IAF EARTH OBSERVATION SYMPOSIUM (B1)
Earth Observation Data Systems and Technology (4)

Author: Mr. Juan Carlos Lopez
United States

Dr. Nelli Babayan
United States
Mr. Rob Emanuele
United States
Mr. Taylor Corbett
United States
Mr. Minh Nguyen
United States

# NATURAL LANGUAGE EXPLORATION OF SATELLITE DATA WITH ARTIFICIAL INTELLIGENCE, LARGE LANGUAGE MODELS, AND PLANETARY COMPUTER

## Abstract

Satellite data is a valuable source of information for understanding the Earth and its changes. However, accessing and analyzing satellite data can be challenging for non-experts, as it requires specialized tools, knowledge, and skills. In this paper, we present a prototype that uses artificial intelligence (AI) to enable natural language exploration of satellite data, including data from NASA, ESA, and others. The prototype integrates two Microsoft technologies: Planetary Computer, an open cloud platform with petabytes of Earth observation (EO) data, and Copilot, an orchestration engine for AI-powered large language models (LLMs). The prototype allows users to ask questions about satellite data using natural language prompts, such as "chat with the data about the changes in the planet". The prototype then uses Copilot to process the queries and generate responses through the LLMs, such as GPT-4. The responses include visualizations, summaries, or insights, depending on the query and the data. The prototype also uses the Planetary Computer's API to search for and access relevant data from the data catalog. The user can interact with the prototype through an interactive display in Planetary Computer, where they can see the text and visualizations generated by the LLMs. The prototype also provides links to data sources and external sites to explore and learn more. The authors will describe the architecture that is powering this prototype so that users can recreate some of these capabilities in their own cloud environments. The prototype demonstrates the potential of using AI and natural language to make satellite data more accessible and understandable for a wider audience. We discuss the challenges and opportunities of this approach, as well as the future directions for research and development.