W011: Setting the Correct Data Management Level for Model Centric Workflows

J.B. Kozman (Schlumberger Information Solutions)

The transition from UNIX-based data centric interpretation applications to powerful laptop-based model centric workflows began over 5 years ago (Shirley, 2002). It has recently been accelerated by the availability of 64-bit operating systems and extended memory. This has led to unique challenges in data and information management. Interactive workflows and flexibility are what make the new generation of interpretation tools attractive to end users, yet it is difficult for PC-based applications to support the robust multi-user database functionalities that could be designed for workstation-based products. This means that the new architectures will always be dependent on some kind of binary or flat-file data structure.

These limitations require data managers and information management organizations to adapt existing strategies for using these new tools in enterprise-wide deployments. The progress of these initiatives can be tracked using the industry-accepted Data Management Maturity Model (DMMM) first proposed in 2000 (D'Angelo and Troy, 2000), and data managers can select the proper level of management for the size and complexity of workflows within their organization (Fig. 1). Data management for model centric solutions can be implemented from a Base Level I, recognizable by flat file export and import workflows performed by individual users, to Predictable Risk at Level IV, characterized by six-sigma data quality methodologies. The industry now has substantial experience with using Standardized Data Management Site Assessments to analyze maturity levels and recommend strategies for movement along maturity and complexity scales (Hawtin, 2007). Successfully matching the correct level and strategy to an application deployment will result in documented reductions of cost and effort (Kozman and Hawtin, 2008).

E&P DATA MANAGEMENT MATURITY MODEL

Maturity levels (I-V)	Process performance	Technology support	Quality, predictability of results	Value determination	Comments
V) Fully optimized Exploration success rate is close to development success rate.	Processes are almost entirely automated.	E&P data mining; expert systems.	Almost complete certainty of results is achieved.		There may be no commercial market for this level of performance.
IV) Predictable risk Ability to routinely reduce uncertainty and data-related risk.	Statistically stable processes routinely measured against industry standard performance metrics.	Automation and background performance of processes, tasks; automated decision support services.	Reliability and predictability of results is significantly improved (i.e. six-sigma vs. three-sigma).	Lower ROI on investments in data management accepted in exchange for reduced risks.	This level may offer diminishing returns on investments; for many, it might be more cost effective to accept somewhat uncertain results and execute.
III) Corporate competency Capabilities are institutionalized within company; enabled by mature technology.	Standard, consistent, statistically capable, measurable processes; standardized process performance metrics begin to evolve.	Integrated technology designed to enable emerging best practice processes; technology suppliers are partners in defining how technology accomplishes best results.	Good quality results within specified tolerances most of the time; poorest individual performers improve towards best performers; more leverage achieved on best performers.	Measurable; able to recognize costs and benefits, perform cost-benefit analyses, maximize ROI; more good results faster and with fewer people.	Evidence of co-evolution of best practice processes and advanced technology; deployment of standardized processes and technology across multiple locations to leverage investments (economies of scale).
II) Managed Standardized tasks and roles; introduction of advanced echnology begins.	Individuals develop and follow processes that work for them; processes not common among individuals or across locations.	Unintegrated point solutions designed for specific tasks; individuals' primary responsibility is to figure out how to integrate and use technology to accomplish results.	Variable quality with some predictability; best individual performers put on business critical projects to reduce risk and improve results.	Anecdotal; based on individual performers' capabilities and specific memorable events.	Individuals' performance varies, but some may be highly effective. This level is effective with a small number of people in single location, managing small-moderate data volume flowing in at limited rate .
I) Base Capable people and heroic efforts	No defined processes; individual performers may follow a different process each time.	General purpose tools (i.e. Excel, Access) or none at all; data management is mainly personal function - not corporate.	Corporation depends entirely on individuals; little or no corporate visibility into data management cost or performance; variable quality, low results predictability and repeatability.	Subjective; gut feel for performance, costs and value received.	Craftsman level of performance - prior to specialized technology and known best practices, only way to accomplish task..

Figure 1. The Data Management Maturity Model (from D'Angelo and Troy, 2000)

Level I Base solutions can be deployed by letting each individual end user follow a different process each time data is moved in or out of the target application. At this level, data management is a person centric function using ASCII or flat-file imports and exports, and the organization does not attempt to define or enforce any standard processes. Capable people and often heroic efforts are required to sustain this type of deployment and to implement or recreate data flows within or between teams. General-purpose productivity tools such as Excel® are used to document metadata. In this scenario the enterprise does not have or desire visibility into data management functions, leaving it to individuals to produce their own results. Without any specialized technology or best practices, there is only a "gut feel" for the cost and value of geotechnical data being used in model centric workflows. Performance and quality of data processes vary greatly from user to user, and results are difficult to predict or repeat. A general rule is that during a standardized site assessment, if users are asked about data-loading processes or datastores and the answer involves a specific person's name ("Where do you store your production data for the North Basin Field?" "Oh, Bob keeps that in his desk."), a Level I organization is indicated.

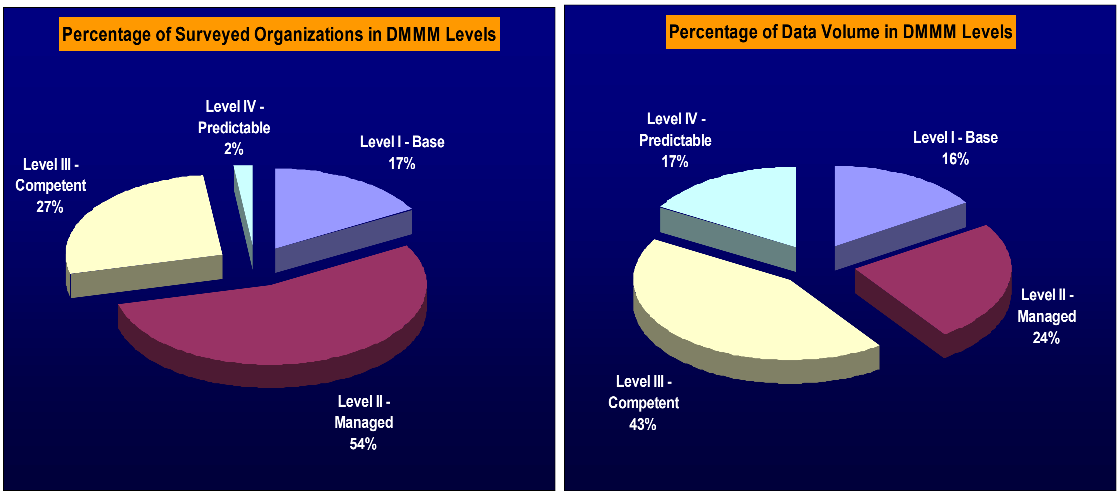

Figure 2. Maturity levels determined from site assessments.

In Standardized Site Assessments (Hawtin, 2006), 17% of surveyed organizations fell into this level. This is most certainly a low estimate because organizations that do not recognize the value of data management seldom participate in such assessments (Fig. 2).

An organization moves to Level II, Managed, on the Data Management Maturity Model for model centric workflows when they begin to introduce advanced technology and to standardize tasks and roles for handling data. Solutions such as point-to-point transfers and data and application integration frameworks are being used for specific tasks, but processes are not consistent between individual users or across different locations. At this stage it is still the responsibility of the geotechnical user to develop the best way to integrate technologies and produce results. This stage sees the deployment of data-sharing, data-transfer, and data-link third-party solutions to support desktop applications, as well as the beginning of standardized storage locations for reference projects or shared data. A Level II organization will begin to incur some cost for data management in either IT services or administration of software. Risk can be reduced using this level of maturity and results improved for business-critical projects by leveraging the best data management performers, but the value of the data management solution is still determined largely by anecdotal capture of memorable events. These events are usually negative, such as having to redo a basin study because results of the original cannot be located, or purchasing the same seismic survey that is already loaded in another application. This level of data management can support an initial deployment of model centric tools, but it is quickly overwhelmed by increases in the number of users, rollout to multiple locations, or increases in data volumes or data acquisition rates.

At Level III, identified as Corporate Competency, the first standardized and consistent methodologies begin to evolve along with mature technologies for integrating model centric technologies and enabling best practices. These technologies may include validated corporate datastores, search engines for unstructured data, application programming interfaces (APIs), and links to Geographic Information Systems (GIS) and plug-ins that allow data management tasks to be completed in context without leaving the native application. The existence of corporate repositories for well data at most large exploration companies is reflected by the higher percentage of managed *volumes* in Level III and above implementations (Fig. 2). This level of Data Management Maturity for model centric workflows is recognizable by measurable processes and institutionalized capabilities; a Level III organization can usually report as a process performance metric the percentage of well and/or seismic data that is loaded into its corporate repository, and there are documented plans to improve performance. At this level there is a designated data management role charged with enabling standard and consistent processes and evaluating the cost and benefits of integrated technology across multiple locations. Level III organizations start to specify tolerances for volumes and completeness of data loaded to applications, but return on investment (ROI) measures still involve only percentage reduction in time spent on data management tasks versus analytic and interpretation activities (Fig. 3). The practice of using reduced cycle time as a proxy for an increase in quality results produced by a reduced number of personnel has persisted for the last decade without improvement (Beham, Brown, et al., 1997).

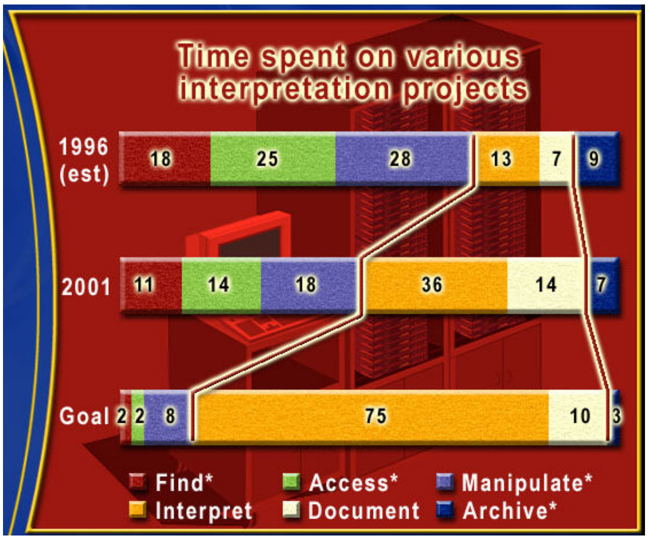

Figure 3. Typical analysis of data management for model centric workflows from 1996 (from Beham, Brown, et. al., 1997)

A critical indicator for Level III is that technology suppliers become partners in planning and measuring the success of deployment strategies for model centric workflows, allowing client organizations to leverage investments from multiple implementations. Level III organizations are still unable to produce quantitative measures for consistency of data across multiple interpretation platforms or between asset teams. This ability to enable a data quality methodology marks a transition to Level IV implementations of model centric solutions.

Some international oil and gas organizations have achieved a Level IV of data management maturity, either for specific geographic areas or data domains or with individual asset teams during targeted rollouts. The key differentiator for a Level IV deployment is that the reliability and predictability of data can be measured using six-sigma methodologies. This requires rigorous and consistent application of business rules for completeness, consistency, and accuracy of data being used to develop the geological and geophysical models. Processes for loading, quality-controlling, and synchronizing data in multiple applications must be performed automatically as background tasks. The organizations must be able to routinely demonstrate a reduction in uncertainty and risk associated with data, while measuring themselves statistically against industry standards and employing advanced technologies such as automated decision support services. The ROI for investments in data management at this level must be measured by quantifiable reductions in risk, such as increased exploratory drilling success or improved production rates.

Measuring against industry standards means that Level IV organizations planning on deploying model centric workflows must be able to compare their maturity level against peers based on company size, diversity of geographic operations, and corporate cultures. This is accomplished by plotting maturity against a complexity metric (Hawtin, 2007) derived during a standardized site assessment. Analysis of the location of a deployment in this information management workspace determines the most effective ways to implement model centric workflows on an enterprise level (Kozman and Hawtin, 2008). A third dimension of analysis is used to customize model centric deployment strategies based on the expected change Trajectory (McGahan, 2004) for both the end-user organization and the technology provider. This allows data management and information technology support groups to select the best tools as user communities change (Fig. 4), and to maintain the most advantageous relationships between suppliers and consumers. Based on the threats to core activities and assets, oil and gas organizations deploying model centric workflows are undergoing creative change while the suppliers of analysis software and data management technologies are on an intermediating change trajectory. This also impacts the choice of strategies for deploying these technologies, and the preferred relationship with technology providers.

Core assets	Core activities: Threatened	Core activities: Not Threatened
Threatened	**Radical Change** *Everything is up in the air.* Examples: makers of landline telephone handsets, overnight letter-delivery carriers, and travel agencies	**Creative Change** *The industry is constantly redeveloping assets and resources.* Examples: the motion picture industry, sports team ownership, and investment banking
Not Threatened	**Intermediating Change** *Relationships are fragile.* Examples: automobile dealerships, investment brokerages, and auction houses	**Progressive Change** *Companies implement incremental testing and adapt to feedback.* Examples: online auctions, commercial airlines, and long-haul trucking

Figure 4. Classification of change trajectories (from McGahan, 2004)

By utilizing available analysis and planning tools, oil and gas organizations can now select the best strategies to reduce the cost and effort of successfully deploying model centric solutions at an enterprise level.

References

Beham, R., Brown, A., Mottershead, C., et al. [1997]. Changing the Shape of E&P Data Management. *Schlumberger Oilfield Review*, Summer 1997

D'Angelo, T. and Troy, B. [2000] Data management helps create value. *The American Oil and Gas Reporter*, and *Oil and Gas Journal*, July 2000, PennWell Corporation, Houston

Hawtin, S. [2006] Experience from IM Assessments. 10th International Conference on Petroleum Data Integration, Information & Management, Houston, Texas, 2006

Hawtin, S. [2007] Adding Value with IM Assessments: E&P Data Management Today. Schlumberger White Paper, www.hub.slb.com/Docs/ofs/sis/06_IS_201_im_asses_wp.pdf

Kozman, J. and Hawtin, S. [2008] The Main Sequence: Matching Data Management Change to the Organization. 12th International Conference on Petroleum Data Integration, Information & Management, Houston, Texas, 2008

McGahan, A.M. [2004] How Industries Change. *Harvard Business Review,* October 2004, Harvard Business School Publishing Corporation

Shirley, K. [2002] PCs' Muscle Makes It a Contender. *AAPG Explorer*, June 2002