Data Quality, An Enterprise Principle and Practice

Chris Michel

July 2007
Objectives of this presentation

- Contribute to the research on Information Quality (IQ), providing a working example of Enterprise Data Quality
- IQ Framework Methodology and Process
- IQ Roles and Responsibilities
- IQ Quality Scorecard
- Discussion
Data Quality Framework

1. Data Standardization
   - Define enterprise data standards

2. Data Monitoring
   - Measure application conformance against enterprise data standards

3. Data Improvement
   - Identify business improvement opportunities related to fixing data issues

- Key Data Subjects (8)
- Key Data Attributes for each Data Subject (300+)
- Key Business Rules for each Data Attribute (1000+)
- Key enterprise applications (8)
- Audit business rules in RoR applications
- Audit RoR compliance in Non-RoR applications
- Significant project recommendations
- Minor grooms/fixes
Data Quality Metrics Methodology

Metric Components:

Business Rules
- Is the value null or blank?
- Does the value match a list of specified valid values?
- Does the value conform to specific NCR business rules?
- Does the value cause a referential integrity concern?
- Does the value conform to postal standards? (address elements only)

RoR Conformance
- Does the value in the application match the value in the RoR

Metric Calculations:

Component Level
- Number of Rows Passing Check / Number of Rows Checked

Aggregate Level
- Sum(Number of Rows Passing All Checks) / Sum(Number of All Rows Checked)
Data Quality System/Process Flow

1. Develop Audit Scripts
2. Audit Compliance Against Standards
3. Generate Scorecards and Detail Reports
4. Root Cause Analysis and Recommendations
5. Fix Problems
6. Annual IT Plan

**Oracle/Progress/SQL Server**

**Teradata**

**EDW**

**Enterprise Applications**

**Enterprise Data**

**Audit Results**

**Rules**

---

The MIT Information Quality Industry Symposium, 2007
Data Quality Roles & Responsibilities

**Data Subject Experts (Business and/or ITS)**
- Provides business subject matter expertise to define data standards, and to determine the root cause of data issues

**Development Teams**
- Determines key attributes & business rules
- Identifies SMEs to engage
- Scopes data improvement opportunities into appropriate annual plan projects

**IT Partner**
- Determines key attributes & business rules
- Identifies SMEs to engage
- Scopes data improvement opportunities into appropriate annual plan projects

**Global Process Teams**
- Approves and prioritizes projects

**Enterprise Data Management (EDM)**
- Facilitates definition of data standards
- Develops and maintains the repository of data standards
- Monitors and measures application compliance with data standards
- Monitors and measures application compliance with data standards
- Identifies business improvement opportunities related to fixing data issues

**Data Quality Roles & Responsibilities**
# Data Quality Scorecard

<table>
<thead>
<tr>
<th>DATA SUBJECT</th>
<th>CSMS</th>
<th>EHRMS</th>
<th>ERP</th>
<th>GSDB</th>
<th>X.500</th>
<th>NOL</th>
<th>PSA</th>
<th>SPOT</th>
<th>TEAM WEB</th>
<th>WOT</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>CUSTOMER</td>
<td>94.26 %</td>
<td></td>
<td>98.59 %</td>
<td></td>
<td></td>
<td>81.51 %</td>
<td></td>
<td>99.71 %</td>
<td>99.94 %</td>
<td></td>
<td>97.37 %</td>
</tr>
<tr>
<td>TEMP CUSTOMER</td>
<td>7.15 %</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7.15 %</td>
</tr>
<tr>
<td>EMPLOYEE</td>
<td>73.67 %</td>
<td>99.86 %</td>
<td>98.42 %</td>
<td>99.36 %</td>
<td>99.54 %</td>
<td>94.00 %</td>
<td></td>
<td>99.96 %</td>
<td>99.92 %</td>
<td></td>
<td>99.20 %</td>
</tr>
<tr>
<td>NON EMPLOYEE</td>
<td>93.73 %</td>
<td>99.48 %</td>
<td>69.37 %</td>
<td></td>
<td>82.92 %</td>
<td>99.97 %</td>
<td></td>
<td>99.95 %</td>
<td>99.95 %</td>
<td></td>
<td>86.53 %</td>
</tr>
<tr>
<td>PRODUCT</td>
<td>98.74 %</td>
<td>98.55 %</td>
<td>99.97 %</td>
<td>99.28 %</td>
<td></td>
<td>99.96 %</td>
<td></td>
<td>99.90 %</td>
<td></td>
<td></td>
<td>99.83 %</td>
</tr>
<tr>
<td>OPPORTUNITY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>99.93 %</td>
<td>99.40 %</td>
<td></td>
<td></td>
<td>99.93 %</td>
</tr>
<tr>
<td>QUOTE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>99.88 %</td>
<td></td>
<td></td>
<td>99.88 %</td>
</tr>
<tr>
<td>ORDER</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>98.90 %</td>
<td></td>
<td>98.90 %</td>
</tr>
<tr>
<td>INVOICE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>99.48 %</td>
<td></td>
</tr>
<tr>
<td>INSTALLED BASE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>71.93 %</td>
<td></td>
</tr>
<tr>
<td>CONTRACT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>99.15 %</td>
</tr>
<tr>
<td>TOTAL</td>
<td>95.22 %</td>
<td>99.43 %</td>
<td>99.20 %</td>
<td>99.97 %</td>
<td>90.90 %</td>
<td>99.28 %</td>
<td>89.46 %</td>
<td>99.96 %</td>
<td>99.89 %</td>
<td>99.89 %</td>
<td>99.30 %</td>
</tr>
</tbody>
</table>