Formulating a GPT-4 assisted exploration framework

R. Daruso[1], D. Gamble, S. Pattanaik, B. Russell, J. Taylor and G. Heinson
[1]The University of Adelaide, rumi.daruso@adelaide.edu.au

SUMMARY

The exponential accumulation of mineral exploration data presents an opportunity for the integration of large language models (LLMs) into the mineral exploration framework, enhancing cost and time efficiency in all stages of data assembly, analysis, and distribution. Focusing on gold prospectivity within the underexplored regions of the Woomera Prohibited Area in central South Australia, a comprehensive pre-competitive dataset—including the Gawler Phase 2 aeromagnetic and magnetotellurics dataset from the Geological Survey of South Australia—was compiled into a 250 x 250 km data cube and analyzed using GPT-4. Clustering models showed great similarity to current inferred solid geology maps, providing more detailed lithological boundaries. While extrapolated data maps might have low accuracy, they indicate areas where additional data is needed. One of the most beneficial functions of GPT-4 is format conversion, both through its server and through generation of Python code to convert formats it cannot directly work with. For instance, it successfully converted geophysical data into DICOM, a medical imaging format. The continuous development of GPT-4 and related OpenAI products suggests promising integration into future mineral exploration efforts.

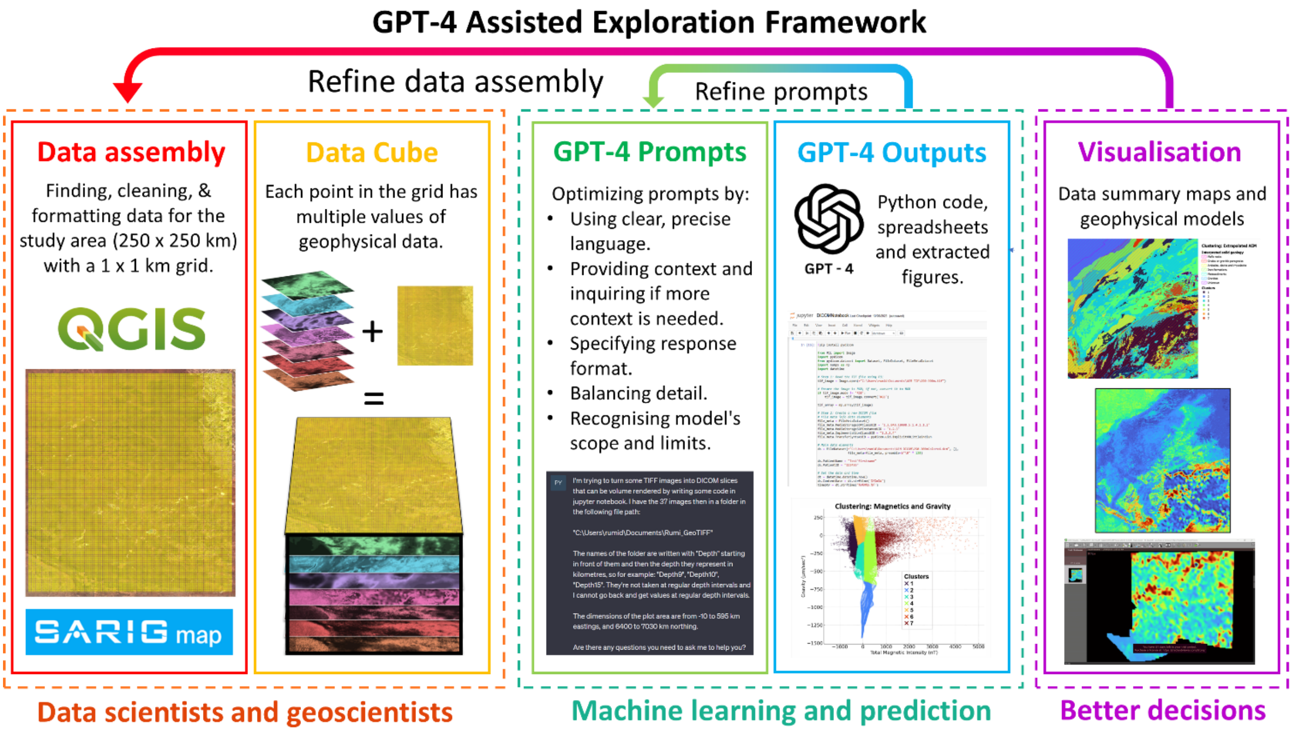

Figure 1. Visual summary

Keywords: GPT, Large Language Models, Artificial Intelligence, Data Analysis

INTRODUCTION

Amidst a growing global demand for critical minerals, extensive exploration is becoming imperative, particularly as data accumulates exponentially. This scenario offers both challenges and opportunities, paving the way for GPT-4 to serve as an essential tool in data collection, processing, and analysis. To this end, a GPT-4 assisted framework was developed under the New Generation Explorers

Award. The framework was applied to the largely unexplored Woomera Prohibited Area (WPA) in central South Australia, a region with high potential for gold due to its proximity to the Central Gawler Au Province and notable deposits like Olympic Dam, as in Figure 2. While the southeastern parts of WPA feature relatively high data resolution, the focus was on the northwest with lower data resolution, identifying it as a critical area for gold exploration. This methodological approach helps pinpoint potential zones for further exploration, facilitating more directed and effective mining efforts in promising yet challenging terrains.

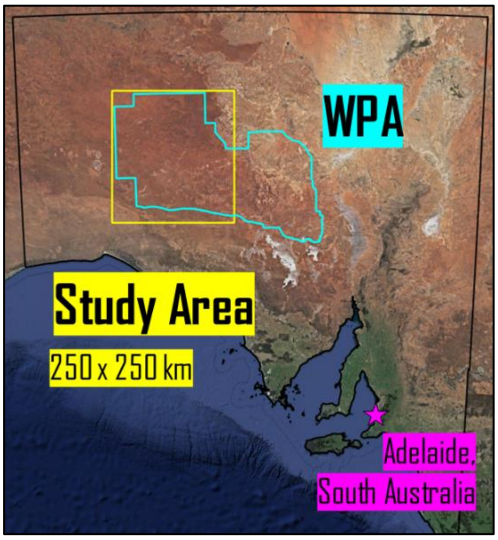

Figure 2. Location of the study area in South Australia.

METHODS

The version of GPT-4 used for the study was available on the OpenAI site between July – October 2023. Compared to the free version of ChatGPT, the $35 USD/month subscription on GPT-4 allows users to upload files and query them. Files below 512MB can be used within the server, while python scripts can be requested to work with larger files. There have been many advancements in GPT-4 since the study was done, and therefore these processes are only a preliminary look into its current capabilities.

Data in the region WPA is sparse, incomplete, of different spatial coverage and have varying levels of depth information. The solution to integrating these datasets was to create a data cube as seen in Figure 3. The data cube, in the format of a CSV file, was then uploaded onto GPT-4 and analyzed. Prompts could be optimized by 1) using clear language, 2) providing context, 3) inquiring if more context is needed, 4) specifying response format, and 5) recognizing the model's scope and limits.

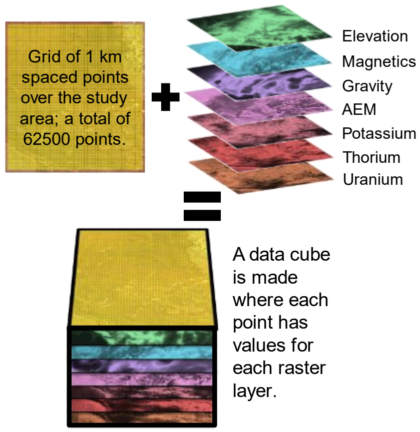

Figure 3. Diagram of the data cube creation process.

Clustering

Clustering is a data science method that assigns similar-valued points into groups. In geology, it can be used to group rocks with similar geophysical signatures together, as seen by the clustering of magnetics and gravity data in Figure 4. Another clustering process was done from elevation and concentrations of uranium, thorium and potassium. All extrapolated AEM data were also clustered to show possible patterns in the dataset.

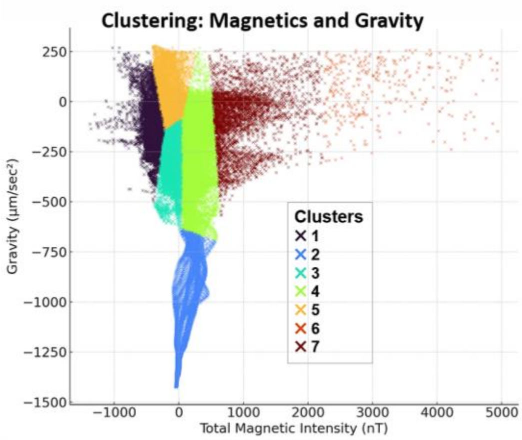

Figure 4. Clusters of magnetics and gravity values.

Extrapolation

The AEM data only overlaps the south-east quarter of our study area. We used GPT-4 to extrapolate this, as seen in Figure 5.

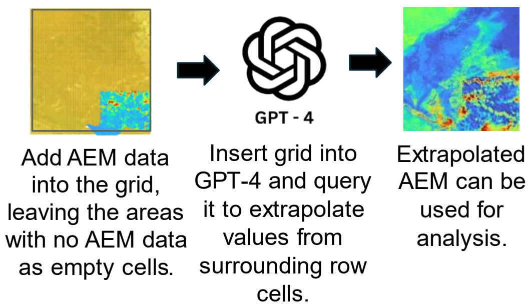

Figure 5. AEM extrapolation process using GPT-4.

Format Conversion

Geophysical data comes in a variety of formats that are often not compatible with each other. GPT-4 proved to be very useful in converting between formats, even sucessfully converting geophysical data into DICOM, a medical industry file type which can store both 2D and 3D data. DICOM files can be opened with any DICOM viewer, many of which are free and open-source. The proccess is displayed in Figure 6.

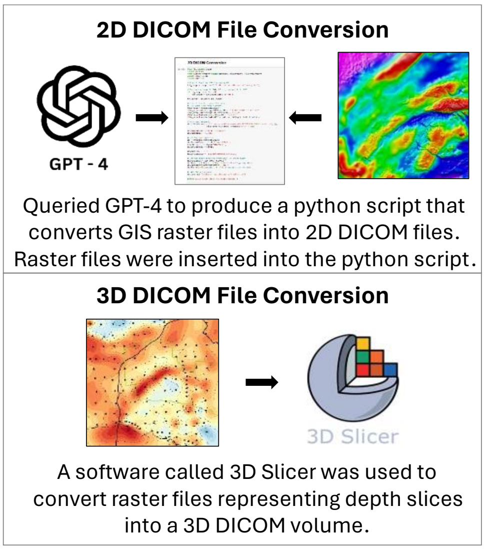

Figure 6. Conversion process to DICOM format.

Considering the success of the 2D conversion, it is likely that with enough time, we would have been able to create a code with the help of GPT-4 to do 3D conversion purely through Jupyter Notebook.

RESULTS

Clustering

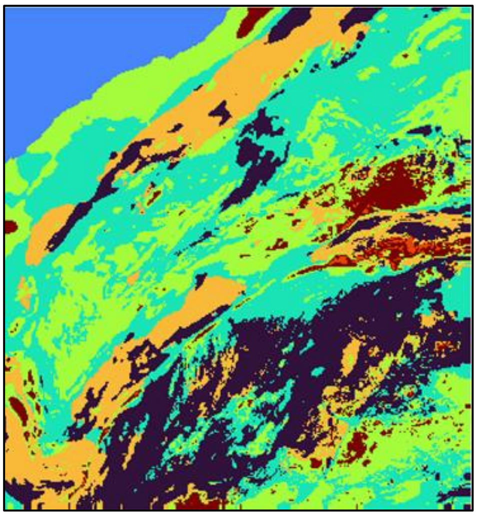

Figure 7. Clusters of magnetics and gravity data.

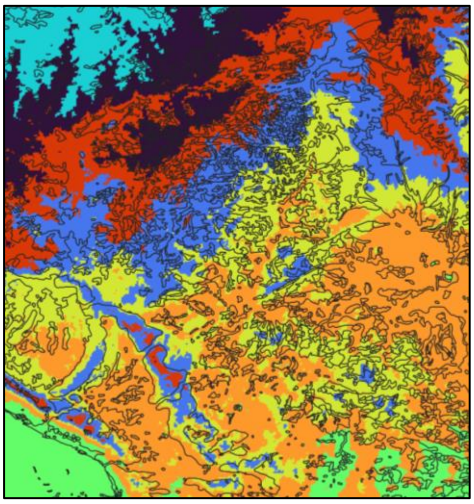

Figure 8. Clusters from elevation, U, Th and K. Black lines represent boundaries between surface lithologies.

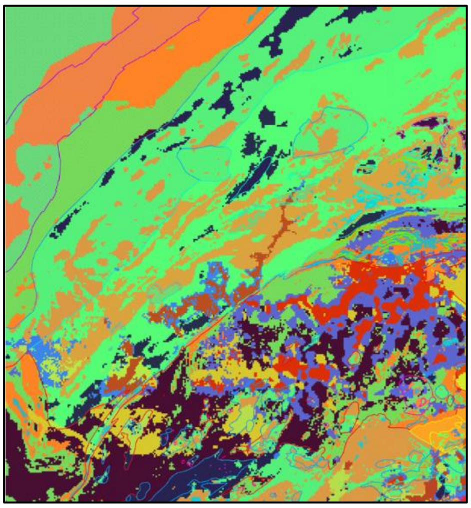

Figure 9. Clusters from extrapolated AEM data at all depths. Lines represent interpreted solid geology from SARIG as visualized in Figure 12.

Extrapolation

The extrapolated AEM shown in Figure 10 are purely speculative, but conveniently show relationships between elevation and paleochannels.

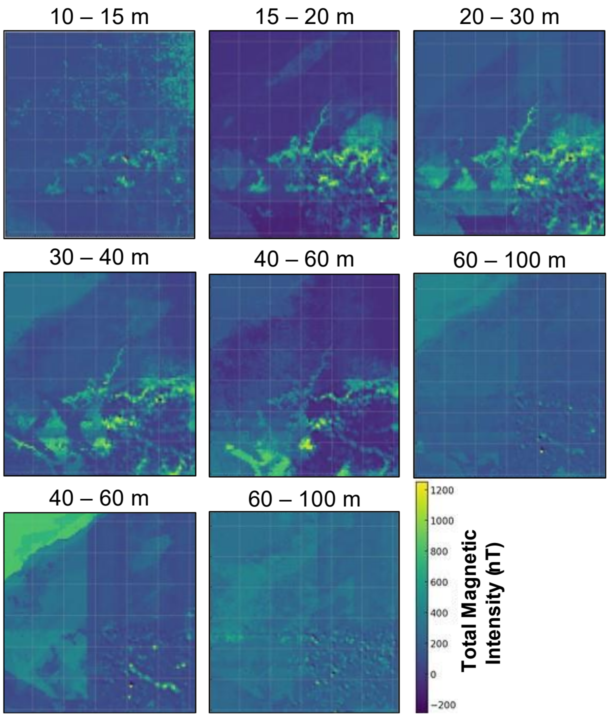

Figure 10. Extrapolated AEM data over the study area at various depths.

Format Conversion

2D and 3D models of geophysical data were successfully converted into DICOM, as shown in Figure 11. Large rasters can be viewed as a single DICOM slice. Data that are collected for different depths can be represented as DICOM slices that have depth and ordering metadata associated. Depth slices can be turned into 3D DICOM models can be rotated, moved, zoomed in and out, and rendered to different levels to identify anomalies.

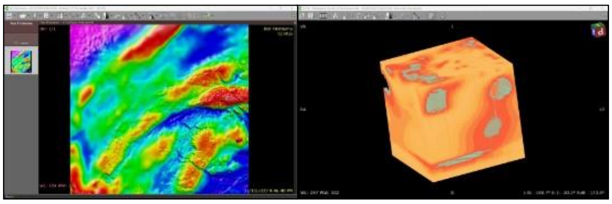

Figure 11. (Left) Gravity data for the studey area viewed on RadiANT DICOM viewer. (Right) MT Depth slices rendered into a 3D model in RadiANT DICOM viewer.

DISCUSSION

Comparison of the magnetics and gravity the clusters with solid geology interpreted by GSSA shows correlations between clusters and solid geology, as seen in Figure 12. The clusters seem to pick up more defined geological boundaries. K-means clustering on extrapolated AEM slices show paleochannel patterns, with red clusters being the deepest part of the paleochannel and violet clusters being the larger extent. Clustering K, U, Th concentrations and elevation gives a good representation of surface geology.

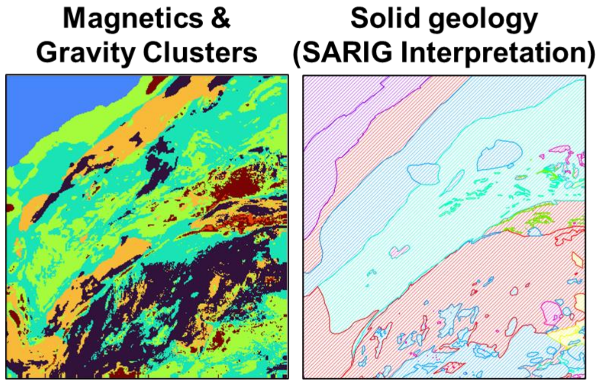

Figure 12. Comparison of generated clusters to open-source data on interpreted solid geology.

During data assembly, GPT-4 was useful to search for possible sites that host geophysical data, data cleaning, and providing software support for various QGIS functions.

CONCLUSIONS

The benefits of using GPT-4 in exploration include:

1. Making data analysis more accessible to individuals outside the data science profession.
2. Supporting a non-invasive exploration framework of detailed desktop studies that can efficiently target smaller hotspots before larger ground exploration.
3. Supporting a time and cost-efficient method of sharing data with various audiences through easy format conversions and quick data visualizations.
4. Bringing more awareness to emerging AI industries such as GPT-4 and training the AI literacy of future explorers.

ACKNOWLEDGEMENTS

Thank you to Dr. Carmen Krapf and Andrew McCulloch for mentoring the team alongside Professor Graham Heinson. Thank you to Philip Heath, Tim Chen and Russell Menezes for answering the team's queries.

We acknowledge the Maralinga Tjarutja, Anangu Pitjantjatjara Yunkunytjatjara, Antakirinja Matu-Yankunytjatjara, Arabana, Gawler Ranges and Kokatha peoples as traditional custodians of the lands in the WPA.