Building a Digital Toolkit to Transform Access to Subsurface Knowledge

Karen Heyburn[1], James Scotchman[1], Matt Booker[1], Nicholas Thompson[1], Thomas Jewell[1], Owen E. Sutcliffe[1]

[1]Halliburton

Abstract

Difficulty accessing accumulated subsurface knowledge across an enterprise is a common problem. Whilst we have robust management systems for measured data (e.g., petrophysical logs or seismic) the same cannot be said for knowledge. Knowledge is data, given meaning through interpretation and integrated to derive insight. As we create knowledge, we move away from established data standards towards more unstructured formats – reports, presentations, emails – which have advantages of speed and flexibility at the time but are poor storage formats for the future. All too often geoscientists sink disproportionate amounts of time in finding, ingesting, and applying existing internal knowledge. Worse still, knowledge is often unused as it is simply too difficult to access.

We have, over many years, created an exploration-focussed subsurface knowledge management system, which aggregates data and knowledge from the public domain, and provides on-demand interrogation and analysis through geologically aware visualization tools. The challenge now is to adapt this concept for the complex, multi- generational repositories that are typical within energy companies. Successful subsurface knowledge management relies on three main approaches:

- Identify, extract and store useful information from legacy and unstructured sources.
- Standardise against a robust geological framework for classification of various geological parameters.
- Visualize through smart user-defined integration and analyses of related information.

To meet these challenges, we have created a digital toolbox comprising:

- ML and Natural Language Processing (NLP) tools to extract information from legacy reports, including lithology, age, depth, rock quality, stratigraphy, depositional environment etc.
- ML and NLP tools to map free text lithology and depositional environmental descriptions to our hierarchical frameworks, achieving an accuracy of 97% at the highest level.
- Geoprocessing tools, and global dictionaries to map lithostratigraphic terms to sequence stratigraphic and chronostratigraphic markers to define geological age.
- Dashboards and map-based visualizations which use the geological framework attributes for on-demand analyses of a range of subsurface properties.
- Integration tools to connect the original, un-edited data repositories with derived framework attributes and the visualization applications.

We have combined these tools into an integrated knowledge management system on top of a complex repository of subsurface information. This provides access to insights that are typically found by time-consuming search and digestion of unstructured reports or multiple separated databases. Geoscientists can immediately gain answers to questions such as porosity and permeability distributions of, for example, paralic sands in a basin, or the spatial distribution of lithological proportions in a particular lithostratigraphic unit. The process of combining ML, NLP

and geoprocessing tools with geological ontologies and visualization technologies liberates information from inaccessible repositories and enables its application in subsurface characterization workflows. This allows organisations to fully leverage the data and expertise they have invested vast sums in obtaining.