



## FEDERAL RESERVE BANK *of* NEW YORK

### MIT 2010 Information Quality Industry Symposium

Massachusetts Institute of Technology (MIT), Cambridge, MA, USA

### *Information Quality and the Financial Crisis*



The following presentation represents the views of the author and not necessarily the views of the Federal Reserve Bank of New York or the Federal Reserve System.

John Bottega  
Chief Data Officer  
15 July 2010

Modified: 31-May-11

## Quality Defined

### What determines quality?

#### Quality

According to the Praxiom Research Group's translation of the ISO 9000 standard\*.

"The *quality* of something can be determined by comparing a set of inherent characteristics with a set of requirements. If those inherent characteristics meet all requirements, high or excellent quality is achieved. If those characteristics do not meet all requirements, a low or poor level of quality is achieved."



What determines the quality of a suit?

- ✓ Fit
- ✓ Style
- ✓ Durability

How is quality achieved?

- ✓ Materials used
- ✓ Workmanship employed

#### The "Manufacturing Process"

Process by which you construct the final product – combine the **best materials** with the **best workmanship** through **best practices** to achieve a quality end-product

#### What can go wrong?

When any element of the manufacturing process is allowed to degrade, the weakest link affects the quality of the whole...

- Shuttle "O" ring
- Toyota Brakes

\* <http://www.praxiom.com/iso-definition.htm>

### What determines the quality of data?



**Data Quality is based on how your data “A-C-T-S”**

- ✓ Accuracy
- ✓ Completeness
- ✓ Timeliness
- ✓ Adherence to **Standards**

### How is quality achieved?

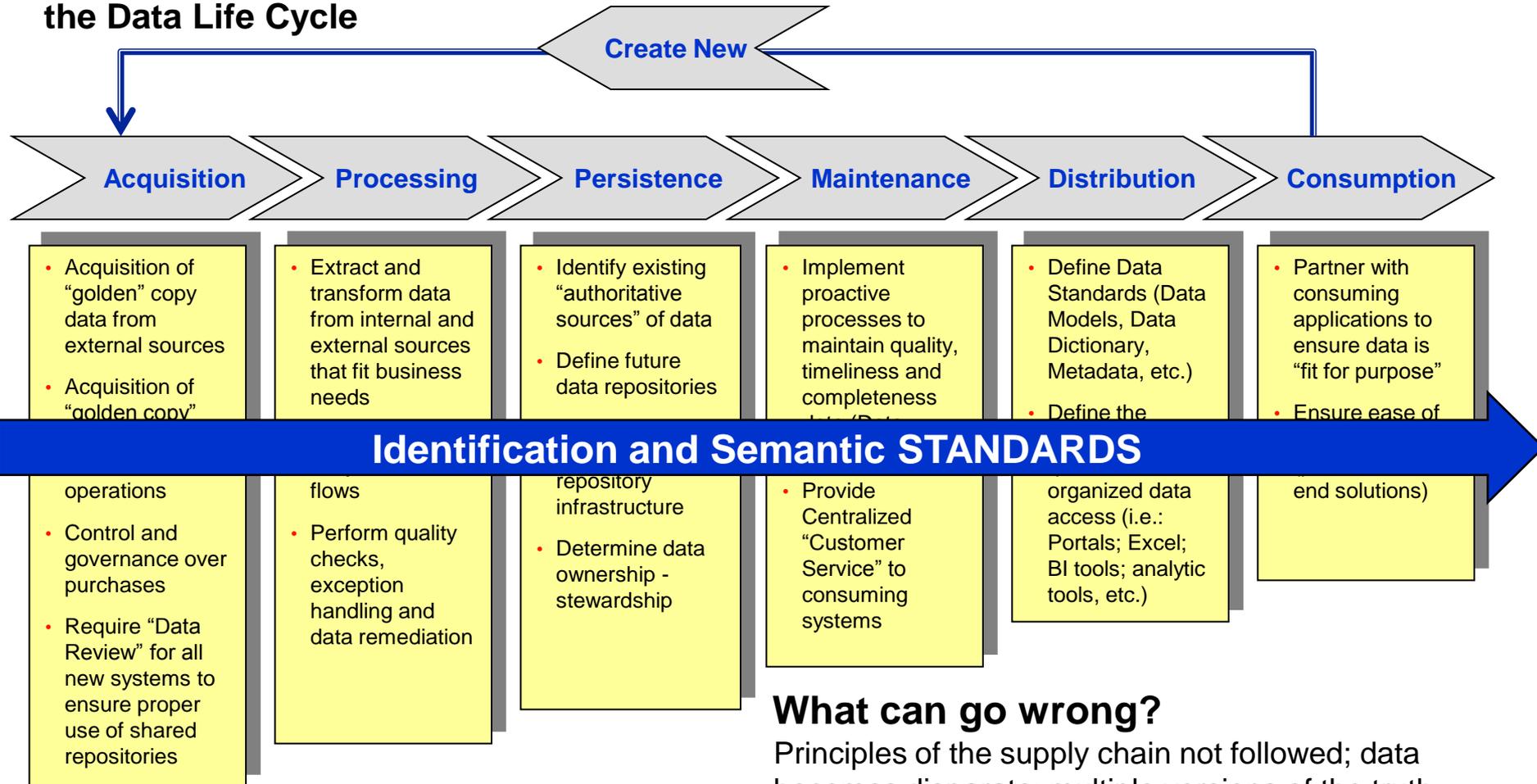
- Quality of the raw material – data captured “@ source”
- Quality workmanship – skill set of the data stewards
- **The “Manufacturing Process”** by which data is collected and maintained

## Data Supply Chain

*It’s about getting the right information, to the right people, at the right place, at the right time*

# Data Supply Chain

Data Supply Chain is made up of a series of 'component disciplines' that makes up the Data Life Cycle



## What can go wrong?

Principles of the supply chain not followed; data becomes disparate; multiple versions of the truth emerge. If the "manufacturing process breaks down; Quality diminishes...

### Data Management in Finance – A retrospective...

- **30 years ago, data management was simpler**
  - Finance industry simpler
  - The products were less complex
  - The data was less disperse
  - Data was maintained in one place – “The Mainframe”!
- **Mainframe**
  - Single version of data
  - Single owner – single stewardship
  - Downside: Access was limited and controlled
  - Lived in a “**Data Dictatorship**”
- **We were liberated...**
  - The PC and distributed computing were introduced
  - **A chicken in every pot, and a PC on every desktop!**
  - Data could be created, acquired, stored and used by any individual, in any department, anywhere in the firm
  - In one leap – we moved from “Data Dictatorship” to “**Data Freedom**”

**OR DID WE???**

### Data Management in Finance – A retrospective...

- **As the world changed, so did the Data Management Landscape**
  - Data Management went from centralized to disparate
  - Silo's emerged
  - Ownership was blurred – everyone owned it, so no one owned it
  - No guideline, no governance of data acquisition and data usage
  - No “rules of the road”

***“Data Dictatorship” became “Data Anarchy”***

- **But this was OK for the person on the desk because...**
  - Data Freedom outweighed data anarchy
  - Profits outweighed efficiencies

**But nothing stays the same...**

### Data Management in Finance – A retrospective...

- **Markets Began to Change**
  - Financial Products were becoming more complex
  - Markets, economies were becoming more interdependent
- **Bank's Focus was Changing**
  - Repeal of the Glass-Steagall Act removed the separation that previously existed between Wall Street investment banks and depository banks
  - Investment Banks discovered cross-selling
    - Banks were becoming “customer-centric” vs. “product-centric”
    - GM vs. General Motors
- **But then – significant shocks hit the system**
  - Orange County, CA
  - Asian Financial Crisis / The Russian Financial Crisis
  - Euro Conversion
  - Y2K
  - Terrorism

**Banks began to realize the importance of data and data management in understanding the interdependencies of the financial markets**

### Data Management in Finance – A retrospective...

- **How did the industry respond?**
  - Made Investments in data programs
- **But 3 critical mistakes were made...**
  - **Threw it “over the wall” to technology**
    - Did not recognize the significant Business component of Data Management
  - **Did not anticipate resistance from the Business**
    - Had not anticipated the unwillingness of the business to “give up” their new-found data freedom
  - **Assumed short-term investment to solve the problem**
    - Had not realized the amount of unraveling that needed to be done
    - Had not realized the dependency of current processes on the existing fragmented infrastructure

**The Art of the Long View: Planning for the Future in an Uncertain World  
– Peter Schwartz**

**Suffered from the “Curse of the Short View”**

### Data Management in Finance – A retrospective...

- **No commitment to a long-term fix**
  - Firms viewed data as a “project” and not a “program”
  - Many were canceled after 1-2 years, leaving a landscape of unfinished projects - like shiny new bridges spanning **half-way** across the rivers
- **Data fragmentation continued unresolved**
  - Some progress was made, mostly within firms – exacerbated the fragmentation
  - Most industry standards efforts stalled (ex: GSTPA a non-starter)
  - Data standards never got the right level of commitment from “the Board room”
- **Making of a “Silent Killer”**
  - While the issues of data quality and data fragmentation persisted, the industry and its products grew more complex, more opaque, more interdependent
    - When the health of an organism is weakened through continuous bad practices, exposure to a contagion or extreme stress can cause the organism to break down.
- **Data and the Economic Crisis of 2008**
  - Did the quality of data available to the decision makers impact their ability to properly analyze the state of the economy?
  - Did data quality (or lack thereof) have an effect on the economic crisis?

## Observations of the Financial Crisis

---

### Basic Observations from the crisis (through the “data” lens)...

- **We saw historic market turmoil...**
  - This resulted in unprecedented number of mergers, acquisitions, divestitures and bankruptcies – putting immense pressure on the need for timely and accurate data about entities.
- **We saw increasingly complex investment vehicles...**
  - Resulting in increasing complex data sets with layers upon layers of abstracted data and data relationships putting pressure on our already weakened legacy of financial instrument infrastructures and architectures
- **We saw pricing and valuation discrepancies**
  - Demanding better public transparency of pricing and pricing methodologies`

### According to the National Academy of Sciences (2009)...

“Rapid change in the financial system driven by innovation and deregulation...has altered the mechanisms and pace of financial intermediation to such an extent that regulatory tools, processes and data have fallen behind”.

## How did we get here???

---

### Some theories as to the cause of the financial crisis...

#### Lenders

#### Perspective:

- Some believe the financial crisis stemmed the desire to increase home ownership
- Others believe it was caused by the false assumption that home values (and salaries) would continue to rise ...the “LTV promise”
- While others believe it was due to the creation of specialized mortgage programs which lured people in with incredible “teaser rates”, but never fully explained how the payments would adjust, nor explained the impact on the home equity.

#### Investors

#### Perspective:

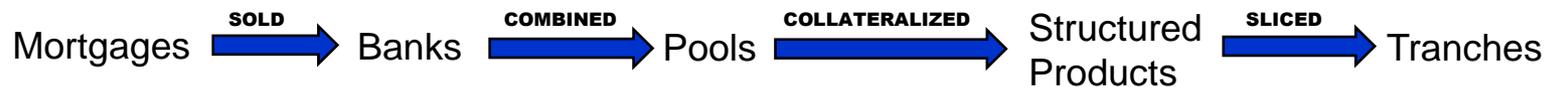
- Put blame squarely on the loan originators (“risk” based to “fee” based)
- Others claim that the creation of all this mortgage backed debt was done to satisfy the high demand for quality investment opportunities (viewed US mortgages as sound investments)
- Others point to the glut of foreign investment dollars “looking for a home”
- Some blame the rating agencies
- Others blame the GSEs (Government Sponsored Enterprises)
- Others simply blame greed

**Whatever the exact mix of ingredients was, it resulted in an environment that flooded the financial markets with toxic, unstable, risky and opaque investments**

## Through the Lens of Data Management

### Given this influx of toxic instruments, how did data quality play a role?

#### Looking at the Collateralization Process...



#### Data issues along the way...

- Bits and pieces of descriptive data are not carried through each step of the process
- Linkages to the underlying loans become broken or obscured
- Slicing and dicing creates a wedge of abstraction between the original loans and the newly created investment vehicles.

#### The result...

- Critical information about the underlying loans became so abstracted from the financial product in hand that it became increasingly difficult, if not impossible, to truly assess the true value (and risk) of that instrument.
- And when the perfect storm materialized – home values dropped and adjustable mortgages adjusted up – loans began to default leaving decision makers without the proper information needed to assess and react to changing market conditions.

### Is this impact over-exaggerated?

#### Remember how the banks communicated their exposure to the subprime crisis?

- We know we have a problem
- We've assessed the damage, we know where we stand
- We're ok

#### Reality...

- Banks didn't really have all the information they needed about these toxic instruments
- There was no standard way to identify these instruments, no agreed upon business rules
- Whatever data that existed was either never captured, or lost through the data supply chain

#### Quality of data was compromised

#### What happen next...

- Banks began to restate increased exposures, often double or triple original estimates
- Financial institution's stock prices came under tremendous downward pressure
- Then the unthinkable started to happen – banks started to fail...

### Consider the events of the Lehman collapse

- On that Saturday in September, word was spreading that Lehman Brothers may fail
- Tech and Ops teams streamed into their respective banks and the data crunching began...
  - What was our aggregate exposure Lehman?
  - Who were all the Lehman subsidiaries?
  - Which Lehman entities actually declared bankruptcy – and which did not...

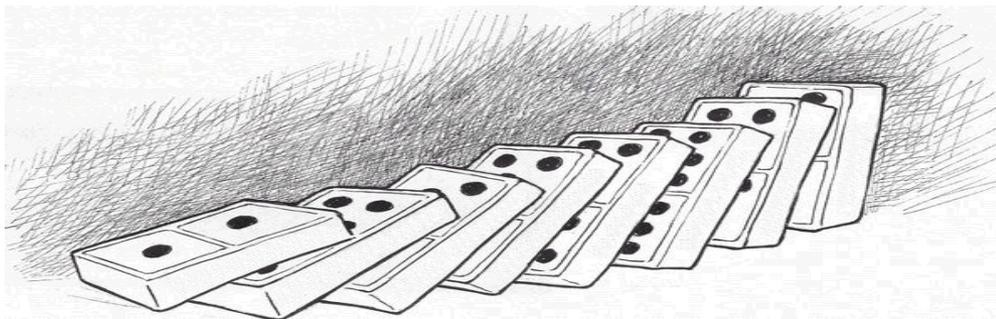
**Not moving cash to legitimate entities was just as bad as moving cash to bankrupt entities**

### Unique Entity Identification and Legal Hierarchy are critical

- Must be able to uniquely identify financial entities
- Must understand the complex relationships of parent to sub, entity to sub-entity
- Must have an understanding of the effect of one entity on another

### The realization – the data was not readily available....

- The “Raw Materials” were not there
  - There was no standard way to identify Lehman and her entities cross the industry.
  - There was no clear understanding of the Lehman organizational structure
  - There was not way to fully understand and predict the impact across the industry



John Liechty, associate professor of Marketing and Statistics at Penn State University, co-founder of the Committee to Establish National Institute of Finance said the following in a Business Week interview last August 18<sup>th</sup>...

*“Does anybody have the data in place to really deal with systemic risk”?  
Had this collection of data and analytics existed last fall, “regulators could have modeled the repercussions of a Lehman collapse using actual data...”*

***If the quality of data had been better, would different decisions have been made?***

### What do we conclude about the quality of data and the crisis?

- Quality was poor...
  - Lacked accuracy, completeness, timeliness
  - Did not adhere to standards (disparate; inconsistent; silo'ed)
  - Critical components (lineage; linkages; hierarchies) were not present
  - Process to collect and normalize data (Data Supply Chain) malfunctioned

### What do we conclude about data and the crisis?

#### **More is NOT better...**

We had data but it was not comparable. It did not satisfy requirements. It was not collected and captured AT SOURCE in a methodology and format that would enable analysts to effectively utilize this critical financial data to perform their analysis.

*Paraphrase the Rime of the Ancient Mariner:  
“Data, Data Everywhere, nor any a drop to drink”*

Data may not have been the cause, but “gaps” in the quality and completeness of our data may have contributed to the crisis by leaving the decision makers often without the important and timely information they needed to make sound decisions.

## How do we address this problem?

### Two Levels:

1. Institutional (firm) level
2. Industry level

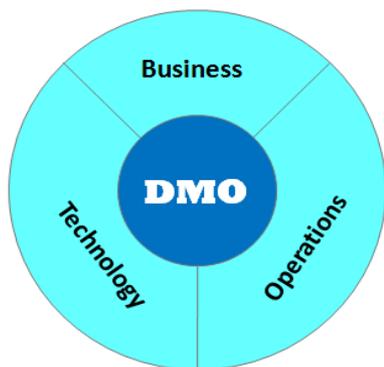
### Institutional Level

- **Organizational alignment within a firm focused on Data Management**
  - Establishing data as a critical corporate discipline
  - Assigning clearly defined roles and responsibilities
  - Defining and enforcing strong data program governance, recognizing data as a critical corporate asset
- **Support from Senior Management that makes Data Management a priority...**
  - Formally established
  - Commitment to the long-term
  - Willingness to change corporate culture with respect to data management

**“It's not your father's data management anymore”**

# How to Address The Data Challenge Going Forward

## How are (should) firms implement?



### Establish a “Data Management Office”

- Act at the “organizational glue”, bringing together the key Stakeholders, to ensure a successful Data Management program
- Represents a “Paradigm shift” – It is not just a technology problem anymore - Establishing a Business Driven, Technology enabled, Operations supported organizational discipline

### The Role of The Data Management Office

- Vision & Ownership
  - Assume accountability of the Data Management challenge
  - Establish and communicate the Data Management vision
- Drive Organizational Alignment
  - Define the roles of business, technology and operations
  - Define the Operating Model
- Establish Data Management Governance
  - Define Data Management program governance
  - Define Data Management “content” governance – define metadata policy

## How to Address The Data Challenge Going Forward

---

### How do we correct it at the industry level?

#### We have to drive the components of the “Data Supply Chain”

- Unique identification of all objects (symbology)
- Standardization of descriptive data (semantics)
- Standardization of reporting (standard markup)

**This is not a new approach! This has been looked at for years, seemingly always falling short at the finish line. Why? Seems it was always missing the catalyst – until now...**

**“You never want a serious crisis to go to waste”**

Rahm Emanuel, President Obama’s Chief of Staff

Wall Street Journal Conference, November, 2008

#### **Changes must be driven by mandatory compliance**

- Imposition of basic rules of data management and data quality at inception to ensure proper object identification, definition and capture
- Map to common semantics. All financial agreements are legally binding – we need to capture and standardize this information when it is created.
- Impose reporting standards on all financial entities to ensure data flows consistently and unambiguously throughout the data supply chain

## How to Address The Data Challenge Going Forward

Testimony by **Mr Daniel K Tarullo**, Member of the Board of Governors of the Federal Reserve System, before the Subcommittee on Security and International Trade and Finance, Committee on Banking, Housing, and Urban Affairs, US Senate, Washington DC, on “**Equipping financial regulators with the tools necessary to monitor systemic risk**” February 12, 2010

“The recent financial crisis revealed important gaps in data collection and systematic analysis of institutions and markets. Remedies to fill those gaps are critical for monitoring systemic risk and for enhanced supervision of systemically important financial institutions, which are in turn necessary to decrease the chances of such a serious crisis occurring in the future. “

“Greater standardization of data than exists today is required. Standardized reporting to regulators in a way that allows aggregation for effective monitoring and analysis is imperative.”

“Legislation will be needed to improve the ability of regulatory agencies to collect the necessary data to support effective supervision and systemic risk monitoring.”

## How to Address The Data Challenge Going Forward

---

### In Conclusion...



Data will not prevent the disease. However, accurate, complete and timely data, that can quickly and effectively analyzed, will enable the decision makers to prescribe the right medication to minimize the impact of an illness and perhaps even prevent it from occurring again.

We have an opportunity as an industry to implement change that will improve our ability to protect and safeguard the financial industry and our economy!

Data and Data Quality stand at the forefront of the tools needed to accomplish this mission.



**FEDERAL RESERVE BANK *of* NEW YORK**

**THANK YOU!**

**John A. Bottega  
Chief Data Officer  
Federal Reserve Bank of New York  
(212) 720-5922  
[John.Bottega@ny.frb.org](mailto:John.Bottega@ny.frb.org)**