

## Improving Your Business IQ and Governance

### ABSTRACT

---

This presentation stresses a pragmatic approach to creating and implementing quality information & governance for a business unit that must service hundred of users and ensure speed & accuracy. How do you identify and extract data from all data sources and ensure the quality of the data is high and relevant? The principles and approaches outlined will enable the attendee to utilize a pragmatic approach to getting your data ready for that big project of integration(database & spreadmart) and data sharing. This presentation will outline

- Data quality for a package data warehouse implementation
- Inserting the data quality process in an Excel world
- Data Hops by Role – Inserting Data Governance
- Challenging top-down from the bottom (a business data model approach driven from a bottoms up process)

### BIOGRAPHY

---

#### **Linda Kresl**

Data Warehouse Manager  
Global Customer Care Division  
Yahoo!



Ms. Linda Kresl has held a variety of professional and management positions with world-class companies such as The Boeing Company, Hewlett Packard, PriceWaterhouseCoopers, and Nike. From 2001-2007 Ms. Kresl established her own consultancy specializing in BI & Enterprise Data Architecture. Her professional experience of more than 20 years includes development of enterprise Business Intelligence, Enterprise Information Management and Data Quality Improvement. Ms. Kresl is currently the Data Warehouse Manager for the Global Customer Care Division at Yahoo!

Ms. Kresl has been a speaker on Information Quality Management at conferences such as the MIT IQ Industry Forum and the Information & Data Quality Conference. She is a member of the Data Warehouse Institute (TDWI), & is a certified DW 2.0 Architect. She sits on the board of DAMA International (Global Chapter). She has published in DMReview, and her articles have also appeared in Oracle Toolbox.



**Improving your Business IQ & Governance**  
MIT Information Quality Industry Symposium  
July 14-16, 2010



**Linda Kresl**  
Data Warehouse Manager

**YAHOO!**



The MIT 2010 Information Quality Industry Symposium



## Outline

- **Data quality for a package data warehouse implementation**
- **Inserting the data quality process in an Excel world**
- **Data Hops by Role – Inserting Data Governance**
- **Challenging top-down from the bottom (a business data model approach driven from a bottoms up process)**



The MIT 2010 Information Quality Industry Symposium



## What we do



### OUR VISION:

To be the center of people's online lives.

**FUN FACT:** There are more Yahoo! Games users in the U.S. than people living in Tokyo, Japan.

Unleash the Power of Your Information



The MIT 2010 Information Quality Industry Symposium



## Who are we?

### Websites

**YAHOO!.COM**



**YAHOO!.FINANCE**



Unleash the Power of Your Information

### Products

**YAHOO!.MESSENGER**



**YAHOO!.GROUPS**





The MIT 2010 Information Quality Industry Symposium



## Who are we?

**Retail**

**Subsidiaries**

**Divisions**

**match.com** an YAHOO!

**flickr**

**Customer  
Advocacy**

**YAHOO!** SMALL BUSINESS

**YAHOO!** hotjobs

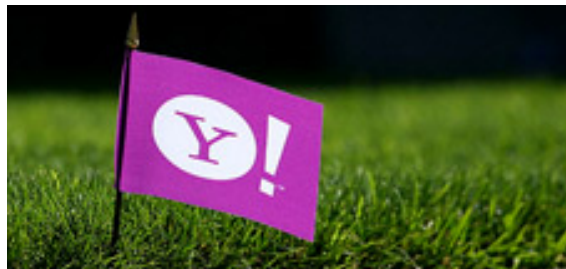
**Engineering**



The MIT 2010 Information Quality Industry Symposium

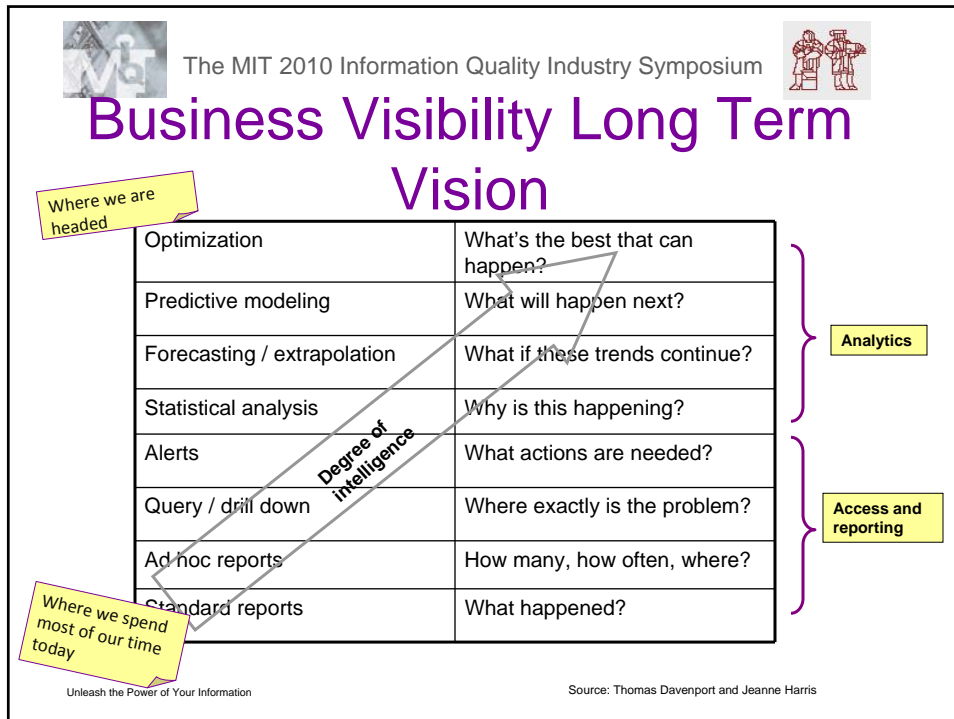



## FY09 Total Yahoo, Inc. Revenues



**\$ 7 Billion**


Unleash the Power of Your Information





The MIT 2010 Information Quality Industry Symposium

**Category** **Business Requirements Specification**



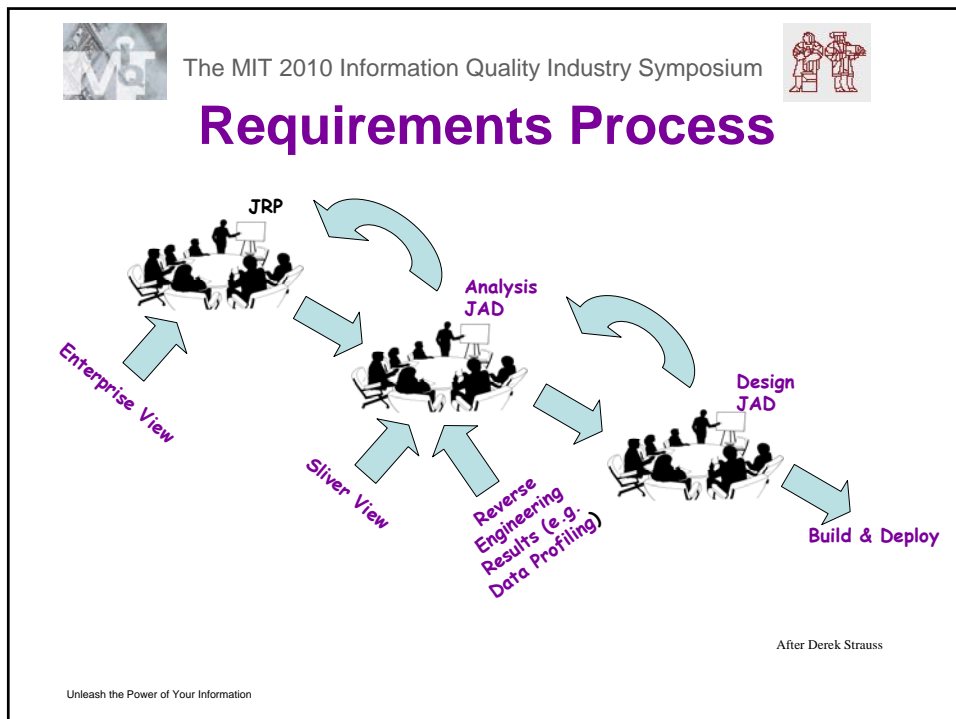
<b>Functional Requirements</b>	<ul style="list-style-type: none"> <li>Build Strategic Plan &amp; Framework</li> <li>Requirements Gathering including Business Rules, Use Cases, System/User Requirements and Business Data Dictionary.</li> <li>Detailed Plan (Project Plan/DW Prototype)</li> <li>Business Data Model</li> <li>Report Inventory/Data Source Inventory for input into Source to Target Mapping</li> <li>Business &amp; Process Metadata creation</li> </ul>
<b>Technical Requirements</b>	<ul style="list-style-type: none"> <li>Data Models, Database layouts, schemas</li> <li>Technical Metadata dictionary</li> <li>Design Technical Infrastructure</li> <li>Source to Target Mapping</li> <li>Tool requirements for design, ETL, &amp; reporting</li> <li>Conversion/ Maintenance Plan</li> <li>Data Quality Testing and UAT</li> <li>Presentation Layer Design</li> </ul>
<b>Non-Functional Requirements</b>	<ul style="list-style-type: none"> <li>Training Guide Design and End User Training</li> <li>Data Access (Permissions, roles) /Privacy laws for domestic and International Policies</li> <li>Legal/Regulatory Policies for data retention</li> </ul>

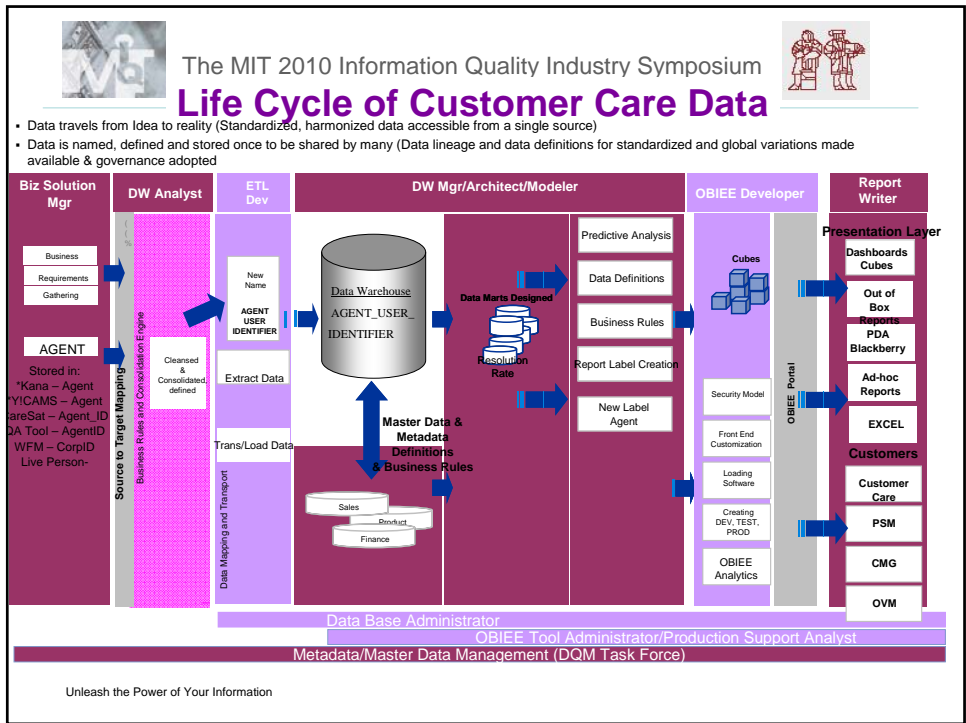
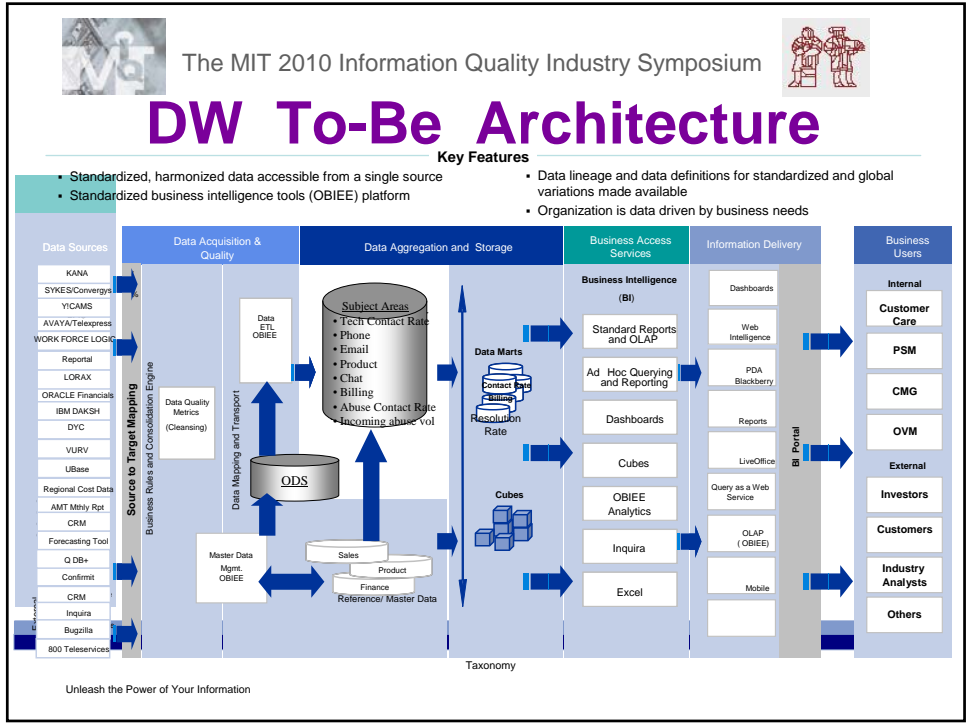
Unleash the Power of Your Information

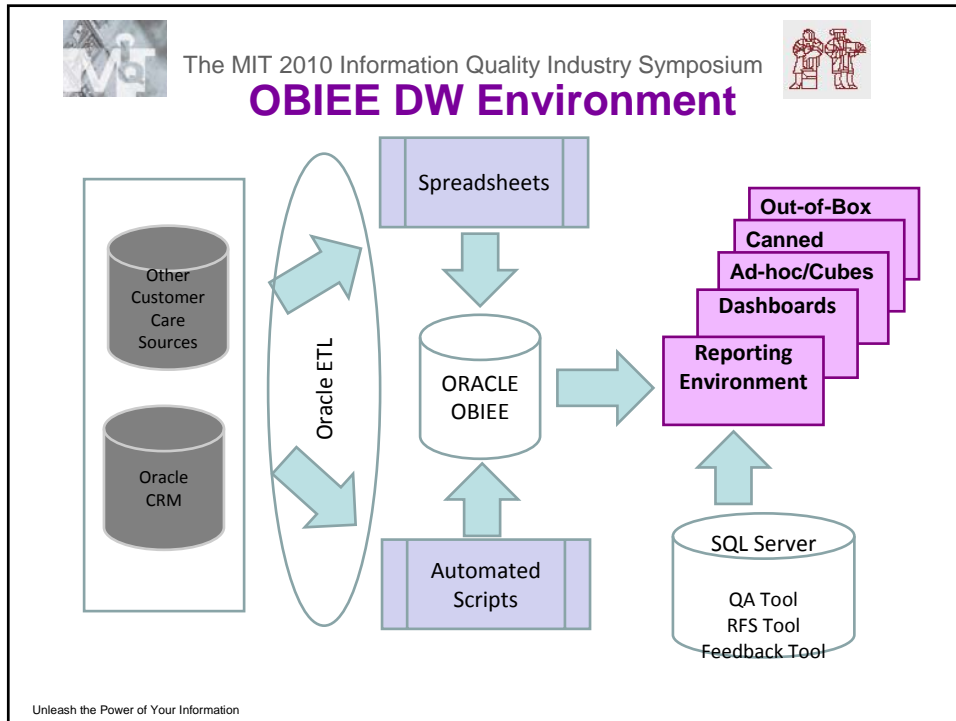
Legend

Biz Viz Team

IT









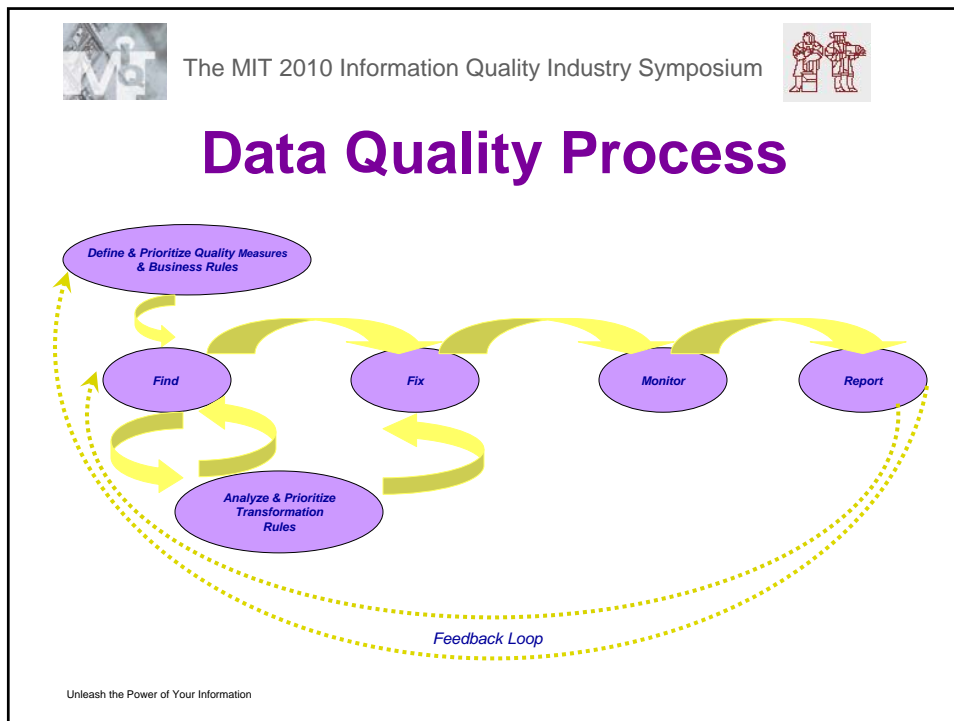
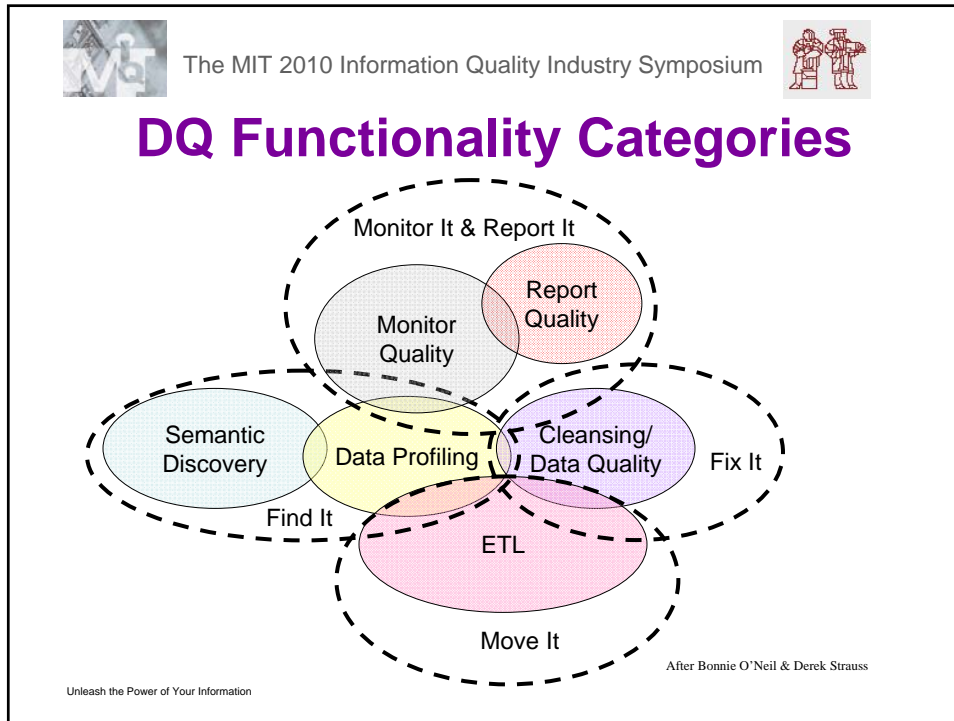
The MIT 2010 Information Quality Industry Symposium

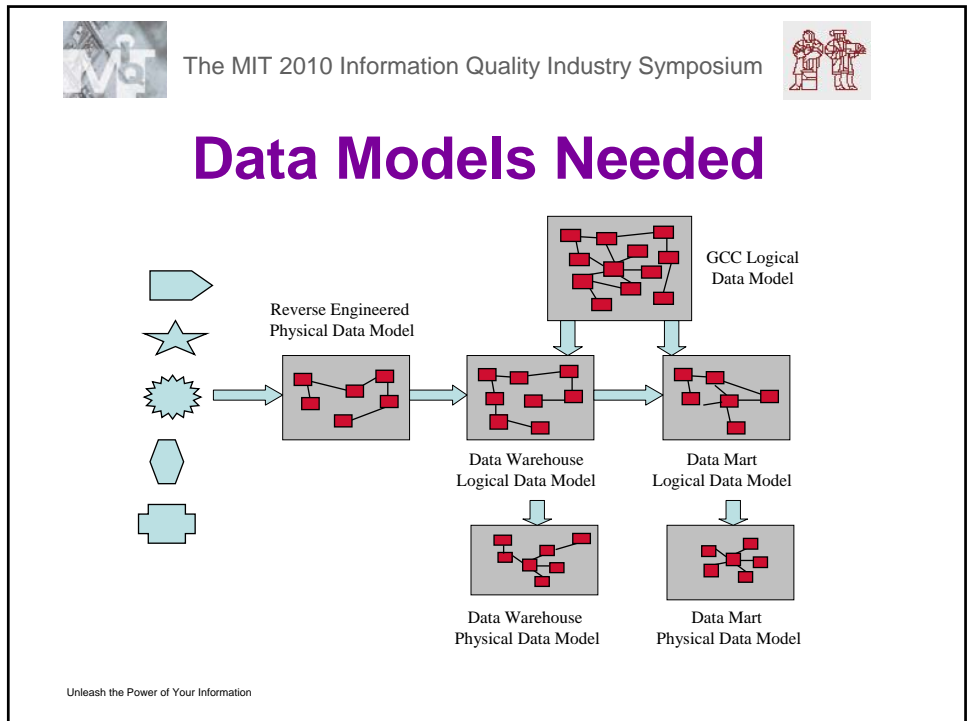
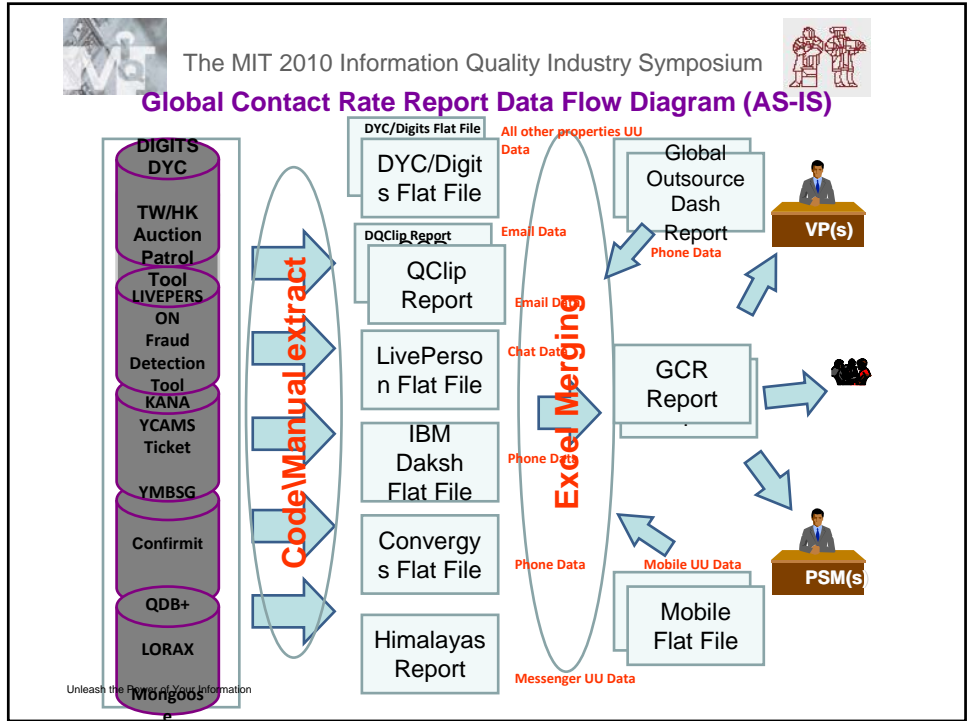
## Data Quality for GCC Data Warehouse

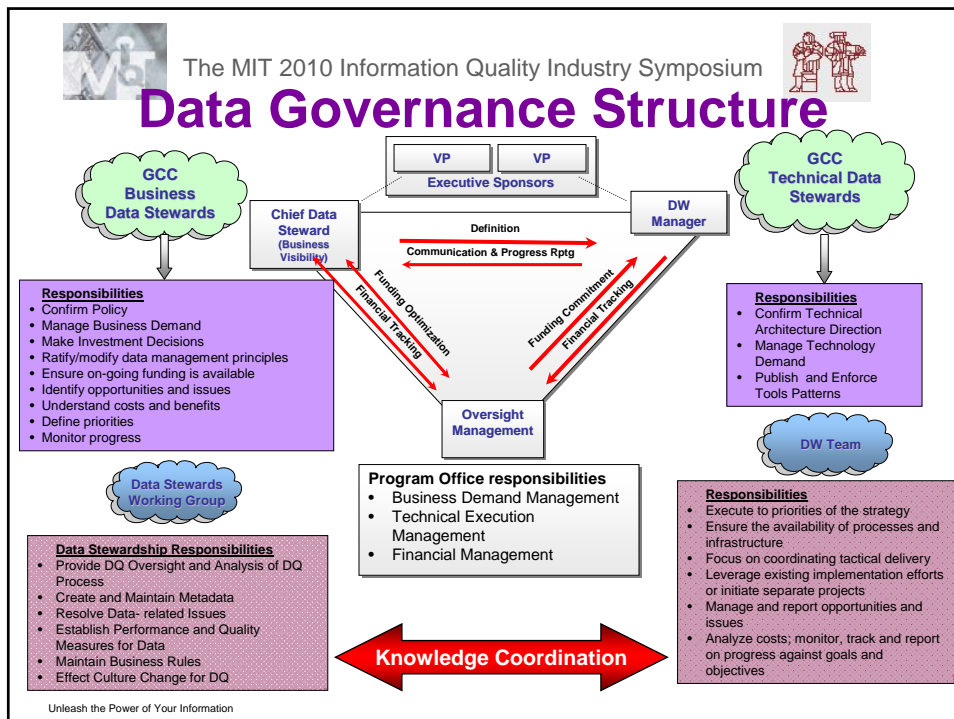
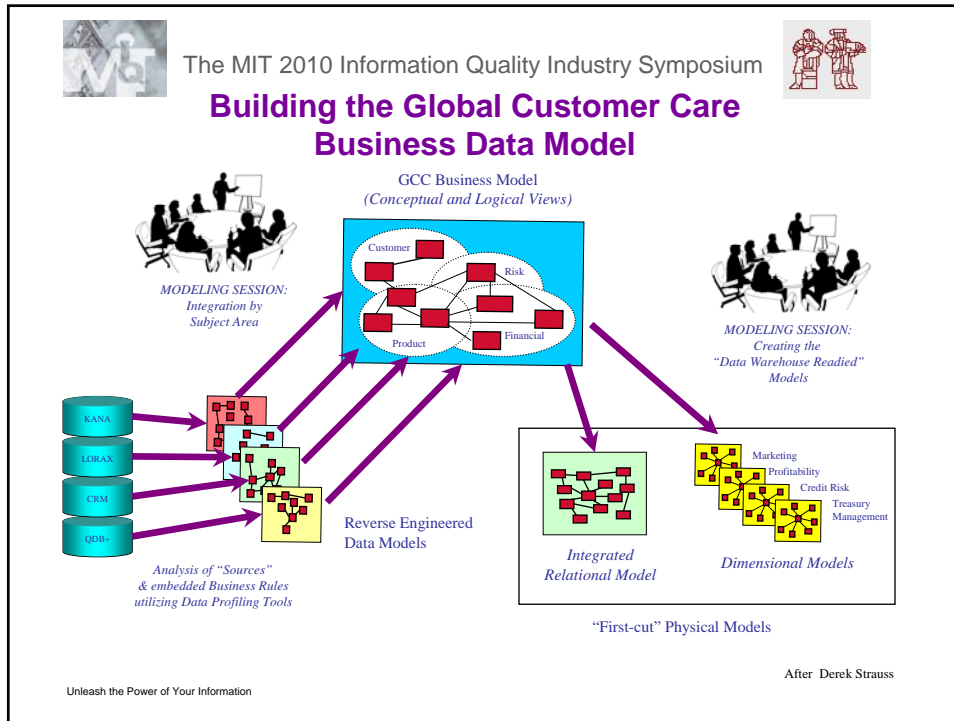
- Getting away from “code, load and explode” 
- The “data scouts” (team with business and IT representation) 
- On finding significant DQ problem, choose from strategies such as:
  - fix the data at the source (actually go into the data store and physically zap the data)
  - fix the program the source (apply the correct edits in order to validate the data)
  - fix the business process (a broken business process is very often the main cause of poor quality data)
  - recognize and resolve situations where data attributes are being used for a purpose other than their original intent (e.g. a gender code, which has more than two distinct values)
  - transform the data on the way into the data warehouse (this is the most common of strategies, but should not be the only strategy employed)\*

Unleash the Power of Your Information











The MIT 2010 Information Quality Industry Symposium



## It's all about the People

We share an infectious sense of mission to make an impact on society and empower consumers in ways never before possible. We are committed to serving both the Internet community and our own communities.

- Jerry

Unleash the Power of Your Information



The MIT 2010 Information Quality Industry Symposium

