Data Quality – Decision Quality

Dr. Frank Block, CEO, FinScore AG
Swiss Statistics Meeting
17.11.-19.11.2004, Aarau
Agenda

- The effect of bad Information Quality (IQ) on decisions (examples)
- Where does bad data come from?
- A framework for managing IQ
- Conclusion
We’re in the information age!

- Data: telco, banking, science, web...
- Data bases
- Data warehouses
- Data marts
- Operational data stores
- ....

→ Information is the raw material of our time!

But…
Managing Information is not easy – some examples

- US Elections 2000 - "Lazy matching" identified too many voters as criminals (not allowed to vote)
- Mars Orbiter - Loss of Mars Climate Orbiter in 1999 due to a conversion error from metric system cost $125 millions
- Bank - found it has $x bn credit exposure in the animal care sector. The data migration project did not consider for this type of use of the data.
- Another Bank – Published earnings were withdrawn and corrected by ~CHF 200 millions
Managing Information is not easy – some examples

- Yet Another Bank
  - Some clients born in the year 945, and still single...
  - Yet others died before they were born
  - Unexpectedly many clients born on 1.1. or 31.12.
  - Some have a profession ")(#$*ksd."

- Large retailer
  - Club members with negative bonus points and negative purchase volume

- Swiss Statistics
  - Publication of wrong inflation rates (3-4 years ago)
Are we in the Information Age?

- If this is the information age...it must be its pre-industrial phase.
- Information Technology is still in its infancy, still decades behind car manufacturing, pharma, etc.
- Requirements always getting tougher: CERN-LHC will process data equivalent to 20 times the telephonic traffic of the world, 12-20 PetaBytes/year.
- Why is it so difficult to produce good information?
The Importance of IQ for Decision Support

- „Garbage in – Garbage out“ or „B___ s____ in – B____ s____ out“
- TDWI
  - Bad IQ costs US companies alone $600+ bn
  - Many major companies are making important decisions routinely on remarkably inaccurate data…These [IQ] problems cause wasted labor and lost productivity that directly affect profitability…
  - 25% of data companies use is of bad quality
  - Projects fail due to underestimation of IQ
The origins of bad IQ

- Wrong manual data capture
- Migration of systems, mergers and acquisitions of companies
- Multiple, independent input and storage of the same attributes
- Erroneous interpretation and aggregation of data
- Lack of standards, business definitions, metadata, etc.
By enhancing IQ companies get more competitive

- Increase customer satisfaction, reduce attrition
- Enhance process throughput and performance
- Enhance quality of strategic and tactical decisions
- Reduce number of lost market opportunities
- Reduce risk of project failure
- Enhanced degree of compliance with regulations (Basel II, AML, Data Privacy, Sarbanes Oxley…)
- Enhance brand image
Driving in the dark

- **Language**
  - How many **customers** do we have?
    - #customer id's? Active customers?
  - Who are our **best** customers?
    - Buy expensive products, are profitable, have most potential?

- **Processes**
  - Documentation, owner, roles, responsibilities, automation, media breaks, quality controls

- **Systems**
  - Unclear specification, metadata incomplete/inaccessible
  - Data models too complex, data not “fit for purpose”
  - No IQ measurement ("Our data is bad", When would it be good? Value of good IQ?)
“Can I trust that data?”

- Absence of IQ indicators, automation, workflow
- Lack of business rules and object definitions
- Reporting not standardized
- Too many reporting, analysis tools
- Lack of business oriented data
- Difficult to integrate, consolidate, validate data
- Incomplete data
- Instability of operation
- Lack of quality controls
- Data too complex to use
- Transaction oriented data
- “Never touch a running system”
Getting out of the dark

- IQ Project Framework
- Apply an(y) IQ model
- Integrated assessment of information and process
  - Two sides of the same coin
  - The “Top-down Bottom-up” approach
- Model cost impact of bad IQ
  - Define + prioritize IQ actions
- Define and measure IQ indicators - “What get's measured get's done!”
IQ Project Framework

Assessment and IQ Business Case
- Identify processes and priorities
- Assess quality of information specification + content
- IQ Business Case
- Identify quick hit initiatives

IQ Quick Hits
- Enhance performance of key processes
- Measure impact of IQ
- Evaluate IQ tools
- Develop recommendations for “Managing IQ” phase

Managing IQ
- Specify IQ management system
- Specify and put in place IQ organization (processes, functions, people)
- Put IQ tools and system in production

Actions

Benefits
- Rapid understanding of IQ issues and organizational impact
- Guidance, priorities, and roadmap for enhancing IQ
- Early ROI
- Internal buy-in
- Productive IQ prototype
- Automated IQ management
- Central IQ knowledge base
- Enhanced business process performance
### Example of an IQ Model – PSP/IQ

By R. Wang, MIT

<table>
<thead>
<tr>
<th>Product Quality</th>
<th>Conforms to Specifications</th>
<th>Meets or Exceeds Consumer Expectations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sound Information</td>
<td>Useful Information</td>
</tr>
<tr>
<td></td>
<td>- Free-of-Error</td>
<td>- Appropriate Amount</td>
</tr>
<tr>
<td></td>
<td>- Concise Representation</td>
<td>- Relevancy</td>
</tr>
<tr>
<td></td>
<td>- Completeness</td>
<td>- Understandability</td>
</tr>
<tr>
<td></td>
<td>- Consistent Representation</td>
<td>- Interpretability</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Objectivity</td>
</tr>
<tr>
<td>Service Quality</td>
<td>Dependable Information</td>
<td>Usable Information</td>
</tr>
<tr>
<td></td>
<td>- Timeliness</td>
<td>- Believability</td>
</tr>
<tr>
<td></td>
<td>- Security</td>
<td>- Accessibility</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Ease of Manipulation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Reputation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Value-Added</td>
</tr>
</tbody>
</table>
Link Process and Information to maximize impact of IQ Initiatives

- **Top-down**
  - Business driven
  - Process oriented
  - Prioritization, strategy
  - Economic impact
  - Identify applications
  - Identify data systems
  - Important data domains

- **Bottom-up**
  - Measure quality of data model and content
    - Completeness
    - Flexibility
    - Robustness
    - Timeliness
    - Complexity
    - Consistency
  - Measure quality of data services
    - Timeliness
    - Security
    - Credibility
    - Interpretability
    - Accessibility
Bottom-up: Detailed IQ Analysis

- **Count** over **Time**
- **Birthday Distribution**

Graph 1: Count vs. Time
Graph 2: Birthday Distribution (Year vs. Number of Samples)
Bottom-up: IQ Visualisation and diagnostics

Duplicate delivery in January 2003

Quarterly data related to clients from segment K1

Quarterly data related to clients from segment K2

1 = May 2001, ..., 36 = April 2004
Model Cost Impact of bad IQ

- **Hard Costs**
  - Customer attrition
  - Error detection
  - Error rework
  - Error prevention
  - Customer service
  - Fixing customer problems
  - Delays in processing
  - Delayed or cancelled projects

- **Soft Costs**
  - Difficulty in decision making
  - Time delays in operation
  - Organizational mistrust
  - Lowered ability to effectively compete
  - Data ownership conflicts
  - Lowered employee satisfaction
Model Cost Impact of bad IQ

- Impact on operational costs
  - Detection costs
  - Correction costs
  - Prevention costs

- Impact on strategic and tactical costs
  - Delay in decision making
  - Ad hoc integration of data
  - Difficulty in accessing and using data
  - Organizational mistrust
Manage IQ

Discover IQ problem ("manually")

Automate

IQ Knowledge Base

Diagnose

Hot fix and correction
Manage IQ – Visualisation

- IQ of database
  - Customer table
  - Orders
  - Currencies

- IQ of customer table
  - customer ID
  - Country
  - Birth date

- IQ of birth date
  - FORMAT
  - MISSING
  - BUSINESS

- Useless
- Limited usability
- Useful

Birth date: MISSING Test

- country
- time
Systematically managing IQ is key for good quality decisions – today more than ever!

Strategic frameworks and methodologies for managing are available today

Automation of the IQ function and knowledge repository

Expect benefits!

- Enhance operational productivity
- Enhance quality of customer intelligence
- Reduce resource allocation
- Enhance project planning and reduce risk of failure
Thank You!

Contact Info

Frank Block
Email: frank.block@finscore.com
Tel.: 021 647 77 44

FinScore AG
Chemin de l’Orio 6
1032 Romanel sur Lausanne
Switzerland