

The Rock Whisperer: Integrating Automated EDA Workflows with LLM

Putra Sadikin¹, Steve Prager, Arthur Sedek, Pouria Mehrani

¹Imdex, Melbourne, Australia

Traditional exploratory data analysis (EDA) in geoscience benefits from visual computation of large, multivariate datasets. Augmenting EDA with dimension-reducing workflows helps geoscientists derive insights and identify structures, such as correlations or similarities. This process integrates data interpretation with geological context, enhancing applications like exploration vectoring and rock characterization, which are crucial for ore body knowledge development.

Further enhancement involves the templating of exploratory data analysis through recordable workflows. Recent developments have enabled these workflows in Headless GAS, a headless cloud-native environment of IMDEX ioGAS, allowing for automation and standardisation of complex geochemical data analysis processes, ensuring consistency and reproducibility. Headless GAS retains all the functionality of ioGAS, enabling cloud-based workflow processing and handling unlimited data sizes, ideal for "big data" from real-time sensors and scanning platforms.

Headless GAS is limited to recording workflows within the program, restricting workflow creation and dissemination to SMEs. This presentation explores integrating a headless cloud-native environment with LLMs to automate and scale workflow creation. Natural language requests replace traditional workflow recording, dynamically translating into actionable workflows. Domain knowledge and workflow paths are captured in knowledge graphs for LLM fine-tuning. This simplifies workflow recording, leveraging a domain-trained LLM for sophisticated analytics in ioGAS. Geoscientists describe desired outcomes, with the LLM determining optimal paths, shifting focus to goal-oriented analysis. Embedding an LLM enriches result context, interpreting visualizations and recommending further analysis steps to achieve desired outcomes.

Another key advantage is the potential for creating workflows without a traditional interface, which is particularly beneficial in coreshed applications where hands-free operation is crucial for safety. This method ensures reproducibility and scalability, with the LLM acting as a cognitive layer that bridges human-centric commands with the functionalities provided in the headless GAS environment. This foundational step paves the way for natural conversational workflows to interact with traditionally complex desktop applications.