-world geospatial applications. This includes methods for model compression, optimization, and deployment on **resource-constrained** edge devices.

Finally, investing in the development of LSMs trained on vast amounts of geospatial data, similar to how LLMs are trained on text corpora, is a transformative research direction. This involves developing **new training paradigms**, **data augmentation techniques**, and **model architectures** specifically designed for geospatial data.

10.3.2. Education and Training

Equally important is to focus investing in education and training for developing geospatial AI expertise. Training programs should emphasize climate resilience applications, teaching professionals **how to leverage GenAI** for flood prediction, evacuation planning, and adaptive infrastructure design.

Moreover, encouraging public-private partnerships can leverage synergies between sectors. Public organizations can provide access to data and expertise, while private companies can contribute technological resources and market expertise, driving innovation and accelerating adoption of GenAI in geospatial applications. Together, they can interact with academia and educational programs to promote GenAI integration into geospatial science.

10.4. Security, Privacy and Ethical Considerations in GenAI for Geospatial Applications

The integration of Generative AI (GenAI) in geospatial applications offers significant benefits for climate resilience and disaster management. However, responsible deployment requires addressing key security, privacy, and ethical concerns to build stakeholder trust.

Data privacy and protection must comply with regulations like GDPR, ensuring the security of sensitive geospatial data, particularly in disaster-prone regions. Algorithmic bias and fairness must be mitigated by using inclusive data practices and fairness-aware algorithms to prevent discrimination. Transparency and explainability are critical, especially in high-stakes scenarios like disaster response, where Explainable AI (XAI) enhances user trust.

Ethical considerations require GenAI to align with principles of equity, inclusivity, and sustainability while avoiding misuse or exploitation. Strong security measures are necessary to protect against cyber threats and data breaches, ensuring infrastructure resilience. Additionally,

accountability frameworks, regular audits, and oversight mechanisms must be established to maintain ethical AI deployment.

To address these concerns, organizations should develop clear data policies, promote inclusivity in data collection, invest in explainability research, enhance security protocols, and foster collaborative oversight. By implementing these measures, GenAI can be deployed responsibly in geospatial applications while ensuring fairness, security, and trust.



BIBLIOGRAPHY





BIBLIOGRAPHY

- [1] 6 Steps to Success with Generative AI. In: **Amazon Web Services (AWS) Resources**, 2024. <https://anz-resources.awscloud.com/transform-your-business-value-with-gen-ai-ml/6-steps-to-success-with-generative-ai>.
- [2] Generative AI tools can enhance climate literacy but must be checked for biases and inaccuracies. In: **Communications Earth & Environment**, Vol. 5, Article 226, 2024. <https://doi.org/10.1038/s43247-024-01392-w>.
- [3] Pangeo Forge: Crowdsourcing Analysis-Ready, Cloud Optimized Data Production. In **Frontiers in Climate** (Vol. 3), 2022. Frontiers Media SA. <https://doi.org/10.3389/fclim.2021.782909>
- [4] Leveraging AI in Emergency Management and Crisis Response. In: **Deloitte Insights**, 2024. <https://www2.deloitte.com/us/en/insights/industry/public-sector/automation-and-generative-ai-in-government/leveraging-ai-in-emergency-management-and-crisis-response.html>.
- [5] The paradoxes of generative AI-enabled customer service: A guide for managers. In: **Business Horizons**, Vol. 67, Issue 5, pp. 549–559, 2024. Elsevier BV. <https://doi.org/10.1016/j.bushor.2024.04.013>.
- [6] Generative AI: A systematic review using topic modelling techniques. In: **Data and Information Management**, Vol. 8, Issue 2, Article 100066, 2024. Elsevier BV. <https://doi.org/10.1016/j.dim.2024.100066>.
- [7] Decoding the AI Virtual Assistant Design Architecture: An In-Depth Look into Design Components. In: **Medium**, December 7, 2023. <https://medium.com/@senol.isci/decoding-the-ai-virtual-assistant-design-architecture-an-in-depth-look-into-design-components-73fabba31de8>.
- [8] OGC CDRP 2024: Moving beyond SDI to GKI to bring geospatial awareness to LLMs to support climate and disaster resilience. In: **OSF**, 2024. <https://doi.org/10.17605/OSF.IO/CE34A>.
- [9] Enabling Knowledge Sharing by Managing Dependencies and Interoperability Between Interlinked Spatial Knowledge Graphs. In: **The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, Vol. XLVIII-4/W7-2023, pp. 117–124. Copernicus GmbH, 2023. <https://doi.org/10.5194/isprs-archives-xxviii-4-w7-2023-117-2023>.
- [10] AI Beyond Words: How Large Spatial Models Could Revolutionize 3D Understanding. In: **Medium**, November 6, 2024. <https://medium.com/@daniil.rossy/ai-beyond-words-how-large-spatial-models-could-revolutionize-3d-understanding-0b292df400cd>.

- [11] AI hallucination: towards a comprehensive classification of distorted information in artificial intelligence-generated content. In: **Humanities and Social Sciences Communications**, Vol. 11, Article 1278, 2024. <https://doi.org/10.1057/s41599-024-03811-x>.
- [12] OGC Climate and Disaster Resilience Pilot IV: D-123 Generative AI in Wildland Fire Management Engineering Report. In: **Zenodo**, 2024. <https://doi.org/10.5281/ZENODO.12721058>.
- [13] ChatClimate: Grounding conversational AI in climate science. In: **Communications Earth & Environment**, Vol. 4, Article 480, 2023. <https://doi.org/10.1038/s43247-023-01084-x>.



ANNEX A (NORMATIVE) ABBREVIATIONS/ACRONYMS



ANNEX A (NORMATIVE) ABBREVIATIONS/ACRONYMS

API	Application Programming Interface
ARD	Analysis Ready Data
C3S	Copernicus Climate Change Service
CAMS	Copernicus Atmosphere Monitoring Service
CDRP	Climate and Disaster Resilience Pilot
CDS	Climate Data Store
CEOS	Committee on Earth Observation Satellites
CGDI	Canadian Geospatial Data Infrastructure
CoT	Chain of Thought Reasoning
CRIM	Centre de recherche informatique de Montréal
DMSMM	Data Management and Stewardship Maturity Matrix
DSMM	Data Stewardship Maturity Matrix
EO	Earth Observation
FAIR	Findable, Accessible, Interoperable, Reusable
FCU	Feng Chia University
GCM	General Circulation Models
GenAI	Generative Artificial Intelligence
GKI	Geospatial Knowledge Infrastructure
JSON	JavaScript Object Notation
LLM	Large Language Model

LVM	Large Vision Models
ML	Machine Learning
MLaaS	Machine Learning as a Service
NOAA	National Oceanic and Atmospheric Administration
OGC	Open Geospatial Consortium
RAG	Retrieval-Augmented Generation
RCMs	Regional Climate Models
SKG	Spatial Knowledge Graph
SPI	Standardized Precipitation Index
UTCI	Universal Thermal Climate Index
WEkEO	WEkEO Environmental Data Hub



B

ANNEX B (NORMATIVE) DEMO ILLUSTRATIONS FOR VIRTUAL AI ASSISTANTS (D100 – GEOLABS)



B

ANNEX B

(NORMATIVE)

DEMO ILLUSTRATIONS FOR VIRTUAL AI ASSISTANTS (D100 – GEOLABS)

B.1. Demo Illustration 1: Searching data source

User Query: *"Where can I get climate data?"*

Illustration

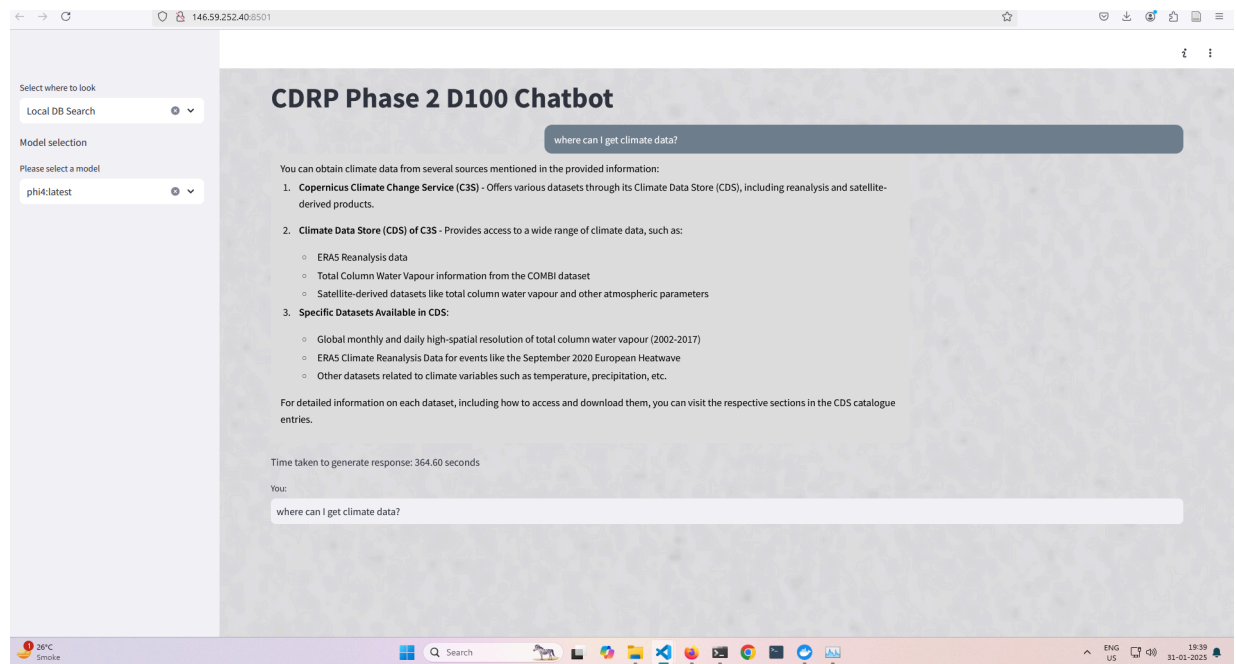


Figure B.1 — D100 (GeoLabs) - Demo 1

B.2. Demo Illustration 2: General climate-related query

User Query: *"How to do climate trend?"*

Illustration

Deploy ⓘ

CDRP Phase 2 D100 - Data Discovery Bot

how to do climate trend?

To analyze climate trends based on the provided text, you can follow these general steps:

1. **Access Climate Data:** Use data from reliable sources such as the Climate Data Store (CDS) of the Copernicus Climate Change Service (C3S). This involves searching for and downloading relevant datasets.
2. **Calculate Climatologies:** Compute climatological averages over a specific period to understand typical seasonal patterns. For example, calculate the monthly climatology of total column water vapour or global surface temperatures.
3. **Analyze Trends:** Identify trends in climate data by examining time series for long-term changes. This can involve looking at overall increases or decreases in variables like total column water vapour (TCWV) or surface temperatures over years.
4. **Decompose Time Series:** Break down the time series into components such as trend, seasonal cycle, and anomalies. For instance, decomposing TCWV data to understand its trend and seasonal variations can help identify periods of significant change related to climate phenomena like El Niño or La Niña events.
5. **Compute Anomalies:** Calculate anomalies by comparing current values against the climatology. This helps in identifying deviations from typical patterns, which might indicate unusual climate behavior.
6. **Visualize Data:** Use visualizations such as time series plots and anomaly maps to better understand and communicate the trends and changes observed in the data.
7. **Analyze Specific Events or Regions:** Focus on particular events (e.g., record-breaking temperatures) or regions (e.g., North Atlantic sea surface temperature trends) for more detailed analysis.

These steps are derived from the methodologies described in the text, which involve using climate data to study past, present, and future climate conditions.

Time taken to generate response: 44.22 seconds

You:

how to do climate trend?

Figure B.2 — D100 (GeoLabs) - Demo 2

B.3. Demo Illustration 3: Generate code example for climate data access and analysis

User Query: *"How can I do climate trend using Copernicus data?"*

Illustration

Select where to look

Web Query

Model selection

phi4.latest

CDRP Phase 2 D100 Chatbot

how can i do climate trend using copernicus data?

To perform a climate trend analysis using Copernicus data, you'll need to follow several steps. Here's a general guide:

1. Access the Data

- Copernicus Climate Data Store (CDStore):** Visit the [Copernicus CDS](#) and create an account if you don't have one.
- Select Datasets:** Choose datasets relevant to your analysis, such as ERA5 for atmospheric reanalysis or other specific climate variables like temperature, precipitation, etc.

2. Download the Data

- Use the CDS API or web interface to download the data. You can specify parameters like time range, geographical area, and variable of interest.
- For scripting, you might use Python with the `cdsap` library to automate downloads.

3. Data Preprocessing

- Load Data:** Use libraries such as `xarray`, `pandas`, or `netCDF4` in Python to load the data into your environment.
- Handle Missing Values:** Clean the dataset by handling missing values appropriately, either through interpolation or removal.
- Temporal Aggregation:** Aggregate the data if necessary (e.g., monthly averages) to match the temporal resolution needed for trend analysis.

4. Trend Analysis

- Statistical Methods:** Use statistical methods like linear regression, Mann-Kendall test, or Theil-Sen estimator to identify trends.
 - Linear Regression:** Fit a linear model to your time series data to estimate the slope (trend).
 - Mann-Kendall Test:** A non-parametric test that assesses whether there is a monotonic upward or downward trend of the variable over time.
 - Theil-Sen Estimator:** Robust method for estimating the slope of a trend line.
- Software Tools:** Use Python libraries such as `statsmodels`, `scipy`, or `pandas` for implementing these methods.

5. Visualization

- Plot your data and trends using visualization libraries like `matplotlib` or `seaborn`.
- Create time series plots, scatter plots with trend lines, or maps if spatial analysis is involved.

Figure B.3 — D100 (GeoLabs) - Demo 3

Plot your data and trends using visualization libraries like `matplotlib` or `seaborn`.

Create time series plots, scatter plots with trend lines, or maps if spatial analysis is involved.

6. Interpretation

- Analyze the results to understand the significance of the trends.
- Consider external factors that might influence the observed trends and validate your findings with other datasets if possible.

7. Documentation and Reporting

- Document your methodology, assumptions, and findings.
- Prepare a report or presentation summarizing the analysis and its implications.

Example Code Snippet

Here's a simple example using Python to perform a linear trend analysis on temperature data:

```

import cdsapi
import xarray as xr
import pandas as pd
import numpy as np
from scipy.stats import linregress
import matplotlib.pyplot as plt

# Download data from CDS (example for ERA5 monthly mean surface air temperature)
c = cdsapi.Client()
c.retrieve(
    'reanalysis-era5-pressure-levels',
    {
        'product_type': 'reanalysis',
        'format': 'netcdf',
        'variable': ['2m_temperature'],
        'year': [2020],
        'month': [1, 2, 3], # Example: first quarter of 2020
        'day': [1, 15, 31],
        'time': ['00:00', '12:00'], # Midnight and noon
        'pressure_level': ['surface'],
    },
    'download.nc'
)

# Load data using xarray
ds = xr.open_dataset('download.nc')

# Convert to pandas DataFrame for easier manipulation
df = ds['t2m'].to_dataframe().reset_index()

```

Figure B.4 — D100 (GeoLabs) - Demo 4