A Framework for Improving Business Intelligence through Master Data Management

Ray M. Kekwaletswe\textsuperscript{1} and Tshegofatso Lesole\textsuperscript{2}

\textsuperscript{1}University of the Witwatersrand
Johannesburg, South Africa

\textsuperscript{2}University of Pretoria, Pretoria, South Africa

Correspondence should be addressed to: Ray M. Kekwaletswe; Ray.kekwaletswe@wits.ac.za

Received date: 12 March 2015; Accepted date: 16 July 2015; Published date: 18 April 2016

Abstract

For reporting and for analytics purposes, financial banks highly demand real-time delivery of quality operational information. This paper argues that for banks to improve their decision making and performance management, business intelligence and data management ought to be core activities. Conversely, Business Intelligence (BI) and Master Data Management (MDM) are activities with most challenges as financial banks grapple with ways to make profitable informed decisions that align with the projected business targets and goals. In this paper, the argument is that a profound understanding of master data management, as an activity, is important to the delivery of real-time quality information. Subsequently, business intelligence should improve. To this point, this paper explores and describes how an understanding of MDM activities may lead to improved business intelligence in a financial bank. In this paper, MDM and BI are seen as activity systems. However, this paper focuses on the MDM activity and not the BI activity. Activity theory is used as a lens to analyze the activity system. The paper concludes by conceptualizing a framework for improving business intelligence through a sound master data management.

Keywords: Activity theory, Business Intelligence, Master Data Management, Data Governance

Introduction

For reporting and for analytics purposes, financial banks highly demand real-time delivery of quality operational information. This means working with high volumes of data regularly, that would need to be managed to produce the best quality of data for analytics and reporting purposes. Through the empirical study, it was found that banks are still faced with challenges of data management, maintaining and creating quality data. Business Intelligence (BI) entails timely analysis and decision
making with relevant, accurate, adequate, correct and complete data. Business Intelligence is a discipline that draws business decision making through performed analytics and reporting based on the data an organization has.

Wu et al., (2007) define BI as a business management term used to describe applications and technologies that are used to gather access to, and analyse data and information about the organization to help management to make business decisions. Ghazanfari et al., (2011) add on to say that the theme of BI has two 'divisions' from technical, system-enabler and managerial viewpoints - tracing two broad patterns, managerial and technical approaches to BI. Thomsen (2003) posits that BI as a term replaces decision support, executive information systems and management information systems. On the other hand, Nelson and Phillips (2003) believe that Business Intelligence (BI) communities bring clarity and reasoning to unreadable data to empower good decision making. This is through data that have been rationalized, centralized and mastered for reporting on the business (Rudin and Cressy, 2003); Langseth and Vivatrat (2003) say that there are essential components that fulfill BI namely: (1) Real-time data warehousing, (2) Data mining, (3) Automatic learning and refinement, and (4) Data visualization. Decision making is aided by data; therefore, it is valuable for organizations to manage their data to gain desired results when performing business analytics and reporting. Data are of high quality if are fit for their intended use in operations, decision-making, and planning (Juran, 1964).

Data Management as defined by Mosley (2007), is the development, execution and supervision of plans, policies, programs and practices that control, protect, deliver and enhance the value of data and information assets. The effectiveness of BI lies in the ability to present business information in a timely manner and this is dependent on the data availability (Clack et al., 2007). Data Management helps in controlling and coordinating the usage of relevant and reliable data.

Swanton et al., (2007) suggest that Business Intelligence can take advantage of data management and further more Master Data Management (MDM) to improve the latter's practice. They go on to say the ultimate goal of MDM strategies is to harmonize data to facilitate analytics for business intelligence. They are of the opinion that the analytics that MDM could provide include (1) Measuring data quality and improving on metrics reporting throughout all levels of business, and (2) Creating a sift that will assist in identifying problems and getting them resolved. The next section discusses master data management and approaches.

**Master Data Management**

Master data management is the technology, tools, and processes required to create and maintain consistent and accurate lists of master data (Haselden and Wolter, 2006). Some of the processes in MDM include source identification, data collection, data transformation, normalization, rule administration, error detection and correction, data consolidation, data storage, data distribution, and data governance. The tools include usage of data networks, file systems, a data warehouse, data marts, an operational data store, data mining, data analysis, data virtualization, data federation and data visualization. Master Data Management is also seen as a collection of the best data management practices (Loshin, 2009). It should be a system of business processes and technology components that ensures information about business objects, such as materials, products, employees, customers, suppliers, and assets (Swanton et al., 2007). These approaches have different actions that can happen in them. Figure 1 shows the different practices that pertain in each MDM approach.

---

To have a more solid MDM strategy that would assist improve BI, two (2) MDM approaches should be practiced simultaneously. The Business Approach would look at a plan of action on MDM activities and processes with the analogy of the organization and what it needs. It will focus on the Business Objectives, Business Process Policies, Business Rules, Change Management and Resource (people) Roles and Responsibilities in ensuring and controlling a set MDM strategy. The Technical Approach would look at the technologies, tools and technical rules and guidelines that will ensure a single truth of the data ‘mastered’ and an efficient MDM. The advantage of using the two approaches simultaneously is the commonality that they both promote data governance and data quality, which is a mandatory outcome for a good MDM strategy. Combining the two approaches will tackle the people, processes and technology aspect of MDM.

This paper looks at BI as an activity that could be improved through a thorough understanding of MDM. The rest of the paper is organized as follows: first the background to the research problem is discussed, followed by the theoretical framework and research methods. Thirdly, the case study findings are discussed. This is followed by conceptualizing a framework for improved business intelligence through master data management. Lastly, the paper is concluded.

Figure 1: MDM approaches and actions
(Lesole and Kekwaletswe, 2014; Haselden and Wolter, 2006)

Background to the Research Problem

Context and study Location

The study was conducted at a financial bank whose main business is to allocate funds from investors/savers to borrowers efficiently. The bank provides personal, commercial and corporate banking services to more than 6 million customers across South Africa. Remaining on the cutting edge of banking by offering customers’ innovative services is one of the bank’s missions. Currently, with the exponential growth of data and innovative services provided to customers in the banking environment, there is a high demand in the for real time delivery of quality operational information for reporting, analytics and to enhance customer service.

Challenges and issues experienced by the bank

A preliminary observation showed that the bank is faced with challenges of managing data. Banks are posed with the challenge to gain insight on how to make business agile, through data usage. The demand to use
data to deliver accurate information in real time for Business intelligence is also as crucial but challenging. This demand is increasing exponentially. South African banks generally have the fastest growing use and demand for quality data; however, they are often faced with data management challenges that tend to hinder business intelligence. The commonly experienced issues are data inaccessibility, data duplication, poor data quality, data inaccuracy, data incompleteness and data irrelevance are being experienced, particularly in business decision making and reporting.

The challenges experienced at the bank included inadequate data governance that causes redundant data storage, inconsistent data due to duplicate/multiple data storage in the organization and incorrect data sourcing for reporting which is normally due to the multiple data storage locations in the organization; secondly, the business needs and the chosen MDM methodologies and technologies used are not in sync. Inappropriate data methodologies and technologies to assist with data analytics are often used.

The next section discusses the underpinning theory and the research methods used to address the research problem.

**Theoretical Framework And Research Methods**

This section discusses the theory underpinning the study and the research approaches undertaken for the study. It discusses Activity Theory, research paradigm, strategy, design, as well as the data collection techniques.

**Theoretical Framework**

This paper argues that Master Data Management as an activity may improve business intelligence in banks; therefore, Activity Theory (AT) was seen as an appropriate theoretical framework to underpin the study. Morf and Weber (2000) posit that AT is a conceptual framework based on the idea that activity is primary, that doing precedes thinking and that goals, images, cognitive models, intentions and abstract notions like ‘definition’ and ‘determinant’ grow out of people doing things. Activity Theory has four elements which helped guide the study: tools in use, subject of study, objective of study and the outcome. Informed by Engestrom (2001), this paper used elements that include definition of rules, community and division of labour.

These activity theory elements, also iterated by Kaptelinin et al (2005), are defined in this paper as follows:

a) Activity towards an objective (goal) - MDM actions done by the data community for business intelligence purposes.

b) Action towards a specific goal (conscious), done by an individual or a group possible goals and sub-goals, critical goals - answers the "What" question.

c) Operation structure of activity normally automated and not conscious concrete way of implementing an action in accordance with the specific conditions surrounding the goal – answers the "How" question.

In the study, the major activities include Data Management, Master Data Management and Business Intelligence. Empirical data were collected around these activities; which formed the objectives of the study as themes. These themes were: Identifying data management activities and processes within the case study; Analysis of the current MDM practices and actions within the studied bank; Determining an efficient MDM approach that would best suit the bank’s current environment; Analysis of BI activities in the case study.

**Research Approach and Paradigm**

In research, there are two (2) types of research approaches that could be used, Qualitative and Quantitative. Shank (2002) defines qualitative research as “a form of systematic empirical inquiry into meaning”. Ospina (2004) adds on to say, by
systematic it means "planned, ordered and public", following rules agreed upon by members of the qualitative research community. Maxwell (2013) says, qualitative is research that is intended to help one better understand the following: the meanings, perspectives of people you study – seeing the world from another point of view not the researcher’s point of view and how these perspectives are shaped by – their physical, social and cultural context and the specific processes that are involved in maintaining the phenomena and relationships. Anderson (2006) summarises the difference between qualitative and quantitative research by simply saying that qualitative research is subjective and the other is objective.

A qualitative approach was chosen for the study. As highlighted by Conger (1998), Brynam et al (1988) and Alvesson this approach added the following advantages to this study:

- flexibility to follow unexpected ideas during research and explore processes effectively;
- sensitivity to contextual factors;
- ability to study symbolic dimensions and social meaning;
- increased opportunities
  - to develop empirically supported new ideas and theories;
  - for in-depth and longitudinal explorations of leadership phenomena; and
  - for more relevance and interest for practitioners.

This means qualitative approach allowed us to dive deeper into the case to understand the meaning and significance of MDM and BI. It also gave us the ability to follow unexpected leads during the study, explore currently carried out processes effectively, and to study representative elements of the case and its social meaning. This study approach also ties well with the interpretive paradigm. Obrien (1998) explains interpretive paradigm as an emphasis on the relationship between socially-engendered concept formation and language. Interpretive paradigm is underpinned by observation and interpretation, thus to observe is to collect information about events, while to interpret is to make meaning of that information by drawing inferences or by judging the match between the information and some abstract pattern (Aikenhead, 1997). Interpretive paradigm was found appropriate to follow since the intention was to understand the subjective meanings experienced by the participants of the case study.

Research Location and Strategy

This was a case study of one of the leading banks in South Africa. Yin (2003) defines a Case Study Research Method as an "empirical inquiry that investigates a contemporary phenomenon within its real-life context. The studied bank possesses large volumes of data with a complex BI environment; this study also helped draw views of a banking industry in South Africa and provided a more realistic approach to understanding issues and problems with respect to MDM and BI activities. A case study was seen appropriate as it was seen to be a more interactive research approach. It allowed the researcher to spend time at the study location and interact with participants to understand and observe how business intelligence and data management is conducted in the banking industry.
Research Design

Participants

The study participants were drawn from the following units: (1) The Data Warehouse Team - which provides data to the bank’s reporting team and should ideally hold the master data, (2) The Business Intelligence Team – these are the report generators/creators, (3) The Analytics Team from the varied business units (core banking services) – this is a group of individuals whose mandate is to analyse how the bank is financially and operationally performing and give suggestive activities to improve business, (4) The Production Team - manages end of day data feeds into the warehouse or warehouses, (5) The Branch Consultants – they capture data on to the bank’s systems at branch level.

Data Collection Techniques

In collecting empirical evidence for the study, the following were the data collection techniques used:

Interviews: For this study, semi-structured face to face interviews were individually conducted with 14 participants whose average number of year that the respondents held at their current role and within the bank was seven (7) years that work and were affected by Data Management and Business Intelligence in the bank.

Observation: A naturalistic observation was done in terms of thick description of MDM and BI activities, e.g., how data are captured, stored, managed and used for reporting and analytics, including who is acting in which role, when and where, etc. There was engagement with limited interaction and intervention with research participants and events; interventions only occurred when clarity of activities and actions was needed.

Discussion and Interpretation of Findings

The discussion and interpretation of findings is done in this section, per the three themes: (1) Data Management Activities and processes within the banking environment; (2) current MDM practices and actions in a banking environment; and (3) Business Intelligence activities in a banking environment. The three activities are discussed below:

Analysis of Data Management Activities and processes within the banking environment

Data Management is an activity performed by a community of employees in difference teams within the bank. Kuutii (1991) states that an activity contains a number of actions, with the same actions featured in different activities to show their significance. Figure 2 shows the different actions to creating a new profile and a new account. It shows the data management process and procedures involved in the activity.
In a data management activity, a branch consultant performing ‘Customer Profile Opening’ action could simultaneously be satisfying data quality, business process rules, data standardization and conforming to the needed data structures. The different identified functions of data management, figure 3, shows the actions in the activity and how they are interlinked and coordinated. The data management actions include: (1) data entry, (2) follow data policies developed for standardization, (3) through given data security measures granted to them perform this role, (4) data quality measures that are put on the system restrict how the data are captured – user authentications and log in credentials created for all given access to those that use the data, (5) at the end of creating a customer profile and account, the data are stored onto their respective repository. Some of the actions that could be performed at the back-end of the system include data quality assurance, which could be initiated by different set rules on the batches and on the data storages (tools), data security measures (rules and data policy measures) that are set by the data stewards and owners to assist in sharing and accessing data in the data storages, data sharing, data ownership, and data standardisation, including how they are stored in the different data storages and accesses.

Figure 3 shows the actions that are performed at the back-end of the system. These include data quality assurance that could be initiated by different set rules on the batches and on the data storages (tools); data security measures (rules and data policy measures) that are set by the data stewards and owners to assist in sharing and accessing data in the data storages; data sharing, data ownership, data standardisation and the different data storages and data accesses.
The data management activity included data entry, data storing, data sharing, data security and data quality assurance actions. All the participants emphasized the importance of data quality in the bank, particularly for decision making and reporting; they also mentioned that data quality is compromised in many ways than one. It was highlighted in this theme that data quality assurance needs to be continuously monitored and improved in order to get the best quality data constantly – which currently is not the case. Most data storage business units do not conform to the same rules in regulating the data they store and/or maintain.

**Analysis of the current MDM practices and actions in a banking environment**

Master data management is a collection of the best data management practices (Loshin, 2009). The practice includes the people that perform the data action, the tools used to perform the action, the policies and rules defined to govern the action, the infrastructure used to sustain the activity and its actions e.g. the business applications (data entry points) and integration and sharing the use of accurate, timely, consistent and complete master data.

There are multiple identified data storages that keep different but associated data in the bank; (1) Integrated Deposit Storage – part of the book keeping system that stores all customer and deposit account systems and feeds the deposit and demand deposit front-end systems; (2) Integrated Loan Storage – also part of the book keeping system that stores customer and loan (short & long term loans) accounts data and feeds the Integrated loan processing and online collection front-end systems, (3) Card Storage - also part of the book keeping system that stores credit card account data, it feeds the front-end card system; (4) Relationship Profitability Storage - this data storage holds all customer static information and feeds the customer information, Relationship and Product Management front-end systems; and (5) Online Delivery System - this data storage holds all data that are captured on external data entry points, which feeds front-end systems like ATM, Switch Transactions, Branch Automations and Online Applications.

All the preceding data storages are, in their rights, treated as master databases. Each business unit takes its data storage as the master database. This is in contradiction to what Karel (2011) sees MDM to be; he posits MDM to be the business capability charged with finally delivering that elusive single trusted view of critical enterprise data. Loshin (2009) also says master data can also be referred to using these terms like “critical business objects”, “business entities”, and “business concepts”; this essentially referring to common data themes. The bank needs to have a single view of their “business concepts” – customer, account, inventory and so on- not the current multiple views of account data.

The studied bank currently follows only one MDM approach – the technical MDM - and this proves to be a challenge for the bank. As business objectives and goals change, the technology does not necessarily change. And if it does, there is no formal communication downstream to the technical teams about these changes. MDM should be a system of business processes and technology components that should ensure information about business objects, such as materials, products, employees, customers, suppliers, and assets (Swanton et al., 2007). The two MDM approaches have different actions that can happen in them. Business/Managerial MDM Approach focuses on the processes required to create and maintain consistent and accurate lists of master data (Haselden and Wolter, 2006), and this is the MDM approach that the studied bank did not implement. The Technical MDM Approach focuses on the tools required to create and maintain consistent and accurate lists of master data (ibid., 2006).

With the technical MDM approach, the middle management and general staff (these could be the ETL Developer, data analysts) are forced to constantly analyze the data that are stored in the different...
data storages manually. The data are checked for correctness and validity – also known as data quality assessment and assurance. The employees would determine "like" data in the data storages and integrate these data onto a central point. However, without a buy-in or communication from top management – those that know and understand organizational objectives - these desires are as good as nothing and may lead to incorrect capturing and storage of data. A combination of the business and technical MDM approaches could work for the studied bank to help control the data management activities performed.

Analysis of Business Intelligence activities in a banking environment

To recap, this paper aimed to conceptualize a framework for an improved Business Intelligence. This means articulating ways to improve in delivery of quality operational information for reporting and analytics (outcome of activity). BI actions that are carried out at the studied bank include: (1) Analyzing data sources and integration– drawing adequate and relevant data from the multiple data sources timely is very difficult and time consuming, for the BI developers and data analysts; (2) Data profiling and quality – these actions are informed by assessing data and making sure it is accurate for what it is needed for. Analysis of the candidate data sources for a data warehouse clarifies the structure, content, relationships and derivation rules of the data (Loshin, 2009). In the bank, data profiling and quality is only done when the BI team has sourced the data and the data is prepared reporting instead of at the initiation phase of when the data is captured by bank employees/consultants.

A Framework For Improved Bi Through Mdm

This paper conceptualizes five actions for the master data management activity. The actions, signifying the MDM activity, are Business Strategy alignment with MDM Strategy; Data Governance; Data Integration; Availing data for Use (analytics and reporting) and Continuous Improvement of action 1 -4, respectively. These actions are each succinctly discussed next.

Action 1: Business Strategy alignment with MDM Strategy - This would be done by aligning the organization's key business fundamentals, business objectives, business process policies and business rules, to the rules that are built to capture and store data. In understanding the results hoped for by the organization, the business objectives could be the guidelines that become the foundation of the Master Data Management activities and practices.

Action 2: Data Governance – the activities to govern and standardize data into the different data entry systems and data storages within the bank that could be followed are: (a) Data Quality – There should be measures on the organization's front-end systems, data storages and data transformation rules that are aligned to the business needs and rules; (b) Data Stewardship - Build a community of data stewards from different subject areas of the organization – EXCO, IT and general employees at branch. These data stewards will be responsible looking at data integrity issues and measures within the entire bank. They need to look at data needs at a strategic level (business level) and what rules or ways would enable the bank to get these data; (c) Data Profiling - all the organization’s current data sources must be examined.

The quality of data values within these data sources should be evaluated by comparing it to the desired data profiles. A uniform way of storing data in the organization especially where the data entry systems are the same should be introduced. This could be done through data structure definition of all the data sources; (d) Data Standardization - Profiling the organization’s data can help promote and improve data source standardization. Data profiling and standardization rules should be in sync and should be enforced mainly at all data entry systems. This could be done by building a standardization job (function)
on extraction of data from the different data sources to “master” data sources to perform a data quality assessment of data in each data source based on the set rules that are also in line with business rules (to retrieved “expected” data). This is to ensure the “master” data source only receives the same kind of data content and format; and (e) Metadata Management - have meta-models define the structures of metadata. Centralization of metadata encourages the reuse of data and avoids data user confusion of usage of data.

**Step 3: Data Integration** - The entire organization’s data should be integrated and stored at a central point within the bank; this is by sharing and merging common and key business function data from existing repositories.

**Step 4: Availing data for Use (analytics and reporting)** - This is the security measures and intelligent way of availing high quality and single view of data to the entire bank’s community for both internal and external use.

**Step 5: Continuous Improvement Process of Step 1- 4** - To ensure continuous return on investment on master data management and improved Business intelligence, it is important to continuously evaluate the above given steps and enhance them when gaps are identified. The business objectives, needs and rules should constantly be reviewed to ensure that the ‘master’ data in place are still relevantly matched to them.

The levels which an efficient MDM activity system should consist of are shown in figure 4. The varied actions need to be continuously updated and followed in order to improve business intelligence in a banking environment.

---

**Figure 4: A Framework for Improving Business Intelligence through Master Data Management (Adapted from Lesole and Kekwaletswe, 2014)**

The framework, conceptualised as figure 4, shows the convergence of variedactions (MDM activity) which, when performed correctly, may lead to improved business intelligence. It is inferred that the data challenges and issues observed during the case study, and which are also typical in most banks, could be what hinders efficient business intelligence. These challenges could be leveraged by understanding the actions and processes as unraveled in the study. Master Data Management should not
be only about technology, processes and policies to be followed in integrating business data; it should also be about people that will follow these processes and policies.

Conclusion

This paper argued that a sound and rigorous master data management activity is vital to improving business intelligence. In other words, it is imperative that the master data management activity, which has different actions and procedures, be analyzed and understood before business intelligence could be improved. Analysis of the MDM, as social activity, was done using Activity Theory as a lens. The paper concludes that for any master data management actions to successfully work, the business should be the driver in choosing the appropriate data management procedures and the technologies that would leverage business intelligence; that is, consistently aligning business objectives, policies and rules within the actual master data management activity. Thus, an initial profound understanding of master data management, as an activity system, is vital to the delivery of real-time, on demand, quality data and information. The delivery of on-demand quality data may subsequently ensure improved business intelligence.

References


6. Langseth J. and Vivatrat N. (2003), 'Why Proactive Business Intelligence is a Hallmark of the Real-Time Enterprise: Outward Bound', Intelligent Enterprise, Pg 34-41


15.Morse JM. (1999), *Qualitative Generalizability*, Quality Health Research, Pg 5-6


