

## **Eighty Years of Legacy Exploration Well File Data: Restructure and Retrieval**

**Ahmed M. Bakhsh<sup>1</sup>, Ahmed F. Saleh<sup>1</sup>, and Abdulrazaq N. Salamah<sup>1</sup>**

<sup>1</sup>Exploration Technical Services Dep, Saudi Aramco Company, Dhahran, Eastern, Saudi Arabia.

### **ABSTRACT**

The Kingdom of Saudi Arabia's long history of hydrocarbon exploration has generated a large volume of well information. Over the last fifteen years, Exploration well files have been captured digitally and made available to geoscientists through in-house data retrieval applications. This wealth of data and resources have huge importance for future prospect evaluation. Recently, in an effort to unlock value from their unstructured content, a remediation project was established to address geoscientist's difficulties in retrieving well data. The existing workflow is comprised of three processes. The first process extracts a well file from the repository, using database query language. The second process segments this file into forty taxonomical categories, based on the well pre-drill to post-drill lifecycle (e.g. Drilling Program, Post Drilling Analysis, Biostratigraphy, Geochemistry, Well Test, etc.) using an exploration keyword matrix. Finally, the third process indexes the resulting files with extended metadata (e.g. Well ID, Field, Reservoir, Document Type, and Proponent) to enhance search efficiency. This conventional, manual workflow is being automated with the introduction of analytical machine learning methodologies; thus saving resources. A combination of both assisted and supervised machine learning methods are being employed to identify well file content. This is done by using logic statements derived from the exploration keyword matrix and algorithms based on a supervised training set of completed and segmented files.