



November 3rd, 2025

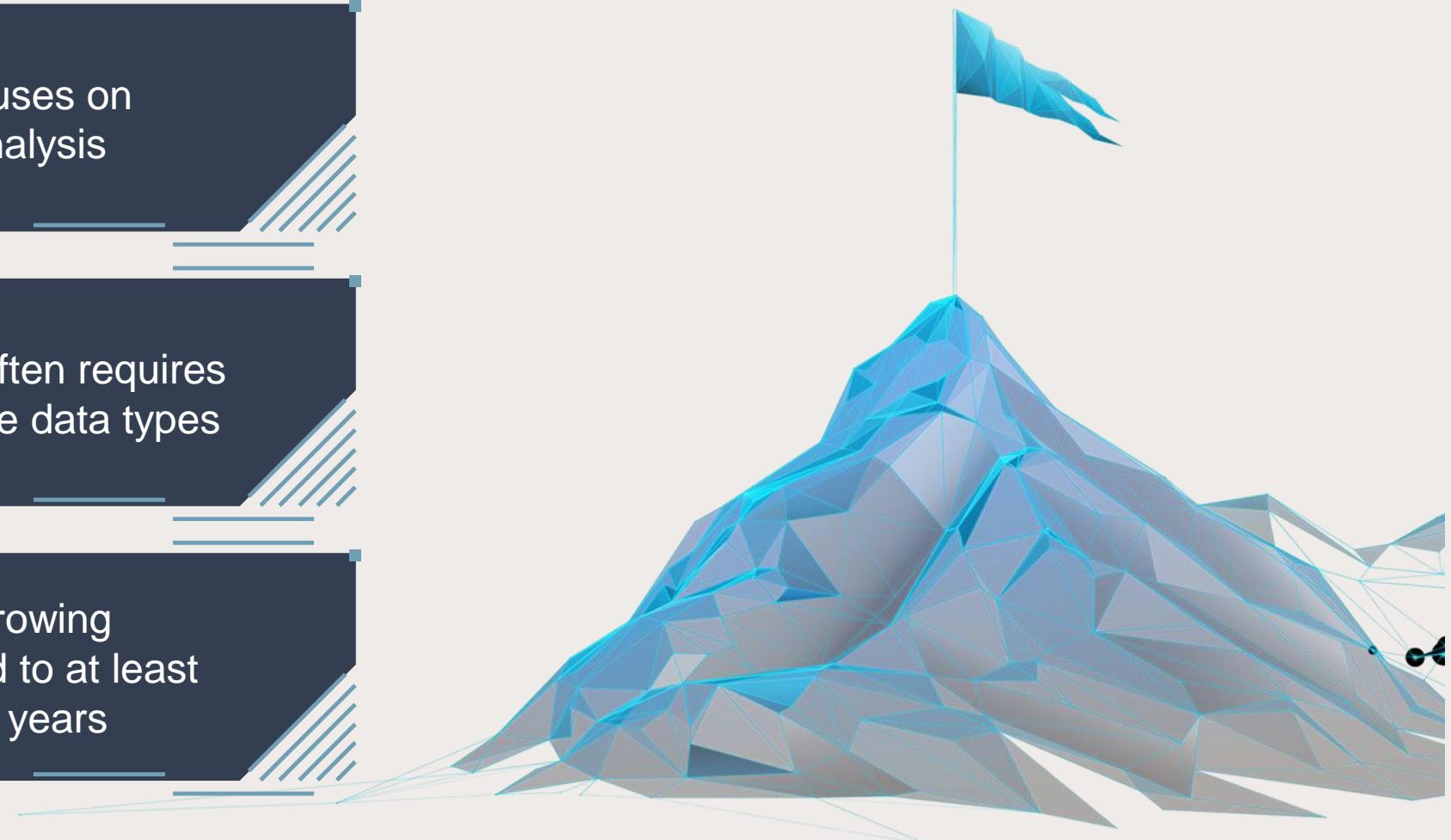
# Multimodal LLMs for GEOINT Looking Towards the Future

Pedro Uria, David Siamon, Nicholas Harvey  
Vibrint ML Team



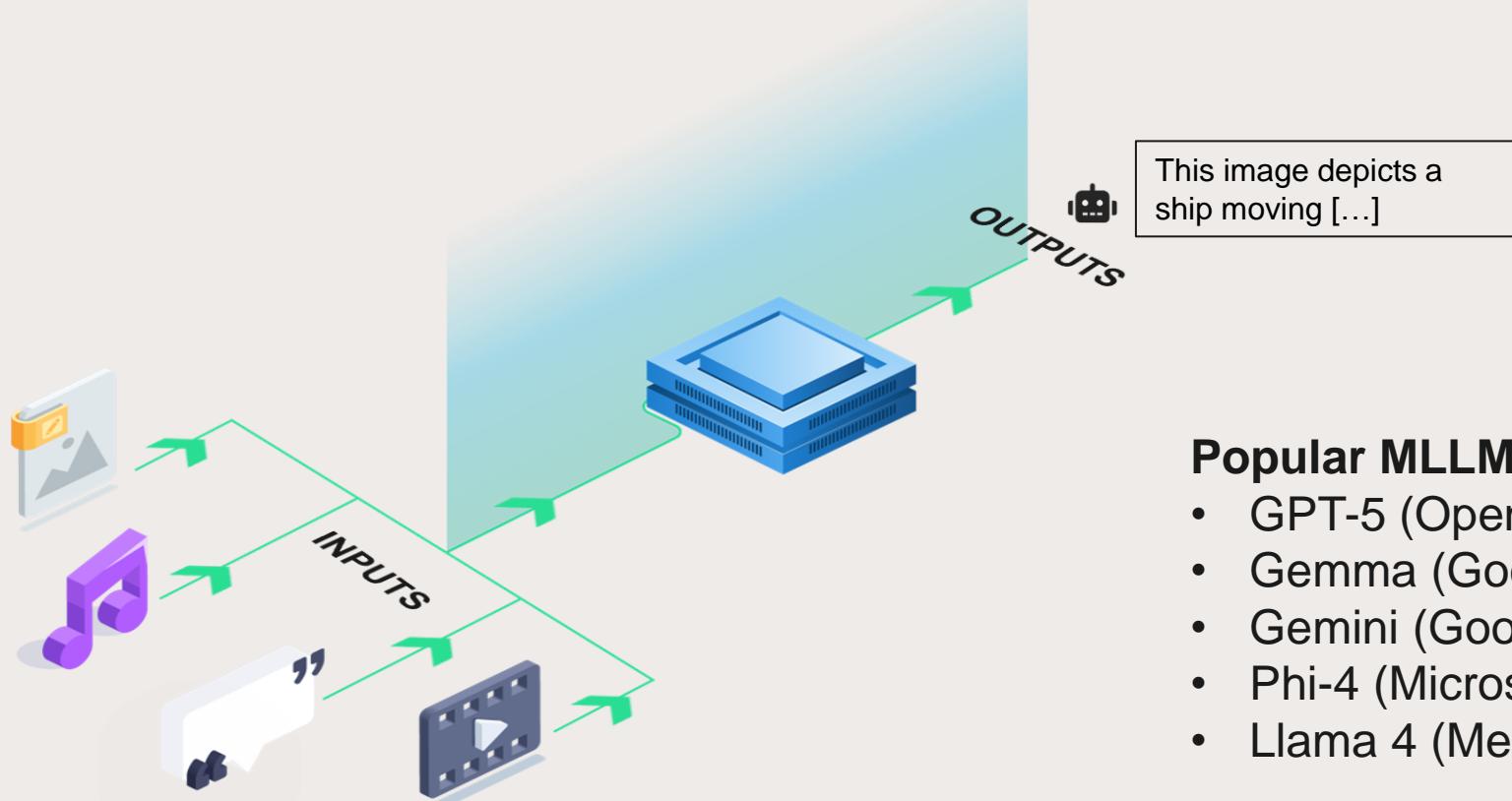
# Challenges with traditional GeoAI approaches

- GeoAI typically focuses on single-modality analysis
- Actionable intelligence often requires the integration of multiple data types
- GEOINT data is growing exponentially, expected to at least triple in the next 5 years



# Multimodal Large Language Models (MLLMs)

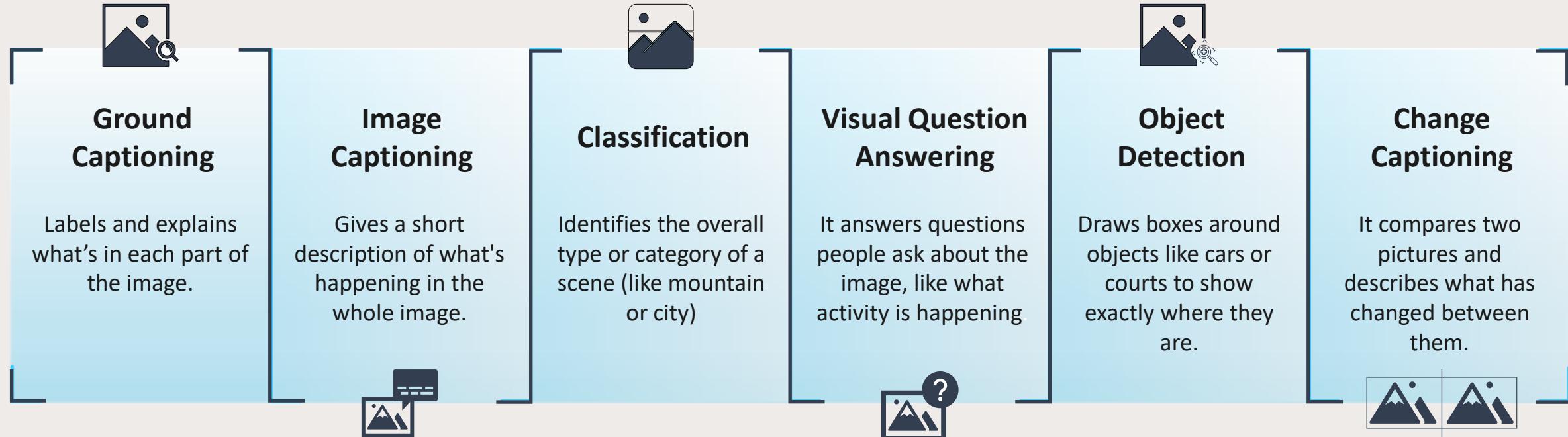
MLLMs are advanced AI models designed to process and integrate data from *multiple modalities*—such as text, images, audio, and video—with a single architecture.

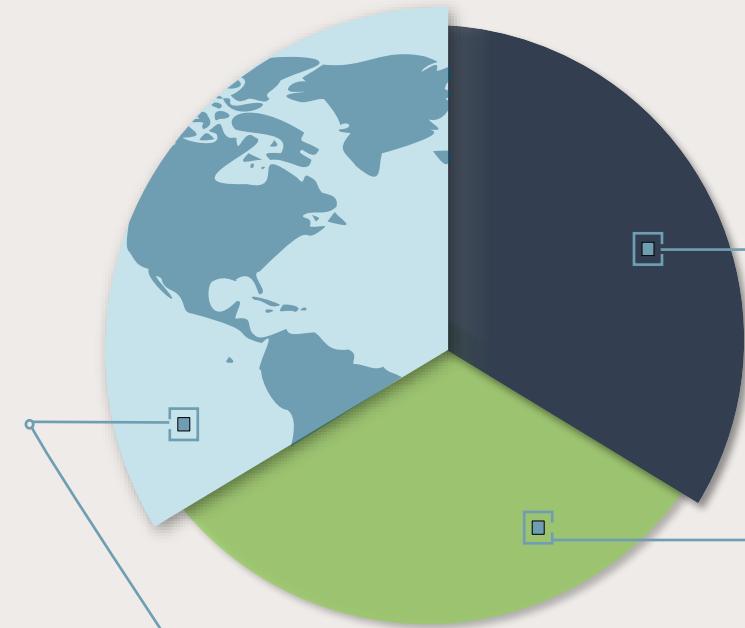


## Popular MLLMs:

- GPT-5 (OpenAI)
- Gemma (Google)
- Gemini (Google)
- Phi-4 (Microsoft)
- Llama 4 (Meta)

# Remote Sensing MLLMs Capabilities





## Automated Imagery Triage and Object Recognition

Automatically detect, label, and prioritize objects or areas of interest  
Go beyond bespoke object detectors by describing and contextualizing

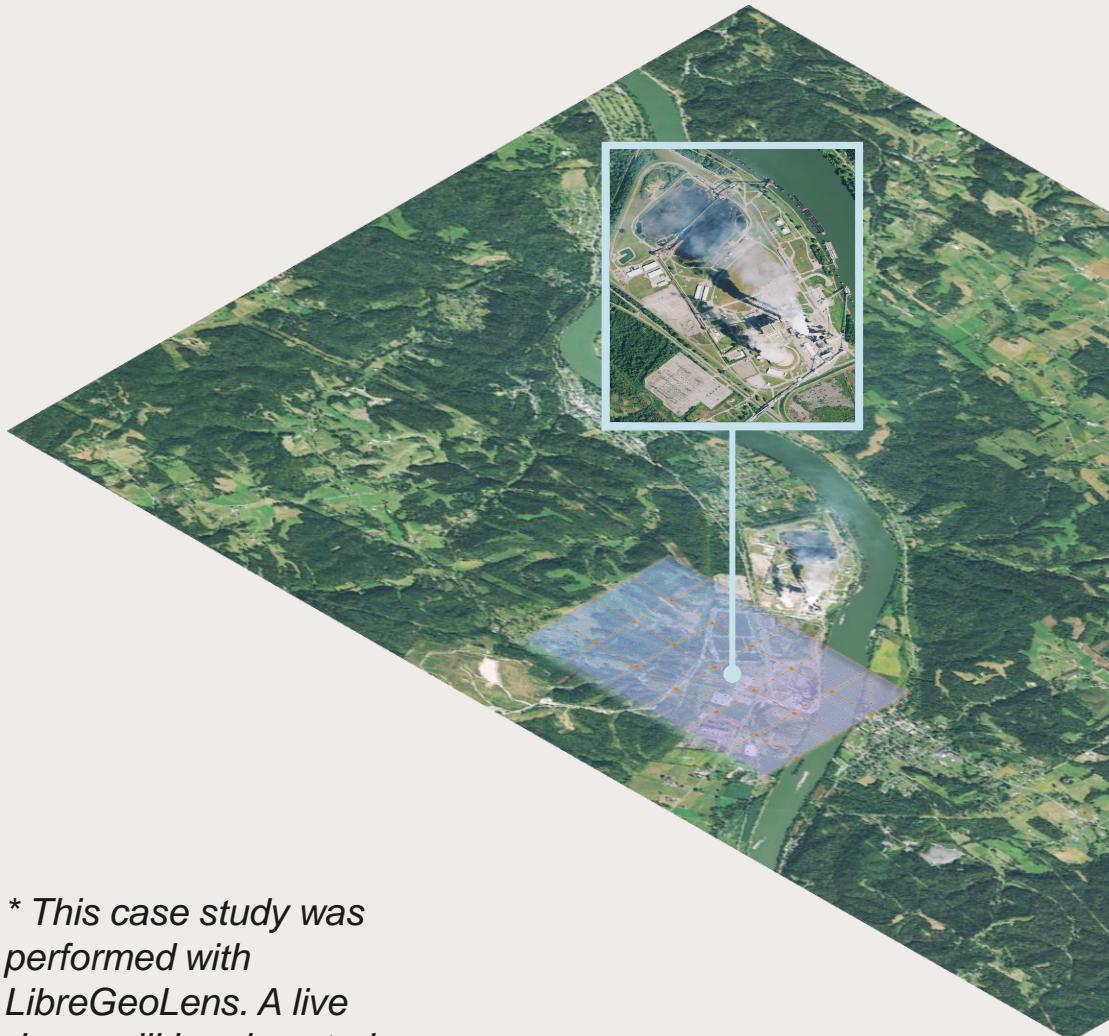
## Multi-source data fusion

Combine overhead images, social media posts, textual reports, GIS layers, etc.  
Incorporate situational context

## Interactive Q&A and Analysis

Interactive analyst's assistant  
Capable of answering questions about geospatial data in natural language

# Case Study | A Multi-Tiered Image Processing Approach for Contextual Analysis with MLLMs



## Layer 1: Wide-Area Search and Identification (Macro Analysis)

Identify potential power plants and tell me their locations in the image



The image shows potential power plant located near the center, adjacent to the river. It can be identified by its industrial features and infrastructure typically associated with power generation facilities.



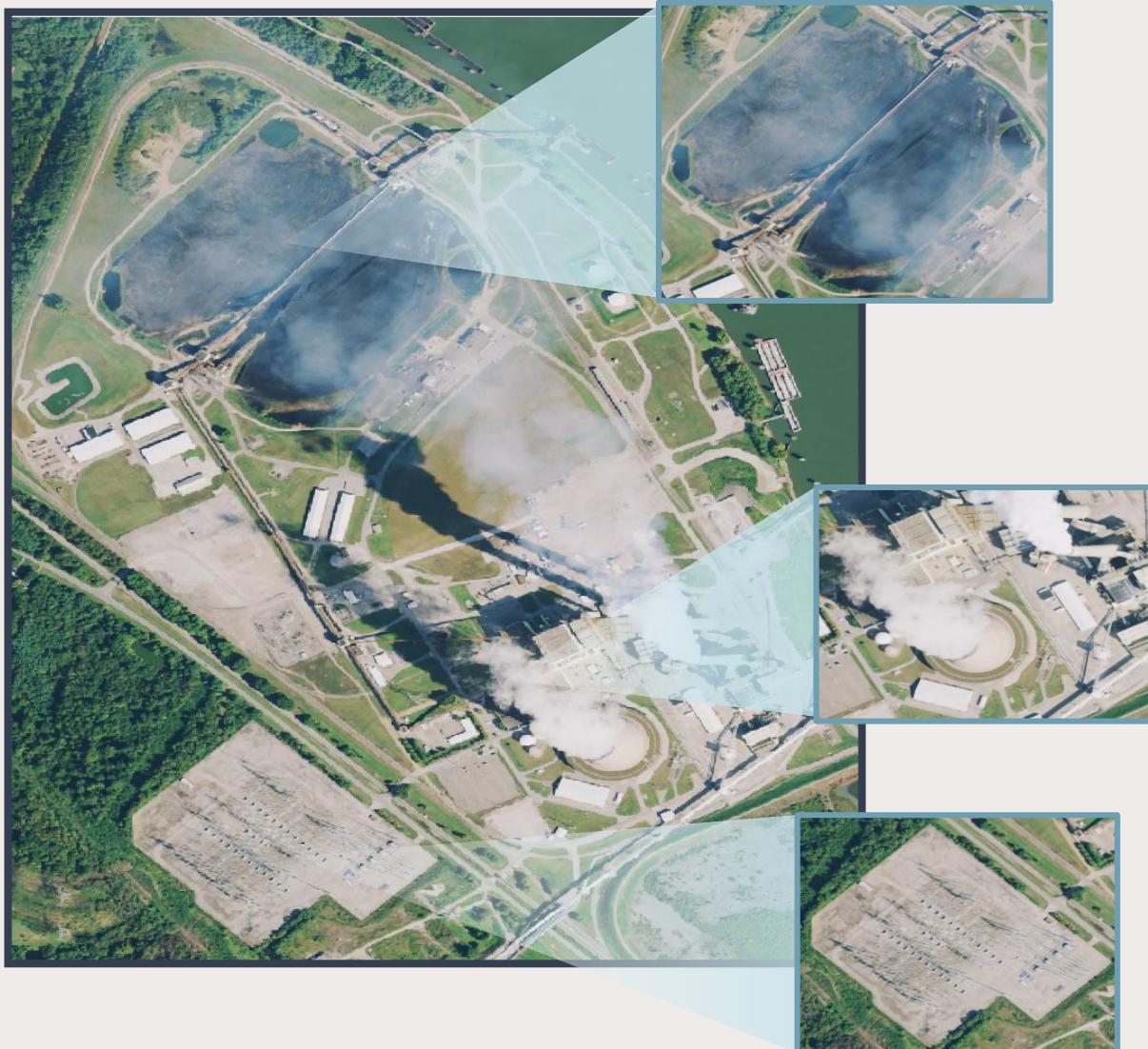
Without specific labels, it's challenging to determine the type of power plant, but the location along the river suggests it may rely on water for operations, potentially indicating hydroelectric, coal-fired, or gas-fired generation.



 Send

\* This case study was performed with LibreGeoLens. A live demo will be given today.

# Case Study | A Multi-Tiered Image Processing Approach for Contextual Analysis with MLLMs



## Layer 2: Facility-Level Delineation & Segmentation (Meso Analysis)

This is a closer look at power plant you identified in the previous image. Describe it in greater detail and adjust your previous analysis accordingly.



The closer view at the facility indicates it is likely a coal-fired power plant. Key features include:

- Coal Storage: (...)
- Emissions: (...)
- Infrastructure: (...)
- River Access: (...)

Overall, this characteristics support the identification of the facility as a coal-fired power plant.



 Send

# Case Study | A Multi-Tiered Image Processing Approach for Contextual Analysis with MLLMs



## Layer 3: Focused Activity Analysis (Micro Analysis)

These are relevant subregions that you've identified.  
Describe them in greater detail.



Certainly! Here is a detailed description of each subregion:

1. Coal Storage Area
  1. (...) contains large coal piles (...) fuel for (...)
  2. Conveyor belts run across the piles (...) transport
2. Electrical Substation:
  1. (...) grid-like (...) electrical equipment.
  2. (...) connection point for electricity generated (...)
  3. Presence of transformers and switchgear is likely.
3. Main Plant Structure:
  1. (...) cooling towers, chimneys (...) steam & smoke.
  2. (...) heat dissipation.
  3. (...) turbines, generators (...)
  4. (...) core operational area of the power plant.



## Layer 4: Contextual Integration and Querying

Taking all this context into account, can you summarize your findings about this region into a detailed report?



The region examined is a coal-fired power plant situated along a river, utilizing the water source for operational needs. (...)  
Key components: (...)

- Operational Insights:
- River Proximity: (...)
  - Environmental Considerations: (...)

This analyzed region is a notable example of an operational coal-fired power plant. Its strategic riverside location and comprehensive infrastructure showcase the typical features and challenges associated with coal-based energy generation. This assessment provides a foundational understanding of the plant's role in regional power supply and its environmental and logistical dynamics.



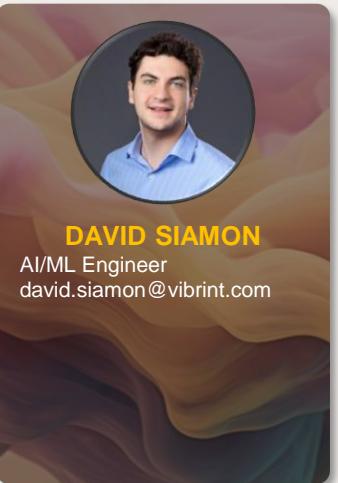
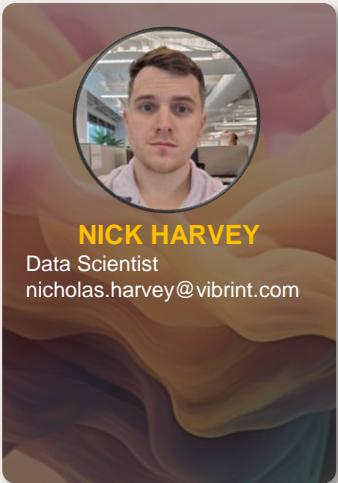
## Challenges

- Grounding via retrieval / prompting is essential to reduce the amount of vagueness and hallucinations
- Without metadata like image resolution or location, MLLMs struggle to adapt to different environments or scenarios
- Can't ingest full-scale imagery, so external infrastructure is needed to manage context

## Limitations

- High compute/storage costs and ongoing maintenance impact scalability and deployment
- Fine-tuning with expert input might be needed for reliable performance in specialized use cases
- MLLMs struggle to convey numerical confidences in their outputs

# Contact Us



## Open-Source QGIS Plugin

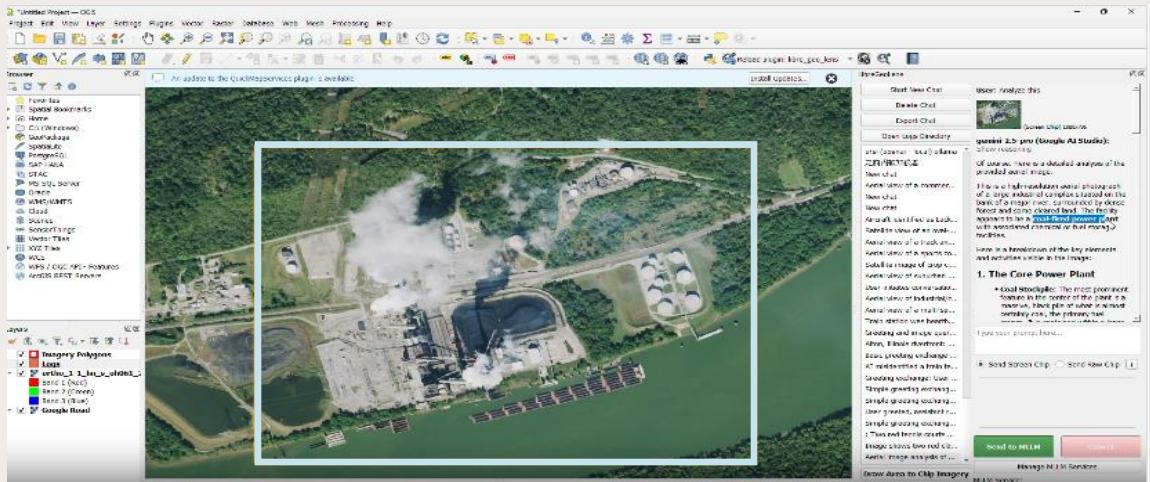
 **LibreGeoLens**  
by Ampsight

Experiment with MLLMs to analyze remote sensing imagery

 **Experimental**

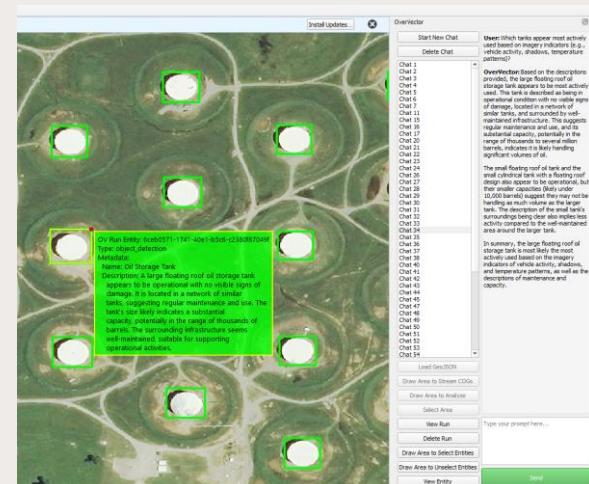
 Oct 15, 2025  (6)

 Download  1,258



# OVERVECTOR

An AI assistant to help you analyze  
remote sensing imagery



# Thank you