

Observations from ECIM & the rise of data science



Up Front

This being my first visit to [ECIM](#) meant I could go with a completely open mind, no stereotypes or preconceptions to bog me down. And odd as it may seem I'm glad I'd somehow avoided attending throughout my 20+ years in Oil & Gas data and information management. This year's event had a real sense of excitement and optimism for the shift changes taking place in the industry and how data and data management is now becoming pivotal to that change.

Data Science - just another buzz word?

From the outset it was clear that data science was going to be the hot topic. This

was easily recognisable by the 50+ attendees who attended the inaugural data science bootcamp run by [Atti Balazs](#) (Sword Venture Chief Data Scientist) and [Ruairi Dunne](#) (BP Challenger Graduate and data science specialist). A hugely well received introduction to data science, its practical application within E&P and a click along session accessible to all levels.

The pervasiveness of data science throughout our industry is clear, with a genuine sense of the step changes taking place and the part data science techniques are playing in this; whether through expediting data pipelines, supplementing quality, accelerating data discovery or augmenting geoscientists. That said, it is estimated that less than 10% of companies are extracting the true value of data at a scale to make significant change, so the journey has clearly only just begun.

The work stream sessions highlighted the varying approaches being developed to solve the unstructured data challenge, crystalising our view that there is no prescribed answer to a problem with such depth and diversity. Many organisations are still in the process of delivering PoC or MVP's, with few progressing to deployed production systems, even those working in collaboration with major technology providers. What is clear, and advocated here at [Sword Venture](#), is that where there is an open framework for collaboration, there is a greater chance of success.

Data Science, Artificial Intelligence and Machine Learning are terms all too often used when businesses are looking for a solution, sometimes even before the challenge is clearly understood, but is this a bad thing? In the sense that business leaders are beginning to recognise the value that data management and data science can deliver, then no - but if we believe that data science, AI and ML as examples are the sole solution, then yes, as although these may be providing a technological accelerator the key to successful adoption still lies with the organisations ability to adapt existing mindset and process to meet technology. And then there's the data.....

Is data centricity the key to unlocking value?

The rise in adoption of Data Science techniques to solve geoscience and data management challenges often expose the need for high quality input data to train and test models, a point regularly raised through the conference sessions. The growing influence, and value of creating high quality, standardised open data

platforms were recognised as a key component to the data pipeline, feeding the rise in tangible, achievable applications of Data Science.

Data is the key compliment to analytics and prediction, creating the quality, volume and velocity of resources data science needs to consume is fast becoming the real challenge. Developing pipelines to feed the machine are becoming essential, removing the accidental, often siloed architectures which have “developed” will need dedicated data engineering. So does this mean we all need to invest in data lakes, warehouses and master data management? To a greater or lesser extent yes, there is no doubt that the ability to access at speed any data type, whether static or streaming is essential to engineering a sustainable data pipeline – it’s how we respond to this challenge that will set us apart.

Call me old fashioned....

Data quality and governance were also keen topics discussed through the work-streams, highlighting once again that data quality is still key to the generation of high value knowledge and insight. The need for strong governance is neither a new concept nor a new requirement, but with the ever increasing volume, pace and diversity of data both in terms of creation and consumption the need for applying robust governance has never been more acute. And guess what, we can even use data science to do that too!!

Paul Newey, Head of Digital Enablement

September 2019