Data Quality for Successful Data Governance

INFORMATICA

Ivan Chong
EVP and General Manager
Data Quality Business Unit
Informatica Corporation

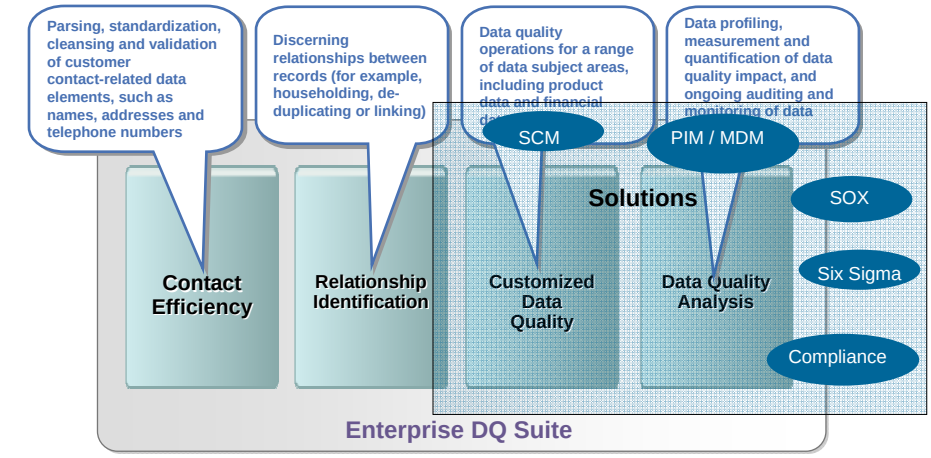
Informatica Confidential - 2008

1

Agenda

- **Introduction**
- **A Pragmatic Approach**
- **Common Pitfalls and Misconceptions**
- **Examples / Scenarios & Case Studies**
- **Q&A**

Introduction: Data Quality Evolution



Introduction: Impact of Poor Data Processes & Data Quality

Poor Data Quality or inconsistent data causes defects in the value chain and is a momentum killer for any Data Alignment, Synchronization and Collaboration project, Master Data Management, CDI, Application Implementation etc …

The MIT 2008 Information Quality Industry Symposium

Introduction: One Perspective to Governance, top down

Governing corporate data may be slightly easier than governing nations, but companies can benefit from adopting similar strategies, *according to a Gartner analyst*.

Executive level. sponsorship, strategic direction, funding, advocacy and oversight

Judicial level. planning activities and to enforce governance activities or corporate policies. Mediating disagreements

Legislative level. chaired by a senior business leader designated by the executive team and may include business and technology leaders from Finance, IT, Data Management and Operations

Administrative level. Implement data governance on a day-to-day basis; responsible for developing data models and corporate data vocabularies, implementing master data management best practices, organizing content

Top Down – often initiated by Exec initiative – Compliance / Audit / BPR / Six Sigma

The MIT 2008 Information Quality Industry Symposium

Introduction: Data Quality Perspective to Governance

Lack of a comprehensive **Data Quality Management Strategy** and poor **Data Quality** is causing *Master Data Management* ***(MDM)*** *initiatives to* ***fail*** *or be significantly delayed.*

key elements of effective Data Quality Management:

- capable data processes
- ownership and accountability
- standards and metrics
- measurement and control

key stages in any migration or MDM implem

- discovery and assessment
- planning and preparation
- cleansing, alignment, enrichment and enhanc ent
- operational impact and support in prepar tion for, during and after implementation
- **Transition management**
- post implementation measurement and control

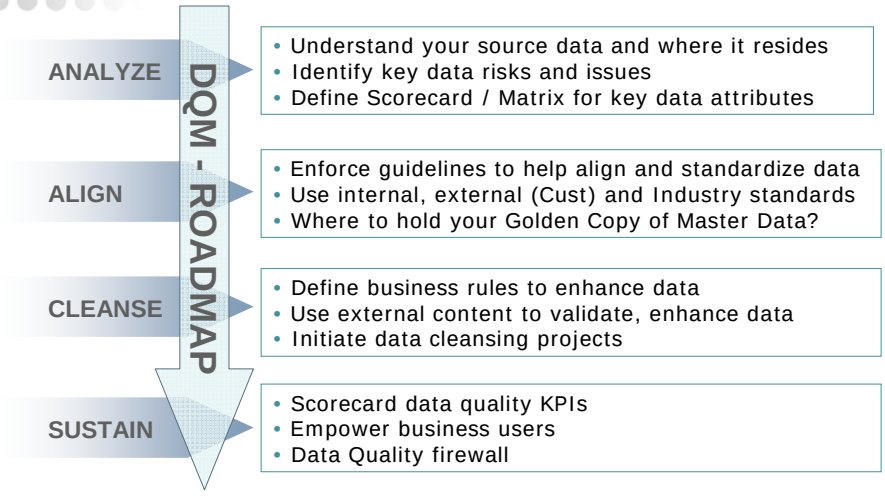
Operational Risk

Bottom Up – often initiated by project – DQ / Migration / App Implement / Six Sigma

Agenda

- Introduction
- A Pragmatic Approach
- Common Pitfalls and Misconceptions
- Examples / Scenarios & Case Studies
- Q&A

Data Quality Management – A Good Starting Point

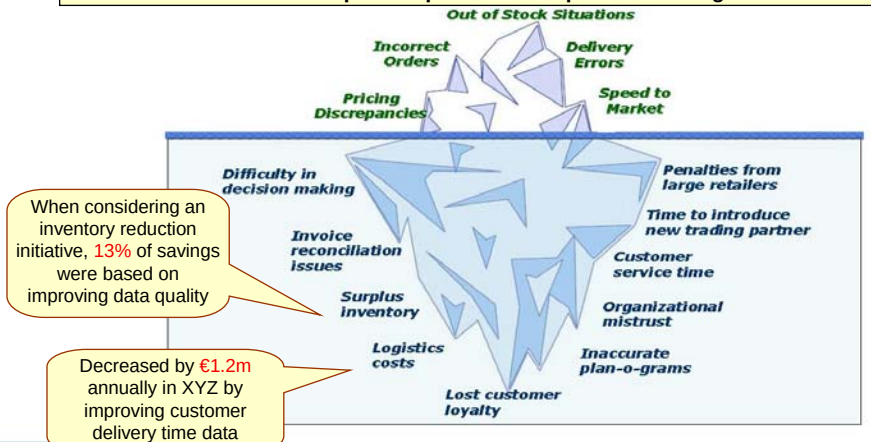


Agenda

- Introduction
- A Pragmatic Approach
- Common Pitfalls and Misconceptions
- Examples / Scenarios & Case Studies
- Q&A

Background: Focus on value early

The impacts of poor quality data are not always highly visible. Using Supply Chain as an example here, many impacts are hidden 'below the waterline'. The visible impacts represent the 'tip of the iceberg'



When considering an inventory reduction initiative, 13% of savings were based on improving data quality

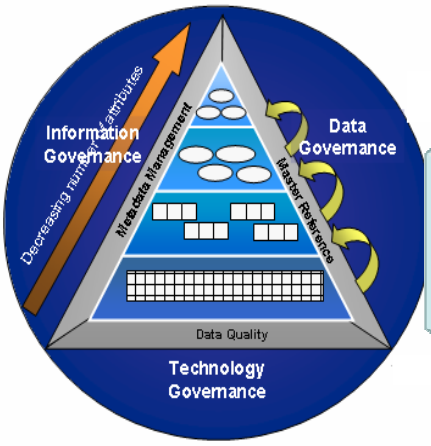
Decreased by €1.2m annually in XYZ by improving customer delivery time data

The MIT 2008 Information Quality Industry Symposium

Information Governance Framework

Be clear of what is in scope for 'DATA GOVERNANCE'

Information Governance
- Data Protection
- Licensed 3rd PTY Data
- Legislation

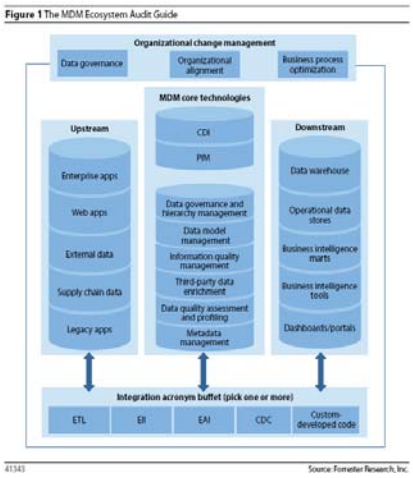


Data Governance
- Business Processes
- Data Quality
- Roles and Response'
- Business context

Technology Governance
- Data Security
- Anti-virus
- Disaster Recovery

The MIT 2008 Information Quality Industry Symposium

MDM – Generic concerns



MDM Applications / Solutions

•Decouple Master and Reference Data capture from Legacy

•All Master Data captured in one application process across the enterprise

•Operational Data Stores and Data Warehouse Risks

Implementation Risks

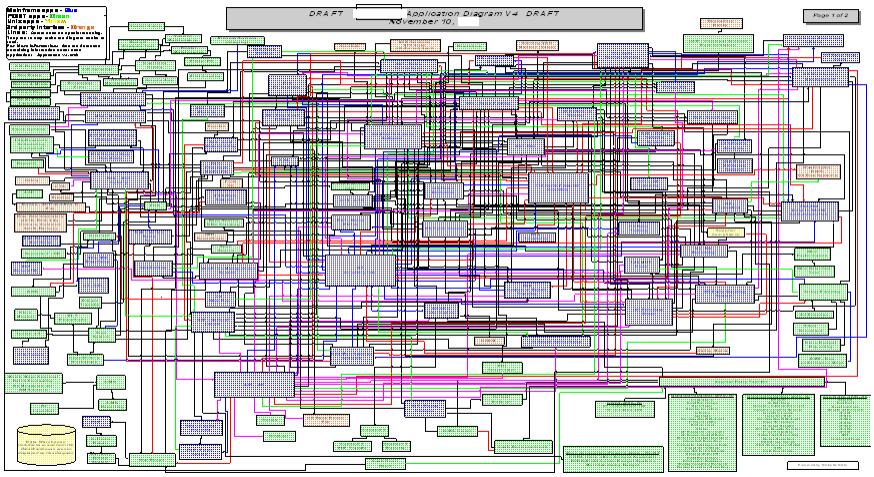
• Registry / Repository / Hub

• Focus on quick wins and value

• Operational Risk, Internal & External

The Challenge

Data Integration Complexity



The Challenge

Early Analysis is critical – Obsolescence / Duplicates

Data Alignment means focus on Data Quality

ECR Europe

- Results of worldwide Data Cleansing:

	Obsolete	Duplicate	Active	Cleansing rate
Material	724'804	29'529	544'383	58%
Customer	1'728'298	105'033	1'733'809	51%
Vendor	1'079'660	23'266	632'952	64%
Total	3'532'762	157'828	2'911'144	56%

- Cleansing includes completeness and correctness checks on active data records

One Customers Experience - Large US Bank

Lessons Learned

- Lessons Learned #1
  - Too much, too fast. Don't try to solve all the problems at once
  - Give the stakeholders enough to start with, but do not overload them
  - Only give the stakeholders the defects that apply to them & separate defects by type:
    - Data Entry defects → Front line stakeholders
    - Data Movement defects → Extract/Load Technology teams
    - Data Enrichment → Transformation Technology teams
- Lessons Learned #2
  - Create data quality business rules that align with Policies and Procedures
  - Don't create rules that "seem to make sense". Create rules that are "Actionable"

Agenda

- **Introduction**
- **A Pragmatic Approach**
- **Common Pitfalls and Misconceptions**
- **Examples / Scenarios & Case Studies**
- **Q&A**

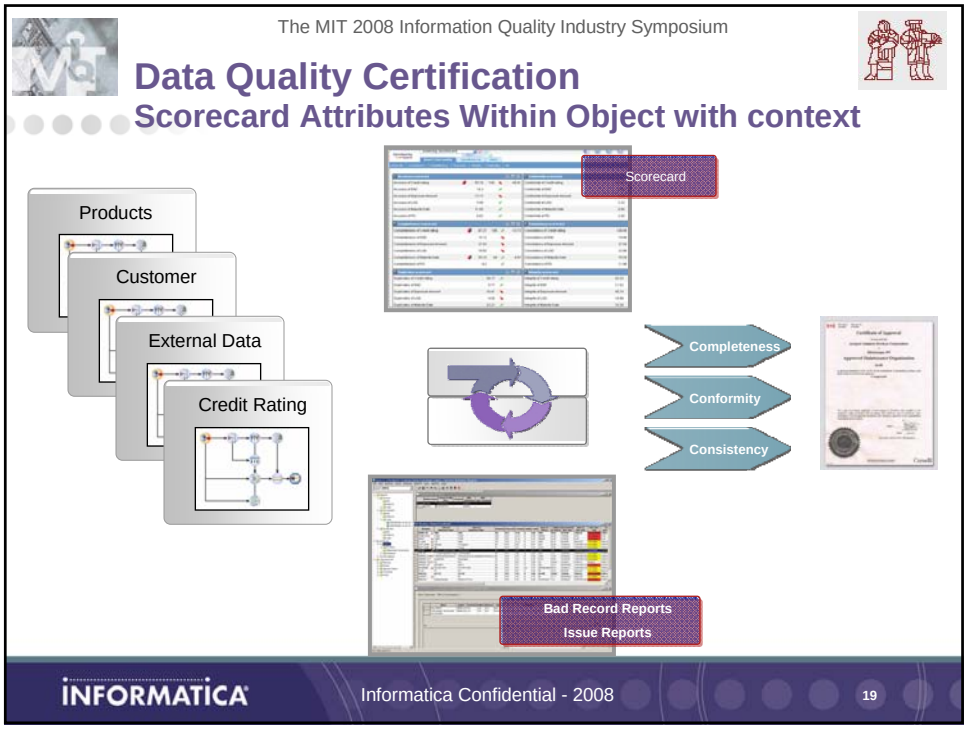
INFORMATICA

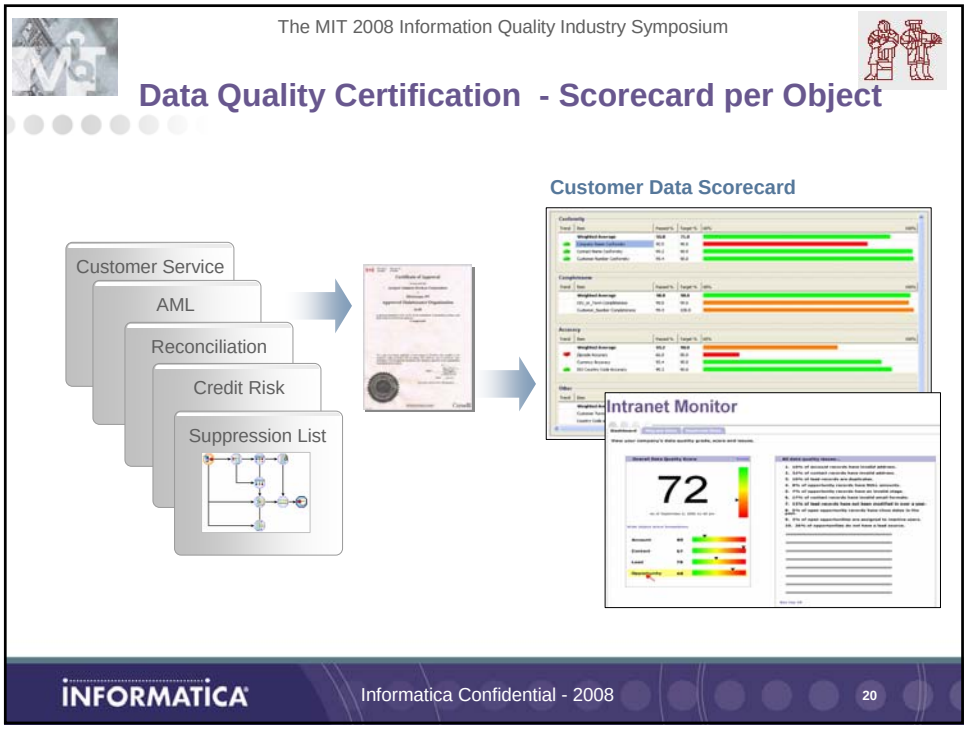
Metrics and Measurement

Informatica Confidential - 2008

Six Types of Data Quality Problems

Completeness	What data is missing or unusable?
Conformity	What data is stored in a non-standard format?
Consistency	What data values give conflicting information?
Accuracy	What data is incorrect or out of date?
Duplicates	What data records or attributes are repeated?
Integrity	What data is missing or not referenced?


The MIT 2008 Information Quality Industry Symposium
Data Quality Certification
Scorecard Attributes Within Object with context
Products
Customer
External Data
Credit Rating
Scorecard
Completeness
Conformity
Consistency
Bad Record Reports
Issue Reports
INFORMATICA
Informatica Confidential - 2008
19


The MIT 2008 Information Quality Industry Symposium
Data Quality Certification - Scorecard per Object
Customer Data Scorecard
Customer Service
AML
Reconciliation
Credit Risk
Suppression List
Intranet Monitor
72
INFORMATICA
Informatica Confidential - 2008
20

Example – CPG - What is Data Quality?

Completeness	Required values electronically recorded
Standards Based	Data conforms to industry standards
Consistency	Data values aligned across systems
Accuracy	Data values are right, at the right time
Time Stamped	Validity timeframe of data is clear

Data Quality

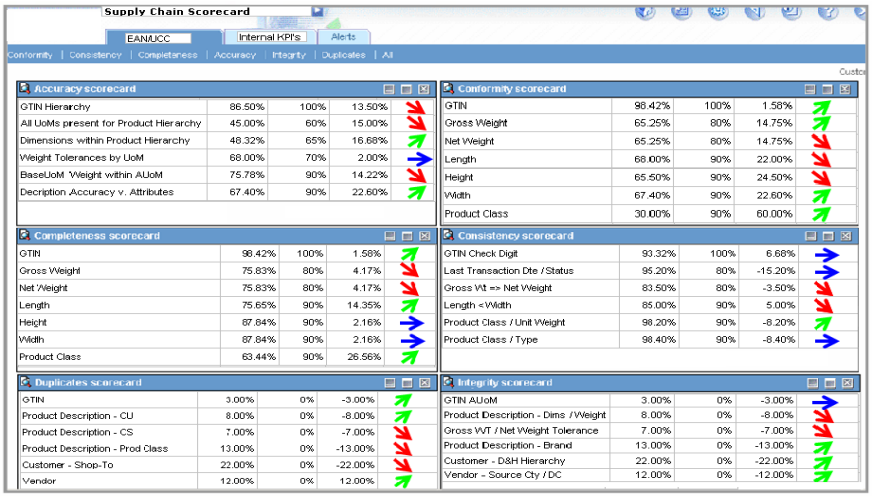
Customer Example: AML

CIP/KYC- Reporting Definitions

- **Completeness -** A measure of the degree to which the requested information field contains data and is non-blank.
- **Logical -** A measure of the degree to which the requested information field contains data that has been determined to be sound and reasonable within parameters defined by the responsible AML Process team.
- **Verification –** A measure of the degree to which the requested information field contains data that agrees with an appropriate authoritative source as determined by the responsible AML Process Team.
- **Accuracy –** A measure of the degree to which the requested information field contains data that agrees with the source or is consistent with fact or reality.

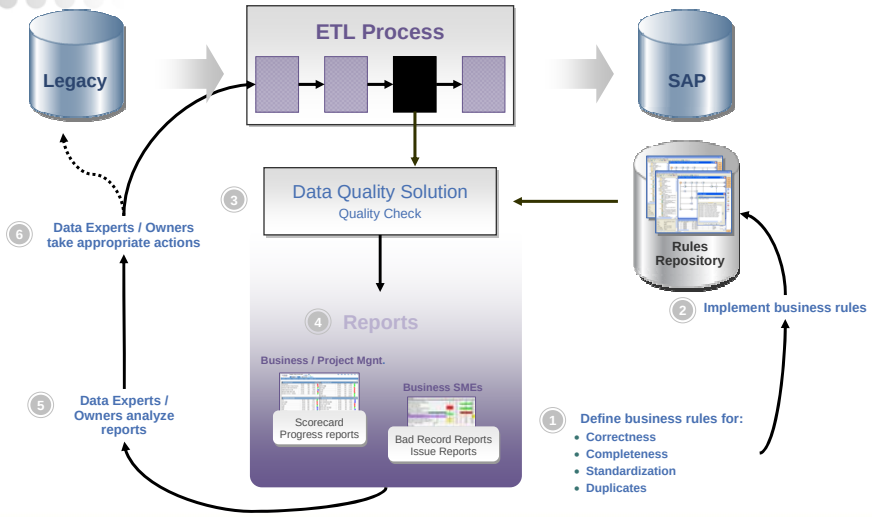
The MIT 2008 Information Quality Industry Symposium

Scorecard: Using 'Business' terms

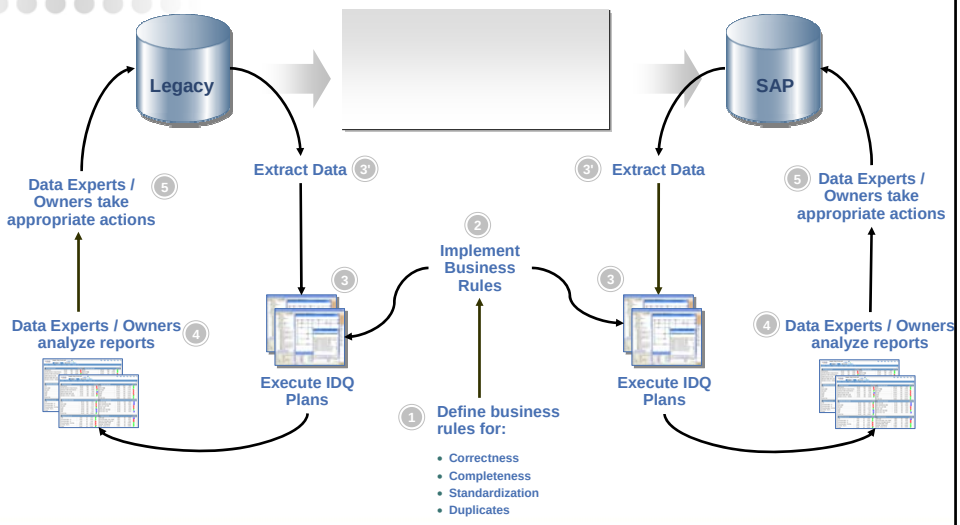


The MIT 2008 Information Quality Industry Symposium

Global CPG Organization: Quality Check in ETL



Global CPG Organization: Quality Check in ETL



DQ must deliver value

Data Quality Framework Business Value Drivers

- Improve time to market
- Support business vision in GCI area
- Increase customer/consumer trust
- Support organizational flexibility
- Enable leverage from decision support as competitive advantage
- Strive for flawless execution
- Enable leverage of globe benefits
- Reduce cost and rework through efficiency

DQ underpins Data Governance

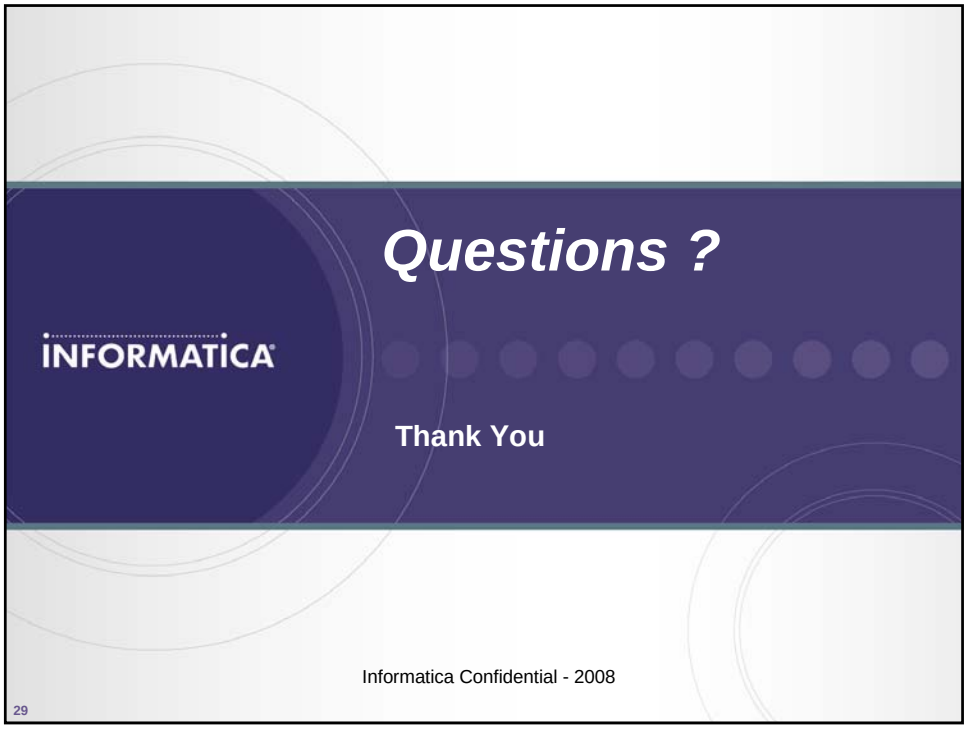
What we are doing at A.N.Other CPG Org.

Achieving Data Governance thru the monitoring of the critical business rules.

- Defining business rules to support data quality per business object (Material, Customer, Vendor, Employee, Banks etc.)
- Defining and implementing Global Data Quality and Data Management KPIs and exception reports based on the pre-Defined business rules.
- Providing Global visibility of Data Quality to the whole organisation at all levels:
  - Publishing the Data KPIs monthly on the operation site on the intranet accessible by A.N.Other organisation at all levels.
  - Presenting the Data KPIs at the Operational Steering Committee (Top Management Level) and at the market management level each month.
- Monitoring of the KPIs and the exception reports by the markets to take corrective actions.

Agenda

- **Introduction**
- **A Pragmatic Approach**
- **Common Pitfalls and Misconceptions**
- **Examples / Scenarios & Case Studies**
- **Q&A**


Questions ?
INFORMATICA
Thank You
Informatica Confidential - 2008
29