Agenda

• Introduction
• About Devon
• OKC G&G Data Management Program
• Data Governance
  – Why do we need Data Governance?
  – What are the basic components of Data Governance?
• Business Rules
  – Why do we need business rules?
  – What are the basic components of Business Rules?
• How Data Governance and Business Rules work together
Devon History

- Founded as a private company in 1971
- Became a public company in 1988
- Currently listed on the New York Stock Exchange under the ticker symbol DVN
- Has grown from 185 employees in 1981 to more than 5,000 employees today
- Established a portfolio to provide stable production and a solid platform for future growth
Devon Today

Proved reserves: ≈3.0 billion BOE
(42% liquids)

Q2 2012 production: 679 MBOED

Production mix: 22% oil
15% NGLs
63% natural gas

Significant midstream business
2012e operating profit: ≈$385 million

Enterprise value: ≈$25 billion
Larger than you might think...
Enterprise value

In billions of U.S. dollars

Source: Enterprise value as stated on Yahoo! Finance on Feb. 1, 2012.
New Building

- Started Construction Oct 2009
- 50 Stories
- 850 feet high
- 1.8 million square ft
What is the OKC G&G Data Management Program?

The **OKC G&G Data Management Program (OKC DMP)** was established to enable agile business operations by treating data as a corporate asset and providing timely, qualified, integrated and harvested data from the rig to the Geoscientist desktop. This Program, led by the Business, is a fully funded multi-year initiative broken into three waves consisting of more than 40 roadmap initiatives.

This program consists of three projects.

- **Governance**: Establish an organization that can provide policies and procedures to ensure sustainable data management practices
- **Master Data Management (MDM)**: Establish a solution to deliver trusted and reliable information to core G&G business processes and applications
- **Data Quality Management (DQM)**: Improve the management of data quality between business applications to ensure that it is timely, consistent and accurate
<table>
<thead>
<tr>
<th>Business Rules</th>
<th>Governance</th>
<th>Governance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>As-Is State</strong></td>
<td><strong>Why change is needed?</strong></td>
<td><strong>To-Be/ Benefit</strong></td>
</tr>
<tr>
<td><strong>Data Quality Uncertainty</strong></td>
<td><strong>Lack of Data Quality / Order</strong></td>
<td><strong>Can find and use critical data.</strong></td>
</tr>
<tr>
<td>- Timeliness</td>
<td>- Lose Data / repurchase</td>
<td>- Understand completeness of data</td>
</tr>
<tr>
<td>- Accuracy</td>
<td>- Can’t find data for look backs</td>
<td></td>
</tr>
<tr>
<td>- Relevance</td>
<td>- Decisions from questionable data</td>
<td></td>
</tr>
<tr>
<td>- Completeness</td>
<td>- Multiple iterations of validation, rework and review</td>
<td></td>
</tr>
<tr>
<td><strong>Difficult to transfer projects between asset teams or divisions</strong></td>
<td><strong>Flexibility in reassigning staff and/or assets</strong></td>
<td><strong>Portability data across asset areas /teams.</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Leverage resources/learnings across asset teams.</strong></td>
<td><strong>Common set of business processes</strong></td>
</tr>
<tr>
<td><strong>No single authoritative source for G&amp;G data.</strong></td>
<td><strong>Current tools do not support data management.</strong></td>
<td><strong>Simplify integration across applications.</strong></td>
</tr>
<tr>
<td>(e.g. Studies Database/ Barney)</td>
<td>- Too much time spent searching for quality data</td>
<td><strong>Reduce G&amp;G data redundancy</strong></td>
</tr>
<tr>
<td><strong>No cohesive Data Management capability</strong></td>
<td>- Data is stored in unmanaged repositories</td>
<td><strong>Improved data sharing automation</strong></td>
</tr>
<tr>
<td></td>
<td>- Excel / access</td>
<td><strong>Ease to gather data for analysis / lookbacks</strong></td>
</tr>
<tr>
<td></td>
<td>- E-mail</td>
<td><strong>Gold standard sets of data</strong></td>
</tr>
<tr>
<td></td>
<td>- Often easier to repurchase rather than to find it</td>
<td></td>
</tr>
<tr>
<td><strong>Non-standard reference data</strong></td>
<td><strong>Inability to share data across G&amp;G applications.</strong></td>
<td><strong>Reduce the redundancy of G&amp;G data</strong></td>
</tr>
<tr>
<td><strong>Inconsistent naming standards</strong></td>
<td><strong>Duplicated information stored across applications.</strong></td>
<td><strong>Simplify integration across applications</strong></td>
</tr>
</tbody>
</table>
Benefits

• **Quality**: Improve the trust, consistency, relevance, and timeliness of data that is used on a daily basis to support business decisions

• **Process**: Establish business rules and standards to improve data flow and manage the data lifecycle

• **Standards**: Improve data consistency and sharing across applications and data stores

• **Availability**: Establish an integrated system that delivers reliable information and single source of truth to support master data management

• **Governance**: Institute better overall data management process, procedures, and policies to adhere to business and data quality rules

Overall make it easier for the business to manage data on their own
The Vision

Enabling agile business operations by treating data as a corporate asset and providing timely, qualified, integrated and harvested data from the rig to the Geoscientist desktop.
# Roadmap

<table>
<thead>
<tr>
<th></th>
<th>2011</th>
<th>2012</th>
<th>2013+ ongoing Term</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Structured Data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well Master</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interpretation PM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Quality Management</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standards</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Application Development</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Work Flow</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business Rules</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Unstructured Data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Naming Conventions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>File Automation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Work Flow</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taxonomy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cross Functional</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business Process</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Governance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Application Transition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daily Rig Data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G&amp;G BIDW</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Legend:
- Wave 1
- Wave 2
- Wave 3
- Enterprise Initiatives
### Waves

<table>
<thead>
<tr>
<th>Wave 1</th>
<th>Wave 2</th>
<th>Wave 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Future</strong></td>
<td>• Where are we going? What do we need to accomplish?</td>
<td>• New Ideas, New Technology, New Process</td>
</tr>
<tr>
<td><strong>Foundation</strong></td>
<td>• How do we build a foundation that we can grow on over the next 3+ years?</td>
<td>• Implement designs from Wave 1, new technology &amp; processes.</td>
</tr>
<tr>
<td><strong>Focus</strong></td>
<td>• What does the end state look like?</td>
<td>• New Data Types, Better Data Quality Management &amp; Reporting, Better Tools with more capability</td>
</tr>
<tr>
<td><strong>Innovation</strong></td>
<td>• Fixing anything from Wave 1 or 2 that wasn’t implemented as well as it should’ve been.</td>
<td>• Completing the roadmap</td>
</tr>
<tr>
<td><strong>Implementation</strong></td>
<td>• Correcting anything from Wave 1 or 2 that wasn’t implemented as well as it should’ve been.</td>
<td>• Fixing</td>
</tr>
<tr>
<td><strong>Increased Functionality</strong></td>
<td></td>
<td>• Finishing Well</td>
</tr>
</tbody>
</table>

**Wave 1**
- Wave 1 is about fixing anything from Wave 1 or 2 that wasn’t implemented as well as it should’ve been.

**Wave 2**
- Wave 2 focuses on completing the roadmap.

**Wave 3**
- Wave 3 emphasizes new ideas, new technology, and new processes.
Why do we need Data Governance?

FACEPALM

Because expressing in words how badly the data is messed up just isn’t possible.
Data Governance

Common issues faced by Oil & Gas companies

- Massive amounts of data
- No standard way to store and access data
- Every asset or site uses different standards
- Can’t agree on what data is important
- No way to establish authoritative data standards
- No authoritative data sources
- Data producers do not take data consumer needs into account
- No clear way for employees to work together to fix data issues
Data Governance

What is data governance?

- **Governance**: (1) the act or process of governing; (2a) the office, power or function of governing; (2b) controlling or directing influence: AUTHORITY. -- Webster's

- **Information (Data) Governance**: the act or process of leading, directing, controlling, and assuring that information is effectively managed as an enterprise resource, including resolving information conflicts, across the enterprise. -- Larry P. English
• **Stewardship**...is the willingness to be accountable for the well-being of the larger organization by operating in service, rather than in control, of those around us. Stated simply, it is accountability without control or compliance. -- Peter Block in Stewardship: Choosing Service Over Self Interest, 1993.

• **Information Stewardship** is the willingness to be accountable for a set of business information for the well-being of the larger organization by operating in service, rather than in control, of those around us. -- Larry P. English
Data Governance

Common data governance roles

- **Data Steward:** has overall responsibility for ensuring data is managed in a way that serves the full value chain, supports all data roles across the division

- **Data Administrator:** carries out day-to-day coordination, monitoring, remediation activities, and assists in developing division and project level data policy

- **Data Custodian:** In a manner that serves the full value chain, Data Custodians prepare data for use by others

- **Data Consumer:** consume data in order to carry out operational or decision making activities

- **Data Producer:** Data Producers create, update, and occasionally delete data, typically as a part of carrying out their primary business function

- **Subject Matter Expert:** Employees that may be consulted with specialized knowledge concerning key data types
Data Governance

Roles in real life

IT

• Data Stewards?
• Data Custodians
• Data Administrators
• Subject Matter Experts
• Data Analysts
• Data Modelers
• Data Architects

Tries to clean up data
Makes it worse.

Vs.

Business

• Data Owner
• Data Producer
• Data Consumer
• Subject Matter Experts
• Data Stewards?
• Data Administrators
• Data Custodians
• Data Analysts

HAS A LOT OF DATA TO CLEANUP

SCREW IT, I'M TOO BUSY

Where does data stewardship belong? In the business or in IT?
Data Governance Components

Enterprise Drivers:
- Enterprise Priorities: Accountable, Reliable, Measurable, Transparent, Scalable, Sharing
- G&G Business Drivers: Efficiency, Data Quality, Master Data Management, Data Management Program

G&G Business Drivers:
- Data Governance Charter
- Governing Principles
- Governing Bodies
- Decision Authority

OKC DMP Focus

Policies & Procedures

Data Governance Execution:
- People: Geoscience Data Governance Council, Stakeholders, Stewards, Meeting Schedule
- Process: Metrics, Measurements, Standards, Procedures, Processes
- Mechanisms: Quality Monitoring, Data Profiling, Tools

Information Management:
- Data Architecture
- Master Data Management
- Metadata Management
- Data Quality Management
- Data Administration
- Data Security & Retention

Enterprise Governance Oversight
Data Governance

- Address 3 key areas
  - People
  - Process
  - Technology
- Clearly define roles & responsibilities
- Decision authority
- Maintain accountability

Components of a good data governance model

- Develop standards
- Enable the entire company to work together in an effective manner
- Flexible to organizational changes
- Scalable (should be able to shrink as well as grow)
- Champion data quality initiatives
- Addresses organizational data management needs for operational teams, projects and governance structure
OKC G&G DMP Organization

Program Leadership Team

Executive Team

Steering Team

Stakeholder Group

Project Team 1

Project Team 2

Project Team N

Support
(DBA, Infrastructure, IT Support Organization)
Geoscience Data Governance Council - Impact

Geoscience Data Governance Council will:

• Single point of contact
• Improve data quality
• Manage data as enterprise asset
• Improve business processes
• Framework for data management
• Strategic initiatives
• Address elephants in the room
• Share best practices
• Develop data standards
• Make data recommendations
• Select applications
• Responsible for data initiatives
• Oversee projects

Geoscience Data Governance Council will not:

• Slow the business down
• Take ownership of data
• Implement what the EAC is not willing to enforce
Members

**Chair:**
- Business leader operates the GDGC

**Facilitator:**
- Understands data governance

**Members:**
- Core team: business members
- Advisory team: technical people with systems expertise
- Enough authority to make enforceable decisions
- Other disciplines participate as required

**Members:**
- Stewards from business, experts in data subject area
- IT members contribute knowledge of systems

**Members:**
- Managers, Geoscientists & Geotechs from the Asset Teams
- Define needs, requirements and provide feedback on effectiveness

---

EAC Representative:
- EAC member with decision authority
- Represents GDGC within EAC and Strategic Services

---

Governance organization
Centralized v. Decentralized

Business should be decentralized for nimbleness

Data Governance should be centralized for control and effectiveness

Centralize or decentralize the Data Administrators?

Operational team organization
Option 1: Decentralized Data Administrators

**Strengths**
- Uses existing staffs time
- Speed of change
- More customization allowed

**Weaknesses**
- Business resources focused on data
- Conflicting standards
- Inflexible resource allocation

Option 2: Centralized Data Administrators

**Strengths**
- Resources freed to focus on business
- Scalable standards and processes
- Flexible resource allocation

**Weaknesses**
- Dedicated resources required
- Asset Teams have less control
- Less customization allowed

**Recommendation**
## MDM - Decision Considerations

<table>
<thead>
<tr>
<th>Decision Factors</th>
<th>Strategy Approach</th>
<th>Enterprise Data Management Solution (Top down)</th>
<th>Business Unit Data Management Solution (Bottom Up)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adoption</td>
<td>Typically driven by IT which may make business feel the solution is not truly geared toward their issues. Gaining consensus across divisions can be challenging.</td>
<td>Driven by the business tends to lead to better adoption during the early stages of the tactical and strategic solution.</td>
<td></td>
</tr>
<tr>
<td>Quick wins/early benefits</td>
<td>Deployment strategy focuses mostly on broad cross-sectional areas of the business with the expectations of achieving large enterprise benefits</td>
<td>Has the opportunity to address both local pain points for quick wins, while also building the groundwork for broader, more significant benefits.</td>
<td></td>
</tr>
<tr>
<td>Probability of success</td>
<td>Risk of not succeeding in achieving the business case can be high due to lack of adoption by the business, and addressing the business issues and pain points.</td>
<td>Risk can be lower since the solution can be tactical and seen by the business as something they want to support.</td>
<td></td>
</tr>
<tr>
<td>Consistency across the organization</td>
<td>One standard solution used by all divisions creates more consistency across the enterprise.</td>
<td>Greater possibility of each division wanting to tailor the solution to their existing ways of business process, thus missing the benefits of standardization</td>
<td></td>
</tr>
</tbody>
</table>
We need business rules because people do not know how to manage the data properly.
What are the basic components of business rules?

- Business rules are difficult to create
- Tied to a business term
- Application and database agnostic
- Only implemented for business terms worth maintaining
- Produce a measurable result
- Produce an actionable result
- Easy to understand
DQM - Reference Architecture

Data Profiling & Quality Monitoring
Short Term - Conceptual Model

Reporting & Delivery Layer

Quality Monitoring & Metrics Data Collection Layer

Attribute Index → Data Flux → QC Pro

Data Layer

IHS, WellView, Finder, PETRA
• Requirements are translated into Business Rules.
• Would be nice to have a Business Rule Manager similar to Business Terms Manager.
**Business Rules**

- Finder stores locations as Latitude and Longitude

- Issue where an integration was storing lat/lon as ft with 2 decimal places

- $0.01^\circ \sim 3,641\text{ft}$

- $0.00001^\circ \sim 3.6\text{ft}$

---

### Attribute Business Rules: Location Decimal Precision

<table>
<thead>
<tr>
<th>Rule ID</th>
<th>BR00120</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rule Name</td>
<td>Location Decimal Precision</td>
</tr>
<tr>
<td>Description</td>
<td>Location values should always have more than 5 decimal places. In Finder for each well's NODE_ID the NODE_X or NODE_Y value must contain more than 5 decimal places.</td>
</tr>
<tr>
<td>Notification</td>
<td>Immediate</td>
</tr>
<tr>
<td>Weight</td>
<td>4</td>
</tr>
<tr>
<td>Quality Dimension</td>
<td>Accuracy</td>
</tr>
<tr>
<td>Business Process</td>
<td>Identify Prospective Areas - Mapping; Plan A Well - Proposed Well; Complete A Well - Validate final data; Perform Data Management - Load Data From External Sources</td>
</tr>
<tr>
<td>Integration</td>
<td>Finder-to-Petra; Petra-to-Finder; Finder-to-SDE</td>
</tr>
<tr>
<td>Application(s)</td>
<td>Finder</td>
</tr>
<tr>
<td>Data Type</td>
<td>Well Header; Well Borehole</td>
</tr>
<tr>
<td>Reason</td>
<td></td>
</tr>
<tr>
<td>Applications Applied</td>
<td>DataFlux</td>
</tr>
</tbody>
</table>
Implementation checks to see if the lat or lon has 5 or more digits after the decimal.
**Attribute Business Rules**: For wells with surveys, the well’s bottom hole should be exactly the same as the last point on the survey

<table>
<thead>
<tr>
<th>Rule ID</th>
<th>BR00144</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rule Name</td>
<td>For wells with surveys, the well's bottom hole should be exactly the same as the last point on the survey</td>
</tr>
<tr>
<td>Description</td>
<td>The bottom hole location must match the last point on the survey exactly.</td>
</tr>
<tr>
<td>Notification</td>
<td>Request</td>
</tr>
<tr>
<td>Weight</td>
<td>2</td>
</tr>
<tr>
<td>Quality Dimension</td>
<td>Validity</td>
</tr>
<tr>
<td>Business Process</td>
<td>Complete A Well - Validate final data</td>
</tr>
<tr>
<td>Integration</td>
<td>Finder-to-Petra; Finder-to-SDE</td>
</tr>
<tr>
<td>Application(s)</td>
<td>Finder; Petra</td>
</tr>
<tr>
<td>Data Type</td>
<td>Well Header; Well Borehole; Survey-Directional</td>
</tr>
<tr>
<td>Reason</td>
<td></td>
</tr>
<tr>
<td>Applications Applied</td>
<td>DataFlux</td>
</tr>
</tbody>
</table>
Location > Accuracy > BASE_NODE_X and Y must equal Survey End Point (BR00144) (NODES Fields Validation)

Score

99.67%

Weight: 1.00
Trigger percent: 0.33%
Trigger count: 00557
Rows processed: 81329729

Details
Rule name: BASE_NODE_X and Y must equal Survey End Point (BR00144)
Rule label:

Task name: NODES Fields Validation
Task label:
Code: 502A22418120B18

There are critical BPls concerning Location Terms and need to be run frequently

Trend
### Monitor

**Triggered Rules:**

<table>
<thead>
<tr>
<th>Role</th>
<th>Date</th>
<th>% Triggers</th>
<th># Triggers</th>
<th>Rows Processed</th>
<th>Importance</th>
<th>Status</th>
<th>Assigned User</th>
</tr>
</thead>
<tbody>
<tr>
<td>BASE_NODE_X and Y must equal Survey End Point (BR00144)</td>
<td>06/07/2012 08:29:02 AM</td>
<td>0</td>
<td>0</td>
<td>20863</td>
<td>564795</td>
<td>UNRESOLVED</td>
<td></td>
</tr>
<tr>
<td>BASE_NODE_X and Y must equal Survey End Point (BR00144)</td>
<td>06/07/2012 10:05:17 AM</td>
<td>0</td>
<td>0</td>
<td>20863</td>
<td>664236</td>
<td>UNRESOLVED</td>
<td></td>
</tr>
<tr>
<td>BASE_NODE_X and Y must equal Survey End Point (BR00144)</td>
<td>06/07/2012 20:54:03 AM</td>
<td>0</td>
<td>0</td>
<td>19206</td>
<td>502390</td>
<td>UNRESOLVED</td>
<td></td>
</tr>
<tr>
<td>BASE_NODE_X and Y must equal Survey End Point (BR00144)</td>
<td>06/07/2012 07:49:26 AM</td>
<td>0</td>
<td>0</td>
<td>2625</td>
<td>6388</td>
<td>UNRESOLVED</td>
<td></td>
</tr>
</tbody>
</table>

**Summary**

- **Rows per page:** 100
- **Summary:**
  - BASE_NODE_X and Y must equal Survey End Point (BR00144)
  - BASE_NODE_X and Y must equal Survey End Point (BR00144)
  - BASE_NODE_X and Y must equal Survey End Point (BR00144)
  - BASE_NODE_X and Y must equal Survey End Point (BR00144)
  - BASE_NODE_X and Y must equal Survey End Point (BR00144)
  - BASE_NODE_X and Y must equal Survey End Point (BR00144)
  - BASE_NODE_X and Y must equal Survey End Point (BR00144)
  - BASE_NODE_X and Y must equal Survey End Point (BR00144)
  - BASE_NODE_X and Y must equal Survey End Point (BR00144)
  - BASE_NODE_X and Y must equal Survey End Point (BR00144)
WHAT IF I TOLD YOU

DATA MANAGEMENT IS JUST
GLORIFIED JANITORIAL WORK
Questions?