Quality Control Practices at Central Bank of Brazil

Paula Oliveira
Data Quality – an issue for everyone!

The Data Warehousing Institute (TDWI) estimates that poor quality customer data costs U.S. businesses a staggering $611 billion a year.
Data Quality – an issue for everyone!

A recent industry survey among 500 medium-size corporations with annual sales of more than $20 million has identified that more than 60 percent of the surveyed firms have problems with data quality.
Information Governance

- **Intrinsic Data Quality**
  - Accuracy and objectivity
  - Consistency

- **Contextual Data Quality**
  - Relevance
  - Timeliness
  - Completeness
  - Appropriate quantity

- **Representational Data Quality**
  - Interpretability
  - Validity

- **Accessibility Data Quality**
  - Accessibility
  - Access security
Created in 2013

Benchmarks:

- Costs reduction
- Articulation between the creation and the use of information
- Deactivation and disposal of useless data (529 DB)
- Exemption of outdated reports to BCB
- Master data (datos maestros) management
- Creation of the Financial Entities Corporative Register
- Data exchange through information integration
Information Governance at BCB

Information Governance Committee (CGI)

Curator’s Council

DEINF

Information Governance Office

BCB Departments

Curator 1

Curator 2

Curator 2

Curator n

...
Information Governance Model

Identification of information need
  CGI Approval
  Data reception
  Corporative database upload
  Integrity and consistency checking
  Departmental database upload
  Creation of metrics and indicators for the monitoring and supervision of SE
  Creation of tool for data extraction and analysis
CGI defines who is responsible for the Data Quality information

**Information Governance at BCB**

- **Intrinsic Data Quality**
  - Definition and monitoring of quality indicators
    - Accuracy and objeventivity
    - consistency

- **Contextual Data Quality**
  - Maintain ongoing process to assess: Relevance, Timeliness and Completeness

- **Representational Data Quality**
  - Metadata registry
  - Validity rules

- **Accessibility Data Quality**
  - Definition of public
  - Definition of channels to share information
  - Definition of access rules
  - Identification of risks events in database
IDQ - Quality Control must control all sources of information
1. Templates Modelling
   • Negotiation
   • Configuration
   • Validation

2. Verification
   • Critics

3. Loading
   • Database

4. Integrity
   • Data Quality
   • Tuning

5. Consistency
   • Internal x internal;
   • internal x external.

6. Data Quality outputs
   • Questions
   • Reports and indicators
   • Statistical Modelling

IDQ - Overview of the Integrity and Consistency Process
Intrinsic Data Quality (IDQ)

Channel and format standardization

1 Format → XML
1 Channel → STA

Negotiation and Template’s Disclosure

• Negotiate new acquisitions with the SE
• Template’s disclosure
  ✓ layout and instructions
  ✓ FAQ
  ✓ Examples

The pre-processing

• Creating pre-processing rules
• Creating xsd or other type of validator software
The information curator: defines the rules
the TIC department is responsible for the implementation

- **Source Integrity**
  - Every file is encrypted and digitally signed

- **Contents Integrity**
  - Syntax validation (e.g. xsd)
  - pre-processing rules validation

In case of error, the entity must replace the document.
IDQ – Data Integrity Process

Implemented using:
• Informatica Analystics
• ETL process (scripts in PowerCenter)

Basic tests:
• Verify the presence of required information
• verify non permitted values
• verify the values Range
• verify the data format
• verify the existence of nulls
• Cross – checking between columns
• Unicity of a key

In case of error:
• The entity must replace the document or adjust the data (SCR)

Out of process:
• Reports of integrity
**IDQ – Internal Consistency Process**

1) **Cross - Checking**

Compare the same data informed by the SE in different documents.

**Implemented using:**
- Informatica Analytics
- SQL Scripts developed by core departments
- PowerCenter scripts developed by TIC department

**Beyond the quality control, this process may identify:**
- Internal processes issues
- Internal inconsistencies on TI processes

**E.G.**
- Balance Sheet x Credit Information
- Balance Sheet x Operational Limits Report
2) Statistics Modelling

The most used techniques are:
• trend Analysis to find outliers and common range
• Peer group analysis with clustering technics
• Bendford’s law technics

Implemented using:
• Informatica Analytics
• Software R
• SAS
• PowerCenter scripts developed by TIC department

3) Disclosure of information
• Internal and external public
External Consistency
Compare the same data informed by different sources.

Implemented by
- SQL Scripts developed by core departments
- PowerCenter scripts developed by TIC department

Beyond the quality control this process may identify:
- Internal process issues
- Illegal operations
- Irregularities

E.G.
Obituary data x Credit Information
- vehicle registers x Credit Information
- Government Employee register x credit information
Integrity Process – Main Issues

- Difficulties to establish agreements for data sharing among government entities
- The implementation of loading process takes a long time
- Highly bureaucratic process to acquire new information

- The business is much more complex than the models developed to represent them
- Detecting quality issues using internal consistency process is very complex
- Usually, it is necessary to require more information to improve data quality

- Lack of unique keys in several registers complicates the establishment of relations among them. Databases need remodeling and keys standardization
- Quality process takes long time and needs a robust IT infrastructure

Information Governance

- Information reception
- Data Quality
- Consistency and reports
### Results

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SFNI</td>
<td></td>
<td></td>
<td>19,222</td>
<td>0,535</td>
<td>231,39</td>
<td>13,39</td>
<td>198,89</td>
<td>1,401</td>
<td>698</td>
<td>23,49</td>
<td>13,85</td>
<td>19,87</td>
<td>0,01</td>
<td></td>
</tr>
<tr>
<td>CASA ECONOMICA FEDERAL</td>
<td>00.360.305</td>
<td></td>
<td>1,553</td>
<td>1,503</td>
<td>36,81</td>
<td>24,55</td>
<td>29,99</td>
<td>55</td>
<td>55</td>
<td>1,28</td>
<td>0,47</td>
<td>1,14</td>
<td>0,00</td>
<td></td>
</tr>
<tr>
<td>AYMORÉ GFI S.A.</td>
<td>07.707.650</td>
<td>30378</td>
<td>1,123</td>
<td>1,195</td>
<td>27,54</td>
<td>2,39</td>
<td>26,96</td>
<td>29</td>
<td>55</td>
<td>1,19</td>
<td>1,66</td>
<td>1,17</td>
<td>0,12</td>
<td></td>
</tr>
<tr>
<td>BV FINANCEIRA S.A. CFI</td>
<td>01.149.953</td>
<td>51011</td>
<td>1,059</td>
<td>905</td>
<td>19,02</td>
<td>0,30</td>
<td>17,54</td>
<td>599</td>
<td>502</td>
<td>11,03</td>
<td>0,25</td>
<td>10,09</td>
<td>0,09</td>
<td></td>
</tr>
<tr>
<td>BCO DO BRASIL S.A.</td>
<td>00.000.000</td>
<td>49996</td>
<td>709</td>
<td>674</td>
<td>18,89</td>
<td>6,53</td>
<td>17,38</td>
<td>67</td>
<td>67</td>
<td>1,16</td>
<td>7,07</td>
<td>0,93</td>
<td>0,00</td>
<td></td>
</tr>
<tr>
<td>BCO BRESILFORD FINANC S.A.</td>
<td>07.207.996</td>
<td>10045</td>
<td>441</td>
<td>441</td>
<td>9,99</td>
<td>3,99</td>
<td>9,55</td>
<td>11</td>
<td>11</td>
<td>0,11</td>
<td>0,00</td>
<td>0,11</td>
<td>0,03</td>
<td></td>
</tr>
<tr>
<td>BCO BRASIL S.A.</td>
<td>00.746.940</td>
<td>10045</td>
<td>540</td>
<td>540</td>
<td>9,67</td>
<td>7,53</td>
<td>8,09</td>
<td>55</td>
<td>55</td>
<td>0,73</td>
<td>23,00</td>
<td>0,50</td>
<td>0,00</td>
<td></td>
</tr>
<tr>
<td>BCO ITAÚCARD S.A.</td>
<td>17.192.451</td>
<td>10989</td>
<td>704</td>
<td>683</td>
<td>9,42</td>
<td>9,37</td>
<td>8,81</td>
<td>101</td>
<td>101</td>
<td>2,22</td>
<td>96,21</td>
<td>0,09</td>
<td>0,01</td>
<td></td>
</tr>
<tr>
<td>BCO SANTANDER (BRASIL) S.</td>
<td>90.400.880</td>
<td>30379</td>
<td>500</td>
<td>492</td>
<td>9,29</td>
<td>4,07</td>
<td>8,60</td>
<td>52</td>
<td>52</td>
<td>0,69</td>
<td>0,88</td>
<td>0,05</td>
<td>0,00</td>
<td></td>
</tr>
<tr>
<td>BCO MERCEDES-BENZ S.A.</td>
<td>60.814.191</td>
<td>51282</td>
<td>391</td>
<td>90</td>
<td>6,31</td>
<td>7,35</td>
<td>5,73</td>
<td>57</td>
<td>57</td>
<td>0,00</td>
<td>0,00</td>
<td>0,00</td>
<td>0,06</td>
<td></td>
</tr>
<tr>
<td>ITAÚ UNIBANCO BM S.A.</td>
<td>60.701.190</td>
<td>10089</td>
<td>851</td>
<td>582</td>
<td>7,93</td>
<td>33,87</td>
<td>5,29</td>
<td>7</td>
<td>7</td>
<td>0,07</td>
<td>0,00</td>
<td>0,07</td>
<td>0,00</td>
<td></td>
</tr>
<tr>
<td>BCO VOLKSWAGEN S.A.</td>
<td>59.109.165</td>
<td>438</td>
<td>438</td>
<td>426</td>
<td>5,97</td>
<td>32,31</td>
<td>4,88</td>
<td>98</td>
<td>98</td>
<td>1,09</td>
<td>46,77</td>
<td>0,59</td>
<td>0,03</td>
<td></td>
</tr>
<tr>
<td>OMMI SA CFI</td>
<td>92.223.410</td>
<td>50940</td>
<td>3,152</td>
<td>1,340</td>
<td>7,11</td>
<td>24,31</td>
<td>3,55</td>
<td>25</td>
<td>25</td>
<td>0,31</td>
<td>17,76</td>
<td>0,23</td>
<td>0,21</td>
<td></td>
</tr>
<tr>
<td>BANCO PAN</td>
<td>59.205.411</td>
<td>31323</td>
<td>556</td>
<td>553</td>
<td>3,35</td>
<td>11,83</td>
<td>3,08</td>
<td>6</td>
<td>6</td>
<td>0,07</td>
<td>21,27</td>
<td>0,06</td>
<td>0,02</td>
<td></td>
</tr>
<tr>
<td>HSBC BANK BRASIL SA BCO M</td>
<td>01.701.201</td>
<td>51152</td>
<td>272</td>
<td>164</td>
<td>2,63</td>
<td>0,40</td>
<td>2,73</td>
<td>25</td>
<td>25</td>
<td>0,20</td>
<td>18,76</td>
<td>0,16</td>
<td>0,00</td>
<td></td>
</tr>
</tbody>
</table>

Banco Central do Brasil
Results – Sharing benefits

In 2012, BCB informed the INSS about quality issues in the obituary register system and suggested improvements to the quality process. After implementation, the quality of information has significantly improved.

Participação dos CPFs no total de Óbitos
Thanks

Contacts:

Andrei Vanderlei
E-mail: andrei.vanderlei@bcb.gov.br
Tel.: +55 61 3414-1834

Paula Oliveira
E-mail: paula.cristina@bcb.gov.br
Tel.: +55 61 3414-3299
<table>
<thead>
<tr>
<th><strong>Mercado de Credito</strong></th>
<th><strong>Mercado de Cambio</strong></th>
<th><strong>Mercado de capitales, dinero y derivados</strong></th>
</tr>
</thead>
</table>
| • Recibe mensualmente de **1.500 IF**:  
  • Operaciones de crédito de **64 millones clientes**.  
  • **480 millones de operaciones por mes**  
  • **20 mil millones de operaciones gravadas** desde 2004. |
| • **207 instituciones**  
  • **31.000 operaciones por día**  
  • **7.800.000 operaciones primarias en 2014**  
  • **70.000 registros diarios** |
| • **Datos de: Selic, CETIP, BM&F Bovespa, SPB and otros**  
  • **1 mil millones de registros por mes**  
  • **Mas de 900 docs procesados cada mes (DRL y DRM).** |

<table>
<thead>
<tr>
<th><strong>Informaciones Contables</strong></th>
<th><strong>Consórcio</strong></th>
<th><strong>Sistema de Pago</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>• recibe mensualmente <strong>1.136 docs contables e 2.267 informes de limites operacionales</strong>, cada trimestre <strong>mas de 7.300 docs</strong>, representando <strong>mas de 2.2 millones de entradas</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| • **11 millones de participantes** distribuidos en **20.000 grupos** y datos de **9 millones de clientes**  
  • Recibe **787 millones de registros cada trimestre** |
| • **4.5 millones de operaciones mes** |