

Data Governance at Chevron GOM

A Case Study

Authors: Dave Blosser – Chevron North America Exploration and Production
Paul Haines – Noah Consulting

PNEC 17th International Conference

On Petroleum Data Integration, Data and Information Management

May 2013



Table of Contents

<u>1.0 Abstract:</u>	<u>3</u>
<u>2.0 Problem Statement:</u>	<u>4</u>
<u>3.0 Approach to Data Governance:</u>	<u>5</u>
<u>4.0 Implementation of Data Governance:</u>	<u>12</u>
<u>5.0 Building Sustainment Through a DG Organization:</u>	<u>24</u>
<u>6.0 Beyond Data Governance as a Project:</u>	<u>28</u>
<u>7.0 Summary:</u>	<u>31</u>
<u>8.0 About the authors:</u>	<u>33</u>
<u>9.0 About the companies:</u>	<u>34</u>



Data Governance at Chevron GOM – A Case Study

1.0 Abstract:

Data governance (DG) is one of the pillars of data management (DM) and is often considered hand-in-hand with data quality efforts and master data management programs. In our exploration and production (E&P) community, initiatives to implement data governance are often seen as either organizational efforts or technology efforts and are often spun up from the corporate office with a top-down emphasis. That approach, although well intended and with valuable recommendations, is often met with internal resistance and suspicion and falls short when it comes to implementing and sustaining data governance to make a truly significant and lasting impact within the business.

The co-authors present this case study to demonstrate how data governance can be implemented in a unique and pragmatic federated manner. Our approach emphasizes a repeatable and sustainable methodology focused on supporting key business processes. We will review the methodology, components, and stages developed to implement data governance for specific data types through several proof of concepts. We will present lessons learned, challenges encountered, and business benefits realized through our efforts to date.



2.0 Problem Statement:

Many papers have been presented in this conference that relate to data issues, data quality challenges, data inconsistencies, untrusted data, etc. At Chevron’s Gulf of Mexico Business Unit (GOM or GOMBU), end-user surveys showed general dissatisfaction with data quality. There were known issues related to data access and data delivery to the business user. Initial assessments that led to this project showed inconsistent systems of record (SORs) and/or too many sources of data. Additionally, historical data tended to get lost or disorganized during post merger or acquisition activities. Essentially, Chevron GOM had problems similar to many other E&P companies.

This section will not be the focus of this paper, but rather a short introduction so that the audience can relate to the historical situation. Chevron GOM had well sponsored data projects and data clean-up projects to address the challenges stated above. Data projects started but often lacked specific goals that ensured sustainment. Data projects were performed to “make it clean” and the data quality was improved, but it didn’t stay clean. There was not a good understanding of the root cause; there was not a good understanding of how to “keep it clean.”

Data cleanup can happen by brute force and short term improvements will be seen, especially if coupled with improved business processes, but process improvement is not enough. The proper implementation of DG makes it work; makes it sustainable. That is the conclusion that Chevron GOM came to. Over time and through evangelical meetings, the business came to agree that data cleanup efforts (remediation) were not enough. You can improve Data Quality for a short period of time, but without data governance, it will not last.

There was a point in time where IT and the business said, “We’ve been talking about data governance for a long time – let’s do something.”

Through initial efforts and as a DG project was launched, it became clear that a majority of the business did not understand what DG really entailed. Business leaders knew what they wanted (they wanted trusted, clean data), but they had a misguided concept on how to get there. Business, management, and IT did not understand the fundamentals of “how to do” DG. Also, the project sponsors and senior business managers needed to see something “real;” something that would help them realize the value DG can provide. That realization is what led to the decision that formal training in the fundamentals, concepts, and framework was required – a DG Academy was the first step.

3.0 Approach to Data Governance:

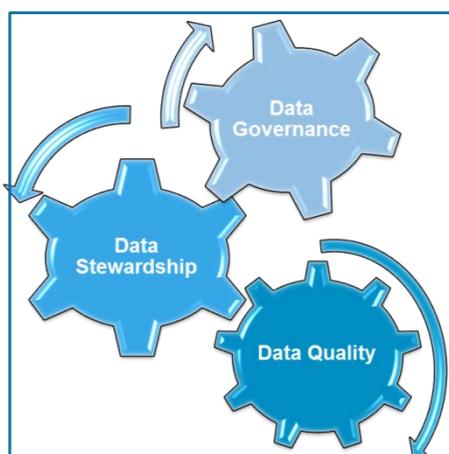
This section will describe the overall DG process, the thinking that went into the DG Academy, and the project structure.

Data governance along with data stewardship and data quality together provide the foundation for properly managing and sustaining information as a key asset in an E&P company.

DG provides a framework of principles, policies, standards, roles and responsibilities, and processes to enable effective information management. DG lets the business take proper and responsible ownership of their information where everyone clearly understands their role and the data they are responsible for. Data governance includes clear descriptions of the tasks and activities that the business must perform to achieve and maintain known data quality. As will

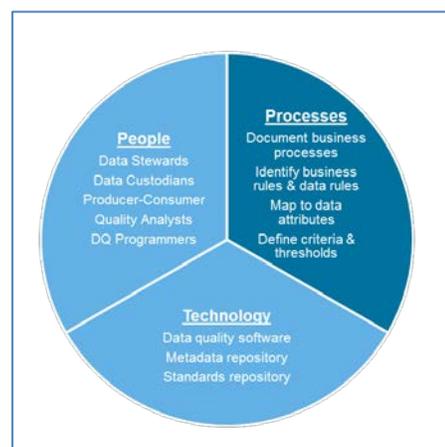
be shown, data governance refers to people and organizational capability, processes and controls, and technology and architecture. A proper balance between people, process, and technology is what enables data governance to be successful.

There are clear and specific DG tasks and activities that can be described and documented and DG roles that need to be defined. Data governance activities include:



- Apply data policies
- Resolve cross functional or cross regional data issues

- Define data and business rules
- Maintain end-to-end responsibility for at least one subject area
- Monitor and sustain quality through data governance policies
- Define standardized data-related business performance measures
- Work with data quality (DQ) groups to ensure compliance with established standards and definitions
- Provide input into DQ program prioritization





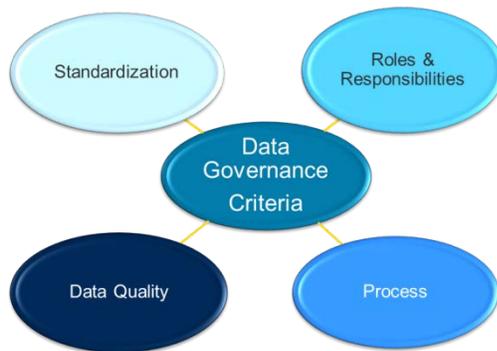
Additionally, in each client organization specific DG roles need to be defined and responsibilities documented. This is essential to establish a DG organizational capability. An example of a listing of roles might include:

- Data content owner
- Data steward
- Data custodian
- Data architect
- Data quality analyst
- Data technical analyst
- Data business analyst
- Data producer
- Data consumer

Additionally, there are roles that provide sponsorship, steering, guidance, and/or pace and priority to the DG program.

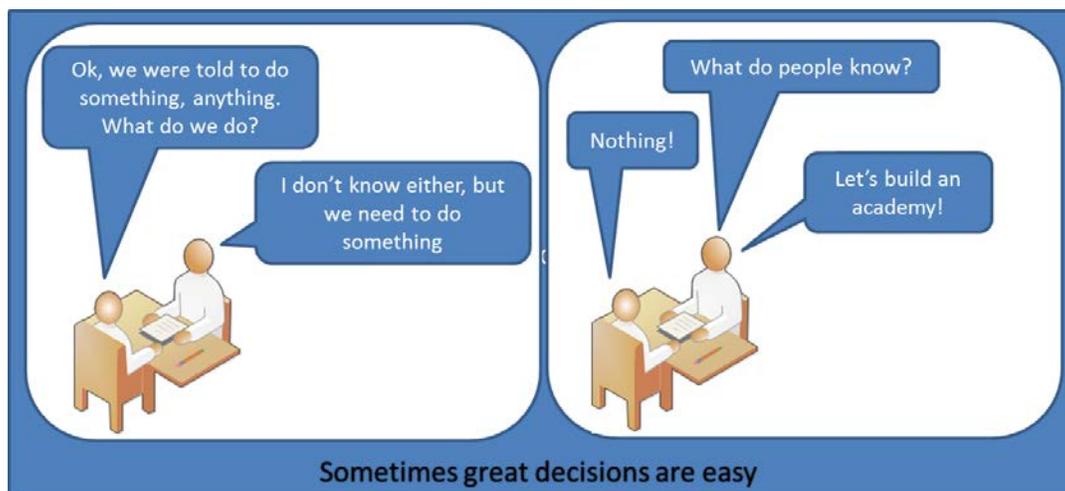
What does it mean for data to be governed? Since no one seemed to have a good definition, we developed our DG criteria to include those elements which, if true, give us confidence that our data is well governed. Thinking of this in terms of a check list, the four key areas to describe include:

- Standardization
 - Are SORs identified and in place?
 - Have standards been established and documented?
 - Are data models architected and documented?
 - Is secure access managed?
- Roles and responsibilities
 - Are data producers and data consumers identified?
 - Are data stewards identified, trained, and their role communicated?
 - Are data analysts and data custodians engaged?
- Process
 - Is master data managed?
 - Have business processes, DG processes, and business/data rules been applied?
- Data quality
 - Are data rules tied to business processes?
 - Are data rules tied to, mapped, and documented to appropriate repositories?
 - Is data monitored to a standard?



These criteria have been established during this project to determine when data is governed to an agreed upon quality standard.

The establishment of a DG Academy began as a way to bring the business up to a proper level of understanding of DG, what value it could provide, and how Chevron GOM would deliver DG and its associated standards. The DG project team wanted to create a reality based understanding of DG among IT project teams that were delivering data-related IT projects. The DG Academy was thought of as a way to effectively communicate what DG is, why it is important, and how to effectively implement it at a high level. Another question contemplated by the project team was how to administer the DG Academy.



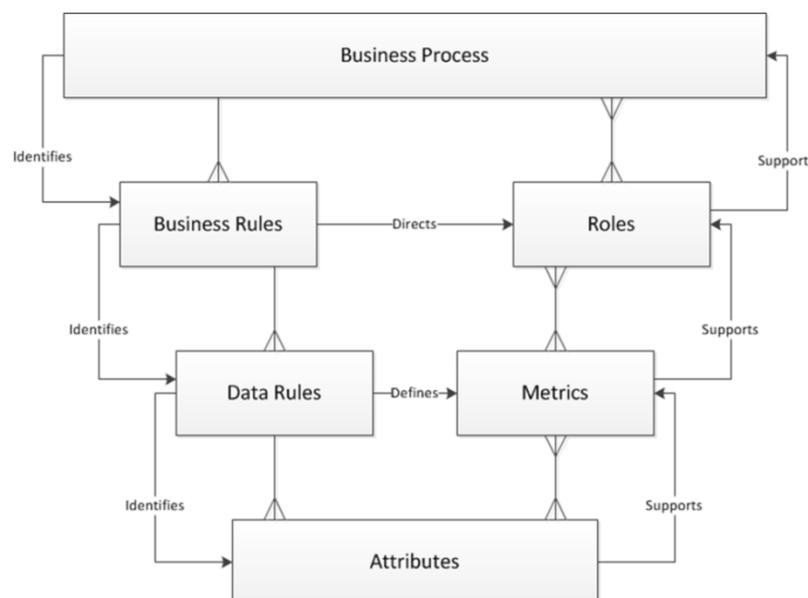
The strategy relating to the administration of the DG Academy included:

- Employ a “light touch” methodology.
- Create a classroom setting to establish DG concepts.
- Engage the participants in workshops to provide hands-on practice in the concepts.
- Through the DG Academy, present the DG concepts in a consistent manner.
- Use specific collateral relating to the data in the DG project.
- Use the data project team to bring out existing DQ experiences, anecdotes, and usage needs that would boost the value of the resultant standards.
- Leverage industry standards such as standard definitions created and accepted by the PPDM Business Rules Workgroup; of which both Chevron and Noah were contributing members.
- Create 1st draft collateral during the workshops to accelerate the development of the standards and engage the participants from the data projects.



Another strategic decision made for the development of DG within Chevron GOM was related to the overall approach. The overall approach focuses on key business processes as the primary dimension for developing data governance to sustain data quality. Key business processes are important because they can be used to reflect real business workflow problems and inefficiencies and allow us a pragmatic method to tie low level data issues to high level business concepts. It has been seen that poor quality data, missing data, etc. cause key business processes to break, fail, or be sub-optimal. The methodology includes:

- Describe key business processes that can be related to data flow.
- Map out these key business processes in standard business process workflow diagrams.
- Identify business rules that support the business data and/or business decisions where:
 - gaps in process are caused by a lack of standards or lack of business rules;
 - data quality issues are causing the process to fail or not be optimal;
 - business rules affect the design of the workflow;
 - the business rules are at a level that is data model and application independent.
- Identify specific data rules that support the business rules; these data rules will be at a granularity that can be acted upon to identify, understand, and rectify the data quality issues impeding the business process.
- The data rules are specific to a data type or data attribute and can be mapped to the SOR or application.
- The data rules should be such that a test can be created and administered that supports the data rule and can be executed manually or automatically.
- The business processes, business rules, and data rules are fully defined and attributed so they can be managed within the DG environment and DQ framework.
- Leverage industry standards such as the business rule and data rule attribution that came from the PPDM Business Rules Workgroup.





The delivery of the DG Academy was designed to be:

- flexible, but comprehensive;
- taught in compact, segmented timeframe so concepts would build upon each other;
- split into tutorials and workshops (about 60% / 40% ratio). The workshops:
 - helped to solidify concepts into the participants working knowledge
 - helped to begin developing the variety of standards needed for creating the sustainable DG environment and DQ framework

Once some of these (and other) fundamental concepts had been established within the DG project team, decisions needed to be made related to the details of the content and how it would be delivered within Chevron GOM. For that, the DG project team decided to bring in information management consultants who had experience in successfully delivering DG and DQ in the E&P industry and would be able to provide resources, thought leadership, and focus. These resources were added to the DG project team.

There are various ways to begin a DG effort; there were active projects working to establish new SORs and an associated data cleanup effort in progress. The projects did not include an element of DG in their scope. We knew these projects were doomed to ultimate failure. It simply made sense to partner with these projects to enable their success and to provide fertile ground for the DG project team to test their ideas. There were several advantages to this:

- The data group worked closely with the business on their highest priority issues.
- The initiatives were business focused and business led.
- The initiatives were business and IT enabled – both groups had skin in the game
- There was strong business acumen on the project.
- Some of these issues were directly related to data and data quality.
- A deliverable to provide a sustained data quality improvement was put into the group's project charters so the DG project was seen as an enabler to meet the group goals.
- The projects contemplated were data and/or workflow centric.
- It was anticipated that working with and training the data group in the fundamentals of DG and helping them establish a sustainable framework for their data types was a clear win for both sides.

As the methodology was being developed and refined, both the structure of the project and the format of the DG Academy were maturing.

Following after the Chevron project execution process structure, the project was divided into 6 stages corresponding to four of the Chevron process phases represented in the graphic as yellow, green, blue, and purple. The data type for the project is to be prioritized and identified as a precursor to the DG Academy methodology.



- Stage 1 – Harvest
 - In this stage, the DG team worked with the data project to harvest artifacts that represented the current state of that data. These artifacts included business processes, data standards, data flows, data models, data rules, etc. These artifacts included the official or de facto standards relating to the data type. During this stage all the collateral for the DG Academy, including the content for the tutorials and the workshops are prepared. The harvested artifacts provide the team the information needed to build the workshop materials so they are specific to the project. The artifacts become an understanding of the starting point for the DG Academy efforts.
- Stage 2 – Draft
 - In this stage, the DG team actually presents/teaches the school. Following the concepts of flexible but comprehensive, the school can be administered in two days to two weeks, depending on the availability of the participants. The intent is for the first draft of the standard to come out of the workshops.
- Stage 3 – Refine
 - In this stage, the DG team works with the project team to take the results of the workshops and fully develop, expand, and refine them. The inputs for this stage are the business processes, business rules, data rules, data models, data quality metrics, etc. The output from this stage is the “standards pack.” The standards pack represents the final draft of a set of DG and DQ standards for a particular data type with elements required to create and sustain quality data.
- Stage 4 – Approve
 - In this stage, the standards are presented for approval to the Standards Committee. As seen in a later section, this committee is comprised of SME’s, Stewards, and the appropriate representatives who understand the data, understand the processes, and can make informed decisions with regards to the validity of the DG standards and the value proposition they will provide once implemented and executed.
- Stage 5 – Apply the Approved Standards
 - In this stage, the standards are used as requirements and specifications to be put into the DQ framework. The standards are identified in the DG Compendium as approved and sanctioned DG standards. As seen in a later section, the Data Custodians are responsible for applying the data rules to the DQ engine. The DG Compendium will be more fully described in a later section.



- Stage 6 – Monitor and Evergreen
 - In this stage, the DG standards have been applied into the DQ framework and are running against the data type. The DG processes to create, edit, delete, and manage all of the DG standards are in place and have been implemented so that the DG environment is properly maintained, sustained, and is evergreen. Thresholds, metrics, monitoring, and exception reporting are in place and have been implemented so that DQ framework reflects the DG environment. At this point, the DG criteria have been fulfilled and data is being fully governed.

The development of the methodology of the DG Academy and the DG project was a work in progress. To accelerate the learning another strategic decision was to test with a series of proof of concepts. Serendipitously, there was plenty of food for the fodder:

- A variety of data projects already in-flight and at different phases in their project
- Projects representing differing functional groups within Chevron GOM
- A variety of data types
- A variety of data format structures; structured, unstructured, and temporal

POC Process In-flight Data Projects				
POC	Stage I Harvest	Stage II Draft	Stage III Refine	Stage IV Approve
Reservoir Pressure	√	√	√	√
Seismic Meta Data	√	√	√	√
Facility Docs	√	√	√	
Process Historian	√	√		

The development of the DG methodology and actual preparation of the DG Academy material leveraged consultants experience and a proven methodology, along with Chevron GOM's thoughts and experience. The resultant collateral was a blend of structure from academia, thoughts from Chevron GOM, and conceptual collateral from Noah Consulting.

4.0 Implementation of Data Governance:

Once the DG standards pack has been created and approved, those standards must be properly leveraged in a data governance implementation for a given data type. The components of a DG implementation for a data type are building blocks for sustained data quality and include:

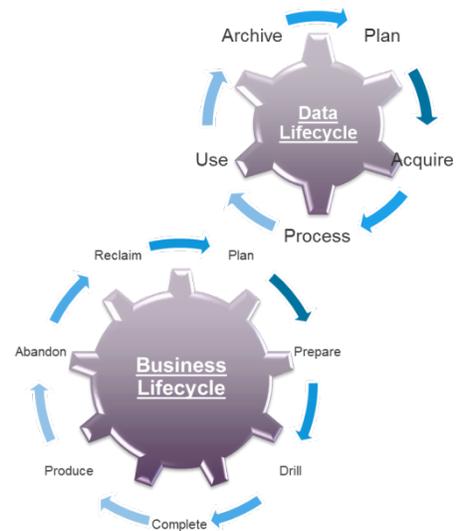
- Business lifecycle
- Data lifecycle
- Business processes
- Business rules
- Data rules
- Standards repository
- Data repository
- DG environment
- DQ framework and DQ engine



For any data type there is a need to map how the data type interacts and influences decisions made throughout the business lifecycle or business value chain. Additionally, the data lifecycle should be mapped out to understand how the maturity of the data through its lifecycle affects business processes, business rules, and data rules.

Additionally, the responsibility, content ownership, stewardship, etc. can/will change depending on both the business lifecycle and data lifecycle.

The business and data lifecycles are independent but inter-related. As data matures and improves through its data lifecycle, the business decisions made based upon that data mature and improve as well. With proper data management and data governance the data quality, confidence, and trustworthiness improve. By the nature of our upstream industry, data acquired early in the business and data lifecycles is used throughout the life of the asset.



The creation, approval, socialization, management, and sustainment of the DG standards are keys to success as defined in the DG methodology. The development of a standards repository was seen as critical to the success of the DG project. Historically, in other DQ or DG efforts, the resultant collateral (e.g. the standards) are often drafted and finalized as unstructured documents with no indexing, no metadata, and no technology to systematically tie them



together and keep them together as managed data. There are processes that need to be created to determine how the DG standards are stored, edited, added to, deleted, etc. There is technology required to sustain DG standards and deliver the material as needed. A data governance compendium was developed to satisfy that need.

The DG Compendium addresses these failings and provides many benefits. The DG Compendium is the knowledge base for data asset teams and business users to find what they need. It is comprised of four key components:

- The Organization Registry
 - Provides the names and contact information for stewards, custodians, architects, analysts and DG/DQ support personnel
- The Processes Library
 - Describes the processes for remediation and data creation including required training
- The Rules Repository
 - Identifies both the business rules for data asset and the data rules for monitoring and maintaining data quality
- The Data Catalog
 - Contains the data definitions and logical data models for all data types under governance and SOR links and metadata for data content



Besides a managed and maintained repository for the DG standards, the DG Compendium is envisioned to fulfill other important needs:

- It is used as a collaboration point to work with the DQ Support and Development team in IT responsible for taking the data rules and imbedding the rules into the DQ engine.
- It is used to on-board and train new business and IT resources so that data governance concepts become part of the fabric of the organization and help them understand how data-related processes affect their role in Chevron GOM.
- It is used to provide a platform to query, view, to better understand DG standards, request new rules and standards, request changes to existing DG standards, etc.

The DG Compendium acts as the working standards repository as well as the final, approved standards repository. The templates, definitions, and workflows associated with the DG standards and loading up the DG Compendium are also accessed from the DG Compendium.



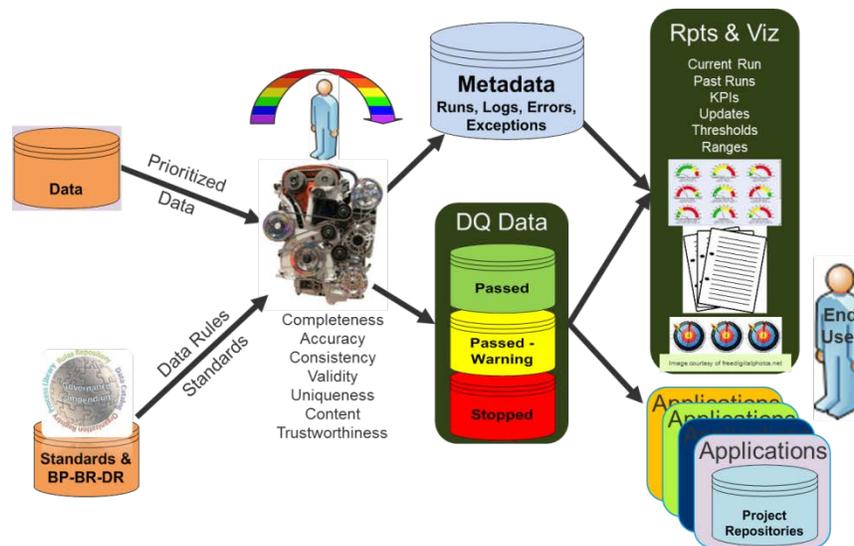
After the DG Academy is completed the work begins to improve and detail the business processes, expand and validate the business rules, and create and map the data rules to the SOR or primary repository for data remediation, and monitor the entire system. During the development of the DG methodology and the DG project we learned that adding properly implemented DG standards and components to an already launched project often resulted in resistance from stakeholders. This was rectified by adding resources on to the data project from the DG project team to support the DG effort. The benefits of this approach included:

- Experienced DG consultants were able to work alongside the data project team to hone the DG standards deliverables.
- The project team met their established project time line.
- The project team gained knowledge about data governance and share/provide knowledge about the data and/or workflows in question.

Creating and properly managing the DG standards through the DG Compendium is not enough, though. A DQ engine (Data Quality Engine) is required to take the specific data rules and act upon the data. Within Chevron GOM, two different technologies are already being used to act as that DQ engine and apply data rules to specific data types and data attributes. Since the DG methodology and standards within the DG Compendium are software agnostic, the DG project team leverages these technologies to act upon data. The collaboration and working relationship between the DG project team and the DQ support and development team (data custodians) is a key organizational aspect to effectively enable data governance and data quality to work together and for the data quality to be sustained.

As depicted in the graphic below, the DG Compendium stores the DG standards required for the DQ engine. The data custodian implements the rules in the DQ engine and relates them specifically to the repository (SOR or application) as it runs through the DQ engine. The DQ engine is also synchronized with the DG Compendium for the data quality dimensions and data rule types:

- Data Quality Dimensions
 - Consistency
 - Accuracy
 - Validity
 - Uniqueness
 - Content
 - Trustworthiness
- Data Rule Type
 - Corrective
 - Detective
 - Preventative



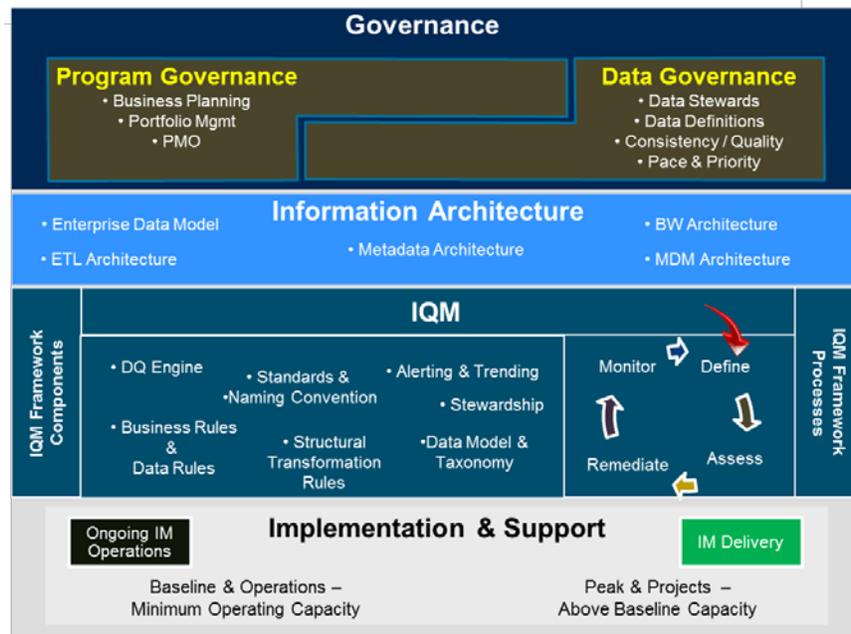
The output of the DQ engine includes results of the execution of the DQ rules and information related to whether the data passed or failed and metadata related to the results. The metadata includes run logs, error logs, exception reports, data quality level, etc.

The metadata, metrics, results, updates, ranges, thresholds, etc. can then drive reporting and visualization related to the “state of the quality of the data” and/or the “state of the data.” Additionally, the data, having been tested against the DQ engine and with known quality can be passed to the appropriate applications, data hubs, etc. depending on the end-user needs and requirements.

The DQ engine is a component of the overall DG/DQ Framework and environment. This Framework is comprised of several inter-related components. They are:

- Governance/Stewardship Model
- Information Architecture
- IQM Components
- IQM Processes
- Implementation and Sustainment

The generic framework depicted below, coupled with the DG methodology not only enables the DG strategy to be successfully implemented, but also achieves and sustains improved data quality. Some of the most important of these components will be conceptually described.



Data Governance and Stewardship

Data governance is the practice of organizing and implementing policies, procedures and standards to effectively manage the quality and usage of data assets. Data governance specifies accountability and responsibility for data quality and data stewardship. Data Stewardship is the implementation of data management processes in accordance with the governance policies, principles, and standards. Data governance and data stewardship are required in order to attain a sustainable data quality capability.

Data governance is about making data management better.

- Availability of data – right people, right data and right time
- Limiting the redundancy of data stores
- Increasing the reuse of data integration and publishing services
- Standard business performance measures - definitions and calculations
- Determination of appropriate use of data, (e.g., business sensitivity, confidentiality, retention)
- Documentation of what data is added to which data repository and when
- Documentation of data sources used for data processing and publishing
- Commonality of business facts - rules, data names, definitions and structures for data integration

Data governance serves many purposes within the DG Framework such as establishing roles and responsibilities for the players within the data quality and data governance community – data stewards, data quality analysts, members of the data governance and/or stewardship communities, and the group that helps to facilitate the DG Framework. But additionally and



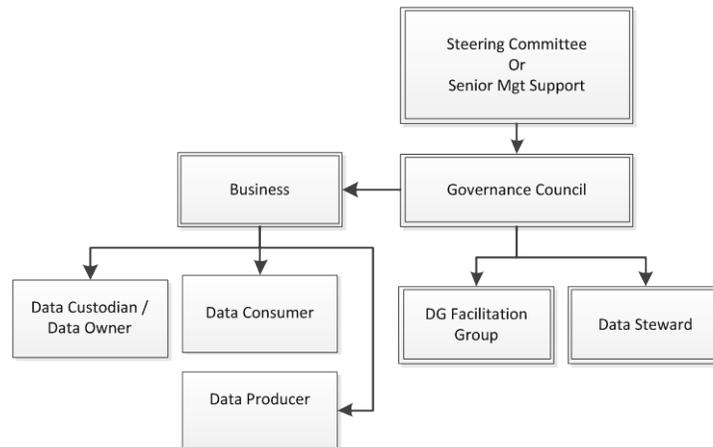
perhaps most importantly, data governance establishes and ensures accountability between data producers and data consumers. A data producer is a person, process, or application that creates or derives data that is stored, accessed, and used by one or more data consumers. A data consumer is a person, process, or application that relies on data to perform its required functions.

A lynchpin to achieving the benefit of sustained data quality is in the understanding of how these two players (often business people) relate to key business processes, business rules, and the underlying data rules that establish the level of data quality that satisfies the business needs. The data steward is the person who holds data producers accountable to the terms of their data quality metrics to fulfill the needs of the business process. When conflicts between a producer and its consumers arise, the data steward is responsible for mediating the situation by working with both parties to identify an appropriate solution.

Additionally, data stewards use the historical quality measures and metrics produced by data quality controls to perform reviews and ongoing analysis to identify the root cause of data quality impairments and identify corrections or improvements that can be made to the producing person, process, or application to eliminate or mitigate the possibility of future impairments being generated. In many cases, data produced in one business function is consumed in one or more other business functions. This is certainly the case with master data and, thusly, shows the importance of the role of a BU or corporate data steward.

Governance Organizational Model

The following model illustrates a generic data governance hierarchy example that supports a DG framework implementation and sustainment program. It is important to note that this is not an organization chart per se, but shows the relationships of the roles to each other within the DG framework.



One of the key benefits of this approach is that it supports an “organic” implementation. Additional DG roles and/or layers of DG responsibility are added as data types are added, as more key business functions are mapped and require support, and/or as the organization matures its implementation of a DG framework. The seed of data governance is planted by identifying subject matter experts, anointing them de facto business representatives and identifying resources that are doing data quality as a major aspect of their daily job and anointing them data quality analysts and/or data stewards.

Information Architecture

Much of the information architecture of the DG framework is developed in conjunction with, but not as a central part, of the DG efforts. The information architecture requires a business focus both on a logical data model for structured data and a primary taxonomy for unstructured data. For proper data governance, the focus is on (a) defining and documenting standards within both the structured and unstructured environment; (b) defining and documenting the data objects and attributes that support the key business processes and data rules; (c) creating and maintaining a data dictionary or data catalog of the standards; and (d) creating and maintaining data governance and data quality metadata.

An understanding of proper capture and management of metadata related to data, data quality, and data governance enables successful data governance implementation. A key work product of the development of a DG framework and DG implementation is metadata – metadata architecture, repository, and management. The success of implementing the IQM framework processes – define, assess, remediate, and monitor – to enable data governance, improve data quality, and sustain data quality is highly dependent on a robust metadata management strategy and system. Metadata is found in the DG Compendium and the DQ engine. Ultimately, there could/should be an effort to look at metadata across systems and functions on its own so that metadata can be modeled, architected, managed, and leveraged to



its best efficiency and highest value for Chevron GOM. But, as previously suggested, in the meantime caring for metadata associated with data governance, data quality, and stewardship requires dedicated effort to understand how metadata will help to manage the standards, rules, exceptions, alerts, etc. that are associated with the DG Framework.

The ability to add metadata in a straightforward and simple manner and access it in a timely manner is important to sustaining the data governance initiative and, ultimately, its success. This was accomplished by designing forms, processes, and worksheets to collect the metadata that directly relate to the DG standards, business processes, business and data rules, etc. The fully vetted and approved the data collected by these forms (etc.) become the DG standards and the metadata that supports the standards.

While a normalized metadata model will make porting to a more comprehensive solution easier, it will make creating reports somewhat more difficult. In most cases, ad hoc reporting will be beyond the average business user. Thus, the DG Compendium also needs a library of reports to meet the business needs. At a minimum, reports should be created and maintained that:

- allow users to view, browse, and query business processes, business rules, and data rules and their relationship to each other;
- allow users to query and browse for any and all data standards, including the organizational perspective of who are the stewards, SME's, custodians, data quality analysts, business analysts, etc. for any given type of data and their associated attributes;
- allow consumers to view quality measures and metrics for their critical data elements; both current and historical;
- allow producers to view their current responsibilities and adherence to their business processes and data rules that govern their data quality;
- allow managers to monitor their goals and objectives related to data quality thresholds, data KPIs, etc.

Ultimately, business will be best served with some type of dashboard technology to present quality measures, metrics, and statistics.

IQM Framework Components

The information quality management (IQM) framework components support the development, vetting, and approval of standards.



- DQ Engine

The DQ engine was previously described.

- Stewardship

Stewardship has already been discussed as it relates to governance and organization. Stewardship as a component here relates more to the function of stewardship, how the steward relates to the consumer and producer, how the steward relates to the Business Processes, Business Rules, and Data Rules, how stewards function to help govern data, etc.

- Business Rules and Data Rules

Governance helps provide consistent and logical business and data rules. A business rule is a statement that defines or constrains some aspect of the business. A data rule is a statement that provides the opportunity to validate compliance to expected conditions or to identify and review exceptions.

The goal of a data governance initiative is not the proliferation of a multitude of rules. Rather, business rules and data rules should be used to validate that a business process is working correctly, prevent new errors from being introduced, and, in limited cases, automatically correct errors.

- Standards

Standards allow data rules and other controls to be more automated and allow different functions or groups to compare assets. Standards include definitions and terminology as well as processes and conventions. For the purposes of DG Framework standards are drafted, vetted, and approved by the appropriate body with the recognized authority. There are no de facto standards when data governance is being implemented; the standards need to be documented, managed, able to be modified, and communicated in the DG Compendium.

- Data Naming Convention

Data naming conventions as components of the IQM Framework are described as business-focused, consistent as possible, human-readable, machine generated, and informative. Data naming conventions can cause a lot of human churn because many people have angst and passion about how things are named.

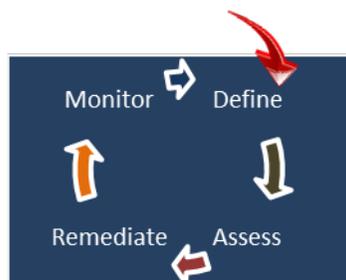
- Data Model

The Data Model as a component of the IQM Framework is described as a business-focused, logical data model to describe data, data attributes, and its appropriate metadata. As part of the DG Framework efforts a data model should be defined that supports the business needs, the business processes, the data rules, and the underlying data elements and attributes that the data rules will be run against.

- Data Taxonomy

Data Taxonomy as a component of the IQM Framework is described as a business-focused, logical, fit-for-purpose hierarchical taxonomy to describe folder structures for properly putting away log files and seismic files for future needs. As part of the DG Framework efforts a taxonomy should be defined that supports the business needs, the business processes, the data rules, and the underlying data elements and attributes that the data rules will be run against.

IQM Framework Processes



The IQM framework processes as defined as part of the overall DG Framework relate to the overall conceptual process of how data quality issues are identified, made clean, and kept clean. That is the goal of the process. Additionally, it is a pragmatic process that can begin as a single data element and grow. The process can be thought of as an upwardly growing spiral.

- Define

The first step in the process is to prioritize and define what is to be governed or quality managed. This prioritization can be made as part of the overall process, run as a separate project, or as part of another initiative. For our DG project, the data types are associated with the POC's selected. The define step includes the pace as well as the priority. The pace can be thought of as how big a bite is to be taken or how many parallel threads of effort are to be undertaken. As suggested by the title, the define step is a discovery step. Artifacts are gathered and analyzed; key repositories for the data type in scope are identified; the appropriate producers and consumers are identified. The geographic scope and/or areas to work are identified. During this step, it is also appropriate to talk with business users, data workers, producers, and consumers to determine their perspective on the “state of the



data”. In the define step the artifacts need to be analyzed to see if they are of sufficient granularity for our DG efforts. If gaps and/or missing key artifacts are discovered then some efforts should be taken during the Define step to shore them up.

- Assess

The Assess step assesses both the data and the artifacts. If there are data quality rules, then during the assess step the rules will be run, warnings and failures will be captured and analyzed, and a more quantitative “state of the data” can be determined. However best determined, an initial assessment of the data quality must be made. This needs to include overall data quality, specific areas for key problems and issues, and an understanding of how the quality of the data is affecting business decisions. The more effort put to establish as baseline then the better the future monitoring will help demonstrate improvement.

In addition to assessing the data and data quality, the standards and artifacts also need to be assessed for completeness, accuracy, business applicability, and capability for automation. In each step, data rules will be documented, improved, and/or edited but in this assess step the business processes need to be complete, vetted, and ready for approval. This will be important when planning remediation. Capturing metadata related to this assessment is also important.

- Remediate

Remediation occurs when the data gets cleaned, and data quality is improved. The assessment shows what needs improvement and modification and remediation performs that function based on priority and capability. During remediation, automation can play a significant role in lowering costs and improving efficiency of remediation. This is a main reason why the “define and assess” steps are so important. Without these steps and establishing certain DQ rules, remediation will be manual. Remediation is best performed by running automated error detection / correction procedures. Ultimately, users will validate the remediated data to ensure it is accurate and suitable for actual use.

- Monitor

The Monitoring of data is an important step in the process. Monitoring data after it has been remediated is the only way to effectively keep it clean and of high quality (or at least known quality). Monitoring data allows for the business users to trust the data and have that trust remain constant and/or grow over time. Monitoring as depicted here includes monitoring of the data and (also) monitoring of the state of the data governance criteria.

Monitoring includes establishing thresholds, ranges, and target KPIs for data quality. Over time, to improve data quality, the thresholds and ranges need to be tightened based on consumer needs and producer abilities. Also, the monitoring is based on the key business



processes continuing to meet the business needs. Monitoring includes historical and current runs of the data quality values, understanding changes that have been made, referring to the metadata related to exceptions and alerts, etc. One key to successful monitoring is good reporting and visualization.



5.0 Building Sustainment Through a DG Organization:

A goal of the DG project is to transform the project into a sustained practice – a functional capability that is part of the core Chevron business. There are certain activities identified to assist in building this sustainment. These include:

- maintaining and improving the DG Academy as a part of the standard data enhancement process;
- establishing a proper organizational capability and building it into a service line;
- maintaining upper management and function level support and sponsorship;
- maintaining flexibility and an outward-focused awareness; looking for opportunities to prioritize efforts to drive towards success;
- documenting, measuring, and reporting on key success factors and objectives frequently, as well as celebrating successes.

An important aspect of the DG project is to understand how to establish a DG organizational capability -- a capability that can begin relatively small and grow organically as the data types under governance grow and as the business benefit of sustained data quality is realized. Properly building the capability into the DG program structure requires that there is a clear understanding of overall process and methodology that are being developed. Organizationally, there are groups that will be directly responsible for certain aspects of data governance in Chevron GOM. These groups need to be clearly identified and named, and their roles and responsibilities clearly defined. DG organizational considerations include:

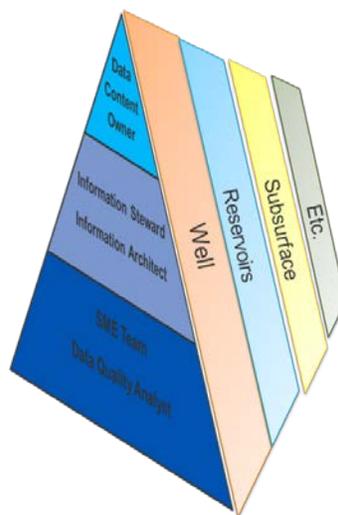
- the need to define and describe the various “Data Governance bodies;”
- the desire to leverage current functional organizations;
- the need to decide what roles are to be within the DG/DQ Framework;
- the need to give each role a label or name that can be agreed on;
- the need to define/describe roles and responsibilities as they relate to the specific data types to govern.

The various data governance bodies will include the following groups and attributes.

- DG standards committee (currently, domain-oriented business leadership teams)
 - A hybrid group from the DG practice, the business, and (possibly) IT
 - Performs portfolio, champion, and standards guidance and approval
 - Approves the DG standards pack
 - Domain oriented
 - Accountable for the data types and information objects within a domain
 - Decisions and work effort facilitated by DG practice
- Data steward
 - Part of business; in the function by discipline
 - Manages a portfolio
 - Performs oversight on standards development

- Monitors data quality metrics and results to agreed standard
- Provides oversight on data remediation and clean-up
- Resolves hard problems; the authoritative source
- Data custodian(s)
 - Part of IT's data quality team
 - Manages the implementation of the data rule base in the DQ engine
 - Manages or facilitates coordination of data remediation and clean-up
- Data analyst
 - Part of IT's data quality team
 - Performs analysis on data issues
 - Performs heavy lifting for data remediation and clean-up
 - Resolves most data problems
- Data governance practice
 - More details below; includes key DG responsibilities, key DG roles
 - DG practice lead
 - DG Compendium librarian
 - DG Academy support

The graphic below shows the hierarchical nature of various DG roles as envisioned by the Chevron GOM DG project team. Additionally, as see in the graphic, these roles will be repeated for the major data domains or functions within Chevron GOM.



As described above, the standards committee fulfills an important role in the approval of the DG standards that come out of the DG Academy efforts and the data project teams. The graphic below shows the structure of this committee.



Setting up the data governance practice is a crucial activity of the DG project.

Building the DG organizational capability begins with establishing the DG practice as a “work team” and the DG practice lead as a role. The DG practice:

- is dedicated to sustaining data quality through proper data governance;
- ensures roles are filled;
- ensures metrics are maintained and communicated;
- ensures alignment and proper leadership support;
- captures and communicates value;
- monitors budget for DG efforts;
- maintains and enhances the DG Academy, DG processes, and DG collateral;
- maintains and manages the DG Compendium;
- maintains and enhances DG and DQ roles and responsibilities;
- guides and/or takes guidance from the sponsors and business leadership for future changes in strategy and direction;
- on-boards new DG resources, business groups, and project teams as DG expands; and
- keeps the vision alive and communicates it effectively.

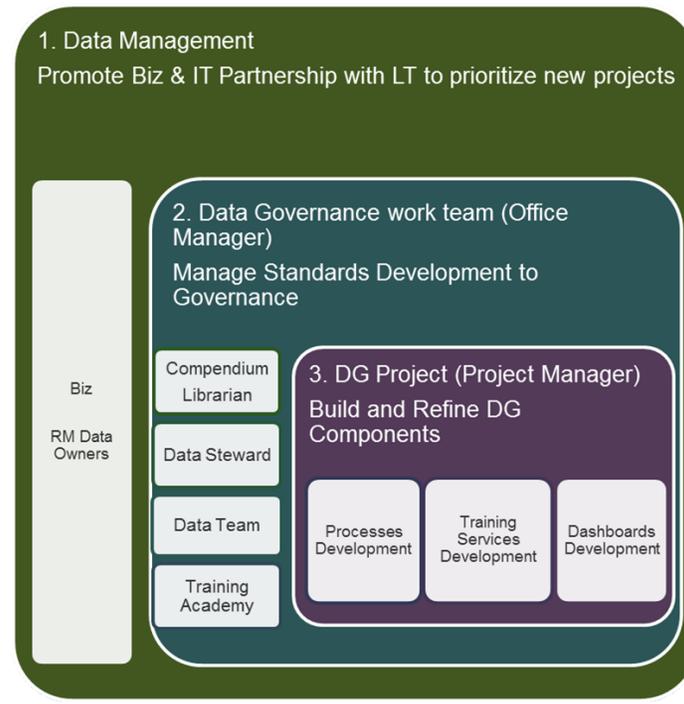
Effective communication includes:

- describing what data governance entails in our vision;
- describing the current thinking related to DG and DQ roles and responsibilities;
- describing the overall DG and DQ organization and how the DG practice fits into it;
- describing the challenges with communicating the concepts related to DG and leveraging the DG Academy to address those challenges;
- being consistent in the naming of and communication of role names – there has been inconsistency and ambiguity in the past; and

- describing the DG practice as it relates organizationally to both the business and IT.

Much of the organizational capability described in the text and graphics above are designed to show the future state of DG in Chevron GOM. During 2013, the DG project and the project team will be very active in continuing to work with data project teams, perform additional POCs and progress through the Chevron project execution process.

Related specifically to organizational capabilities, 2013 will be a ramp up and transition year. The DG project team will continue to work under the Chevron GOM data management (DM) program. The project will continue developing business and IT partnerships to determine how data content owners in the business and DQ support in IT can work together with the DG project team to prioritize, recommend, fund, and implement new projects to bring new data types under data governance.



In 2013, the DG project team will continue to utilize consultants but the DG practice will bring on Chevron GOM staff to ramp up internal capabilities and begin to prepare for DG implementations beyond the end of the DG project. The DG project will support the DG practice and continue to develop components that will be consumed by the DG practice. As new data types come under data governance, whether initiated by the DG project or the DG practice, the results will be handed off to the DQ support team to integrate into the system.

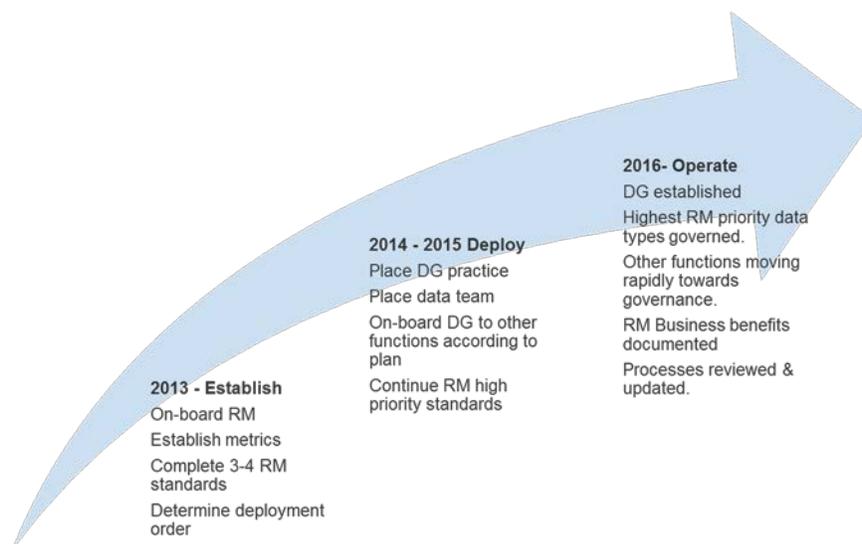
6.0 Beyond Data Governance as a Project:



For Chevron GOM, Operational Excellence (OE) drives everything the employees do – <http://www.chevron.com/about/operationalexcellence/>. There are certain foundational capabilities that are used to help achieve OE such as the standard method used to execute projects, the standard method used to map and change business process, the standard method used to set and measure performance, etc. As the data governance methodology and capability matures and grows beyond a project and into an operational practice, the DG project team sees data governance becoming one of those foundational capabilities: a standardized method used to define, assess, remediate, and monitor data quality into a sustained, high quality state. Data is an asset and must be managed accordingly. The DG project team sees the proper management, stewardship, and governance of these data assets in Chevron GOM as a differentiator; a competitive advantage. This is the path laid out.

Some of the strategies that will help to pave this path include:

- engage full time data stewards by function;
- create and grow a data governance practice;
- remediate data only when necessary and consistent with business priorities;
- minimize change.



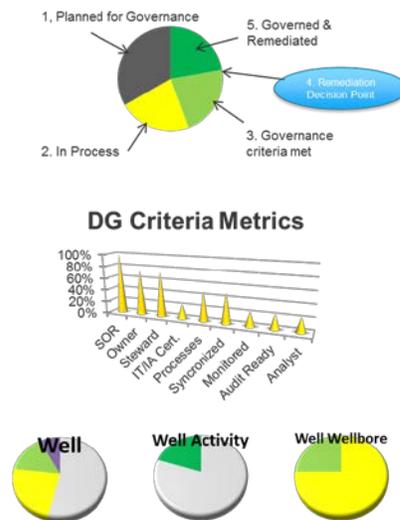
As stated in the previous section, building a proper organizational data management capability with Chevron GOM staff is on the critical path. A viable service or capability cannot be operated effectively over the long term using only consultants. Since data is an asset and the decisions made based on governed data are central to the health and future of Chevron GOM, then data governance must be considered a core competency of the business.

The shaping curve in the graphic above shows, at a very high level, some of the milestones envisioned for the DG project as it moves into an operational mode.



- 2013 – Establish – A year of continued learning, continued development of various facets of the methodology, and continued work towards getting some specific data types under governance. Additionally this year, a focus will be on establishing the ground work for the organizational capability. Working with the business and senior management who are part of the organizational capability, the prioritization of data governance deployment will be ordered – likely the top 10 data types will provide the highest value.
 - Foundational work to establish a sustainable DG environment will be performed in 2013 and includes:
 - an on-boarding process for various functions;
 - a data-type prioritization model;
 - a value/benefit model and investment model;
 - a scalability model;
 - alignment with other Chevron initiatives and corporate groups;
 - alignment and collaboration with other Chevron BU's.
 - Train IT projects so that any and all data-intensive projects, application upgrades or change-outs, and/or system migrations that involve data migrations will follow data management best practices and DG guidelines
- 2014 – 2015 – Deploy – Data governance will be deployed in Chevron GOM. The project will be completed early 2014 and the DG practice will be in place and operating. Two functions will have been on-boarded. The DG practice will grow and be matured and a core team will be established to assist specific data teams on implementing the DG methodology. Data governance will be expanded based on functional need and business priority.
- 2016+ – Operate – The DG components, methodology, and technology will be mature. Data being governed will be expanded in several dimensions. All of the DG components will be reviewed for scalability, maturity, quality, etc. and updated as appropriate.

Monitoring and reporting on the quality of specific data being governed will greatly increase the confidence that data consumers have in the data. Additionally, the ability to monitor and report on what data is being governed in a more holistic sense will provide the business with one aspect of the success of the DG project. It is envisioned that a series of charts, reports, and graphs will be available via a dashboard at varying levels of granularity to show DQ metrics, DG metrics, and the state of overall efforts



The dashboard will show the current state and plans for data that are currently being governed and that data which is “next up” to be governed. The DG team and DQ Support team will work with the business to understand requirements and design appropriate dashboards, reports, and presentations for monitoring data quality and the governance criteria status. They will work with IT to create processes and technology to generate the data and metadata needed for the dashboards. The data project teams will test the dashboards’ ability to accurately represent the progress of their projects and to visually represent the state of the data being governed. Finally, working with the leadership teams, the DG project team will present reports and graphs via the dashboard that will represent measured business value.

Project pace, data priority, and DG spend must be balanced with business needs. A refresh rate will be established that is in line with business needs and timing. These efforts from the DG project through the DG transition and into the DG operated state need to be somewhat nimble, but not haphazard. The refresh or re-prioritization should not be so frequent that efforts cannot get to a final “governed state,” but neither should effort be expended to govern data that is not a priority. Whatever is created as the ongoing DG practice needs to be consistent with the business imperative, IT planning, and Chevron Operational Excellence.



7.0 Summary:

Data governance is a difficult concept to master but sustaining data quality requires good data governance. The data asset in a major oil company like Chevron is huge. The investment to improve this asset should be very large as well. In order to sustain this investment, the data must be treated as the asset it is. That implies there must be recognition of the asset value, an organizational and governance structure to support it, and funding to maintain it.

Chevron GOM embarked on a journey to create and implement a program for attaining and sustaining known quality data through a data governance development process. The process and methodology developed is a fresh approach to managing the quality of information and this paper has described the thoughts and actions behind the process, the methodology, and the implementation. The methodology has been broken into six stages (Harvest, Draft, Refine, Approve, Apply, and Monitor) that correspond to four of Chevron GOM's execution process phases.

As described, this approach is pragmatic and generic and has been designed to present a "light touch" to most of the people directly impacted. It includes a repeatable series of clearly defined activities and tasks to achieve desired goals and consistent results. Achieving a state of data being governed is accomplished with a balance of processes, people, and technology.

The paper has described the methodology with significant detail, but it can be summarized as:

- a data governance and stewardship model that is taught through a DG Academy and is based on placing specific data types under governance, meeting the criteria of what it means to be governed;
- a framework for data quality controls that is based on key business processes that are impacted by data with associated standards built around the business process and supporting the business process through rules for both business and data. The data rules are imposed on the data through a data quality engine;
- a data governance standards repository called the DG Compendium stores and manages the standards and a variety of metadata from the results is monitored and reported;
- an organizational capability to set priority and pace, govern the standards, and mature the data governance practice.

The DG project has so far focused on four data types through POCs and can be thought of as an agent of change. The project:

- realigns the concept of data ownership;
- removes data and the discussions centered on data issues as an impediment to change and innovation;



- breaks down the data fortresses between different groups in the business by clearly establishing relationships such as data producer, data consumer, data content owner, data steward, and data custodian; and
- creates a reality for the sponsors, senior business management, IT, and data consumer so they can better understand, appreciate, and value data governance and the important part it plays in their decisions.

The other critical aspect of the DG project is establishing the DG organizational structure and capability to effectively create, approve, manage, and administer the DG standards. The transition from a DG project to a DG operational capability will transform over the next 18 to 24 months. The establishment of a DG practice and the role of DG practice lead is the first step.

Through the DG project, it has been seen how data governance and data stewardship contribute to information management (IM). Data governance specifies accountability and responsibility for IM. Data stewardship enables IM to be more effective across the company.

Although the DG project is not over yet, there are clear indications of key success factors and lessons learned:

- Key success factors:
 - The sponsors have requested that the project be accelerated so it can go into operational mode more quickly. That is seen as a clear validation of the project's success to date with the DG Academy, the DG Compendium, and the communication with the project sponsors and the business.
 - Data project groups are coming to the DG project team requesting assistance in establishing a DG capability as part of their project
 - The DG project and the DQ support team (data custodians) are getting traction in moving data rules from the DG Compendium into the DQ engine and making them part of the DQ tool set.
- Lessons learned
 - Be focused on why you are implementing standards.
 - Use a phased approach.
 - Leverage standards or de facto standards.
 - Determine who is best positioned to discover/document standards to improve data quality; use a team approach.
 - Involve SMEs to ensure accuracy.
 - Standards are key products (measurable results) of data stewards and DG teams.
 - Remediation (clean up) is difficult. Ensure it is focused to bring business value.
 - Find ways to measure business value.



8.0 About the authors:

Dave Blosser – Chevron North America Exploration and Production - Data Management, Program Manager

Dave has been in his current position for three years. His purview for this project was, “We are having a real problem with our data. Go fix it.” Prior to this position, Dave was program manager for GOM’s Upstream Foundation Program. During that time several leading edge applications were developed resulting in over \$200 million of realized value and leading to the team’s selection for Chevron’s prestigious Chairmanship Award.” Dave has led or been a part of many IT development efforts over the other 18 years of his career with Chevron.

Prior to joining Chevron in 1988, Dave was a working petroleum engineer for a small independent where he gained experience in all aspects of the oil industry from exploration to abandonment. Too often exploration and abandonment happened in the same well, which led him to a new career path with IT.

Paul Haines – Noah Consulting – Partner and Upstream and IM Subject Matter Expert (SME)

As a subject matter expert, Paul has been consulting and advising about information management for E&P companies, both large and small, for the past seven years. Before his entry into consulting, he worked for Kerr-McGee Oil & Gas for five years. There, he was manager of Data Management and Integration. This role included responsibilities over all E&P technical subsurface and operations data.

Prior to mid-2001, Paul spent 23 years with Schlumberger, beginning as a wireline field engineer. During these years, he gained a broad understanding of the industry and the data associated with the E&P industry. He held a variety of positions in areas of operations, technique, sales, marketing, and management.

Paul served on the board of directors of the Professional Petroleum Data Management Association (PPDM) from 2007 to 2012. In 2011, Paul was honored with the **Cornerstone Award** for his contributions to the field of E&P Data Management at this conference. Paul has been a presenter and contributor at PNEC conferences frequently over the years.



9.0 About the companies:

Chevron Corporation

Chevron is one of the world's leading integrated energy companies and conducts business worldwide. Our success is driven by our people and their commitment to get results the right way—by operating responsibly, executing with excellence, applying innovative technologies and capturing new opportunities for profitable growth. We are involved in virtually every facet of the energy industry. We explore for, produce and transport crude oil and natural gas; refine, market and distribute transportation fuels and lubricants; manufacture and sell petrochemical products; generate power and produce geothermal energy; provide energy efficiency solutions; and develop the energy resources of the future, including research for advanced biofuels.

Chevron's business momentum is driven by three enterprise strategies – people, execution and growth. We invest in people to maintain world-class workforce. We follow systems and processes that allow us to execute with excellence. And we use our competitive advantages to grow our company profitably.

Based in Covington, LA, Chevron's Gulf of Mexico business unit has approximately 1,800 employees and approximately 2,000 contractors daily. Chevron is one of the largest producers and second largest leaseholder on the Gulf of Mexico shelf; and one of the top leaseholders in the deepwater Gulf of Mexico.

Noah Consulting

Noah Consulting makes a significant difference in the lives of Business and Data Management professionals in the companies they serve through energy industry specific information management services that transform data into valuable assets. We believe your data should work for you – 24 hours, 7 days a week. Re-creating or re-purchasing data can be costly, error prone and time consuming. As such, excellent data management, optimal extraction and aggregation of data, and the proper analysis of that data are investments in a company's future. Noah Consulting strives to improve the wellness of the industry by helping clients navigate their information. This mission is clear by our alignment with Best-in-Class organizations concerned with industry standards, such as PPDM.

Founded in 2008 by Information Services professionals, Noah Consulting is exclusively focused on the data and information domain within the energy sector. By creating a home for focused experts with a common passion, Noah Consulting has put together a team of professionals who average over 18 years of experience and who have both deep industry knowledge and technical expertise in Strategy and Planning, Data Governance, Program Management, Architecture and Business Intelligence.