The Role of Data Quality in Business Intelligence – An empirical study in German medium-sized and large companies

KARIN HARTL, University of Applied Sciences Neu-Ulm
OLAF JACOB, University of Applied Sciences Neu-Ulm

In the digitalization age, the importance for companies to use internal and external data effectively and efficiently is increasing. Therefore, the application of BI technologies becomes more popular continuously. But German companies are still using BI technologies crudely and one-sided. It is believed, that one reason for this phenomenon is missing confidence in the BI system and tools. BI technologies are dependent on data and data quality as information gained out of the system are only as valuable as the data they are based on. Therefore, data quality is a success factor for BI and understood as the prerequisite for transparency and acceptance of the system. In this research, first BI supporting data quality and provisioning characteristics are defined and afterwards evaluated towards their development in German companies. The results indicate that the backlog demand of German companies in BI can indeed be caused by faulty master data, lack in data consistency and non-transparency of the BI architecture.

1. INTRODUCTION

Annually companies loose huge amounts of money due to poor data quality [Apel 2010; Xu and Hwang 2008]. Besides profit loss, the companies face trust issues of the data and information users which goes hand in hand with insufficient and decreasing usage of the Business Intelligence (BI) system [Apel 2010; Bachmann and Kemper 2011]. BI transformed the role of information technology in companies from a data storage solution into a method for solving strategic and organizational decision problems [Osei-Bryson and Ngwenyama 2014]. It is a process for analyzing data and the business environment and promises companies a more effective as well as efficient performance management in today’s rapidly changing business environment. This support BI provides for the company success and survival is based on the transformation of data into information, knowledge, decisions and actions [Turban 2011]. Therefore, data quality and data provisioning can be considered as BI’s most valuable assets. As the implementation of BI technologies are generally coupled with enormous costs, poor data quality and data provision can lead to a disastrous outcome for companies of all sizes.

The purpose of this research is to identify data quality and data provisioning characteristics which influence the successful deployment of a BI solution and to evaluate their development in mid-sized to large German companies. It is believed, that ensuring data quality and data provision throughout a company is the initial step for both, a successful BI and a successful performance management. First, a literature review has been conducted, identifying items which are considered as suitable to describe the data quality and provisioning components important for BI. Afterwards, these items have been assessed towards their development, analyzing responses from an industry survey. As data quality and provisioning are the prerequisite for successful BI, aim of this research is the evaluation of the BI related data quality status in German companies. Results are believed to support the development and successful implementation of BI in German companies.

2. THEORETICAL BACKGROUND

Through BI, data gathering, data storage and knowledge management are combined with analytical tools to present complex internal and competitive information to decision makers [Negash 2004]. Enterprise data and external data are transformed into actionable information about objectives, opportunities and positions of an organization [Wieder et al. 2012]. Accordingly, data are the main components of the
successful usage of BI systems and need to be provided and of high quality. Otto et al. [2007] defines data as simple and raw facts. The product manufactured out of data is information and this product needs to satisfy consumer needs. Just as regarding every product, the value lies in the eye of the beholder. Within a company, the consumers of information are the decision makers. Therefore, it is not recommendable to declare data quality as an only responsibility of the IT department. Before the data quality criteria – consistency, accuracy, completeness, relevance and timeliness – can be ensured by the IT department, the data and information users need to define their information need [Apel 2010; Otto et al. 2007]. Both IT and departments have to work on the indemnification of high-quality data which then can be used to produce actionable information which serve as the key enabler for efficient processes [Bachmann and Kemper 2011; Otto et al. 2007].

BI is considered the key to leverage this value in data accumulated in and around the enterprise [Nazier et al. 2010]. It is an analytical process providing insights and recommendations for managers and decision makers [Aho 2010]. Therefore, BI is a crucial element of performance management which is also known as Business Performance Management (BPM) or Corporate Performance Management (CPM) [Turban 2011]. In this research, we focus on the CPM concept as it represents the strategic deployment of the BI solution. CPM is defined by Gardner as “an umbrella term that describes all processes, methodologies, metrics and systems needed to measure and manage the performance of an organization” [Becker et al. 2006]. It needs BI to work effectively on accurate, timely and high quality data and BI needs CPM for a purposeful commitment [Becker et al. 2006]. High data quality and data availability are the initial requirement to transform the raw data into actionable information [Bachmann and Kemper 2011]. But information based decisions are only as valuable as the data they are based on [Bachmann and Kemper 2011]. Therefore, data quality and provision are required for successful BI and for an effective and efficient CPM. Accordingly, data quality is the most valuable asset for a successful BI and even a successful company.

3. RESEARCH CONSTRUCT

In order to identify the data quality related key characteristics which impact BI and through BI CPM a review of academic literature was conducted. Besides well-established monographies discussing the subjects Data Quality, BI, CPM and Management Information Systems key related Information Systems and Business Journals have been evaluated. Furthermore, several academic databases such as Business Source Premier (BSP), Emerald Insight, ACM Digital Library and AIS e-library have been searched using subject related keywords as “Data Quality”, “Business Intelligence”, “Data Quality Success Factors”, “Business Intelligence Success Factors”, “Business Intelligence and Corporate Performance Management” and “Determinants of Data Quality”. The subject related elaborations have been assessed regarding their interestingness for this research purpose. Afterwards, the important elaborations have been evaluated and data quality items significant for BI have been identified. In total, 8 items have been detected as most suitable to describe the importance and impact of data quality and provision for BI. Table 1 provides an overview of these items and the corresponding literature they are based on.

The first item Data consistency describes Single Point of Truth storage solutions which provide data quality throughout the enterprise. Data from many data sources gets stored together in an enterprise wide Data Warehouse often referred to as Enterprise Data Warehouse (EDW). This doesn’t imply that the company has only one data
storage solution. EDWs are a consolidation of all data sources into one database system and can consist of several physical Data Warehouses and/or Data Marts. The ideal case is, that EDWs prohibit bypass-reporting, ensure on-time delivery of data and guarantee high data quality and comparability of performance indicators. EDWs are the “Single version of the Truth” in a company.

<table>
<thead>
<tr>
<th>Item</th>
<th>Related Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data consistency (“Single Point of Truth”)</td>
<td>Bachmann and Kemper [2011], Turban [2011], Morabito et al. [2011], Lang [2015], Hartl et al. [2016]</td>
</tr>
<tr>
<td>Data integrity during simultaneous use</td>
<td>Negash [2004], Bachmann and Kemper [2011], Lien Mbp and Jacob [2014], Hartl et al. [2016]</td>
</tr>
<tr>
<td>Traceability of BI master data changes</td>
<td>Seufert and Oehler [2011], Lang [2015], McKnight [2014], Hartl et al. [2016]</td>
</tr>
<tr>
<td>24/7 operation of the BI system</td>
<td>Turban [2011], Morabito et al. [2011], Andriole [2006], Hartl et al. [2016]</td>
</tr>
<tr>
<td>BI architecture is documented</td>
<td>Wieder et al. [2012], Kemper et al. [2010], Lien Mbp and Jacob [2014], Hartl et al. [2016]</td>
</tr>
<tr>
<td>Only mandatory BI tools are used</td>
<td>Akhavan and Salehi [2013], Lien Mbp and Jacob [2014], accenture [2007], Hartl et al. [2016]</td>
</tr>
<tr>
<td>Responsibilities for the BI-development are clearly shared between the company’s departments and the IT and communicated throughout the whole enterprise</td>
<td>Bachmann and Kemper [2011], Hartl et al. [2016]</td>
</tr>
<tr>
<td>User rights to access the BI system are defined</td>
<td>Eckerson [2008], Kemper et al. [2010], Lien Mbp and Jacob [2014], Hartl et al. [2016]</td>
</tr>
</tbody>
</table>

Implementing a BI systems is aimed on making data and the contained information available for many if not all people throughout the company. To ensure acceptability and usage of the BI system all over the enterprise it is an essentiality that reliability, integrity and security of the BI system is given at all times. Employees need to be able to use the BI system simultaneously without facing problems of integrity. Therefore, Data integrity during simultaneous use has been identified as important data quality and provisioning characteristic of BI.

The homogeneity of the business environment is increasing continuously, making it important to standardize common data elements like master data. Master data is quality data. Once defined, it can be used constantly and it ensures the quality of the common data elements used in the BI system. Therefore, it is crucial for the success of the BI system to ensure a consistent master data management which supports system comprehensive integrity, the creation of a single consistent few on the enterprise and uniform reporting. Traceability of BI master data changes comprises the importance of master data for the BI system and problems caused by changes of the BI relevant master data (e.g. the change of dimensions as regions, product groups, article group, and so on). The different requirements departments have on master data and the need for transparency only can be guaranteed if changes in the master data are traceable. Furthermore, transparency facilitates trust in data and in the BI system. The 24/7 usage of the BI system recaps the need for on-demand business monitoring, performance analysis, the permanent availability of performance measures and easy access technology. Only if guaranteed, the users have the opportunity to work on quality data when needed and use BI features like real-time analysis and ad-hoc queries continuously.

The documentation of the BI architecture has two important aims. First, if documentation is demanded in a company, uncontrolled growth of BI applications can
be prevented. Second, transparency is created making it easier to adjust and work with the BI system. Transparency is also improved through the usage of mandatory BI tools. Required tools and reports can be discussed and selected before the implementation of the BI system. As employees tend to prefer familiar products, a standardized set of BI tools supports the acceptance of the BI system by its users.

**Responsibilities for the BI-development are clearly shared between the company’s departments and the IT and communicated throughout the whole enterprise** is a very important success factor of BI. For BI to be effective and efficient it needs cooperation between the departments who are going to work with the solution as well as the IT department, which installs and develops the BI system. The departments have to define specific requirements on the BI solution for the IT department to purposefully deploy it. Without a focused deployment of the BI system, an absence in usage of the BI application by the decision makers is most likely guaranteed.

An important challenge for successful BI is getting the users to adopt and work with the tool. The entirety of the BI systems is usually very complex and therefore discouraging for the users. **Defining user rights to access the BI system** is a counteraction. User get access to the BI system and tools according to their need and ability. Only a few users need so called power-user status to work with complex queries. Most users are casual and don’t need to be familiar with the background on how the BI tool and the databases behind it work. To personalize the user rights might be more complex to begin with, but in the long run it relieves the IT department and supports the acceptance of the tool throughout the enterprise by easier handling.

### 4. RESEARCH METHOD AND RESULTS

The rapidly changing business environment intensifies the need for companies to react in a timely manner on high quality information [Hannula and Pirttimäki 2003]. Therefore, the effective deployment of BI solutions continuously gains in importance. But German companies still have backlog demand on the subject. Just every second company in Germany is using BI related management tools and the companies already using BI have room for improvement [arcplan Information Service GmbH, TNS Emnid 2015]. Data Quality and provisioning is the prerequisite for successful BI and the evaluation of the BI related data quality status in German companies helps to identify the areas for improvement.

#### 4.1 Data Collection

A questionnaire in German was created for data collection consisting of the above mentioned items. Each item had to be rated on a 5-point Likert scale anchored at “totally applies” and “doesn’t apply at all”. An additional definition point for the mid-stage as been added stating “applies half and half”. Furthermore, the respondents have been given the opportunity to choose “no answer”. This minimized the distortion of the sample by not forcing the interviewees to answer the question if they were not aware of the answer.

The data collection has taken place in December 2014 until March 2015 using telephone interviews and an online questionnaire. Subjects were German companies who use BI for supporting their performance management. Hence decision makers from management, controlling and IT were addressed. In total 169 questionnaires were completed resulting in a response rate of 11.3%.
4.2 Results

The frequency distribution of the survey results are visually displayed in Figure 1. Especially Data integrity during simultaneous use, User rights to access the BI System and 24/7 operation of the BI system are evaluated as already applied in the companies interviewed. 75% to 88% of the participating enterprises specify to have no problems in these BI related data quality and provisioning disciplines. But 21% of the interviewees conceded that their BI system is only partially or isn’t at all available 24/7 and 13% specify not having defined user rights to access the BI system.

The areas Data consistency, Traceability of BI master data changes, BI architecture is documented, Only mandatory BI tools are used and Responsibilities for the BI-development are also evaluated rather positive but between 30% and 40% of the interviewees admit problems in these areas. Especially the documentation of the BI architecture seems rather not common in German companies. 40% of the participants recognize that there BI architecture is only documented partially or even not at all. 34% of the respondents acknowledge problems in the traceability of master data changes and 33% admit difficulties regarding the cooperation between the IT and the departments during BI-development. Furthermore, 32% of the interviewees stated that they are not working with mandatory BI tools. The prerequisite of a trustworthy BI is data consistency. The survey results indicate that 30% of the companies interviewed still have difficulties providing reliable data.

In the sample there were nearly the same amount of interviewees from the IT-department (61 persons) and from the controlling department (60 persons). Assuming that controllers are the major user group of BI, it makes sense to compare their answers across the eight BI related data quality items with the responses of the IT-department. The results are shown in Figure 2.
It can be noticed, that the items *Data integrity during simultaneous use* and *User rights to access the BI system are defined* are evaluated similarly by controllers and IT-specialist. Data integrity and user rights are fully applied and obviously no problem areas in the usage of the BI system. The same applies for the item *24/7 operation of the BI system*. The majority of the interviewees states, that this is a given characteristic of the BI system in place.

**Fig. 2** Responses by Controlling and IT department

ICIQ 2016, Article 4, Publication date: June 22nd, 2016
For all the other data quality items the distribution of the answers over the 5 point Likert scale is more widespread in both departments, esp. for the items Data consistency, Traceability of BI master data changes and Responsibilities for the BI development. In general, this indicates that the current status of these items in German companies is differently perceived by the IT-department and controlling.

Regarding the item Responsibilities for the BI development it can be noticed, that nearly 8% of the interviewees answered with “doesn’t apply at all” and 5% answered with “no answer”. This points out that clear roles and responsibilities for the BI development are still a problematic subject in German companies.

The answers to the item Traceability of BI master data changes deliver similar results. Nearly 9% of the interviewees answered with “doesn’t apply at all” or “no answer”.

Major differences in the evaluation between controlling and IT can be seen in the item BI architecture is documented. Controllers evaluate the existence of a documented architecture more positively than the IT department.

5. DISCUSSION

The results show that 3 out of the 8 defined data quality and provisioning characteristics are already applied in the companies interviewed. The other 5 characteristics have a positive tendency as well, but a non-negligible part of the research participants admit to having problems in these areas. Data consistency defines the usability and trustworthiness of the data within a company and is often taken for granted. But in fact, in many companies there are still different versions and truths of the same data between the departments resulting in redundancies. This leads to a lack in confidence in the data and the BI system. Therefore, professional BI systems need a consolidated database providing the same data throughout the company. 30% of the interviewees recognize not having data consistency in their enterprise, making it difficult for decision makers and employees to trust in data and the BI technology.

This problem is accompanied by the insufficient traceability of BI master data changes. Master data is data shared across a company’s systems and used to classify transactional data [Kopcke 2008]. Correctly managed master data improves and supports decision making tremendously by providing information consistency and the reduction of data redundancy. This again creates data consistency and minimizes the mistrust in data, information and the BI system. But the results reflect that managing master data is still shied by German companies. This can be due to the fact, that master data management is one of the most difficult, time consuming and expensive challenges for businesses [Messerschmidt and Stüben 2011]. However, German companies have to face this challenge to ensure data consistency, followed by the effective implementation and usage of the BI system and the competitiveness in today’s rapidly changing business environment.

Another success factor for BI projects still underestimated by German companies are the clearly defined roles between the departments and the IT for development of the BI system. It is not of use to only invest and implement into a BI system if the departments, decision makers and end users are not willing to work with the tool. Therefore, it is widely known, that the implementation of BI technology needs a strong business support throughout the company. Besides providing the purpose for the BI system, the business support is responsible for the acceptance and usage of the solution. This can only be guaranteed if the end-users clearly define their information need and demands on the BI software. A comprehensive cooperation between IT and departments is crucial for the success of a BI project. German companies might consider instating a
BI project team, combined of the IT department and business users. Furthermore, a project leader with business and IT background seems necessary to create a communication framework for mutual understanding.

Another problem area identified is the documentation of the BI architecture, which is a necessity for transparency. The BI architecture is the framework of the BI project organizing data, information management and technological components. It structures the whole BI project towards supporting the CPM. Only if documented, the adaption of the BI system to the fast changing business environment is guaranteed and the long-time support and success of the BI project secured. German companies obviously need to take this task more seriously.

An often underestimated subject to the success of BI projects is the definition of mandatory BI tools for the end users. This course of action is required, because of the complexity of the BI systems and its tools. Not every tool and/or application is appropriate for each user and task. Even though most BI software providers advert with easy to use user interfaces, offering the end users too many tools to work with causes them to be unfocused. They are not sure where to start with and are most likely overstrained which can lead to rejection of the BI system in general. Therefore, it is essential that BI systems provide functionalities and applications meeting the end user requirements. German enterprises need to realize the importance of this task, to antagonize the one-sided and poor usage of BI systems.

Data integrity during simultaneous use is ensured in most of the companies interviewed. As probably every mid-sized to large German company is working with different data based systems, they are obviously familiar with the assurance of data quality during simultaneous use. Still, regarding the above discussed questionnaire results, does this not implicate that data integrity is provided throughout the whole enterprise. The positive evaluation of this characteristic only reveals that German businesses are confident to provide accurate and consistent data during the simultaneous use of the system. This answers are in opposition to the evaluation of the items Data consistency and Traceability of BI master data changes. The clashing of results indicate a positive self-assessment of the companies by trend. Consequently, German businesses should evaluate their BI projects and systems more critically and not close their eyes to avoid challenges.

The given uninterrupted service of the BI system underlines the high technological development of German companies. But it also indicates a shifted focal point for the implementation and usage of BI projects. For CPM to be effective and efficient, it is important that the decision makers can access the information when needed. However, information is only as valuable as the data it is based on. Data consistency and master data management are primary challenges. The results facilitate the conclusion, that German businesses avoid these time and resource consuming tasks in favour of faster established solutions like the 24/7 availability of the system.

Furthermore, the results indicate that the concept of personalized user rights is well established in German companies. True to the saying of Eckerson [2009] “one size does not fit all”, BI tools need to be fitted to the type of person using it. But it is not enough to restrict user rights. For the BI system to be accepted and thoroughly used by the employees it also needs mandatory BI tools and documented BI architecture.

6. CONCLUSION

In general, the analysis of the 8 identified data quality items shows a positive tendency towards their development in German companies. Nevertheless, in 3 of the 8 evaluated data quality items we notice the need for further research to explain the widespread
answers with the tendency to a neutral or even poor assessment of the items. The findings suggest the following propositions for additional research.

Research proposition 1: Traceability of BI master data changes is an ongoing challenge in master data management.

Master data management is a permanent issue in companies. It is seen as time-consuming and the value is often not clear. However, master data management is an important task in data quality management and the BI success, so that new standards as ISO 8000 are continuously developed and offer solutions on that specific task. Our results actually show, that traceability of BI master data changes is evaluated by 34% of the interviewees as not existing or as exiting in a minor degree. Especially the high amount of “no answer” indicates, that the question was either not fully understood or the concept of traceability of BI master data changes is unclear. Traceability is a mandatory feature in BI planning applications (e.g. financial planning) and is necessary to fulfill compliance requirements. Future research should analyse the value of traceability of BI master data changes for BI data quality and should identify issues in achieving a good traceability.

Research proposition 2: Data integrity is a prerequisite for data consistency.

It is interesting to see, that the item Data integrity during simultaneous use is evaluated more positively than the item Data consistency. Trying to interpret this results one can state, that data in many of the interviewed companies still remain consistent in spite of the simultaneous use through different users. But only a few companies have yet achieved the “Single Point of the Truth”. We can interpret this fact as that German companies in the past took major efforts in realizing good data integrity. This priority for data integrity in the past has to shift to data consistency in the future. Coming research should analyse the relationship between data integrity and data consistency in depth and should identify reasons for the prioritization of data integrity. In addition, future research should test the proposition if data integrity is a prerequisite for good data consistency.

Research proposition 3: The different perception of data quality between users and IT is caused though poor communication between those departments.

If we analyse how different interviewees from the controlling department and the IT department answered the 8 data quality items, we see the greatest differences between those groups within the item BI architecture is documented. Interestingly, controllers evaluate the quality of documentation much more positively than the IT-department. One interpretation of that fact can be, that controllers see the documentation more in sense of an end-user documentation and the IT-department more in the sense of a technical infrastructure-based documentation. If so, we identify an issue in achieving a good technical documentation of the BI architecture. Future research should focus on the different perceptions between user departments and IT department and test the assumption that the differences are caused by a poor communication between these departments.

Summarizing the results of the current research, we identified three major challenges of data quality in German companies regarding BI - non-traceable master data, lack in data consistency and non-transparency of the BI architecture. The future research directions identified can lead to an even better understanding on how good data quality impacts the usage of BI applications. Furthermore, companies can use these and future results to increase the success of their BI solutions.
REFERENCES