

## Eating The Data Elephant

Lessons from the trenches,  
or how to spend a million bucks  
without feeling like you got  
nothing for it!

By  
Jim Lawnin, Wipro Technologies,  
John Yurkanin, Petris Technology, Inc.

## EXECUTIVE SUMMARY

According to a children's riddle, the way to eat an elephant is one bite at a time. The same is true of managing upstream data. However, many enterprises go for the whole animal when dealing with data management. This approach has not proven effective, and as a result we are having the same conversations about data management that we have been having for the past ten (or more) years.

The list of issues is familiar to any manager in any organization and includes symptoms like these:

- Business users aren't engaged.
- All users are complaining.
- Business users don't trust the data.
- Excel and Access databases are growing out of control.
- Projects take too long.
- Metrics are poorly measured.
- The business feels too much money is being spent on DM projects.

Each symptom has an attendant list of causes which in turn relate to leading practices that will effectively address them. These leading practices don't apply to every situation; rather, each addresses specific data management symptoms.

# EXECUTIVE SUMMARY

## Leading Practices In Upstream Data Management

### Implement This Leading Practice...

### When..

“Bite-size” your data management strategy into smaller projects that build on one another.

- Business users aren’t engaged.
- All users are complaining.
- Projects take too long.
- The business feels too much money is being spent on DM projects.

Get the business involved and taking data ownership. Stop if the business thinks this is an IT issue.

- Business users aren’t engaged.
- All users are complaining.
- Business users don’t trust the data.
- Excel and Access databases are growing out of control.
- Metrics are poorly measured.
- The business feels too much money is being spent on DM projects.

Start with production data or other data that applies across the business and achieve quick wins.

- Business users aren’t engaged.
- All users are complaining.
- Business users don’t trust the data.
- Metrics are poorly measured.
- The business feels too much money is being spent on DM projects.

Develop a high level architecture.

- Business users don’t trust the data.
- Excel and Access databases are growing out of control.
- Projects take too long.
- The business feels too much money is being spent on DM projects.

Standardize on common applications and integrate workflows

- Business users don’t trust the data.
- Excel and Access databases are growing out of control.

Define master data stores.

- Business users don’t trust the data.
- Excel and Access databases are growing out of control.

Include structure and unstructured data to capture users’ attention.

- Business users aren’t engaged.
- All users are complaining.
- Business users don’t trust the data.
- Excel and Access databases are growing out of control.
- Projects take too long.
- Metrics are poorly measured.
- The business feels too much money is being spent on DM projects.

# TABLE OF CONTENTS

Introduction.....	5
But Different.....	5
Eating the Data Elephant.....	5
The Cycle of Doom.....	6
Industry Issues.....	6
Common Symptoms and Their Causes.....	6
Barriers To Success.....	9
Data Management: Too Much, Too Few, Too Long.....	9
Business Management: Reliability, Time, and Process.....	9
Leading Practices.....	10
Conclusion.....	11
About the Authors.....	11

## Introduction

“It’s like deja-vu, all over again...” -Yogi Berra

### ...But Different

Chances are that this is the latest of a long string of papers on upstream data management that you have read. You have also probably sat through numerous presentations at various conferences over the years where subject matter experts have talked about issues around E&P data management and what needs to be done to correct them.

You may be thinking, “If I have read all of this before, what am I going to gain from reading this again?”

It is not our intent to simply repeat what you already know. Upstream data management issues are not about content; rather, they are about the humans dealing with the data. They are about demonstrating real value and reward all along the way rather than telling everyone that if they stick it out it will be worth all the pain. We offer insight into solutions that target the people dealing with upstream data and are the result of field experience helping E&P clients improve the results of their data management initiatives.

Once the papers are put aside and the conferences are over, E&P data managers must return to the day to day challenges of improving data quality, access and ROI. Through our client engagements, we understand the symptoms and root causes of common upstream data management problems, and we have identified leading practices that pragmatically address specific quagmires in which clients find themselves.

Here is what you will get from this paper: Common sense, applied-in-the-field insight into the range of data management problems that upstream businesses face. We will stay away from high level summaries or abstract theories. We will travel much closer to the ground and offer recommendations that you can put into action appropriately and effectively.

## Eating the Data Elephant

- 60% of CEOs feel that utilization of IT investments (including data management solutions) is inadequate.
- Upstream professionals spend almost 70% of their time on the job search for the best available information from the data they have access to.

Source: David Shipman, IBM (presentation at 2008 Digital E&P Conference)

According to a children’s riddle, the way to eat an elephant is one bite at a time. The same is true of managing upstream data. However, many enterprises go for the whole animal when dealing with data management. This approach has not proven effective, and as a result we are having the same conversations about data management that we have been having for the past ten (or more) years.

Many of the leading practices we have identified over the course of our client engagements focus on taking one bite of the data elephant at a time as well as what and how big each bite should be. It is not a “one size fits all” solution. The best route to effective upstream data management depends on the organization and the issues it is facing.

In this white paper, we consider common industry “symptoms” first, then offer a range of underlying causes to help you identify symptoms/causes that are showing up in your own organization. We then offer leading practices that target one or more symptoms along with recommended actions to solve your biggest data management challenges.

## The Cycle of Doom

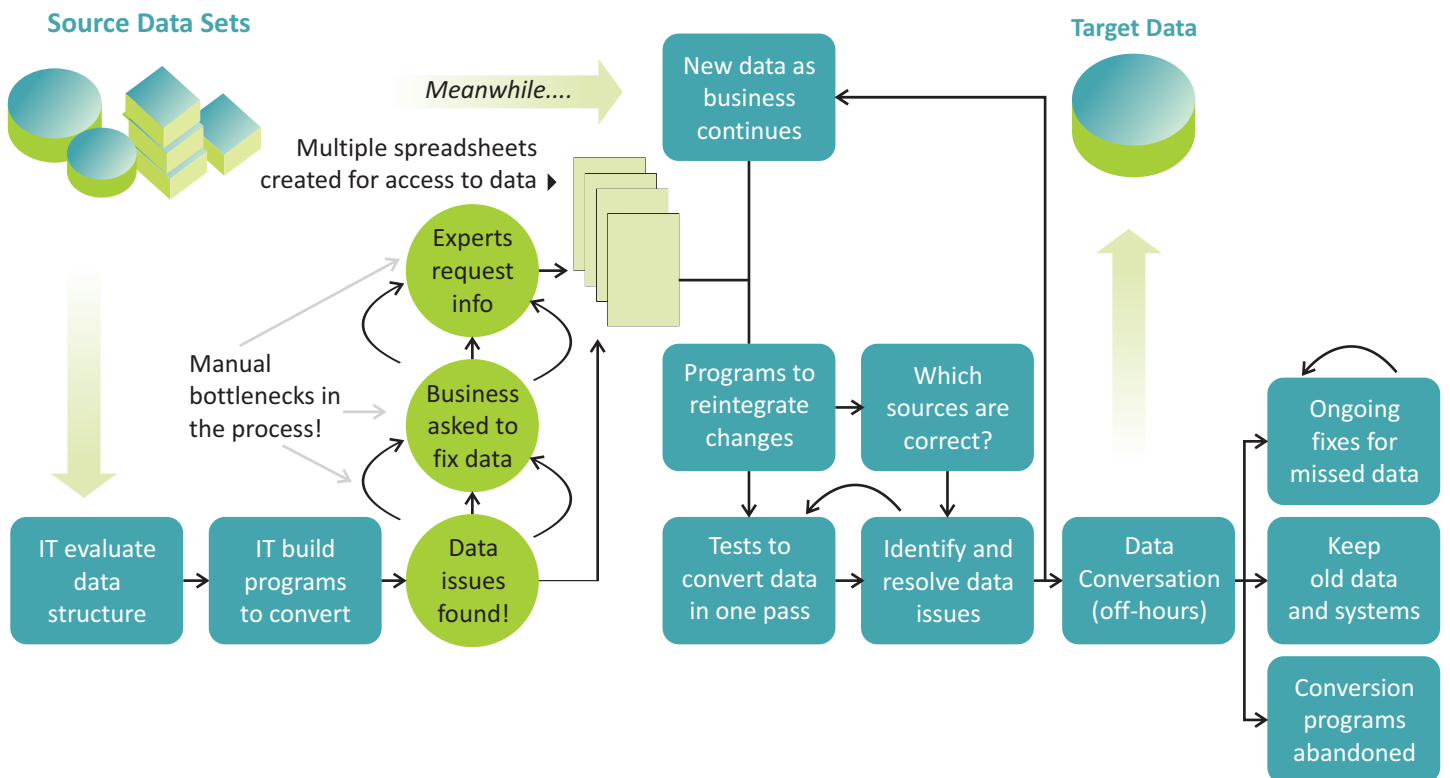
“We must use the data we have because there is no time to stop and wait for the project to complete!”

Attempts to implement change in data management are often static projects in a dynamic workflow. This is the foundation of the cycle of doom. The process can never impact the reality of the workflow situation and cannot provide benefits that keep pace with the data generation. As a result, the changes are perceived as irrelevant at best, or considered a waste of money because they are a drop of water in the ocean.

Most data transformation initiatives focus on moving data from a source data set to a new target for various reasons (e.g., consolidating a number of data bases, moving from an old technology to a new one, improving an existing data set). The manual method begins with IT staff evaluating the existing data sets and, while running scripts, they quickly discover data issues such as inconsistencies and lack of standards. When these issues are found, costly business experts are brought in to fix the problems. They request sound data to be further described in spreadsheets and begin going through the data line by line, reviewing thousands of well, drilling, production or exploration exceptions one by one.

This creates a cyclical process where three separate data sets are being used at the same time (i.e., original scripts derived by IT, spreadsheets which are reviewed and edited by business experts, and new business data generated at source) The data is never properly consolidated and the conversion is completed with an attitude that the result is “good enough.”

In the end, thousands of labor hours and dollars are spent on a project that exceeded resources and where none of the results were maintainable, repeatable or reproducible. The new system’s data quality deteriorates, which further undermines the confidence users have in the information and the technology. In addition, the data still requires ongoing fixes – just as it did before the project began. Eventually, it is realized that the process was a one-time solution to a continuing problem and business decisions for data quality are lost and unavailable to be reapplied in the future.



## Industry Issues

Most major oil and gas companies have invested billions of dollars in data acquisition over a number of years. Today, however, corporate investment in data is down. In spite of the expertise and resources employed to address upstream data management challenges, end users of various upstream data types are still struggling to get accurate and clean information for analysis.

## Common Symptoms and Their Causes

The list of issues is familiar to any manager in an E&P organization and includes symptoms like these:

- Business users aren't engaged.
- All users are complaining.
- Business users don't trust the data.
- Excel and Access databases are growing out of control.
- Projects take too long
- Metrics are poorly measured.
- The business feels too much money is being spent on DM projects.

Digging deeper, each of these symptoms has an attendant list of causes.

When business users aren't engaged, it is likely that:

- They do not have pain.
- The business believes that data is solely an IT issue.
- Management and/or the project team aren't communicating enough to end users.
- They are concerned that their decisions will be questioned because they have been made on poor data - the pain of change is higher than the gain of change
- The data management initiative starts off with a focus on data that is too complicated.
- Users don't see what specific benefits they will get from the results of the initiative.
  - User issues are not confirmed, prioritized or truly needs- often such projects are undertaken without a sponsor or based on wants rather than needs
  - Metrics and milestones are not defined
  - Course checks along the way lose emphasis or users 'assume' the IT group will involve them

### When users are complaining, it is likely that:

- They are experiencing initiative overload.
- They are expected to start with highly complex data that is difficult to deal with.
- Data is hard to access, which fosters loss of productivity and creation of duplicate data in various data sources.
- The data is not clean enough to be of value to the users.

### When business users don't trust the data, it is likely that:

- The same data is inconsistent across databases.
- There is little or no governance process within the business unit to track, maintain, and govern data.
- There is not one source of authority for various data types.
- Base acquired data is handled by several entities and stored in several corporate locations
- There hasn't been sufficient end user training in data management systems.

### When Excel and Access databases are growing out of control, it is likely that:

- Users don't trust any data except what they have on their own computers (see the previous point).
- Data is stored in multiple databases with no integration between them. This is costly to maintain and lacks any ability to impose data standards or governance.
- There are no clearly defined master data stores.
- There hasn't been sufficient end user training in data management systems.
- The business is not finishing past initiatives, and so is not transferring data or creating reports.

### When initiatives are taking too long, it is likely that:

- There are no measurable and/or visible milestones to create a sense of progress. In this case, "too long" can be a feeling rather than a fact.
- The initiative is trying to eat the whole data elephant at one go.
- Users aren't engaged (see above).
- Acquired data is not well tracked, leading to re-acquisition of the same data, perhaps more than once.
- The initiative is taking on too many data sets and/or developing architectures in too much detail.

### When the business believes that too much money is being spent, it is likely that:

- The initiative is trying to eat the whole data elephant at one go.
- There are no measurable and/or visible milestones to create a sense of progress. In this case "too much money" can be a feeling rather than a fact.
- There is little or no data governance within the business unit.
- There are no metrics to measure DM initiative benefits and no measurement of inefficiency costs or risks inherent in the current DM workflow.

### Barriers To Success

*Recently, a client company's Senior Vice President of Exploration and Production expressed his frustration. While there is great benefit in consolidating their information assets, he said, he does not believe there are solutions for this. Nor does he believe they have the resources to even take on such a project. His prior experience with smaller projects never delivered any value that came close to justifying the required investment in time, people and money from this very senior user's perspective.*

There has never been a question about the need to fix underlying causes like these. However, though common sense dictates the objectives to be attained by a data management improvement exercise, actions taken to produce the desired results have often not been effective. The barriers to success reside in two areas: data management and business management.

### Data Management: Too Much, Too Few, Too Long

Barriers to solution in the data management area relate to availability and use of resources:

- **Too much** time is being spent by expensive resources looking for, cleaning-up and managing complex data.
- **Too much** time is spent by geological and geophysical resources on managing seismic data (low value) rather than interpreting seismic data (high value).
- **Too much** money is spent on seismic data acquisition costs because of poor data management practices (data is lost, stored on local drives, purchased twice, stored in non-digital form, not accessible remotely, etc.).
- There are **too few** qualified, experienced resources to evaluate reservoir prospects.
- It takes **too long** to move projects from exploration to development to production.

### Business Management: Reliability, Time, and Process

Solution barriers that reside in business management impede productivity in various ways:

- Too time consuming to analyze data from legacy data systems
- Unstructured data isn't linked to structured data
- Siloed data sources (structured and unstructured) make it difficult to integrate the data to make informed decisions
- Different data results for the same queries from different systems
- No single source of the truth for accurate production data
- Business management abdicates responsibility for data and data quality to IT or projects
- Engineers often store "clean" production data on their hard drive so they can control the data quality
- Inconsistent business processes for common engineering activities (e.g. well performance review process)

## Leading Practices

The symptoms, causes, and barriers to success outlined above are not new; anyone who manages or works with upstream data will be familiar with most if not all of them, either in practice or from all of those books, articles, and presentations on the subject. We have listed them here so that we can tie them to the leading practices we recommend to help overcome DM challenges.

- Following the logic of our riddle, **“bite-size” your data management strategy into smaller projects that build on one another**. Success fosters the motivation and energy to take the next bite.
- **Get the business involved and taking data ownership**. Stop if the business sees data management as solely an IT issue. If business stakeholders do not feel pain from the current data situation, they will remain bystanders with more interest in preserving the status quo.
- **Start with production data or other data that applies across the business and achieve quick wins**. Focus on highly visible data that many of your stakeholders see, use, or build on for their workflows. This will build momentum from the very start and will attract the interest and participation of additional business units.
- **Develop a high level architecture**. This helps guarantee the parts will eventually equal the needs of the whole elephant.. No one can predict the future with accuracy but plan for your growth and enable scaling in you plan from day one.
- **Standardize on common applications and integrate workflows**. Following the industry trends towards open architecture and common standards means that your efforts will support a sustainable solution that can be modified with your changing business needs easily and more cost effectively.
- **Define master data stores**. Use a meta-data catalog strategy with search and strong data governance standards to assure that the solution is logical, efficient and provisions for the unique access needs of your business units.
- **Include structure and unstructured data to capture users’ attention**. Data is forever, but the people who know where the data is come and go. The combination of structured and unstructured data repositories, files and secret drawers is a major asset for any company. Mining for it can be an overwhelming and unsolvable task unless you have the governance and architectural vision to see its value today.

### Conclusion

Where are you in relation to your own upstream data elephant? What symptoms are showing up in your end user community that will likely impede your data management improvement efforts?

Whatever symptoms are stopping or slowing progress, the key to solution is addressing underlying causes. The leading practices we have identified here don't apply to every situation; rather, each one addresses specific causes discussed above. The table below will help you determine which practices to employ in your particular situation.

## About the Authors

### **Jim Lawnin**

Mr. Lawnin is Vice President and leader of Wipro's Global Energy Consulting Practice. He has over 25 years experience oil and gas experience: 15 years as an energy consultant for several large management consulting firms and 10 years experience as a petroleum engineer. Before joining Wipro, Mr. Lawnin led the oil and gas industry practice for a large consulting and outsourcing firm. He also led a global consulting and systems integration practice. He holds certifications as a Project Management Professional (PMP) and a Black Belt 6 Sigma, and has held certifications as a licensed CPA, a Chartered Financial Analyst (CFA), Certified Financial Planner (CFP) and licensed Professional Engineer (PE). He is a frequent lecturer on solving complex oil and gas industry issues through innovative solutions.

### **John Yurkanin**

Mr. Yurkanin is Vice President for Operations for Petris Technology, responsible for all business in the Western Hemisphere. He has over 30 years experience in multiple aspects of the global energy business and has been involved in many emerging trends for deregulation in natural gas and electricity. He was also a practice leader for a large consulting firm serving the strategic business needs of the utility and energy business. Mr. Yurkanin's current focus revolves around the integration and knowledge preservation of the data asset. He holds a degree in chemical engineering and advanced degrees in business and process management.

## ABOUT WIPRO TECHNOLOGIES

Wipro Technologies (<http://wipro.com>) is the first PCMM Level 5 and SEI CMMi Level 5 certified IT Services Company globally. Wipro provides comprehensive IT solutions and services (including systems integration, IS outsourcing, package implementation, software application development and maintenance) and Research and Development services (hardware and software design, development and implementation) to corporations globally.

Wipro's unique value proposition is further delivered through our pioneering Offshore Outsourcing Model and stringent Quality Processes of SEI and Six Sigma.

## ABOUT PETRIS TECHNOLOGY, INC.

Petris Technology, Inc. (<http://petris.com>) is a leading supplier of practical data management solutions and geosciences applications to the global energy industry. Founded in 1994 and with over 500 clients throughout the world, Petris leverages its insight and knowledge to design technology that integrates information from diverse data stores including financial, seismic, borehole, production, drilling, and pipeline to enable continually better decision making and application transparency.