Our Energy Clients' Top 3 Data Management Challenges Today

Despite all the hard work of the last 25+ years, we still face enormous challenges in Oil & Gas data management. Recently, the focus has shifted from managing structured data repositories to the enormous volumes of 'dark data' that most of our clients still have in their digital vaults.

Gareth Smith, Head of Consulting at Sword Venture, explains that this shift is driven by the need to feed large volumes of high-quality data into analytics and data science-driven processes; growing regulatory pressure to report data; and the need to ensure safe, efficient operations (especially for the many assets that have swapped hands). However, most of this data is locked away in collections of documents and legacy proprietary formats, is nearly always poorly indexed and is often not machine readable. We have seen a surge in requests to help our clients do something about this.

Most of these Oil & Gas data management projects break down in to three core challenges:

1. **Find It:** Trawl through large volumes of files and understand what we have.
2. **Sort It:** Classify, tag, de-duplicate, make it machine readable.
3. **Use It:** Search for valuable content and extract it into a form that can be used in 'traditional' E&P (exploration & production) applications and data-driven analytics.

Crucially, this has to be done without years of manual effort. It needs a different approach, using highly automated data science and analytics solutions to tackle these core challenges at scale and pace.

1. Find It

The first step is to build a pipeline that can ingest, store, process and extract/index data at scale. Unless our clients have the required compute resources to hand, we make use of the major public cloud platforms, e.g. AWS (Amazon Web Services) and Azure (by Microsoft), to provide the tools and processing power to tackle this challenge. A re-usable, cost-optimised and efficient data processing architecture based on cloud infrastructure reduces the costs and overhead to the client and allows us to move quickly through this stage of the process.

2. Sort It

The next step is to clean up the data, understand what we have and identify the data that has some value to us. This is where we apply data science and analytics to automate and massively speed up what once was a laborious, labour intensive process. For example, we have developed a model using a machine learning algorithm, or neural network to predict the classification of a given document based on its text content, contained images and structure. Accuracy levels are often in excess of 95%. We combine this with analytics-driven clean-up and cloud-based optical character recognition (OCR) to create a repository of well structured, machine readable content ready for further analysis and processing.

3. Use It

The 'Find' and 'Sort' stages are just a means to an end; the goal is to extract value from data by putting it to work. We use a combination of machine and deep learning to identify and classify specific data within a document or text and extract that data in a machine readable format. For example, identifying deviation survey data (essential to all well interpretation) within documents and scanned images, extracting and making available to engineers.

We are working on the use of natural language processing (NLP) and machine learning techniques to draw meaning out of documents in order to generate greater insight. For example, to automatically recognise the findings and outcomes of final reports for thousands of wells, reducing the requirement for manual analysis.

Beyond the Core Challenges: Insight & Intelligence

Our goal going forward is to develop and deploy techniques such as deep learning and data engineering to synthesise large amounts of information and provide recommendations to assist human decision makers. For example, assisting with decisions on where to target exploration investment or how best to configure engineering parameters to reduce failure rates and maximise uptime. We recognise the value of combining data and knowledge to create unbiased predictive reasoning tools to support complex decision scenarios.

The combination of on-demand cloud computing and advanced data science and analytics techniques is revolutionising the way we manage data and extract value. We can tackle data at scale and pace, reduce manual effort and automate processes in a way that just wasn't possible even a few years ago.

To find out more about how we can help you with your data management challenges, read on about our data and [information management solutions here.](#)