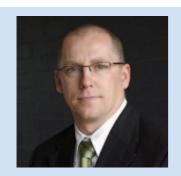# Data trends

Guy Holmes
guy.holmes@katalystdm.com.au

Finding the future: the role of Big Data Analytics in oil exploration

Anyone who has a LinkedIn account, watches the news or simply surfs the internet sees technological innovation occurring all around them. Innovation is regularly on the front page and words like 'Artificial Intelligence (AI)', 'Big Data Analytics', 'Internet of Things (IoT)' etc. are popping up everywhere. One area of innovation that the oil industry would benefit from, but in my view has not been able to adopt or adapt to, is the area of Big Data Analytics.

Big Data Analytics has appeared on Forbes' and Gardners' Top 10 lists of technology trends for several years in a row and every industry that I can think of seems to have found innovative ways to use this technology on their greatest resource – their data. Today there is something called the 'three Vs' of data that are seen as gauges of how effective Big Data Analytics can be on data. The three Vs stand for the variety, velocity and volume of data. An example of how these three Vs can be used comes from Macey's Inc., where the retailer adjusts the price of over 70 million items that they sell in real time based on inventory and sales demand.

Although it is now making inroads in oil and gas production data analytics, the oil exploration industry seems to have struggled to find a way to successfully get the three Vs aligned in order to derive the greater insight into their exploration data holdings that this technology offers. Some of the issues that I think may be responsible for the slow adoption are:

1. The sheer volume of data (Volume). The oil exploration industry generates massive volumes of data. It is not uncommon to see single seismic surveys that are over a Petabyte (Pb) in size. And whilst that is not close to too large for these tools, it is still a huge volume of data to deal with.

2. The fact that Big Data Analytics are really only effective when large volumes of data are combined with other large volumes of data (Variety). Running analytical tools on datasets that have little variety within them, and no comparative datasets for them to connect to, is an issue. More than a single dataset is required to get the best value or results from running these tools and, if that is the case, multiple Petabyte datasets will need to be in play for the tools to work.

3. The oil industry has not adopted older technologies that are required to use this new technology (Velocity). This for me is where it starts to get interesting. The oil and gas industry has been a slow adopter of cloud based storage for exploration data, yet the cloud, in many ways, is the only place that can allow these massive data pools to exist (or persist) and to then be 'touched' by these powerful tools. The more data you add into these pools, and the greater the variability of the content, the more you get out of Big Data Analytics. The velocity of data in the oil industry in my view moves at a snail's pace – and it is not just because there is a lot of data involved.

It seems like the oil industry never cleared the first hurdle when dealing with its exploration data. It creates orphaned datasets all the time. They start life on a tape and are only temporarily alive and digital before information is extracted and they are moved back into storage. In fact, there are entire workflows in data processing that are designed to use and dispose of the raw data as fast as possible in order to avoid bottle necks in storage. This type of workflow effectively stops two of the three Vs from ever getting momentum (Volume and Velocity).

What would be possible in a perfect world – where all exploration data are available and can be readily touched by these new tools? I am sure we will find out one day, and I am equally sure that the consequences will be profound.

EAGE

See you in Barcelona!

4-8 September 2016

22nd	SECOND	FIRST
EUROPEAN MEETING OF **ENVIRONMENTAL AND ENGINEERING GEOPHYSICS**	APPLIED SHALLOW **MARINE** GEOPHYSICS CONFERENCE	CONFERENCE ON **GEOPHYSICS** FOR MINERAL EXPLORATION AND MINING

NEAR SURFACE GEOSCIENCE
CONFERENCE & EXHIBITION

4-8 September 2016, Barcelona, Spain