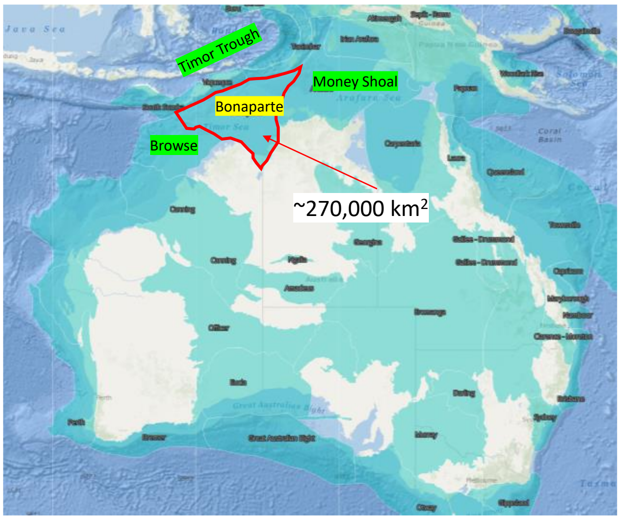
A Case Study of Understanding Bonaparte Basin using Unstructured Data Analysis with Machine Learning Techniques

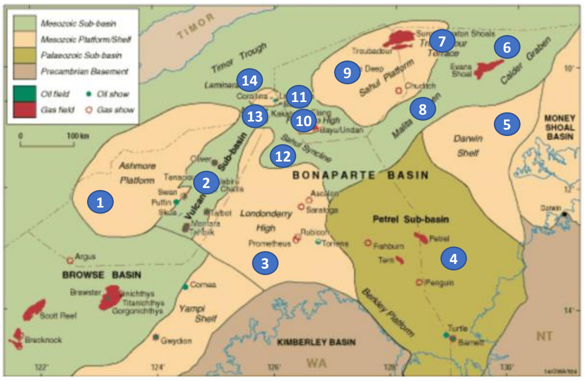
Francois Baillard (Iraya Energies), Nina Marie Hernandez (Iraya Energies)*

Introduction

As part of exploration and production the oil and gas industry produce substantial amounts of data within different disciplines of which 80% are unstructured like reports, presentations, spreadsheets etc and it is expected to grow exponentially. As a result, geoscientists and engineers spend 50 to 80% of their time searching and assembling data and only 1 to 5% of the data is fully utilized. The value of technical work is therefore reduced due to the lack of time available for analysis and critical thinking and the under-utilization of the data. To assist geoscientist and engineers, Machine Learning (ML) and Artificial Intelligence (AI) technologies are applied to process the unstructured data making it possible to perform more accurate analysis and make faster decisions.

In this case study the area of interest covers Bonaparte Basin, which is located north-west of the Australian continental margin (Figure 1). It joins the Money Shoal basin in the north-east and the Browse Basin in the south-west. Furthermore, the Timor Trough defines the northern boundary. The areal extent of the basin is approximately 270,000 sq. km. The objective of this study is to understand and obtain meaningful insights into the Bonaparte Basin based on the substantial amount of information available in previous studies, reports and presentations. The unstructured data of Bonaparte Basin have been ingested in a Knowledge Container through consecutive ML and AI pipelines and analysed using big data analytics tools.





Consist of several structural elements :

1. Ashmore Platform
2. Vulcan Sub-Basin
3. Londonderry High
4. Petrel Sub-basin
5. Darwin Shelf
6. Calder Graben
7. Troubadour Terrace
8. Malita Graben
9. Sahul Platform
10. Flamingo High
11. Flamingo Syncline
12. Sahul Syncline
13. Nancar Trough
14. Laminaria High

Figure 1 *Location of the Bonaparte Basin within the Australian continental margin (left) and 14 structural elements observed within the Bonaparte Basin (right).*

Methodology

As of 2021, the Bonaparte Basin encompasses 440 wells representing 58 years of exploration history summarized in over 270,000 pages of documents and in 250,000 images. It is estimated that billions of dollars have been invested over the year to record and interpreted such data, making it a tremendous source of information for new exploration activities.

The Play Based Exploration (PBE) approach is often used as a traditional framework to refine geologist's understanding from a broad basin level to a narrow prospect focus (Lottaroli et al., 2016). As a start such an approach often involves capturing the current state of the knowledge with massive background resources to understand and analysis the key features of the basin and the major risks associated to it. Such information is primarily available in unstructured data, requiring geologists to process and ingest such information before focusing on a specific play and prospect using structured data. Therefore, we have modified the existing PBE pyramid to introduce an additional dimension associated to the data science identifying the different type of data available at different stage, allowing us to better define the suited ML/AI strategy for a given stage (Figure 2).

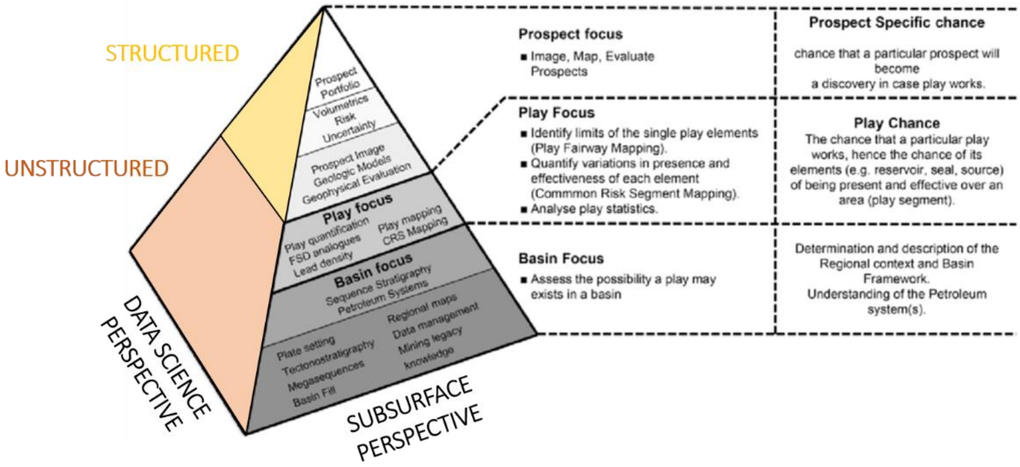


Figure 2 *Customized Play Based Exploration (PBE) pyramid (Modified from Lottaroli et al., 2016) with ML technology*

Focusing on the unstructured data associated to the Basin and Play Analysis, all the data from Bonaparte Basin has been processed through a succession of automated pipeline AI/ML such as Natural Language Processing or Deep Convolutional Neural Network (Hernandez et al., 2019) (Figure 3). The sharable structured data is then further processed through deeper level of analytics to detect trends and anomalies present within the data. Machine assistance is heavily used in repetitive tasks early in the process during the crushing of data, up to 95% of the tasks will be performed by the machine. This provides additional time to the human to focus on critical thinking and cognitive skills to interpret the data.

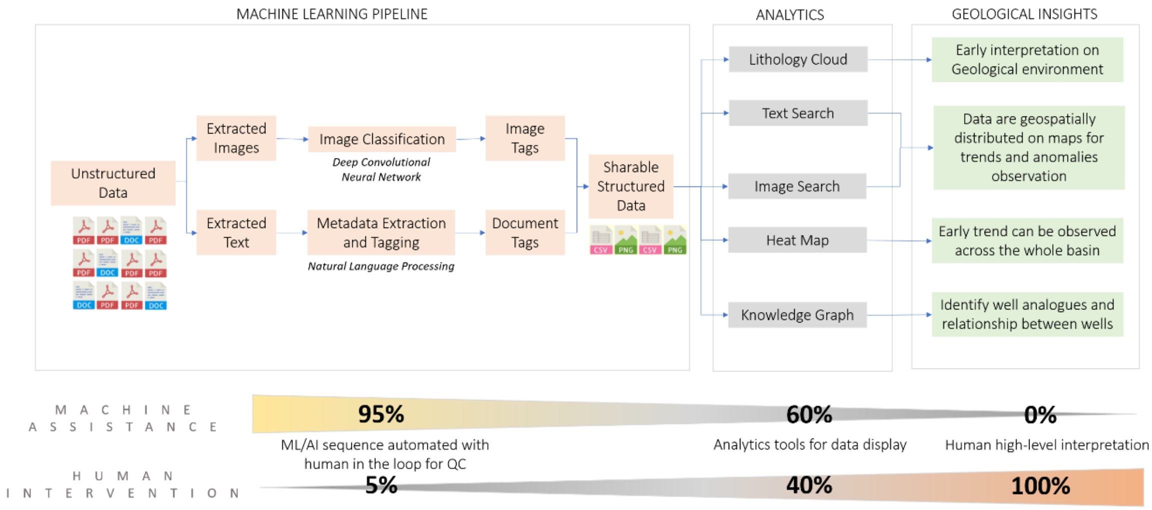


Figure 3 *Unstructured Big Data pipeline*

In this case, interpretation using such a Big Data workflow was used to understand the exploration history, how the basin developed, its petroleum system and the main issue of the dry wells occurrence to avoid repeating the same mistakes during future decision making.

By analysing the data, five potential issues are identified i.e. (i) Discrepancies in Formation Tops, (ii) Limited understanding of Lithology Distribution, (iii) Limited Mineral Composition Understanding, (iv) Fluid Distribution, and (v) Pressure/Temperature Patterns. Each potential issue is tackled by identifying trends and anomalies across the basin using images, tables and plots extracted from the unstructured data corpus.

Results

The analysis of the (ii) Limited Understanding of Lithology Distribution shown in Figure 4 is performed using the heatmaps. The heatmaps show the distribution of clastic and carbonates occurrence across the Bonaparte Basin and identify patterns and anomalies present over the area. The result can be supported by the stratigraphic chart where carbonate environment occurs in the younger formation from Cretaceous to Neogene period, whereas clastic environment occurs in the older formations from Triassic to Cretaceous.

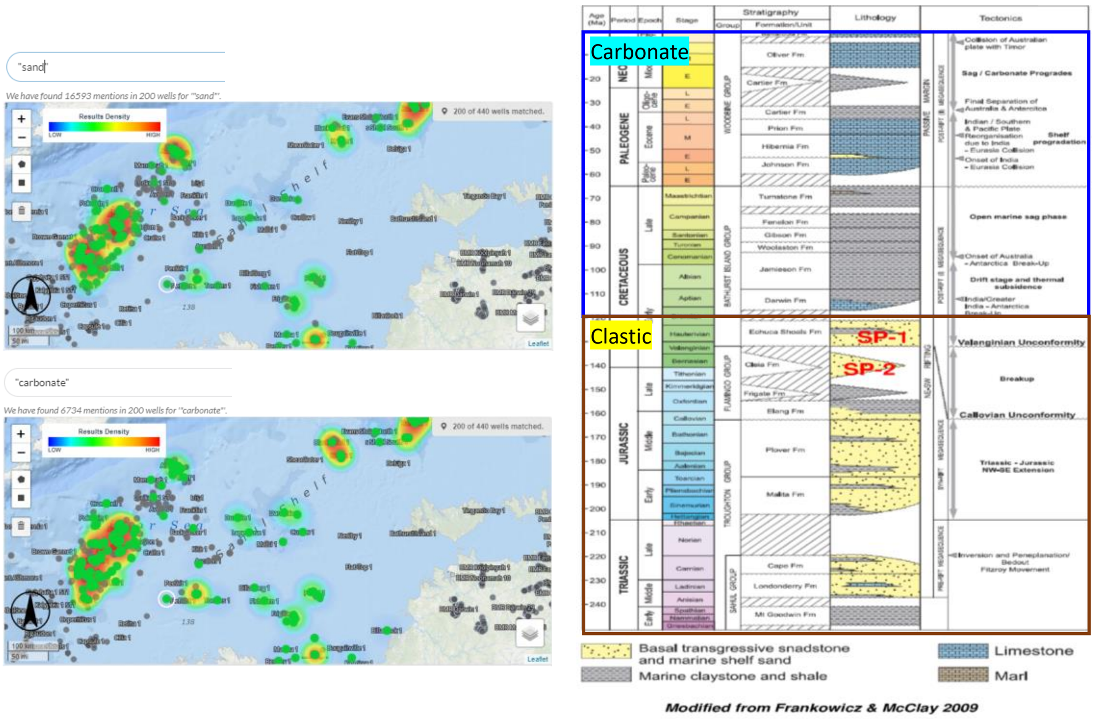


Figure 4 *Lithology distribution on heatmaps (left) and corresponding stratigraphic chart (right).*

The analysis of the (iii) Limited Mineral Composition Understanding shown in Figure 5 utilizes the thin section automatically extracted using ML classification over the full area and suggests that:

- Quartz overgrowth and kaolinite are quite common in Bonaparte Basin
- Mica mineral can be observed at the north-eastern part of the basin
- Highly corroded, skeletal feldspar has been extensively dissolved, which forms secondary porosity, and can be observed in the northern part of the basin
- Some patchy siderites are also observed in the southern part of the basin

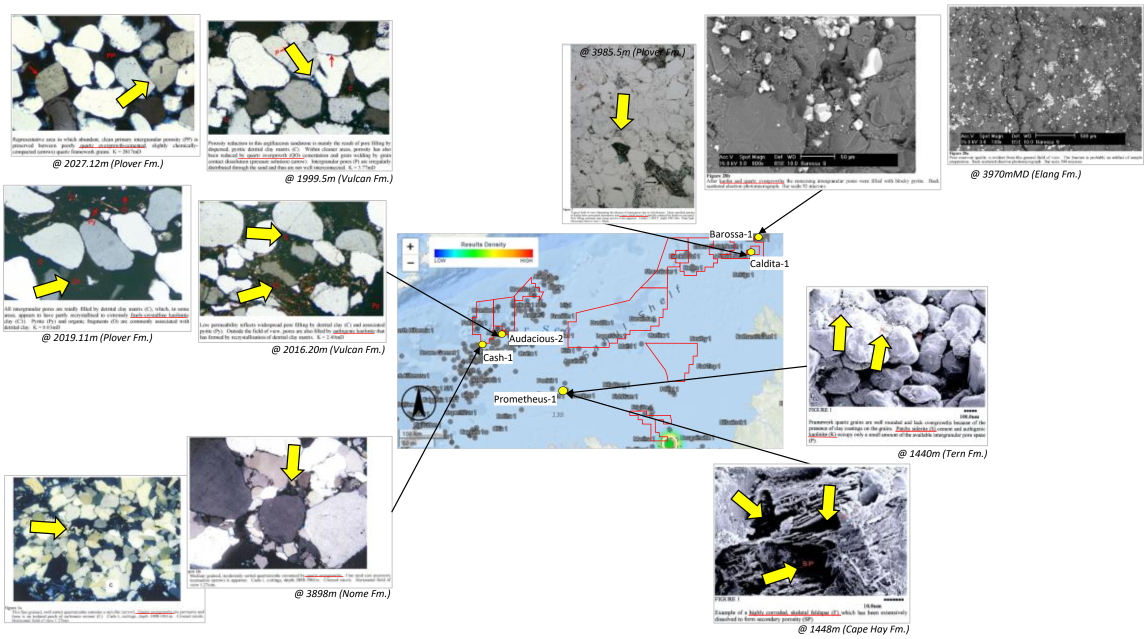


***Figure 5** Thin section images distributed on a map across the Bonaparte Basin*

Conclusion

A regional understanding is critical and time consuming as it involves dealing with a very large data volume. Within a project time frame, based on Play Based Exploration (PBE) pyramid, the time spent at the Basin Focus stage can be reduced, and more time are available to focus on the other project stages. The explorationist will be able to bring more value to the study.

ML applications have proven that it can play a crucial part in order to organize large unstructured data corpuses. This allows faster and accurate decision making within the fast-moving industry.

In this study, some potential issues encountered during exploration of the Bonaparte Basin can be identified. Based on a quick look and gathering all information it can be concluded that most of the production in the Bonaparte Basin is coming from Jurassic and Triassic with observed net pay ~18-60m thick, porosity ~11-29% and ~11-55% Sw.

References

Hernandez N., Lucañas P., Graciosa J.C., Mamador C., and Panganiban L. C. I., 2019: Automated information retrieval from unstructured documents utilizing a sequence of smart machine learning methods within a hybrid cloud container. EAGE Workshop on Big Data and Machine Learning for E&P Efficiency 25 - 27 February.

Lottaroli F., Craig J., Cozzi A., 2016: Evaluating a vintage play fairway exercise using subsequent exploration results: did it work? Petroleum Geoscience , Vol 24, no 2, p. 159 – 171.

Maver K.G., Hernandez N., Lucañas P., Graciosa J.C., Mamador C., Panganiban L.C.I., Yu C., and Maver M.G., 2018: An automated information retrieval platform for unstructured well data smart machine learning algorithms within a hybrid cloud container. EAGE/PESGB Workshop on Machine Learning, 29 – 30 November.