MDM
Challenges and Solutions from the Real World

Toronto MDM Summit 2008

info@adastracorp.com
www.adastracorp.com
Agenda

- What is Master Data Management?
  - Definitions by example

- How can I learn from the experience of others?
  - Industry specific challenges & solutions
Definitions by Example

- **Commercial Furniture Builder**
  - International company with large number of employees
  - Successful & expanding into new markets with healthy growth
  - Over 100 different product families
  - Manufacturing facilities across USA, Europe, and China
  - Competes against larger firms
  - Needs to shorten sales & quotation cycle
  - Needs to shorten order fulfillment cycle
  - Large number of solutions are made-to-order
  - Highly customized products
  - Stringent delivery and installation deadlines

- These are all typical challenges faced by competitors and industry as a whole
Definitions by Example – Business Units

- **Assessment Findings:**
  - Information stored across different data systems
  - Information stored in hardcopy documents
  - Information systems are not integrated, poor data quality
  - Data management workflows are unclear and not standardized

- **Considerations:**
  - Investigate the stated findings and their impact
  - Consider the common source to the challenges presented
  - Consider how any solution must address these challenges
  - Consider what are the challenges beyond data integration
Definitions by Example – IT/IS, Sales

Information Systems Challenges:
1. **Data leakage** – incomplete Bill of Materials, drawings, errors
2. **Data duplication** – suppliers and products have multiple records across product families
3. **Business rules** – conflicting business rules for specification across product families, maintained by different groups in silos
4. **Undocumented knowledge** – domain experts are swamped with work and unable to communicate with other teams and train new hires

Sales Challenges:
1. **Long quotation cycle** – increased number of lost sale opportunities
2. **Long order fulfilment cycle** – client frustration, errors and omissions
3. **Difficult to estimate pricing** – large margin of error that must be absorbed
4. **Difficult to calculate profits** – which numbers to trust, manufacturing vs. sales, vs. purchasing, neither correct
Purchasing Challenges:
1. **Multiple records for the same parts & materials** – information entered multiple times and inconsistent
2. **Vendor management** – difficult to negotiate discounts for batch orders
3. **Shipment delays** – multiple orders for same materials across single and/or multiple vendors, wrong parts delivered to installers
4. **Currency valuation** – unable to respond in timely fashion to currency fluctuations

Manufacturing Challenges:
1. **Multiple verifications** – Bill of Materials must be manually verified and completed
2. **Resource scheduling** – assembly line setup planning and optimization conflicts
3. **Shipment delays** – materials and loose parts required in product assembly
4. **Inventory** – difficult to reuse cut-offs, maintain minimal overhead
Definitions by Example → MDM

- **Master Data identified** – Materials, Products/Parts, Drawings, Vendors, and Clients

- Has impact across all major business areas such as Management, Sales, Purchasing, IT, and Manufacturing

- **Master Data** – a core set of data critical to major business processes and functions

- **Master Data Management** – organizational structures, business processes, culture and technical tools ensuring key (master) data in the enterprise is:
  - **Reliable and Correct** → reliable and stable data sources, managed and provided by reliable and stable systems
  - **Unified** → in content and understanding
  - **Available** → at the right place at the right time
Definitions by Example → The Approach

- Master Catalogue concept: High Visibility, High Quality, Highly Available!

- Beyond the basic information management practices there exist opportunities to lay a solid MDM foundation

- Establish Master Data Governance
- Appoint Domain Experts, Data Stewards
- Balance accountability with empowerment across user groups
- Implement incremental phased approach
European Bank
The Situation

- Large European financial group, offices in many locations, support for personal & commercial clients
- Inefficient sales support, due to:
  - Fragmented data – lack of holistic customer view
  - Application silos across the organization and business units
  - Legacy approach where products drive customer data
  - Business units not talking to each other
  - No single customer view, multiple records exist for different products
- Ineffective contact management, multiple or non-applicable offers mailed to clients
- Low ROI due to increased contact and campaign costs
- Missing campaign management, difficulty of tracking customer reach and conversion goals

- The organization decided to move from product-centric to customer-centric business model
- MDM tools and practices were selected as the methodology to implement the new business model
The Challenges

- **1. Business Involvement**
  - How to include all business stakeholders in the implementation and execution?

- **2. Efficiency of the data quality process**
  - How to achieve high level of data quality improvements?
  - How to apply user assisted data quality improvements quickly?

- **3. Level of primary systems integration**
  - How much to modify primary systems – user interface, master data storage and APIs?
The Approach

- **Architecture:**
  - CDI database designed, implemented for unified customer data, SOA-based master data interface
  - Primary systems gave up their master data storage to common data layer
  - Master data migrated into the CDI database one primary system at a time
  - Original client data management interfaces decommissioned

- **Business Process:**
  - New DQ business process implemented, supported by new DQ workflow
  - Off-line DQ improvement based on documented business rules
  - Multiple business user groups – domain experts, call center, customer relationship managers – made responsible for DQ
  - DQ review and correction tasks assigned to business users
  - KPIs and compensation incentives put in place for timely DQ review and correction such as:
    - Number of records managed daily
    - Number of business users involved
    - Average response time
The Solution

Shared business logic
Data Integration Interface (IBM MQ series)
CDI database
Power user interface
End user interface
Private primary systems interfaces
Private primary systems business logic

Unified front end

Data layer
CDI database
Data Warehouse
Primary systems (without client data management functionality)

Shared code tables
End users

Power user interface
End user interface

End users

Private primary systems interfaces

Private primary systems business logic

Data Quality

f(x1...xn)

f(x1...xn)
The Benefits

- Business processes transformed to customer-centric
- Enabled Sales and Marketing with better cross / up sell opportunities
- Enabling targeted client and event driven campaigns, either for specific clients or specific types of clients
- Streamlined client facing operations
- Utilized integrated and managed customer data platform with unified customer data interface
- Improved data quality and data issue resolution workflow
- Significantly increased business involvement
Insurance Company
The Situation

- Large Central European insurance company
- Independent Life, Property & Casualty lines of business
- Changing the organization from product-centric into customer-centric
- Data Quality is a manual process within each business unit
- Current master client database not providing the expected results
  - No (minimal) Data Quality
  - No DQ escalation paths
  - Incorrect data management processes
The Challenges

- Establish specific data quality level targets
  - How to define realistic DQ level targets and expectations? 80%? 90%?

- Cost estimation
  - How to realistically estimate the effort, cost and timelines to achieve the defined objectives without current state knowledge?
The Approach

- Conducted initial assessment – determine current state for each business unit
- Designed modular (extensible) architecture
- Designed and implemented decoupled systems
- Designed and deployed unified integration interfaces for data services
- Developed multiple design implementation scenarios based on cost, risk, architecture, future enhancements
- Proposed phased approach, gradual automation of manual tasks
The Solution: Data Cleansing

Corrected data is propagated to CDI database

Data Quality Profiling & Cleansing

Data Quality Unification Process & Master Record selection

Combined approach where data corrections are both manually & automatically propagated to primary systems

In CDI system non-unified data is manually reconciled

Clients are prioritized by importance / business value / profitability

CDI

Primary Systems

Candidate Group

Partner Group

Representative Record

Call Center

Contact center, Back Office

Screening

Request

Quality Manager

Data Steward

Existing processes

Automatic processes planned

Reports
The Solution: Data Quality Metrics

Data Quality Metrics

Correct (M1)  Complete (M2)  Consistent (M3)  Valid (M4)  Data Quality

- Correct (M1): 78.20  85  91.24  83.47  90
- Complete (M2): 83.47  90  90.15  91.75  85
- Consistent (M3): 70.36  85  80  91.66  85
- Valid (M4): 45.90  80  69.39  85.50  91.20

Data Source Quality
Acceptable Quality
Final Quality
The Benefits

- The implemented solution provides a standard and documented access to customer data
- Integrated interface for customer profile and reporting
- Implements common corporate strategy and direction for future workflow and architecture enhancements
- Decreased back office costs through combined automated and manual data quality workflow
- Meets business expectations and provides KPIs for quantitative data quality assessment, resolution, and change detection over time
Legacy Systems
HealthCare Organization
The Situation

- There is an existing person-matching process:
  - Running within an OLTP system
  - Using legacy DOS-based matching tool

- Project objectives:
  - Decommission the OLTP system
  - Move its functionality, including person matching, to the EDW data integration layer
  - Decommission the legacy matching tool and introduce a new modern matching tool

- Expected Business Benefits
  - Person-centric reporting and analysis across all enterprise data holdings
The Challenges

- The business users have used the current system for over 20 years and trust its person matching outputs.
- We needed to prove that the new person matching system is better than the previous.
- The results comparison is not a trivial task.
- There are differences in nearly every aspect of the person unification process.
- Yet the results have to be comparable.

How do we compare quality of the legacy and the new matching processes?
The Approach

- Implement comparison plan, document platforms, assumptions, and success criteria

- **Compare the input data:**
  - Profile each data set to be matched by the legacy and the new tools
  - Identify data type, pattern, frequency distribution and collect DQ metrics such as Accuracy, Integrity, Completeness, Validity, Distribution
  - Compare the profiles and identify technical and business reasons for all differences

- **Compare the tools:**
  - Identify a data set representative of the target system data
  - Run both tools and methods independently on the same data sets and compare the matching results
The Findings

- The existence of a high quality legacy system tuned for many years defines a challenging benchmark to be reached from the first increment.
- Existing knowledge and expertise in matching methodologies were leveraged in the implementation of the new data quality tool.
- Existing metrics and KPIs allowed for setting realistic and measurable targets.
- Two major and several minor issues with the data in the EDW were identified and addressed.
- Technical bugs and business processes were fixed to improve the data quality.
- The quality of some data sets was improved by over 40%.
Technology Selection
The Situation

- Multitude of DQ/PIM/CDI software vendors
  Oracle/Siebel & SAP, SAS, IBM, Microsoft, Informatica, Purisma, Siperian, DataFlux, Trillium, Ab Initio, Ataccama, Orchestra Networks, GSX, Riversand, i2 etc…

- Difficult to compare solutions based on different feature sets, architecture and vertical specialization

- No single MDM market exists, market fragmentation and consolidation are creating volatility
The Challenges

- How to select the right MDM tool / technology? Fitting the organisation in terms of:
  - Architecture
  - ROI Horizon
  - Feature Set
  - Usability
The Approach

- Conduct feasibility study / technology selection to evaluate software tools utilizing:
  - Broad set of technology assessment criteria on architecture, features, vendor maturity and cost
  - Downsized but realistic scenarios, pay attention to detail (logging, monitoring, operations, performance, scalability, transactions…)
  - Conduct POCs addressed for specific issues
  - Engage industry experts to leverage existing experience
The Solution

Technology assessment & selection matrix:

<table>
<thead>
<tr>
<th>Features</th>
<th>3.2.1.01</th>
<th>7.5.1</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Version used for basis of Comparison:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Features</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Data Profiling features</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Core features</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metadata Validation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Domain Analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range Analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rule Validation</td>
<td>✔️</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pattern Analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Descriptive Statistics</td>
<td></td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>Frequency Counts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relationship Discovery</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Visualization</td>
<td></td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>Data Validation Rules</td>
<td></td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>Generation of Rules for Cleansing</td>
<td></td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>Trend analysis</td>
<td></td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td><strong>Software additional capabilities</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error Handling, Logging and Tracing capabilities</td>
<td>✔️</td>
<td></td>
<td>✔️</td>
</tr>
<tr>
<td>Format conversion</td>
<td></td>
<td></td>
<td>✔️</td>
</tr>
<tr>
<td>Run in batch mode, Aintegration with Job Schedulers</td>
<td>✔️</td>
<td></td>
<td>✔️</td>
</tr>
<tr>
<td>Job Monitoring</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transformation and Data Reduction Functions</td>
<td></td>
<td></td>
<td>✔️</td>
</tr>
<tr>
<td>Data Generation Capabilities</td>
<td></td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>Ability to Issue SQL Against Non-relational Sources</td>
<td>✔️</td>
<td></td>
<td>✔️</td>
</tr>
<tr>
<td><strong>Usability</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Report Customization</td>
<td>✔️</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
MDM Process

- Formal & transparent methodology:

### Master Data Management Methodology

<table>
<thead>
<tr>
<th>Environment Definition Track</th>
<th>Program Management</th>
<th>Data Governance Track</th>
</tr>
</thead>
<tbody>
<tr>
<td>ED-01: Align to corporate strategy</td>
<td>ED-02: Identify MDM Program owner</td>
<td>ED-03: Identify MDM Program mission / vision</td>
</tr>
<tr>
<td>ED-04: Identify key users / stakeholders</td>
<td>ED-05: Collect key users’ requirements</td>
<td>ED-06: Identify master data and its usage</td>
</tr>
<tr>
<td>ED-07: Define solution concepts (architecture)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
MDM Process

- Adastra has developed a proven methodology comprised of the following subject areas:
  - **Program Management**
    - Environment Definition Track, Data Governance Track, Environment Management Definition Track
    - Definition of: MDM Program objectives, Outputs, Architecture, Management processes / policies, should be performed at the beginning of the MDM Program implementation
  - **Solution Development**
    - Business Track, IT/IS Track
    - Implementation of full Software Development Life Cycle
    - Separation of Business Requirements and Technical Requirements
  - **Operations & Maintenance**
    - Operations Track
    - Measurable value and Metrics for Data Governance, Data Quality, Business Rules validation
    - End-user support framework for education & training, issue escalation, knowledge base & help desk management
  - **Migration**
    - Migration Track
    - Implementation of Scope, Impact, Timelines and Resource plans
<table>
<thead>
<tr>
<th>CANADA</th>
<th>CZECH REPUBLIC</th>
<th>SLOVAKIA</th>
<th>GERMANY</th>
<th>BULGARIA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adastra Corporation</td>
<td>Adastra, s.r.o.</td>
<td>Adastra, s.r.o.</td>
<td>Adastra GmbH</td>
<td>Adastra Bulgaria EOOD</td>
</tr>
<tr>
<td>Le Parc Office Tower</td>
<td>Nile House</td>
<td>Francisciho 4</td>
<td>Bockenheimer Landstrasse 17/19</td>
<td>29 Panayot Volov str.,</td>
</tr>
<tr>
<td>8500 Leslie St.</td>
<td>Karolinská 654/2</td>
<td>Bratislava</td>
<td>Frankfurt a. Main</td>
<td>5th floor</td>
</tr>
<tr>
<td>Markham</td>
<td>Praha</td>
<td>Slovakia</td>
<td>Germany</td>
<td>Sofia</td>
</tr>
<tr>
<td>Ontario, L3T 7M8 Canada</td>
<td>Czech Republic</td>
<td><a href="mailto:info@adastracorp.sk">info@adastracorp.sk</a></td>
<td><a href="mailto:info@adastracorp.de">info@adastracorp.de</a></td>
<td>Bulgaria</td>
</tr>
<tr>
<td><a href="mailto:info@adastracorp.com">info@adastracorp.com</a></td>
<td><a href="mailto:info@adastra.cz">info@adastra.cz</a></td>
<td><a href="http://www.adastracorp.sk">www.adastracorp.sk</a></td>
<td><a href="http://www.info@adastracorp.de">www.info@adastracorp.de</a></td>
<td><a href="mailto:info@adastracorp.com">info@adastracorp.com</a></td>
</tr>
</tbody>
</table>