

The Open Science and Data Platform with Generative AI for Impact Assessment



Mohamed Habbane, Ph.D.

Deputy Director / Natural Resources Canada

Canada

<https://nrcan.gc.ca>



Outline

- 1 Open Science and Data Platform
- 2 AI approach for Impact Assessments
- 3 User case from the Open Science and Data Platform
- 4 Final remarks
- 5 Questions

Open Science and Data Platform

- "Single Window" access to federal, provincial, and territorial data and regulatory information in support of Impact Assessment & Cumulative Effects;
- Built in collaboration and with support from many departments and agencies;
- Increasing usership, and impact: more than 70% of users have accessed OSDP to facilitate participation in impact assessment or regulatory processes, or to better understand/manage cumulative effects;
- “Making data at work” for various users: public, impact assessment practitioners, research communities, and Indigenous communities and governments.

The screenshot shows the Open Science and Data Platform (OSDP) website. At the top, there is a header with the Government of Canada logo and the text "Government of Canada" and "Gouvernement du Canada". A search bar is located in the top right corner with the text "Search Canada.ca". Below the header, there is a "MENU" button. The main content area is titled "Open Science and Data Platform" and includes a description: "The Open Science and Data Platform provides access to science, data, publications and information about development activities across the country that can be used to understand the cumulative effects of human activities to support better decisions in the future." Below this, there are two main search sections: "Search by keywords" and "Search by Map". The "Search by keywords" section includes a search bar with the text "Use quotes to match phrases (e.g., emissions 'air quality')", a "Search" button, and a "Keyword Search Help" link. Below the search bar, there is a row of statistics: "8,505 Datasets", "144,172 Publications", "579 Monitoring Stations", "21,128 Development Activities", and "155 Content Collections". Below the statistics, there is a "Search Options" section with a dropdown menu set to "Exact match on ANY words" and a "Return to previous search" link. The "Search by Map" section includes a "Search by Map" button and a "Explore by area" button. Below the "Explore by area" button, there is a description: "Search by location using the interactive map or add layers to the map." Below the search sections, there are three featured content cards. The first card is titled "Learn about cumulative effects" and includes a description: "Learn why it is important to assess the combined impacts of past, present and future human activities and natural processes." The second card is titled "Learn about development activities" and includes a description: "Consult information and data available on a wide range of topics relevant to development activities." The third card is titled "Curated collections" and includes a description: "Discover collections of datasets that can be used as a starting point for understanding cumulative effects." At the bottom of the page, there is a "Contact us" button and a "Tell us what you think!" button.

AI approach for Impact Assessments



Responsible Adoption

OSDP contribution to enhancing digital services and workforce productivity while ensuring security, accountability and responsible use of AI.



Departmental Strategies

Aligning with federal policies, addressing adoption barriers, and ensuring necessary infrastructure, tools, and training.



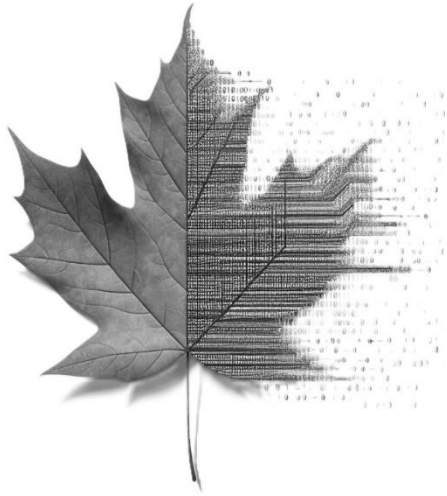
Expertise and Synergies

Collaborating with organizations in Canada in AI initiatives by identifying use cases, supporting the use of AI in impact assessments and regulatory processes.



Implementation Plan

Implementing OSDP-augmented with AI e.g. semantic search and Chatbot Pilot; and participating into the Government of Canada Public Service data/AI challenge “metadata & meta-tagging of scientific information management”.



Use Case from the Open Science and Data Platform

Problem statement:

- OSDP, in collaboration with internal and external partners have been experimenting with AI since 2021;
- Despite the improvements to the platform, users still find it difficult to obtain answers efficiently due to the vast amount of data and the limitations of keyword searches.

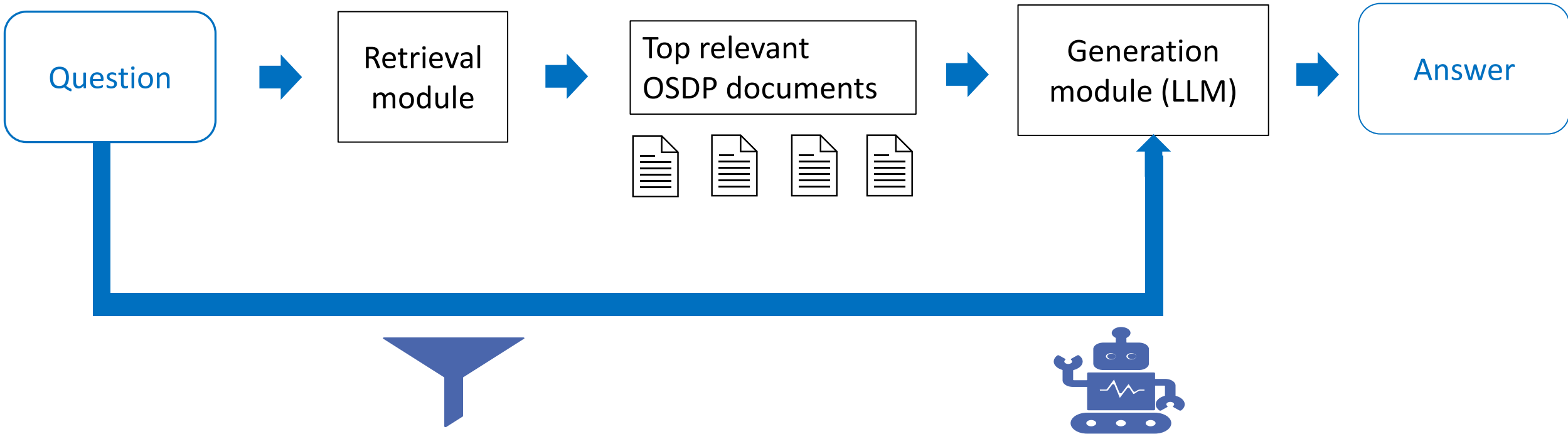
Goal:

- Create a proof of concept for a generative AI chatbot that answers user questions about **Impact Assessments and Cumulative Effects** using OSDP's extensive text content;
- Other consideration: enhance the search (augmented with semantic search); align with AI strategies.

Benefits:

- Help users understand the environment better and faster;
- Discover potential connections / patterns in the data that humans might miss;
- Ease of access to science information for everybody incl. the public;
- Shorten the way for users to get answers;
- Enhance user-friendliness of OSDP as a digital Information service and toolbox;
- Increase usership, usage and users over time;
- Support the Impact and regional assessments as well as understanding of Cumulative Effects.

OSDP Linguistic AI - Workflow



Dataset:

OSDP metadata records*
OSDP history of past user searches
keywords and filters**

*The data analysis / cleaning exercise for the metadata records helped to find issues such as content mismatch, formatting issues, language mismatch

** The history of past user searches (with keywords and filters) was only used to derive questions for the evaluation dataset for the RAG pipeline.

Machine Learning solution:

Retrieval-Augmented Generation (RAG)

The large language model (LLM) answers questions from relevant OSDP documents as opposed to the internet data it was trained on

The solution uses AWS, is secure, and respects data privacy

IMPACT ASSESSMENT AND CUMULATIVE EFFECTS



what is impact assessment and cumulative effects



Impact assessment and cumulative effects refer to the process of evaluating the potential consequences of human activities, such as natural resource development, on the environment and ecosystems. Cumulative effects specifically refer to the combined impacts of multiple activities or stressors on the environment over time, which can lead to significant and long-term changes to ecosystems and species.

In the context of environmental protection, impact assessment and cumulative effects assessment are critical tools for understanding and mitigating the potential harms of human activities. This involves identifying the potential risks and impacts of a project or activity, assessing the likelihood and magnitude of those impacts, and developing strategies to mitigate or manage them.

The Risk Assessment Framework for Cumulative Effects (RAFCE) is a tool that has been developed to support the assessment of cumulative effects in the context of regional assessments. It provides a structured approach to identifying and prioritizing regional risk and cumulative effects issues, and helps to quantify and rank the impacts of different activities or stressors.

Overall, impact assessment and cumulative effects assessment are essential components of environmental protection and sustainable development, as they help to ensure that human activities are carried out in a way that minimizes harm to the environment and promotes long-term ecosystem health.



Semantic Search – Ideal if you are looking for environmental reports and datasets based on meaning rather than exact keywords.



Chatbot Assistant – Best for quick, fact-based answers to environmental queries.

Sources:

- [1] [Cumulative impact assessment: monitoring land surface condition at regional scale using satellite remote sensing](#) (1.0)
- [2] [Rapport sur le Cadre d'évaluation des risques relatifs aux effets cumulatifs \(CEREC\)](#) (1.0)
- [3] [Report on the Risk Assessment Framework for Cumulative Effects \(RAFCE\)](#) (1.0)
- [4] [Risk assessment framework for cumulative effects \(RAFCE\)](#) (1.0)
- [5] [Overview of Cumulative Effects Research at Natural Resources Canada 2018-2023](#) (1.0)
- [6] [Cumulative impacts from anthropogenic activities and stressors on marine ecosystems in Pacific Canada](#) (0.99)
- [7] [Inclusive approaches for cumulative effects assessments](#) (0.99)
- [8] [EO baseline data for cumulative effects, year end report \(FY 2019/20\)](#) (0.99)
- [9] [A Global review of cumulative effects assessments of disturbances on forest ecosystems](#) (0.99)
- [10] [Cumulative Effects of Natural and Anthropogenic Disturbances under Climate Change in Western Canada: Assessing the Trade-Offs Among Long-Term Conservation of Priority Species and Places, Carbon Sequestration and Land Management. Note 7](#) (0.98)

Results to date – Examples (2/4)



show me some results about critical minerals in Canada



Here are some results about critical minerals in Canada:

1. Canada has abundant resources of critical elements, including antimony (Sb), cobalt (Co), indium (In), lithium (Li), niobium (Nb), and rare earth elements (REE).
2. The Critical Minerals Mapping Initiative (CMMI) is a joint research program between the Geological Survey of Canada, the U.S. Geological Survey, and Geoscience Australia that aims to improve understanding of the content and distribution of critical metals and minerals in mineral deposits.
3. The Geological Survey of Canada has developed a web portal that showcases the geochemical composition of ores and ore-bearing rocks, which can be used to facilitate a standardized input of geochemical data.
4. Canada is well-positioned to contribute to the global supply of critical minerals, with a series of innovative policy and research initiatives aimed at generating precompetitive baseline geological, geochemical, and geophysical data for large, underexplored regions within Canada.
5. The Targeted Geoscience Initiative program has contributed to the development of mineral potential models, which can be used to identify new sources of critical minerals.
6. Metasomatic iron and alkali-calcic (MIAC) mineral systems in Canada have the potential to supply a variety of critical metals, including iron oxide-copper-gold (IOCG), iron oxide-apatite (IOA), and affiliated deposits.
7. The Geological Survey of Canada has developed an alteration (paragenetic) model that relates the alteration facies of MIAC systems to their distinct metal associations and deposit types, which can be used to establish vectors to mineralization within MIAC systems.
8. Educational resources, such as the "Critical Minerals" Case information guide, are available to facilitate scientific awareness and outreach activities for geoscientists and the general public.



Semantic Search – Ideal if you are looking for environmental reports and datasets based on meaning rather than exact keywords.

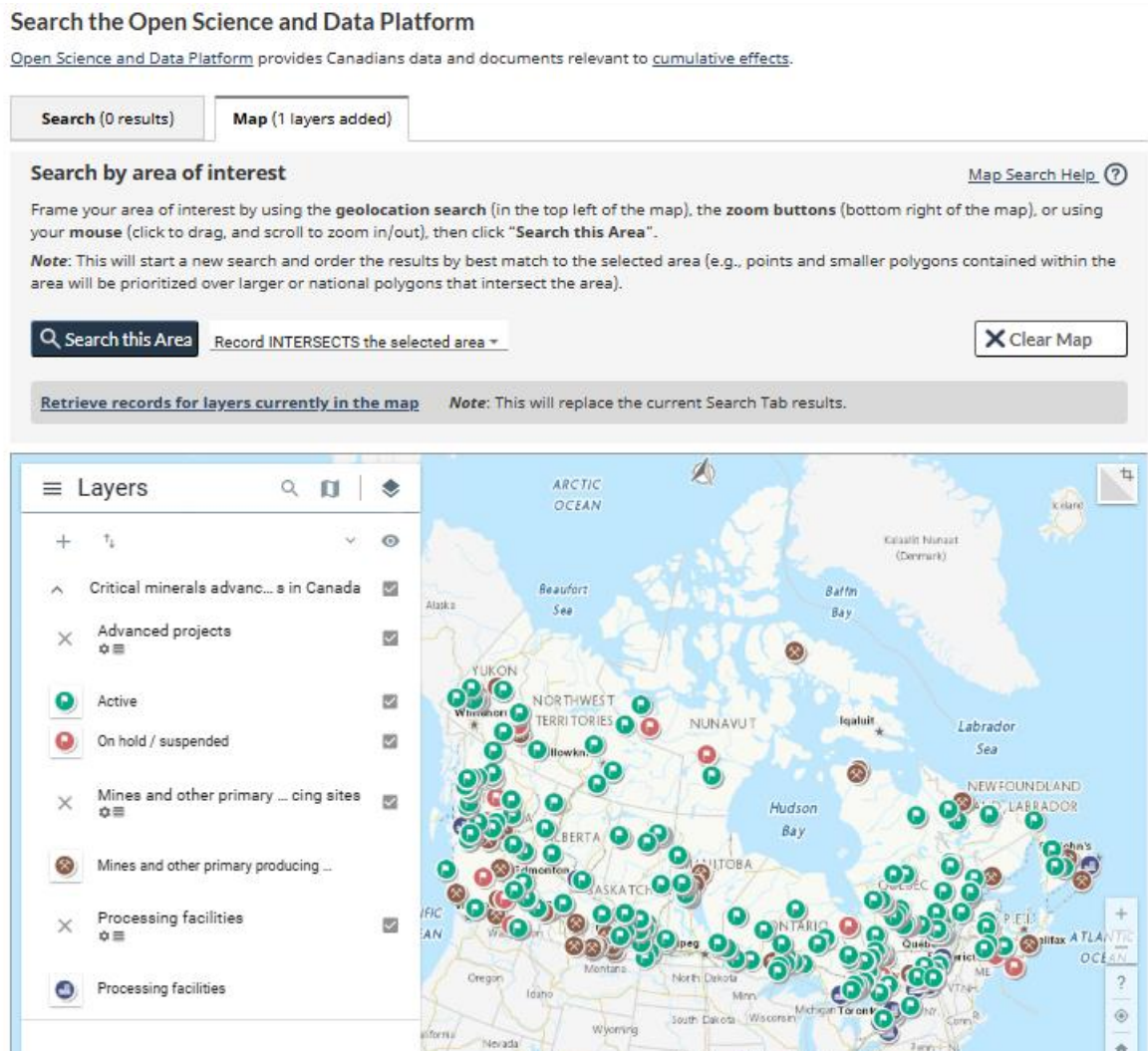


Chatbot Assistant – Best for quick, fact-based answers to environmental queries.

Sources:

- [1] [Critical minerals advanced projects, mines and processing facilities in Canada](#) (0.99)
- [2] [Geoenvironmental characteristics of Canadian critical metal deposits](#) (0.99)
- [3] [Educational case "critical minerals"](#) (0.99)
- [4] [Critical metals in mineral systems with iron oxide copper-gold \(IOCG\) and affiliated deposits](#) (0.99)
- [5] [Public geoscience solutions for diversifying Canada's critical mineral production](#) (0.99)
- [6] [Critical mineral geochemical database development and management](#) (0.99)
- [7] [Toward an effective global green economy: the Critical Minerals Mapping Initiative \(CMMI\)](#) (0.99)
- [8] [Mallette éducative "minéraux critiques"](#) (0.99)
- [9] [Magnetotellurics](#) (0.99)
- [10] [Primary critical metals deposits and their paths to discoveries in mineral systems forming IOCG deposits](#) (0.99)

Results to date – Examples (3/4)



[1] Critical minerals advance projects, mines and processing facilities in Canada

Disclaimer:
This map is for illustrative purposes. The markers represent the approximate locations based on available data. The Government of Canada shall not be held liable for any third party's interpretation of the information. The Government of Canada reserves the right to change or revise the information at any time.

Critical mineral geochemical database development and management

Publications

(Plain Language Summary, not published)

The Critical Mineral Mapping Initiative - a trilateral agreement among Canada, Australia and USA - developed a web portal that showcases the geochemical composition of ores and ore-bearing rocks. This open file report describes Canada's contributions to this online database. Because of the complexities in data structures among different reports, a novel workflow was developed to facilitate a standardized input of the geochemical data. It is anticipated that this workflow can be used in future releases of geochemical data. [less](#)

Source	
Contributor Organization:	Natural Resources Canada (NRCan)
Contributor System:	GEOSCAN Publications
Publisher:	Natural Resources Canada
Author:	Gadd, M. G. Lawley, C. J. M. Lauzière, K. Létourneau, F. Pilote, J. -L. Morin, A. Haji Egeh, A.
Licensed under:	Open Government Licence - Canada
Public availability:	Open Access
Update Frequency:	Not Planned
Citation: <input type="button" value="Copy"/>	Gadd, M. G., Lawley, C. J. M., Lauzière, K., Létourneau, F., Pilote, J. -L., Morin, A. & Haji Egeh, A. (2022). Critical mineral geochemical database development and management. <i>Geological Survey of Canada, Open File, 8883</i> . https://doi.org/10.4095/329924

Time Relevance			
Contributor Dates Use these dates to assess the time-relevance of contributor content.		OSDP Dates Use these dates to find content that has been recently added or updated in OSDP.	
Contributor Published:	2022-04-19	OSDP Published:	2024-04-04T20:38:18.324Z
Contributor Updated:	2024-04-27T01:11:55Z	OSDP Updated:	2024-05-01T06:50:23Z

[6] Critical mineral geochemical database development and management

Results to date – Examples (4/4)



what are the impact of road construction on LNG pipelines



I don't know. The provided context does not explicitly mention the impact of road construction on LNG pipelines. While there are descriptions of various pipeline projects and their environmental assessments, there is no direct mention of road construction and its effects on LNG pipelines.

Sources:

- [1] [Coastal GasLink Pipeline Project](#) (0.34)
- [2] [Merrick Pipeline Project](#) (0.22)
- [3] [Coastal GasLink Pipeline Project](#) (0.21)
- [4] [Temporary Floating Pipeline Use](#) (0.19)
- [5] [FortisBC EGP Project - Pipeline Crossing - Indian River - British Columbia](#) (0.15)
- [6] [Prince Rupert Gas Transmission Project](#) (0.13)
- [7] [Computer analysis of the effects of mine subsidence on pipelines](#) (0.11)
- [8] [Pipeline Blood Timber Limits IR 148A](#) (0.11)
- [9] [Gas Pipeline Installation Along Old Chelsea Road](#) (0.1)
- [10] [Telluric Influence on Pipelines](#) (0.09)

Results to date – Considerations

- Use of Full metadata records, for each record, select metadata fields incl. location.
- Machine learning models can have quite different computational needs too, and the AWS services and costs to deploy them can therefore vary.
 - Aim to select the best ones – results wise e.g. Llama 3 70B, which is a bigger model, so it helps with the quality of the answer.
- We are also considering accessibility incl. availability of the Chatbot in both official languages i.e. English and French.
- Continuous testing and finetuning the embedding and re-ranking models.
- Full production architecture, risk assessment and cost estimation require full some attention.
- Using the solution in support of Impact Assessments and understanding of Cumulative Effects.
- Retaining and augmenting usership.

Final Remarks

- **Alignment with Business Goals:** An effective AI strategy should be closely tied to the organization's overall business objectives to drive tangible value and ensure that AI initiatives contribute directly to growth and competitive advantage.
- **Data Governance and Quality:** Robust data management practices are essential for AI success. Prioritizing data quality, privacy, and security will ensure that AI models are built on reliable and ethically sourced data.
- **Scalability and Flexibility:** AI solutions must be scalable and adaptable to accommodate future growth, emerging technologies, and evolving business needs, ensuring long-term sustainability.
- **Continuous Learning and Adaptation:** AI should be viewed as a continuously evolving field, and an AI strategy should include regular updates, model retraining, and the integration of feedback loops to improve performance and accuracy over time.
- **Collaboration and Talent Development:** Foster a culture of collaboration between data scientists, domain experts e.g. Impact Assessment practitioners, and invest in upskilling teams to ensure the effective implementation and maintenance of AI systems.



HAVE A QUESTION?



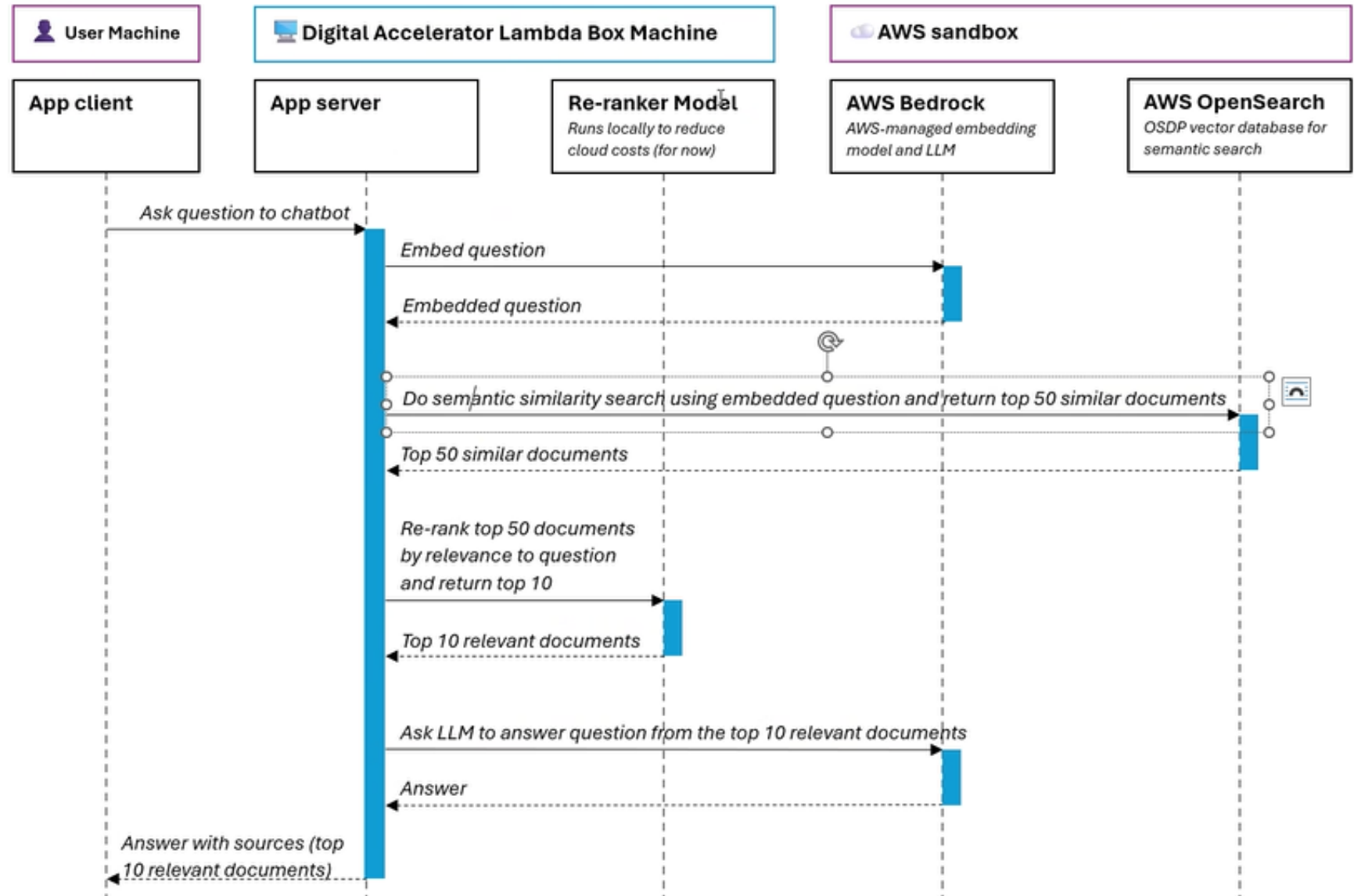
Extra Slides

Open Science and Data Platform meta- Database for AI

Comprehensive data analysis and cleaning were performed on data. Full extract of all the metadata records: **173 551 entries**.

- **Content type:** Most of the records represent publications. The description for this content type can sometimes be a paper abstract, which is useful for question answering.
- **Relevance to environmental cumulative effects:** Not all record descriptions are relevant to addressing questions on environmental cumulative effects. However, there is no specific set of filter values that consistently isolates relevant records. This challenge is partly due to inconsistent tagging practices across entries.
- **Language:** Most records are in English and/or French, with English records outnumbering French records by approximately three to one. Records in other languages are also present but form a minority.
- **Geographical metadata:** Nearly 30% of records lack an associated province, indicating incomplete geographical tagging.

Open Science and Data Platform Chatbot Sequence Diagram



Visit the OSDP:

How to keep in touch?

Web: <https://osdp-psdo.canada.ca>

Email: osdp-psdo@nrcan-rncan.gc.ca

Feedback Form:

Tell us what you think!

<https://www.surveymonkey.ca/r/Y2ZW5K8>

Additional Resources:

About the Open Science and Data Platform:
<https://osdp-psdo.canada.ca/en/about-osdp>

What's New Page:
<https://osdp-psdo.canada.ca/en/whats-new>

Take a quick tour:
https://www.youtube.com/watch?v=zAM_HpKm2Ks

Appreciation and acknowledgement: OSDP team incl. [Digital Accelerator](#)



Let's continue the conversation!

Message me your questions or comments in the IAIA25 app.

Mohamed Habbane, Ph.D.

*Deputy Director / Natural Resources Canada
Canada*

mohamed.habbane@nrcan-rncan.gc.ca

<https://osdp-psdo.canada.ca>

#iaia25