Utilizing Machine Learning to Gain Geological Insights through Unstructured Data for Sustainable Exploration Activities – Case Study Pre-Salt Brazil

Introduction

Understanding the basin regional trends and identifying the anomalies is a crucial background research during basin exploration activities. One way to gain a sound knowledge about the geology and the exploration history is to analyse the vast amount of data accumulated over the years in an unstructured manner. A sustainable data driven strategy leveraging on the latest advancement of Machine Learning (ML) and Analytics is applied on vast amount of unstructured data. By highlighting the data driven strategy, the paper demonstrates such a strategy applied to pre-salt carbonates prospects located in the Campos and Santos Basins, offshore Brazil. The interpretation framework consists of first step, to identify the regional first order trends and second step, to recognize the second order anomalies over the full area providing a holistic picture of the area of interest.

Methodology

The unstructured data comprise of more than 48,000 documents, primarily in Portuguese language, for a total of 330,000 pages related to 50 years of exploration, covering the Campos and Santos Basins. The methodologies are divided into two parts. In the part one methodology, the unstructured data is processed using machine learning techniques such as Natural Language Processing (NLP) for name entity recognition and language translation, and Deep Convolutional Neural Network (DCNN) for auto-image recognition (Hernandez et al., 2019). In part two methodology, ML Analytics leverages the visualization for data relationships to give a holistic view of the whole corpus (Baillard et al., 2021). Three steps study is used here, as explained below with illustration at Figure 1.

- Step 1: Deep Search through text is the ability to use keywords search within the whole corpus and filter the wanted results. Filtered information is easily accessible for investigation. Deep Search is also applied for images. DCNN image recognition and classification techniques classify the images into eight categories: thin section, core, well plot, seismic, stratigraphic structural elements, map, table, and figure. A deep search allows user to search information tagged and identified in these images. This step is to identify the geological and exploration challenges related to the area of interest.
- Step 2: Heat Map highlights the "hot zones" on map based on the frequency of the keywords found. Drawing a polygon on the map allows user to confine the search to focus only in the zone of their interest. On the other hand, instant text search through the whole corpus of English and Portuguese documents is possible with the automated translation. Users able to search through Portuguese documents using English keywords.
- Step 3: Contextual Knowledge Graph illustrates the connectivity of 'related corpuses' based on well name referencing. This is useful to obtain related information from different wells (Hernandez et al., 2019). Besides that, Intuition provides clustered images view, a great approach to discover analogues for alike images such as thin section, core, SEM, and biomarker.

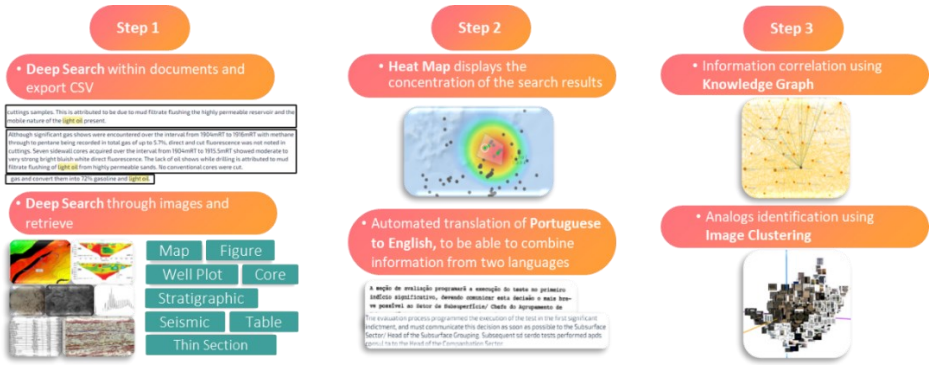

Figure 1 *The data driven search strategy implementing the ML Analytics for visualization of data relationships for a holistic view of whole corpus*

Case study: Understanding the Pre-Salt Carbonate Regional Trends and Anomalies in Campos and Santos Basins

The Campos and Santos Basins, located to the east offshore of Brazil, are some of the most prolific oil and gas basins in the world with significant discoveries such as Tupi, Jupiter and Libra Fields. The investigation of the 50 years of pre-salt exploration history contained in the 48,000 documents processed revealed that most of the challenges during exploration are caused by 1. Fluid distribution 2. The reservoir quality 3. CO_2 and H_2S presence 4. Overpressure patterns.

1. Campos and Santos Basins are showing a variable fluid distribution. The type and quality of the fluid trends are keys to define the best target for a sustainable development.
2. The diagenesis affects the reservoir quality. A good visualization of the lateral distribution of such process allows a better estimation of the porosity and the volume in place.
3. The regional gas issues, such as understanding the presence of CO_2 and H_2S is essential to gauge the reservoir diagenesis properties as well as to plan for production facilities or to avoid regions of sour and hazardous gaseous.
4. The irregularity in the thickness of the salt layer causes variability in the pressure regime with overpressure over the zone of interest and affecting the drilling campaigns.

These challenges are relevant to understand the geology and production issues encountered in distinct parts of the basins to improve sustainable in the exploration risk and reducing CO_2 footprint.

For the pre-salt oil trendings, Campos pre-salt generally has lighter oil compared to Santos pre-salt. As for the anomalies, light oil or condensate discovery at Pau De Acucar, Seat and Gavea Fields in Campos pre-salt, while heavy mostly found in the post-salt in the Southern Campos. In Santos, heavy oil reported in Jupiter Field. Note that heavy oil also discovered in Atlanta and Oliva Fields at post-salt. Illustration in Figure 2.

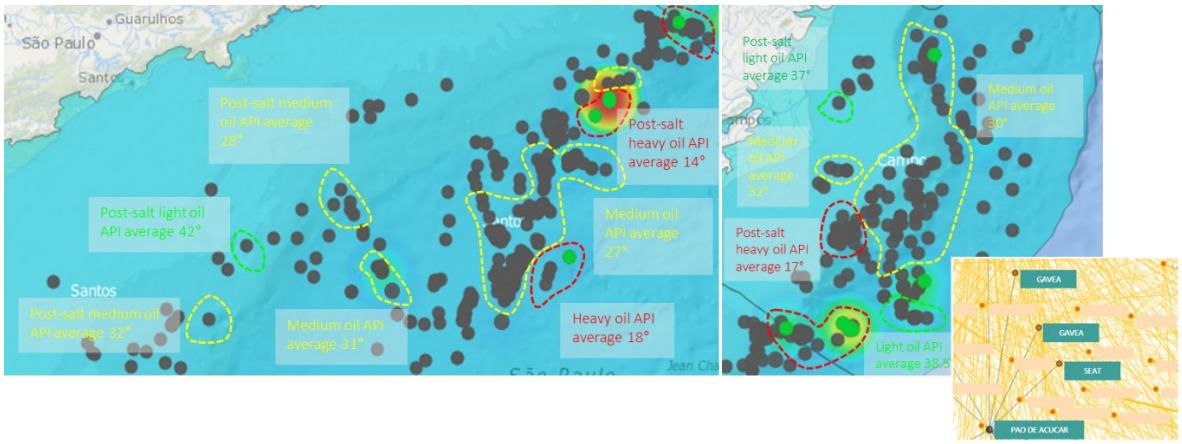

Figure 2 *Heat map highlights the wells with "heavy oil" information. The oil API delineation is based on Deep Search extraction from test reports for Santos (left map) and Campos (right map). Knowledge Graph was analyzed to find a group of fields that share the light oil information – Gavea, Seat and Pao de Acucar Fields.*

Next, we study the facies distribution at pre-salt carbonate. A collection of core and thin section images are retrieved from the Deep Search through images. In general, it is observed that the reservoir formation is in the shrub facies, microbiolites and coquina. For the trendingss, the area with CO_2, leaching activities enhance the porosity, example in the Libra, Buzios Fields, going south to Tupi and Carcara Fields in Santos. Fields proximal to the shore has poorer reservoir quality. One of the anomalies observed is the hydrothermal activity can destroy the enhanced porosity by the leaching. The hydrothermal activity is anticipated in Field Albacora, Carcara, based on the observation on the core. Illustration in Figure 3.

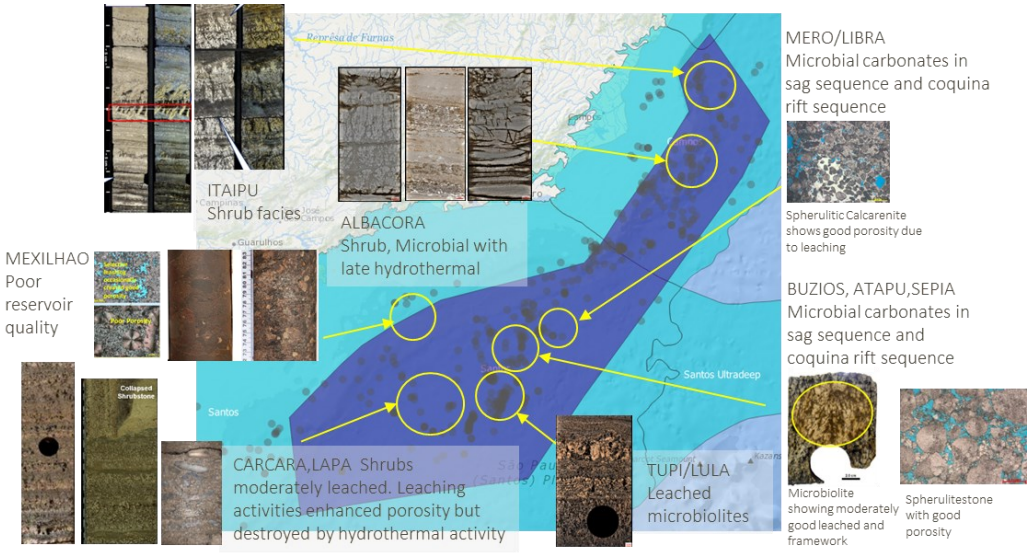

Figure 3 *Pre-salt carbonate facies distribution based on core and thin sections collection from Deep Search through images. Analysing the diagenesis trend has an impact on reservoir quality interpretation. (core and thin section images are taken from BDEP-ANP reports and Core Lab study)*

The following study is to investigate the regional gas issues. The trend is showing that Santos is suffering higher CO_2 contamination compared to Campos, especially in the center deep-water Santos area. Jupiter Field suffers from very high concentration of CO_2 gas, 79% recorded. Center deeper-water Santos is suffering from CO_2 contamination potentially near to the mantel magmatic activities intercepted by deep-seated fault (Luca et al., 2017). Hazardous H_2S gas presence in high amount (exceeding OSHA safe limit 50 ppm) in the Buzios, Iara Fields of Santos and Xerelete, Albacora Fields of Campos. Illustration in Figure 4.

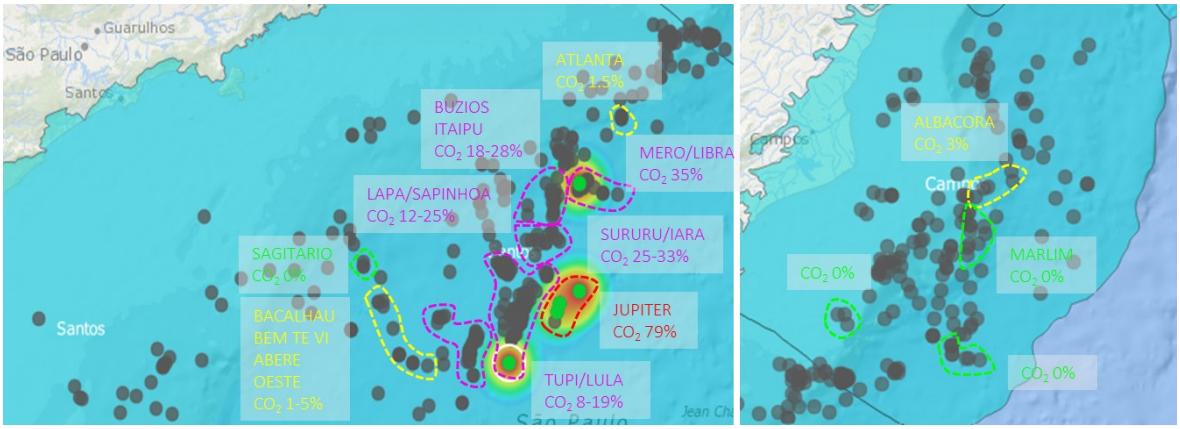

Figure 4 *Heat Map of "high CO_2" information with delineation of CO_2 concentration based on Deep Search extraction from test reports in Santos (left map) and Campos (right map)*

Lastly, is the study of formation pressure pattern. The formation data points are quickly extracted from unstructured data over 80 wells using Deep Search. The points are then plotted on gradient psi/ft ranging from 0.35 to 0.7 psi/ft. Based on the pressure gradient validation, wells in Carcara, Sagitario and Corcovado Fields have exceptionally high formation pressure. An interpretation of fullstack PSDM cross-section shows thick salt area in Carcara and Sagitario, possible with faulting causes saline water intrusion which contribute to high pressure zone. Irregularity in thickness of salt layer causes variability in pressure regime, such as overpressure area to take note as this will have an impact on seismic interpretation and drilling campaign. Illustration in Figure 5.

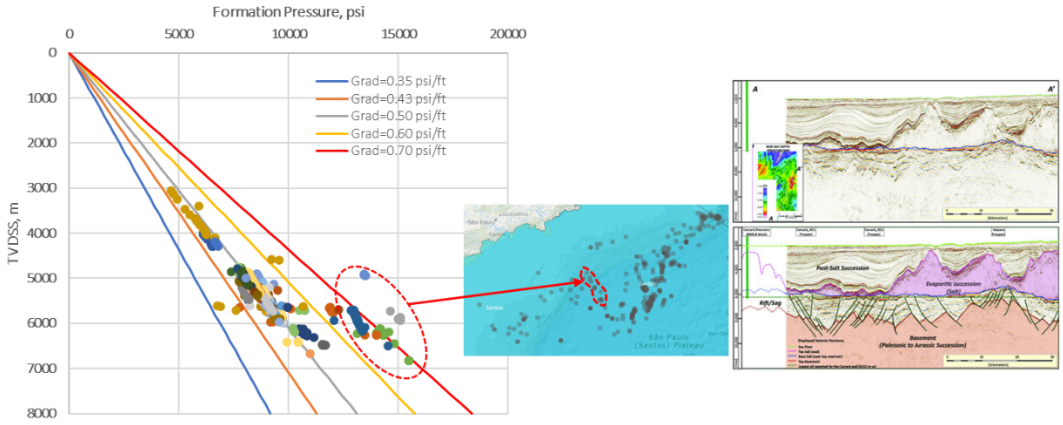

Figure 5 *Formation pressure extraction for over 80 wells (all data points taken from BDEP-ANP reports). Exceptionally high pressure recorded in Guaratiba Group in Carcara, Sagitario and Corcovado, possible due to high salt thickness and faulting causes saline water. Fullstack PSDM cross section shows thick salt in area of Carcara and Sagitario (courtesy of Kattah et al., 2014).*

Conclusion

The research shows the effectiveness of using ML/AI technologies and Analytics to mine through the vast amount of unstructured data and gain insights related to the regional trends and anomalies of important geological, reservoir and production parameters to minimize the risk of exploration and reduce carbon footprint. New tools in ML/AI and Analytics provides a new exploration frontier for subsurface experts allowing them to interrogate and visualize a vast amount of data previously scattered in different format and location holistically.

Acknowledgement

We would like to thank PETRONAS Petroleo Brasil LTDA for the collaboration with Iraya Energies to makes this paper possible. Our gratitude appreciation also goes to Banco de Dadoes Exploracao e Producao (BDEP) of National Agency of Petroleum (ANP) for making the data source available for this study.

References

Baillard, F., & Hernandez, N. (2021). A Case Study of Understand Bonarparte Basin using Unstructured Data Analysis with Machine Learning Techniques. EAGE Annual.

Hernandez, M., & Baillard, F. (2019). An effective G&G exploration strategy inspired by a wolfpack. Force workshop.

Hernandez, N., Lucañas, P., Graciosa, J., Mamador, C., & Panganiban, I. (2019). Automated Information Retrieval from Unstructured Documents Utilizing a Sequence of Smart Machine Learning. EAGE Workshop on Big Data and Machine Learning for E&P Efficiency 25 - 27 February.

Hydrogen Sulfide. (2022). Retrieved from Occupational Safety and Health Administration: https://www.osha.gov/hydrogen-sulfide/hazards

Core Lab [2022]. Pre-Salt Reservoirs of the Santos and Campos Basins, Brazil

Luca, Pedro & Matias, Hugo & Carballo, Jose & Sineva, Diana & Pimentel, Gustavo & Tritlla, Jordi & Cerdà, Mateu & Loma, Rubén & Jiménez, Ricardo & Pontet, Matthieu & Martinez, Pedro & Vega, Victor. [2017]. Breaking Barriers and Paradigms in Presalt

Kattah, S., Balabekov, Y., [2014]. New opportunities evident in the Santos Basin. Hartenergy. 2 September 2022. https://www.hartenergy.com/ep/exclusives/new-opportunities-evident-santos-basin-174870