

Points for Discussion	
Data Management – Data Quality in Bank of Cyprus	
1	What is your organization's definition of 'data quality'? How do you measure data quality today?
2	What subject areas are currently being affected by data quality issues (e.g. customer data, product data, account information, transactional, reference data, etc.)?
3	Do you have a centralized Customer Base?
4	What is the impact of poor data quality to your business (e.g. reliable analysis and reporting not possible, regulatory compliance, profitability of marketing campaigns, effort for additional verification, etc.)?
5	<p><b>Which of these are you experiencing due to poor data quality?</b> (E.g. reliable analysis and reporting not possible, regulatory compliance, profitability of marketing campaigns, effort for additional verification, etc.)</p> <ul style="list-style-type: none"> <li>• Reduced customer satisfaction</li> <li>• Inaccurate or non-existent reporting (regulatory, management, financial, risk, etc.)</li> <li>• Reduced productivity</li> <li>• Inability to produce in a timely fashion management reports</li> <li>• Inability to produce in a timely fashion regulatory reports</li> <li>• Poor customer segmentation for marketing campaigns</li> <li>• Poor reporting accuracy on customer-level risk scoring, payment history, products owned, etc.</li> <li>• Difficulty to automatically assess and filter out bad risk customers or to implement risk adjusted customer level pricing</li> <li>• Poor accuracy of propensity to default scoring models</li> <li>• Incorrect or missing customer data associated with loan product payments which increases bad debt</li> <li>• Inability to report or manage exposure on a customer level</li> <li>• Inability to identify connected parties</li> <li>• Inaccurate billing and invoicing (statements)</li> </ul>
6	Describe the nature of the data quality issues you have been able to identify (e.g. data missing, data not in proper range, field content not appropriate to the business description, dependent fields not populated or in conflict, calculated or aggregated data not correct, data duplication, etc.)?
7	What are the reasons for poor quality of the affected data (e.g. errors during manual data feeds into the system, systems failures, improperly working data transfer or transformation processes, highly heterogeneous data sources, frequent system changes, etc.)?
8	What have you currently implemented to improve the quality of your data?
9	Data Governance practices in place – i.e. Who is responsible for the data content and structure, policies and procedures
10	When you're looking at information coming from your Banking systems and there is something that does not look correct, what do you do to fix it?
11	<p>Who / what department in the Bank is responsible for data quality?</p> <p>Have data quality criteria or specifications been established for the entire company or are they department specific?</p> <p>Is the responsibility centralized (e.g. one department) or is every department responsible for its own data?</p>
12	Are the data quality processes permanent (e.g. data cleansing, data quality monitoring) or are executed on an "as-needed" basis for specific sets of data?

No.	Points for Discussion
13	<p>Do you have a Data Dictionary? Do you intend to build a Corporate Data Dictionary? The below are points to consider:</p> <ul style="list-style-type: none"> <li>• Ensure a common understanding of critical business objects and terms leading to reduced search times and preventing misconceptions.</li> <li>• Has the Information Technology Environment been sufficiently documented?</li> <li>• Detailed data lists, data specifications, and data mappings.</li> <li>• Data Quality Standards (e.g. Naming conventions, data entry guidelines, coding standards)</li> <li>• Business Rules (e.g. Field Optionality, known data quality issues, data integrity guidelines)</li> <li>• Metadata (e.g. Database names, Table and Field names, Validation rules)</li> <li>• Reference Data (e.g. Reference Table Names, Domain and format guidelines)</li> <li>• Usage (Common business usage)</li> <li>• Problems with permissions and access to data been identified.</li> </ul>
14	What is your IT strategy over the next one to three years in terms of Data Management?
15	What are your immediate targets in terms of Data Management?
<b>Data Management Areas of interest to the Bank</b>	
1	<p><b><u>Check and Correction (missing values and range checking, format verification of non-name and non-address data)</u></b></p> <ul style="list-style-type: none"> <li>• Do you have data elements in your source systems that should be checked in regards to proper content (e.g. fields that shouldn't be empty, fields that must be populated and properly formatted – email addresses, telephone numbers, etc.)?</li> <li>• Do you maintain reference or lookup tables and/or business rules for specific data items?</li> </ul>
2	<p><b><u>Parsing (location, isolation and identifying of individual client data elements from data strings/concatenated data elements)</u></b></p> <ul style="list-style-type: none"> <li>• Are source system data elements placed in separate fields or are they kept in a concatenated way (e.g. do you keep name prefix, title, first name, last name in one field or in separate fields, is address data concatenated, multiple phones in one field etc.)?</li> <li>• Would there be occurrences of misplaced data elements in your source systems (e.g. first name in the field for the last name and reverse)?</li> </ul>
3	<p><b><u>Identification (information determination based on available data)</u></b></p> <ul style="list-style-type: none"> <li>• Can business and individual client data be easily separated (e.g. by a dedicated indicator or flag)? Is it necessary to identify or verify if name and address data is business or individual client data?</li> <li>• Can you easily validate CIF, SSN or ID Card No.?</li> </ul>
4	<p><b><u>Standardization (providing customer information in a preferred and consistent format)</u></b></p> <ul style="list-style-type: none"> <li>• Do you use inconsistent name abbreviations (client) in your source systems (e.g. Bank of Cyprus Ltd, Bank of Cyprus, BoC, BoCy)?</li> <li>• Do you use inconsistent address and name elements in your source systems (e.g. Mr. and Mister, Vice President and VP, Doctor and Dr.)?</li> <li>• Do you have misspellings and variant spellings in your source systems (e.g. Konstantinos or Constandinos or Costas, Company or Corporation)?</li> </ul>
5	<p><b><u>Postal Check (checking and correction of address data based on a reference address database)</u></b></p> <ul style="list-style-type: none"> <li>• Do you have null, blank values on the Postal Code field? Do you have a pricing/ mailing process based on this field?</li> </ul>

No.	Points for Discussion
6	<p><b><u>Data Integration (matching = identifying similar data and consolidation = de-duplication or grouping records based on defined rules)</u></b></p> <ul style="list-style-type: none"> <li>• Is there a business need to consolidate records with similar content (e.g. variations of an individual client customer) and create and maintain a master record that represents a cluster of records? -&gt; individual client Consolidation                             <ul style="list-style-type: none"> <li>○ If yes, are there any rules for building a master record? Are there any matching criteria?</li> <li>○ Is there a target master (customer) hub? Where do the data reside?</li> <li>○ Is integration of records based on individual clients with the same addresses, potentially same name or distinct name needed as part of your business processes? -&gt; Consumer Consolidation (House holding)</li> <li>○ Is integration based on relationships between records such as company name needed as part of your business processes? -&gt; Site Consolidation</li> <li>○ Is any custom integration needed as part of your business processes? -&gt; User identified consolidation.</li> </ul> </li> </ul>
7	<p><b><u>Data Augmentation (enrichment of the existing customer data with additional information like geographic data, behavioral data, demographic data, etc.)</u></b></p> <ul style="list-style-type: none"> <li>• What kind of data need to be added to existing data? (Postal code, City, ID, etc.)</li> <li>• What is the structure and level of the external data to be added?</li> </ul>
8	<p><b><u>Data Profiling (understanding the data you have, analysis and auditing)</u></b></p> <ul style="list-style-type: none"> <li>• Is it necessary for your business to receive in batch on demand data statistics such as value frequency reports or phrase and element analysis?</li> </ul>
<b>Data Sources</b>	
1	Which are the main source systems of the Bank? Any other Additional system?(web application for clients, etc.)
2	Which systems are mainly affected by DQ problems?
3	Are there any Data Quality processes applied on those system?
4	Are there interfaces which insert data into the production system? What kind of interfaces? Which systems are affected?
5	In your recent migration projects what kind of issues have you faced?
<b>Data Governance – Organization and Processes</b>	
1	Is there a data management policy aligned with business needs and relevant compliance requirements?
2	Has a data management organizational structure and resource requirements been defined?
3	<p>Are there defined roles and responsibilities to establish and assign agreed authorities for data management roles (Effective governance means establishing clear ownership over data, responsibilities for capturing, storing, managing etc. ) Example roles include:</p> <ul style="list-style-type: none"> <li>• Data Manager</li> <li>• (Business) Data Steward</li> <li>• Data Architect</li> <li>• Data Quality Lead</li> <li>• Data Owner</li> <li>• Business User</li> </ul>
4	Has the organization identified and assigned data owners and has it established standard data ownership accountabilities?
5	Are changes to the way data is defined and used reviewed and approved?
6	Has the organization agreed on common data definitions and types?

No.	Points for Discussion
7	Has the organization developed a master data inventory that identifies data elements, business criticality, associated ownership, and key applications that require the data and a framework to provide master data oversight and concurrence?
8	Has the organization defined and documented data management lifecycle processes based on a policy and that governs the data processes? Examples include: <ul style="list-style-type: none"><li>• Data Maintenance:<ul style="list-style-type: none"><li>○ Create data</li><li>○ Read data</li><li>○ Update and change data</li><li>○ Delete</li></ul></li><li>• Data Migration and Integration:<ul style="list-style-type: none"><li>○ Data aggregation</li><li>○ Data cleansing</li><li>○ Data mapping</li><li>○ Data conversion</li><li>○ Data load / migration</li></ul></li><li>• Quality Assurance and Control:<ul style="list-style-type: none"><li>○ Data dictionary and rules (Define)</li><li>○ Data analysis (Discover and Design)</li><li>○ Data evaluation (Evaluate)</li><li>○ Data remediation (Execute)</li><li>○ Quality assurance and reporting (Control i.e. Quality assurance on behalf of business, post-cleansing)</li></ul></li><li>• Data Archiving:<ul style="list-style-type: none"><li>○ Data retention / archive</li><li>○ Reporting</li></ul></li></ul>