

Data Management: It's About Time!

The Information Technology (IT) world is a buzz today with new jargon that is glitzy and intriguing. Terms like artificial intelligence (AI), machine learning (ML), digital ledger technologies (DLTs), Big Data and the Cloud are all in the news and part of the public's awareness. Digital data has been with us for decades. While many before us have expounded upon the value of proper data management with policies, procedures and best practices for data population, archival and retrieval, the IT world is now starting to understand. **It's about time!** Sound data management practices has become critical in an age of customer personalization, artificial intelligence and predictive analytics. It is the old data management story, garbage in, garbage out, but now it can be done with such speed that the impurities associated with poor data management practices are seen more readily. **It's about time!**

According to the Computing Technology Industry Association (CompTIA), which is a non-profit trade association serving as a voice for the information technology (IT) world with 2,000 member companies, 3,000 academic and training partners and over 100,000 registered users, *"only 25% of companies feel like they are exactly where they want to be with their corporate data management, down from 31% in 2015. Although digital data has long been a part of IT operations, there has not been much focus in terms of job roles or defined components. The hype around big data highlighted a need for comprehensive data policies, and businesses are now beginning to build discipline around capturing, processing and analyzing data. The list of challenges that businesses are facing in building a data strategy shows how important it is to have an established process.*

For example, 38% of companies cite a desire to speed up their data analysis, which involves both infrastructure and process. Similarly, 37% cite the need to integrate data throughout the organization, which involves having the right storage solutions but also the right workflow. Across infrastructure, skills and policies, there are a range of issues that businesses have to solve to properly manage their data.

Only 44% of companies say that they have internal IT employees who are dedicated to data management or data analysis. While there has been a focus on

newer job titles like data scientist, there is also opportunity around more traditional roles like database administrator.” (1)

It’s about time! CompTIA goes on to say that many Chief Information Officers (CIOs) *“reported that their lead database administrator had previously been a software developer who built responsibilities around data. Since software development was not a core competency for many companies in the age of packaged software, it makes sense that the related field of data management is new territory.” (2)*

So, what is happening now? What is the latest trend? With the desire to leverage new technology, there is an awareness of a need for a holistic data management approach. IT personnel are being trained to become experts in understanding the data and putting in place the appropriate policies, procedures and best practices for data management, while they are learning about the data itself. Understanding the data is vital when collecting parameters associated with meta-data capture. Meta-data capture is critical to making use of X analytics.

Some recent trends in the IT world are X analytics, augmented data management and the amalgamation of data and analytical tasks and their inter-relationships. X analytics is an umbrella term used to describe things captured such as text, video or even audio formats. Augmented data management often makes use of AI and ML techniques to support dynamic systems, leveraging meta-data for queries and performance based operational analysis. Historically, data and the analysis of data has been viewed as two separate functions. These two worlds are coming together and amalgamating as one. Of course, none of this is possible without a high quality database, in behind, or shall I say, as the foundation for all of this.

So, in the realm of corporate mergers and acquisitions in the oil and gas industry, what is the power to be able to say, within an hour or so, how many line kilometers of 2D and 3D seismic data you might have from any given speculative survey company that historically subjects its clients to a transfer fee upon license conveyance, and that the dollar exposure to this liability is worth approximately “X” dollars? What is the value of being able to tell a pathfinding new play geologist, how many GeoScout and Accumap projects, done by these authors in the past, in the township and range area defined, that have wells picked with a Wabamun or Slave Point top? This is what ML and AI can do. Being able to

establish search criteria and then be able to query accordingly is the power of this technology. This is augmented data management. This is about making use of and being able to leverage data. Rather than being a static repository of stuff, the database becomes the foundation for all of this. The data manager, (the historical gatekeeper), becomes the guru for serving up more powerful search options and criteria, leveraging the meta-data captured in the database. This is data management plus (DM+).

As I have discussed with colleagues over the years, data management is part of the life cycle of data becoming information and then knowledge. The LAS log digits reside in the database. They get used to build a cross-section, which now constitutes information. The study of that cross-section and what it means, where the porosity in the reservoir develops, is the knowledge.

As I conclude, I wonder what the survey response for CompTIA would be if they asked their member organizations about their successes associated with information and knowledge capture, let alone being able to leverage them. For now, it is important for all of us to know that data management is not just about making sure the database is built with accuracy and completeness, it is about leveraging the meta-data with search and query capabilities quickly. Data, workflows, infrastructure and database compatibility is an issue with multiple search engine criteria. The bottom line is that it starts with a database that has integrity, accuracy and completeness. The IT world now realizes this. **It's about time!**

Doug Uffen, P. Geoph (APEGA), P. Geo (APEGBC)
[Reflection Peak Enterprises Limited](#)

Quotations:

- (1) *Data Management Trends to Watch for in 2020_Big Data_CompTIA, page 3*
- (2) *Data Management Trends to Watch for in 2020_Big Data_CompTIA, page 4*

Doug Uffen is a professional geophysicist in Calgary with a long passion for the integrity of data and databases upon which interpretive information and knowledge is derived. Having worked domestically and internationally, he is acquainted with a barrage of GIS survey and data integrity issues and has become suitably paranoid about data integrity as a result. He is a current Board Member of the Calgary Geoscience Data Manager Society (CGDMS) and a Past President of the Canadian Society of Exploration Geophysicists (CSEG) and a former Board Member of the CSEG Foundation.

References:

- *20 IT Data Management Trends for 2021_Sanity Solutions Inc.*
- *Data Management Trends to Watch for in 2020_ Big Data_ CompTIA*
- *Gartner Identifies top 10 Data and Analytics Technology Trends for 2020*
- *Top 10 Data and Analytics Trends for 2021 – Information Week*