

Treating Data as an Asset – Experiences of the Early Adopters.

Brad Eccles, Director of Operations, ABS Group, Warrington UK
Matt Mowrer, Director of Government Programs, ABS Group, Knoxville, US

Data Analytics and Management (DA&M) is the science of examining available data to inform and improve business and technical decisions. This relatively new area of Business analytics is being driven by the 'Big Data' phenomenon and is becoming commonplace in many sectors where organizations are using historical performance data and predictive modelling to support a wide variety of operational and business needs.

Despite this, Data Analytics (DA) for the more technical applications (e.g., engineering, operations, and maintenance) in oil, gas and chemical (OG&C) sectors are less mature and are, often by necessity, more customized. However, the forward-leaning companies are recognizing that applying incisive analytics has the potential to deliver significant performance improvements by improving safety, environmental protection, operational efficiency and profitability.

Eventually, the application of DA will become the norm, and those who embrace it early are starting to differentiate themselves by how well they are applying it.

In this developing area, ABS Group has been supporting organizations to build innovative digital platforms that use data to monitor, analyse and manage a broad range of operational risks. We have worked with both commercial and government clients to develop data-driven risk models that generate accurate and timely information to base sound decisions on. During that time, we have seen many of the issues that the early adopters have faced; this paper summarises many of these observations, including:

- Stakeholder "buy in" and the need to understand the value of data and recognize DA&M as "good investment" vs. "cost of doing business."
- Treating data as an asset.
- How accurate and timely data could have altered the outcome of process safety incidents.
- Demonstrating management performance to stakeholders and regulators.
- The ability to efficiently present summaries of vast amounts of pipeline component/network data (equipment pedigree, inspection history, current condition, risk attribute data etc.).
- Lessons learned on achieving efficiency and effectiveness in (DA&M).
- How to Get Started – Maturity Roadmaps and Data Management Frameworks.

Keywords: Data Analysis, Date Management, Big Data Process Safety

Introduction and Background

Data Analytics and Management (DA&M) is the science of examining available data to inform and improve business and technical decisions. This relatively new area of business analytics is being driven by the 'Big Data' phenomenon and is becoming commonplace in many sectors where organizations are using historical performance data and predictive modelling to support a wide variety of operational and business needs.

Despite this, Data Analytics (DA) for the more technical applications (e.g., engineering, operations and maintenance) in oil, gas and chemical (OG&C) sectors are less mature and are, often by necessity, more customized. However, the forward-leaning companies are recognizing that applying incisive analytics has the potential to deliver significant performance improvements by improving safety, environmental protection, operational efficiency and profitability.

Eventually, the application of DA will become the norm, and those who are embracing it early are starting to differentiate themselves by how well they are applying it.

In this developing area, ABS Group has been supporting organizations to build innovative digital platforms that use data to monitor, analyse and manage a broad range of operational risks. We have worked with both commercial and government clients to develop data-driven risk models that generate accurate and timely information to base sound decisions on.

Treating Data as an Asset

Data Analytics requires Data Management, the activity of maintaining and increasing the value of an organization's data. Managing underlying data better provides a reliable, high potential platform for effective data analytics. In the past, Data Management was simply a discipline tasked with ensuring the organization could access their data or had simply enough disk space to store it. As the value of data has increased for organizations, so too has the importance of the role of Data Management.

To understand the value of Data Management, the significant value of data must first be understood and fully appreciated. Figure 1 depicts a simple value model for data. It illustrates four ways that data can provide value to an organization (Frost C., 2015).

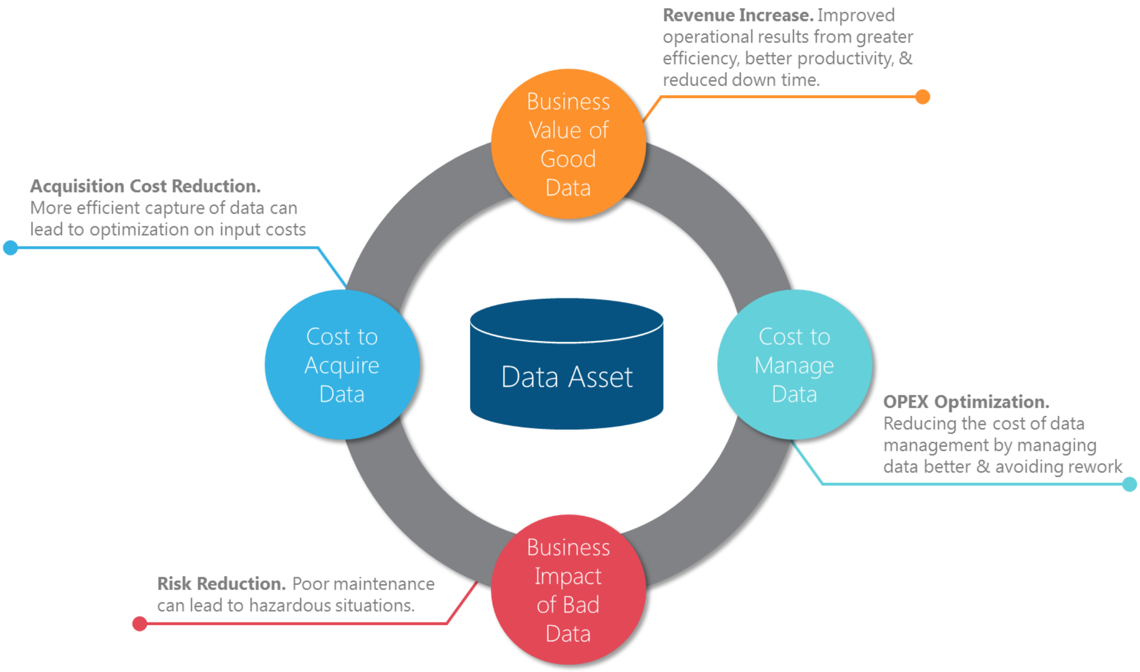

Figure 1 – Data Value Model (Source-ABS Group)

Business Value of Good Data. Critical decision-makers typically spend 50 – 75% of their time looking for the data they need and confirming its validity and provenance before actually using it. Increasing trust in the data makes decision-making much more efficient, focusing on activities which deliver true value rather than just gathering data, and the quality of their analysis work and decision-making will improve significantly as a result.

Business Impact of Bad Data. Making decisions based on bad data exposes operators to significant risk and potentially negative impacts; e.g. mismanaging safety-critical production operations, missing conditions which could cause non-compliance with regulations, misjudging market demand related to new production, etc.

Cost to Acquire Data. A typical integrated oil company invests $2 to $3 billion per year in acquiring data (from seismic to real-time operations data and much more). Applying a straight-line depreciation of the value of that data over 10 years gives an asset which is worth $10 growth to $15 billion. A physical asset worth the same amount would not be neglected.

Cost to Manage Data. A major oil company may spend well over $1 million per annum centrally managing data. If regional costs are added, then annual operating costs for data management can easily exceed $10 million per year. Improving the way data is managed and avoiding re-work can reduce associated costs by at least 2%.

DA&DM are gaining a strong signature on the radar of the OG&C sector because of a number of contemporaneous trends in sensor, communication, storage and processing capabilities. There has been a proliferation of small, intelligent sensors, which measure changes in physical attributes and transmit the resulting data through extensive, easily accessible and fast wide-area communication networks. The data can be stored in massive data centres, and subsequently processed and analysed by extremely powerful computers and processors to deliver critical insights. Those insights can then be acted upon and facilitated by computer-aided interventions. A key point is that today all this can be done at a relatively low cost.

Data management is a key enabler of not only data analytics but also of critical decisions made every day in a wide array of work processes. The collection of enterprise data is an asset to organizations but is often not a recognized or managed one.

DA&DM in Forward-Thinking Businesses

Forward-thinking businesses are increasing the use of DA&DM to improve their business operations. The technology already allows the design of data-driven, self-aware, physical assets, whose performance and state of health are being continually and holistically monitored and predicted. A continuous state of situational awareness is possible now and will likely be the norm in future years.

It is useful to visualise the data management process in the layers presented in Figure 2. The process should start with the identification of the business processes and key decisions that need to be made by each stakeholder, which typically includes organizations related to planning, financial, operations, maintenance, inspection, etc. Once the data-driven needs are known, you can identify the analytics that can provide the best insights to address these needs. DA requires data, which will typically come from existing device management systems (e.g., PLC), SCADA and other control systems (e.g., DCS), manual and electronic logging systems, etc. These, in turn, receive the data from field equipment. The evaluation process will help identify cost-effective modifications to the existing data gathering, processing and storage.

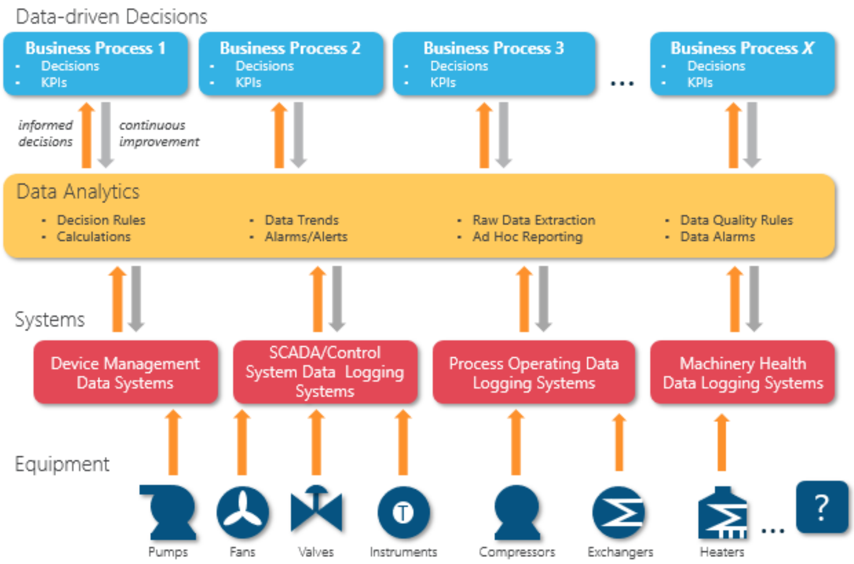

Figure 2 – Data Management Levels (Source-ABS Group)

Ideally, Figure 2 should be applied top-down, but organisations will already have data being generated, recorded and analysed in some way. Improving how these data are managed can produce quick wins.

In the Systems layer, data generation and storage are part of normal plant operations. The value of accurate and timely data at this level is to manage the immediate and short terms risks associated with safety and quality, for example:

- Process Conditions (temperatures, pressures, levels, etc.)
- Plant Status (value positions, equipment status, flow rates, etc.)
- Asset condition (wall thicknesses, integrity operating window (IOW) excursions, etc.)

In the Data Analytics layer, monitoring and trending of the gathered data will add values by creating:

- Current and historical information to present to regulators and other stakeholders
- Data to support incident investigations or process optimisation
- KPIs and trending to enable data-driven decision making

To date, we have seen most of the "quick wins" in the area of asset integrity management, where metrics and sensors can be carefully selected to allow optimization of the asset through reactive and predictive DA. Sensitive 'hot spots', where risks to business and operational objectives are perhaps higher, can be proactively identified. As well as reducing unplanned outages and improved turnaround management, these monitored 'hot spots' can provide continuous predictive insights to enable improved decisions around:

- Turnaround inspection interval extensions
- Sound basis for more efficient risk-based inspection systems
- Greater insights into process changes e.g., managing the implications of consciously operating outside Integrity Operating Windows (IOWs)

These improvements typically provide very high rate of return (e.g., paybacks within a few years.)

Data and Safety Management

The Deepwater Horizon was the state-of-the-art deep-water offshore drilling platform that suffered a blow-out in April 2010. While other factors were also critical, government and independent incident investigations consistently suggested that more or improved DA&DM may help reduce the frequency or consequences of similar events in the future. These quotes from key incident investigation reports illustrate this point (italics and underline added for emphasis):

- The Agency [BOEMRE] should consider promulgating regulations that would *require real-time, remote capture of BOP function data* … Having the data that show which rams have been activated would help analyse intervention options (BOEMRE, 2011, p. 209)

- The simulation was *flawed in that it did not use the most accurate data set available from the well* ... A final simulation on the basis of the most accurate well data followed by a discussion of the results to make a decision on the final centralizer placement would have been prudent (NAS, 2012, p. 37)

- *The real-time data from the rig were being recorded but not monitored on shore.* Even with the negative test having been accepted, subsequent data showing that the reservoir and well were in communication might have been discovered by personnel on shore in time to take the appropriate control action (NAS, 2012, p. 39)

- Although data were being transmitted to shore, it appears that *no one in authority (from BP onshore management or a regulatory agency) was required to examine test results and other critical data* and render an opinion to the personnel on the rig before operations could continue (NAS, 2012, p. 39)

- Overall, the *regulatory community has not made effective use of real-time data analysis*, information on precursor incidents or near misses, or lessons learned in the Gulf of Mexico and worldwide to adjust practices and standards appropriately (NAS, 2012, p. 114)

- Because of the simultaneous offloading and cleaning operations, the mud levels in the tank were changing, making it difficult to monitor whether the well was flowing. As a result, *the recorded flow data is believed to be unreliable* during this period (DHSG, 2011, p. 40)

- This abnormality could have been detected from monitoring the drill pipe pressure data, and could have been *the first clear indication of the flow of hydrocarbons into the well, visible to the crew* (DHSG, 2011, p. 43)

In many process safety incidents, timely access to process information data (pressures, temperatures, levels, status of plant etc.) could have changed outcomes. Even though the data in many cases was available it is often issues with the Data Management system that hampers timely recognition and response. These issues are generally related to: untimely, inaccurate or poorly selected data, data presented in a way that makes recognition difficult, poorly defined triggers and responses.

Data Integration

DA&DM techniques are advancing and becoming more predictive in nature by integrating different information sources using Machine Learning/Artificial Intelligence. Data integration looks to identify trends across different datasets and this is particularly relevant in the area of process safety because of the multi-causal nature of safety incidents which typically involve the undesirable performance of plant, processes and people.

Safety Risk Analysis solutions generally require the integration of multiple datasets, often generated by multiple different operational units and external organizations. The types of datasets that can be used vary by project but usually are related to answering the following questions:

1. What can go wrong?
2. How likely is it?
3. What are the impacts?
4. What are the systems to prevent / mitigate the risk?

Rather than just considering data for single equipment items, multiple datasets can be integrated and connected to provide more insight into potential safety risks. Machine Learning/Artificial Intelligence is starting to be used to efficiently extract data and "join the dots" between such sources as Process Hazard Analysis (PHAs), historical operational data, incident investigations, audits, maintenance logs, occupied buildings risk assessments, culture surveys, etc. to predict potential weakness in process safety management systems.

Data Quality Issues

Data quality has a significant effect on the quality of informed decisions so recognising data issues is important. Input data can consist of varying degrees of quality, which we describe with three (3) facets:

(1) **Data accuracy** is the degree to which the data reflect reality. Data inaccuracy might result from errors in data collection, missing data or random variation in observed values.

(2) **Data precision** is the level of detail expressed in the data. For numerical data, precision might mean the number of digits shown after the decimal point. For non-numerical data, an example of varying precision is whether a location is expressed as detailed latitude/longitude degrees-minutes-seconds pairings, (2) a street address or (3) high-level city or state information.

(3) **Data relevance** describes how closely data fit the purpose for which they are used. Sometimes data are adapted for a use other than their original purpose or are collected prior to the identification of a specific use. If data are only approximately relevant, they might be accurate and appropriately precise but not enough to yield high-accuracy results when used in a model. This might be the case when available data are used as a proxy for more relevant data that are not available.

Theoretically, it may be possible to collect data that are as accurate, precise and relevant as desired. But practically, any data or collection process is limited by analytical constraints. In the field of Big Data Analytics, these fundamental constraints are often expressed as the "three Vs":

1) **Volume** is the problem of having many records and/or fields in a dataset. The higher the data volume, the harder it can be to comprehend features of the data and the longer it may take for computers to process the data.
2) **Variety** is the problem of data being inconsistently formatted or unstructured. For example, requiring investigators to provide spreadsheet summaries of their accident investigations may yield a lot of informative data. However, if each organization arranges its summary in a different format, it may be difficult to make these data useful.
3) **Velocity** of data occurs when new data are continually becoming available and, therefore, must be processed continually to be relevant.

Although continued increases in computing power have made Big Data Analytics less of a challenge than in the past, the three Vs must still be considered when developing data collection and integration programs. However, the pursuit of improved data relevance, data accuracy and data precision often necessitate increasing the analytical load in terms of the volume, variety and velocity of the data collected.

This pursuit must take a balanced approach where the benefits of data quality improvement outweigh the increased analytical load. Examples include:

- Increased data volume improves data accuracy and statistical confidence.
- Collecting a higher volume and variety of data increases the probability of collecting data relevant to future, unknown analyses.
- High-velocity data enable real-time decision-making support.

Increasing the analytical load does not always lead to improved data quality. For instance, poorly standardized data collection (high variety) might create inconsistent levels of precision or accuracy in a dataset. If the inconsistency leads analysts to accept the lowest level of precision or accuracy, then quality has been lost despite the increase in analytical load.

Data Uncertainty

The purpose of collecting data is to support decision making. It is the foundation of the process, but data alone may not be inherently useful. Rather, making data useful is often a progression from data to information to knowledge and finally to wisdom (Figure 3).

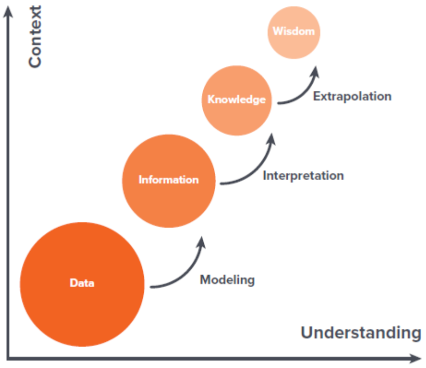

Figure 3 – Data to Wisdom (Source-ABS Group)

Fundamentally, a company needs to understand why it is collecting the data and what it hopes to achieve with it. Companies that have a clear vision, plan and conduct feasibility trails generally progress faster.

Information uncertainty is a function of both data uncertainty (the perception of data accuracy) and model uncertainty. In general, a model's output information will not be of higher quality than its input data. Likewise, a model's analytical scope and complexity will determine how well the model reflects reality or predicts future outcomes. Applying very high-quality data within a model that does not appropriately address key factors may yield highly uncertain results. Whether there are major quality issues with the data or major logic issues within the model, the old adage holds true: garbage in, garbage out.

The progression from data to wisdom is a continuum rather than discrete steps. Highly sophisticated model outputs sometimes approach the threshold.

Lessons from the Early Adopters

The companies who have performed better to date have generally adopted a structured and managed approach, typically having the following attributes:

1) A clear vision, strategy and foundation in terms of competence, enabling technology and organisational readiness
2) Definition of a set of prioritised business opportunities.
3) A clear process for data identification, selection and analysis.
4) Securing "good data" in a timely manner.
5) Effective data presentation to extract the insights.
6) Defined response triggers, workflows and outcome reviews.
7) A continuous improvement review process.

Figure 4 presents the above steps in a workflow (Paula, Arendt & Mowrer, 2016).

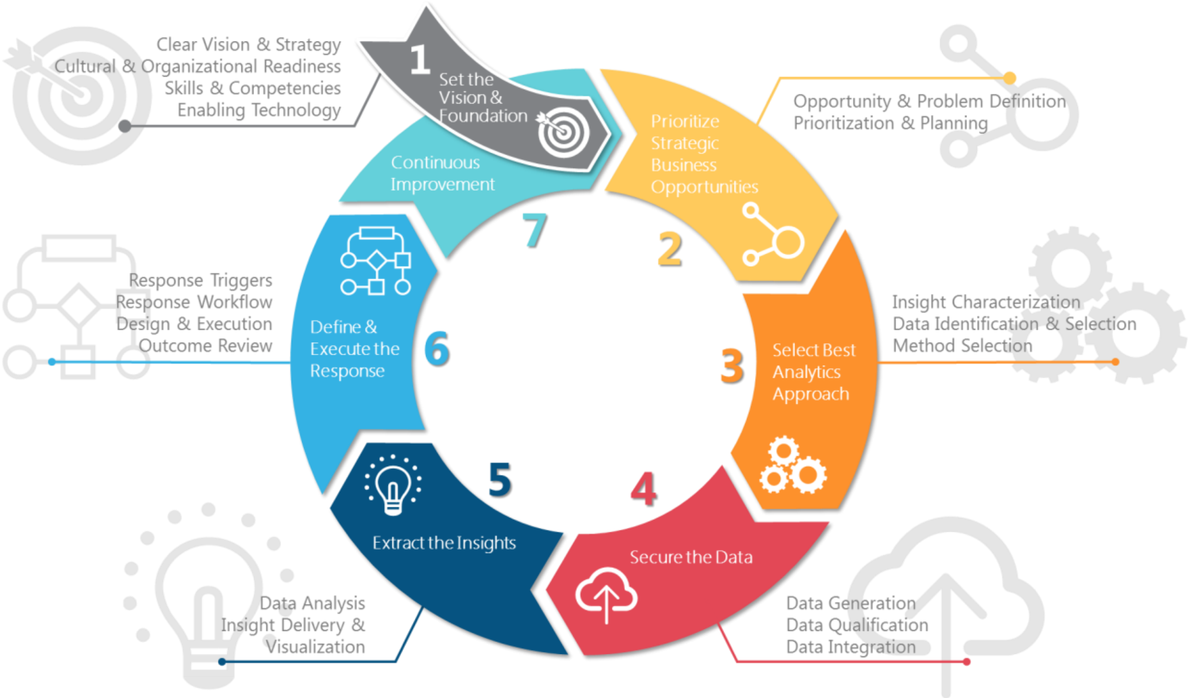

Figure 4 – Data Analytics Workflow (Source-ABS Group)

How to Get Started – Achieving Efficiency and Effectiveness in DA & DM

It is fundamental that organisations consider questions such as "where are we today, what are our asset management needs, what are the regulatory compliance needs, how can we use data as an asset?" A concept of maturity and expertise already exists, which can help Operators to differentiate themselves. Figure 5 illustrates this concept (Frost C., 2015). Operators can assess where they are today and decide where they want to be to benefit most from DA&DM. A typical maturity assessment will evaluate the existing capacity to gain an accurate and deep intuitive understanding of the assets as well as the broadness of the application from component to enterprise level. A fully mature company operates in the greener portion in Figure 5.

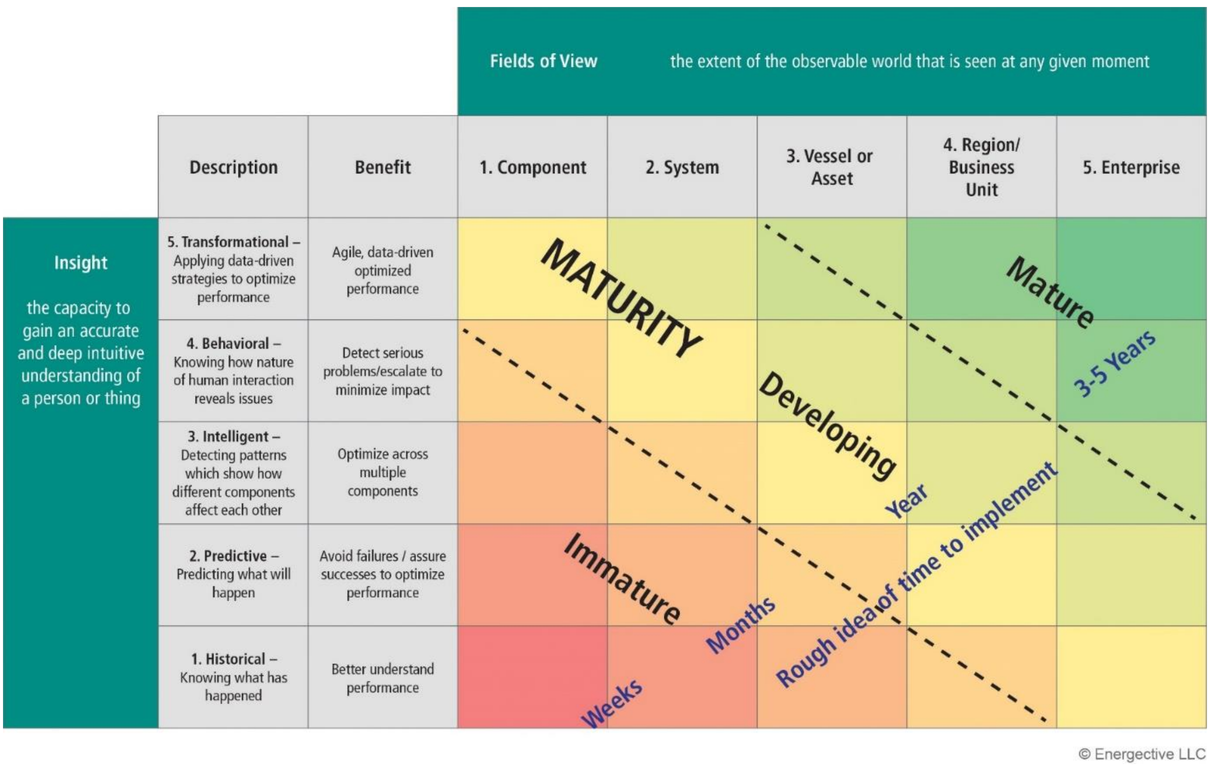

Figure 5 - Maturity Spectrum (Source-ABS Group)

Having understood the current level of maturity of the organisation, and if there is a desire to implement a data management strategy, it is recommended that a data management framework is used to support the workflow presented in Figure 4. Figure 6 presents an example Data management Framework (Paula, Mowrer & Roberts, 2016).

Figure 6 – Data Management Framework (Source-ABS Group)

Each of the elements is described below.

Strategic Alignment of Data with Business Priorities - Formalizing approach to align Data Management activity planning (e.g., 3 – 5 year plan) to business priorities (safety, environmental protection, reliability, regulatory compliance, optimized asset operations, etc.)

Business Ownership of Data and Data Governance - Formalized business executive champion, governance, clear accountabilities, business ownership of data – both structured and unstructured, defining policies and standards.

Defining the role of Data in Key Work Processes - Documented interplay and interdependency between critical work processes and data and clear definitions of the process requirements of Data.

Data Management Organization and People - Clearly defined roles, accountabilities and responsibilities, formally defined skills and competency requirements, career development paths, etc.

Data Management Processes - Putting in place standard definitions of key processes such as Data Cataloguing, Data Migration, Data Security, Master Data Management, Data Quality Management, etc. and ensuring compliance with those processes.

Performance Management - Putting in place performance management KPIs, metrics and associated processes and systems specifically for monitoring and managing performance in Data Management (e.g., around data quality or OPEX spend on DM).

Data Management Systems and Infrastructure - Ensuring there is a formalized approach to promoting standardized and optimized key DM systems and tools e.g., for storing master data, migrating data between systems, cleaning data to standards, archiving to secure storage, ensuring easy access, etc.

Data Security - Systems and human process safeguards put into place to ensure correct data is available only as expected; that confidentiality of data is maintained in all circumstances and that integrity of data is assured through all operations.

Change Management - Recognizing the diversity and abundance of stakeholders who have a dependence on quality data, understanding the impact of changing the way data is managed and ensuring a proactive approach to stakeholder management and change management to secure support and minimize disruption.

Concluding Remarks

The benefits of DA&DM have the industry-wide potential to run into the billions of dollars per year, with applications ranging from basic equipment optimization to enterprise-wide asset performance improvement. However, companies need to understand the value of data and recognize DA&DM as "good investment" vs. "cost of doing business." It is fundamental that they consider questions such as "where are we today, what are the asset management needs, what are the regulatory compliance needs, how can we use data as an asset?" Also, they need to be proactive in getting started or stepping up the pace on the data journey.

References

BOEMRE. (2011). Report Regarding the Causes of the April 20, 2010, Macondo Well Blowout.

DHSG. (2011). Final Report on the Investigation of the Macondo Well Blowout. Berkeley: Center for Catastrophic Risk Management, University of California Berkeley.

Frost, C. (2015). Data Analytics Value Model. Retrieved November 25, 2015, from Energective: www.Energective.com.

NAS. (2012). Macondo Well Deepwater Horizon Blowout: Lessons for Improving Offshore Drilling Safety. National Academy of Sciences. Washington, DC: The National Academies Press. doi:10.17226/13273.

Paula, H., Mowrer, M., & Roberts, B. (2016). Data Management Framework. Retrieved from http://www.abs-group.com/Knowledge-Center/Insights/.

Paula, H., Arendt S. & Mowrer, M. (2016) Data Analytics & Data Management (DA&DM) Solutions to Manage Pipeline Integrity and Comply with New PHMSA Regulations, Mary Kay O'Connor Process Safety Center's 2016 International Symposium.