Wallner Martin[1]

Peráček Tomáš

DATA MANAGEMENT IN INDUSTRIAL COMPANIES: THE CASE OF AUSTRIA

Abstract: Data has become the most critical asset that companies have. Companies often lack the experience to generate the highest possible value from their data. Most current studies on this topic were not restricted to the participants' geographical location or industrial sector. The originality of this research is that a survey is sent to industrial companies in Austria solely. This article intends to answer the following research question: 'What is the current state regarding data management in Austrian industrial companies? It should start a discussion in the community if Austrian and European companies are behind their international competitors in company-wide data management. It should also give an overview of standardized KPIs, responsibilities, problems, data storage technology, and decision-making. While most industrial companies in Austria stated that their KPIs are standardized within the company and strategic decisions are based on data, Austrian companies lack behind their international competitors in terms of C-level positions for data management and a company-wide data management strategy.

Keywords: Data management, Data Quality, Industrial companies, Data, Big Data, KPI

1. Introduction

'Knowledge is power; this quote by the English politician and philosopher Francis Bacon is already 400 years old. However, in times of digitalization, it is more relevant than ever (Bacon, 1597).

According to IBM, data is the most critical asset companies must manage. While they have a long experience managing assets such as properties, plants, equipment, cash, and intellectual property, they often lack experience in managing data successfully to generate the highest value possible from their data (IBM, 2016).

Since data affects more and more areas of our daily lives, such as communication via e-mail and messenger services, documents and

photos, and smartphone apps, 2.5 quintillion bytes are produced by humans daily (Techjury, 2021).

This value is substantial, but even more remarkable is that 90% of the data have been created in the last two years, mainly due to the growing amount of different data sources, which create data with significant frequencies and volumes and a reduction of the storage costs. An important innovation in this area was the cloud infrastructure started by big IT companies that further ease data storage. Companies aim to collect as much data about their processes, machines, competitors, and customers as possible (IBM, 2018).

An end to the constant data volume growth is not in sight, according to leading

---

[1] Corresponding author: Wallner Martin
  Email: wallner4@uniba.sk

companies in the field, such as Oracle, one of the most extensive database software developers. The Internet of Things will enhance this trend (Oracle, 2018).

The objective of the present study is to investigate the current state of the data management process in Austrian industrial companies. It should be investigated if KPIs are standardized in the company, who is responsible for company data, and if a company-wide data management is implemented. This study will also analyze what storage technologies are used and which industrial sectors invest the most in data management. Furthermore, some hypotheses are tested, such as if companies with implemented data management are more successful than those without or to what degree companies with dedicated responsibility for data also start Big Data projects. The results of this study will also be compared with international studies to assess how Austrian industrial companies rank compared to companies in other countries and the main areas where they lag behind their international competitors. This article contains parts employed by the author for his master's thesis.

Following the introduction, section 2 contains literature on data management and studies that engaged with that topic. Section 3 describes the research methodology. Section 4 presents the results, discusses the findings, and Section 5 concludes the work.

2. Literature Review

2.1 Data Management

According to the Data Management Association (2009), responsible for developing concepts and best practices, data management is the company function engaged with developing, implementing, and maintaining rules and standard operating procedures regarding data. This function should control and increase the value of the data and bring data to the proper recipients (DAMA, 2009).

The Data Management Association (2009) subsumed strategic goals of successful data management:

- information should be adjusted to the needs of each stakeholder in the company
- acquisition, storage, and security of data
- ensuring the quality of data
- accuracy
- integrity
- actuality
- relevance
- clarity
- ensuring confidentiality and access security from external parties
- maximizing the effectivity and the value of data

Without a comprehensive data management, it would not be possible to share information on a wide scale between different departments or branches because there would be no central data administration. That would lead to a significant competitive disadvantage because data would not be delivered to the right place or delivered too late. Another problem would be the cost disadvantage because evaluations that have already been made would be repeated, and learnings would be forgotten (Gordon, 2013).

Publications that have engaged with the deployment and experience of data management are the BARC Study Datenmanagement in Wandel [Data management over time] (BARC, 2014) and the BARC Big Data Analytics Survey (BARC, 2014) or the Industrial Analytics Report 2016/2017.

According to the Industrial Analytics Report 2016/2017, 68% of the companies from the worldwide participants had a defined data strategy (IOT Analytics, 2016).

However, in the fast-moving IT world, a survey from 2014 or 2016 could already be outdated and not representative of companies' current state of data management.

Furthermore, these publications were not restricted to the participants' geographical location or industrial sector. This approach is understandable to reach a high number of participants. However, differences between specific industrial sectors or regional differences are tough to define.

2.2 Data Quality

According to the quality management standard ISO 9000 from 2015, quality is defined as the degree of fulfillment of customer requirements. Therefore, quality is no absolute measurable size and can have a product, information, or process as the final result. If we assume that the information reader is the customer, the goal should be to fulfill their needs in terms of quality demand. This makes quality measurable (Brüggemann & Bremer, 2015).

Wang and Strong (1996) added to the definition of the ISO 9000 that the quality has to be assessed based on suitability.

Furthermore, they have defined four categories to assess data quality:

- Intrinsic data quality: accuracy, objectivity, and error-free data
- Contextual data quality: Is the data complete in order to be used reasonable
- Presentational data quality: Is the data arranged clearly in order to be easily understandable
- Access related data quality: Is the data easily findable, accessible, and editable

In times of a sharp increase in data volume, companies need to filter out the data that does not fulfill specific quality criteria and continue the analysis only with high-quality data to ensure high-grade data-based decisions. (Ramasamy & Chowdhury, 2020).

Otto stated that if data is considered a resource, it needs to have a certain quality, life cycle, and value (Figure 1). The life cycle impacts the data quality, which subsequently determines the value of the data for the company. Managing these steps is crucial for companies to achieve data with a high value (Otto, 2015).
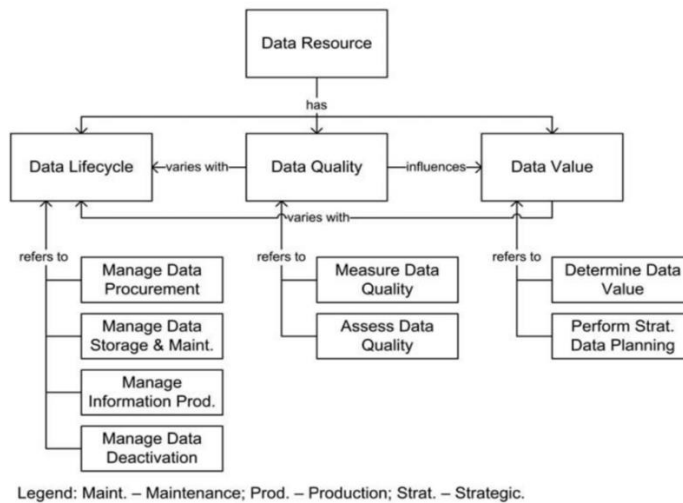
Legend: Maint. – Maintenance; Prod. – Production; Strat. – Strategic.
Source: Otto (2015)
Figure 1. Elements of data

The correct data must be extracted from a massive amount of available data to achieve professional data management. To achieve this, wrong, needless, or data with low quality needs to be separated.

In 2014 integration of heterogeneous data sources was mentioned as the most significant problem in data management. In the same study, 41% of the participants reported that information has a high value in the decision-making process and 61% said that this value will go up significantly in the upcoming years (BARC, 2014).

According to the Bima survey 2017/2018, only 29% of the participants have stated that analyzing parameters are standardized throughout the company.

2.3 Big Data

Garnter, one of the leading information technology research and consulting agencies, defines Big Data as "High-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making (Gartner, 2022)."

The 3 V factors are often used to restrain which data are considered Big Data (Schroeck et al., 2012):

- Volume
- Variety
- Velocity

New definitions add veracity and value. To express that the data quality is essential for further analysis and that data has a very high value for companies (Gandomi & Haider, 2015).

Big Data offers information that has not been available before and allows new applications such as predictive maintenance applications of Big Data, precise analysis of customer behavior, or the automated adjustment of in-line processes (Oussous, 2017).

In the past, those analysis methods were mainly used to describe events that happened in the past with the available data, and new data analysis methods for future events can be predicted or even changed (Figure 2).
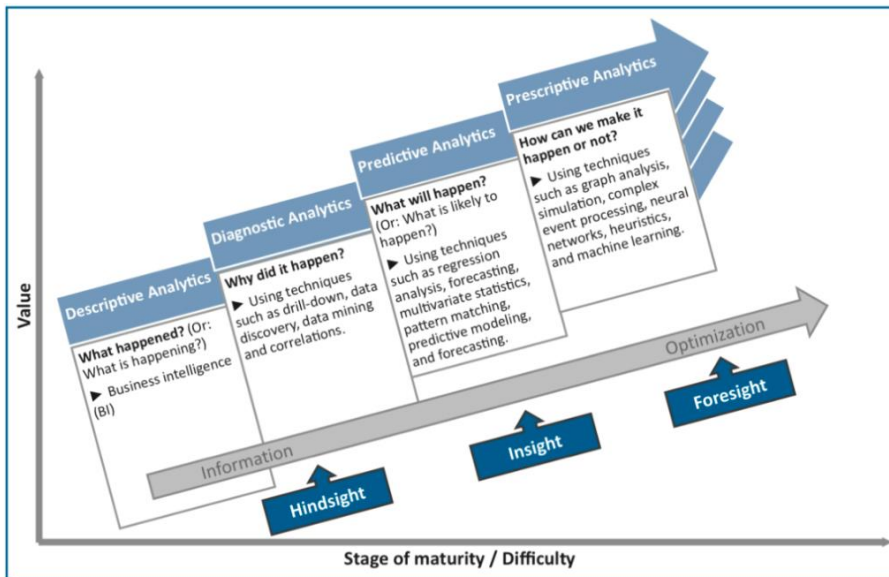

Figure 2. Analysis methods are classified by value and difficulty of analysis
Source: Romeike & Eicher, 2016

2.4 Infrastructure

The correct data must be extracted from many available data to achieve professional data management. To achieve this, wrong, needless, or data with low quality needs to be separated. Although modern data infrastructures (e.g., data lakes) have almost infinite storage space and allow a very fast saving of data, that data must not be included in search queries (Khine & Wang, 2018).

Traditional databases organize and store data electronically. They are typically set up for a specific purpose and department. Data warehouses must be implemented if data from heterogeneous sources are extracted. The main difference between operative databases is that the data is seldomly erased from data warehouses, while operative databases have one (the newest) data value (Freiknecht & Papp, 2018).

Prakash (2020) states that enterprises with Big Data prefer data lake architectures since they are specifically designed for machine learning and artificial intelligence applications. Data lakes work with the schema-on-read principle that allows storing vast amounts of data because not every writing process is checked, and the row data is stored. This increases the speed of data writing but makes it more difficult to access data for regular users. A large pool of data is available and quickly accessed for complex analysis. Using data lakes in combination with a cloud infrastructure is even more flexible since the storage can be adjusted quickly.

41% of the enterprises indicated in 2014 that they use modern data architectures such as centralized DHW or Data Lakes. 10% of the companies stated that a central data strategy exists and that the biggest challenge in data management is insufficient resources (BARC, 2014).

2.5 Responsibilities

In smaller companies, the information technology department is responsible for data management. However, a dedicated data controller in each department must be nominated even in simple organization charts. The primary contact person for all data topics ensures the defined data quality. The data controller should always report to the management to ensure that the information helps the company achieve an advantage over its competitors. If a data manager is assigned, the main tasks are setting up the IT infrastructure for the data management and the responsibility for effectively using the data. In bigger companies, the management of these tasks is done by the Chief Information Officer (CIO), who directly reports to the Chief Executive Officer (CEO) and is also part of the company's general management. His main task is to ensure that the data management efforts align with the company strategy. Two other new C-level positions in big cooperations are Chief Data Officer (CDO) and the Chief Digital Officer (CDO). The Chief Digital Officer's role is to use digitalization for the company's business development. If a Chief Data Officer is assigned, it is a sign that the company takes data management very seriously. In that case, he is the responsible employee for data extraction and processing. The role of the CIO is more of a role as the Information Technology department manager (Krcmar, 2015).

In the 2018 Global Data Management Report, which only considered participants in the USA, 10% of the companies already employed a CDO, 20% a CIO, and 14% had a specified data controller (Experian, 2018).

2.6 Decision-making process and problems with data management

According to DAMA, the Global Management Community, one of the main functions of data management is to ensure

data for decision-making processes. Ghavami states that the possibility of making faster and more reliable decisions is also one of the main reasons for companies to invest in data management (Ghavami, 2020).

The problems with data management can be segregated into two big groups. Technological issues that emerge from complex and large amounts of data and their integration into other systems. Organizational issues are caused mainly by bad data quality (incomplete, missing, or wrong data sets) and the misinterpretation of the found model, according to Fayyad (Fayyad et al., 1996). In 2016 a study amongst data managers found that data quality, transparency over the data usage, and redundant data were the biggest problems in data management at that time (Otto, 2016).

According to the BARC study 2014, the biggest problem concerning data management was insufficient resources and the lack of willingness to change. 41% of the participants of this study have stated that information has a high impact on the decision-making process, but 61% have indicated that this number will have a high increase in the upcoming years (BARC, 2014).

2.7 Comparision to past studies

Surveys such as the Industrial Analytics Report 2016/2017 (IOT Analytics, 2016), the biMa Study 2017/2018 (BARC GmbH, 2018), or the 2018 Global Data Management Benchmark Report (Experian, 2018) were not restricted to the geographical location or industrial sector of the participants. This approach is understandable to reach a high number of participants. However, differences between specific industrial sectors or regional differences are tough to define.

Surveys from the BARC study 2014 or BARC Big Data Analytics Survey 2014

(BARC, 2014) may be outdated in the fast-moving information technology.

The survey of the article focuses on industrial companies in Austria to elaborate on the current state of this sector in a defined country and to allow a comparison to different industrial sectors and countries.

3. Methodology

Considering the results of the literature research and the research question:

'What is the current state regarding data management in European industrial companies?'

A questionnaire was created, which needs to be answered by 40 industrial companies in Austria to be statistically robust. Based on the results, descriptive and inference statistics methods were used to analyze the survey data. It is planned to check the interdependence between variables with the chi-square test. Furthermore, comparing these results with worldwide- and DACH-region studies should point out regional differences and disadvantages of Austrian industrial companies.

To investigate the current state of data management in Austrian industrial companies, the subquestions have been examined:

'What percentage of Austrian industrial companies have implemented a company-wide data management?'

'How strong is the dependence between industrial companies that implemented a company-wide data management and the turnover of those companies?'

'To what degree are evaluation perspectives and Key Performance Indicators (KPIs) standardized?'

'To what extent are Big Data projects not started because of not standardized KPIs?'

'Which technology is used for data storage?

'How strong is the correlation between companies with a modern data infrastructure

and the implementation of Big Data projects by those companies?'

'Which department is mostly responsible for data in Austrian industrial companies?'

'To what degree do companies with a dedicated responsibility for data also conduct Big Data projects?'

'Which industrial sectors in Austria invest the most resources in data management?'

'Which is the biggest problem regarding data that Austrian industrial companies have?'

'To what extent does a company-wide defined data management protect companies from human errors regarding data?'

'To what degree strategic decisions of Austrian industrial companies are based on data?'

'How strong is the dependency between the implementation of strategic decisions based on data and the turnover of a company?'

Definition of an industrial company:

For this questionnaire, companies from the primary or secondary sector that process or extract raw material were considered industrial companies. The tertiary (service offering sector) or the quaternary (information sector) industry branches and commercial companies were not considered (Madhushree et al., 2019).

Data collection:

The survey was sent to industrial companies in Austria via the online survey tool SurveyMonkey. The survey was conducted in German and translated for this publication.

Survey:

The preliminary questions set up relationships between specific questions and basic information about the company, such as the turnover, employee number, or branch.

The first specific section of the questionnaire inquired if companies have a company-wide data management implemented. This question was asked as a closed question to force a clear statement since this is one of the most critical questions of the survey.

To make sure that the participants of the study had the same understanding of this question, the following definition has been added:

'Data management is a company function responsible for developing, implementing, and maintaining rules and operational sequences that control and increase the value of data and information and make sure that it is delivered to the right persons' (DAMA, 2009).

With the stated turnover and the answer to the question about the implemented data management, a correlation analysis between those factors was possible.

The data quality aspect was considered in a closed question about the standardization of Key Performance Indicators and evaluation perspectives throughout the company. with data management problems. It was investigated if there is a relationship between standardized KPIs and experience in Big Data projects.

The type of storage technology was asked as a semi-closed question and allowed to choose given answers of old and new technology or the specification of a custom answer because of the high bandwidth of possible answers. A chi-square test has been conducted to investigate if there is a relationship between the storage technology and the implementation of Big Data projects.

The company's internal organization and responsibilities were queried in the following section. Furthermore, the dependence between a dedicated responsibility for data and the implementation of Big Data projects has been investigated.

The next question investigated if there are single sectors that invest more resources in data management than others.

The following section covered questions about the decision-making process and the problems with data management. Furthermore, the factors companies see as

problems with data management have been investigated. The answers to that question were superficial and covered various organizational and technical areas. There was an option for a custom answer. Subsequently, a yes or no question inquired if the strategic decisions in Austrian industrial companies are made based on data. With the data, two correlation analyses were performed, and it was investigated if there is a relationship between a company-wide data management and the reduction of human errors and how strong the dependency between the decision-making of strategic questions based on data and the turnover of the company is.

Execution of survey:

The survey had 148 participants, of which 141 were counted. Seven participants were not industrial companies according to the specifications of the study. The response rate was 32%. The number of participants was considered enough to represent the industrial sector in Austria.

In the introduction section of the survey, participants were asked in which sector their company is mainly active. The survey was consciously sent to all branches from the primary and secondary industrial sectors to get representative results. The units had an average of 5.88% participants. The only outlier was the metal industry, with 24.82% of the participants. However, this sector also represents one-quarter of the industrial turnover in Austria and consists of 150 big enterprises (Metalltechnische Industrie, 2021).

Subsequently, the participants were asked how many employees their company has at all locations in Austria (Figure 3) and what is last years turnover (Figure 4). The question was asked to only consider employees and turnover from business in Austria so that big international companies with a small representation in Austria did not falsify the results. Most companies (49.26%) that participated in the survey had employees in the range of 201-1000, followed by companies with 51-200 employees (20.59%) and companies with 1001-5000 employees (18.38%). Regarding the turnover, most companies have a turnover of 100-500 million Euro (35%), followed by companies with a turnover of 50-100 million Euro (25%).

Source: Compiled by the authors.
Figure 3. Distribution of employees in all branches in Austria, %
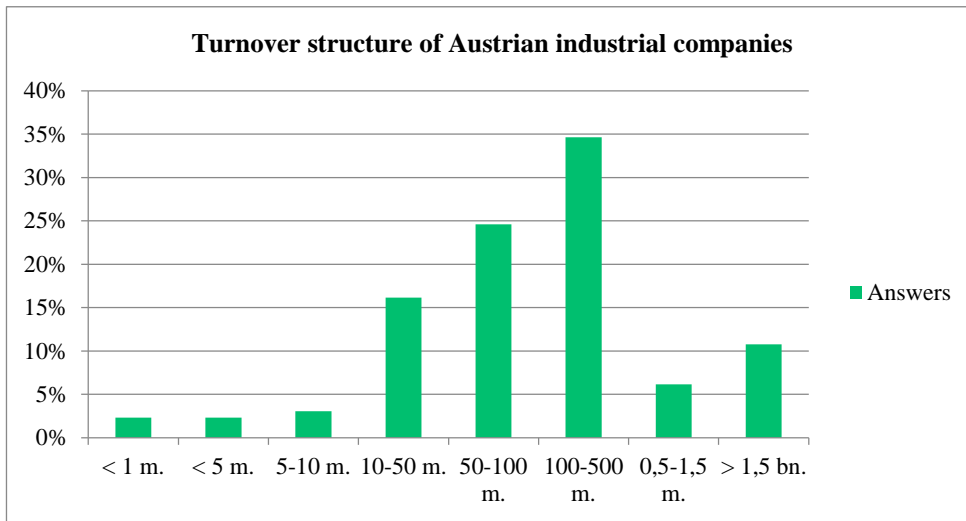
Source: Compiled by the authors.

**Figure 4.** Distribution of companies by realized turnover in Austria, %

4. Results and Discussion

The evaluation of the survey led to the following results, which form the basis for answering the set of questions:

What percentage of Austrian industrial companies have implemented a company-wide data management?

77.05% of the participants had a data management system implemented in specific business units or company-wide, out of those 38.26% had the same standard throughout the company. 3.28% stated that there were no fixed standards and that also, in the future, no standards were planned (Figure 5).

The question about a company-wide data management is answered by 38.26% of the Austrian industrial companies with the answer: 'Standards are already part of the corporate process.' 2014 a similar question about a company-wide data strategy was asked in the BARC study 2014. Only 10% of the companies in the DACH region stated that there is a company-wide data strategy for all data implemented (BARC, 2014).
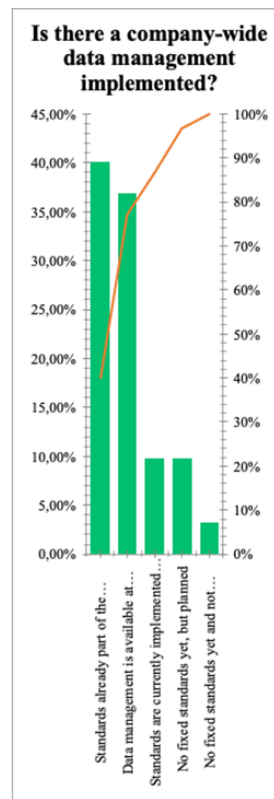
Source: Compiled by the authors.

Figure 5. Questionnaire - Company-wide data management implemented, %

So, in five years, there has been a significant increase of companies that stated that data are handled and organized company-wide. In the Industrial Analytics Report conducted worldwide, 68% of the companies noted that a company-wide data strategy is defined. Austrian industrial companies have caught up regarding a defined data strategy since 2014. However, they are still below the international average (IOT Analytics, 2016).

How strong is the dependence between industrial companies that implemented a company-wide data management and the turnover of those companies?

Forty-four participants reported that a company-wide data management is implemented. Of these companies, 28 had over EUR 50 million in turnovers and were classified as large-scale enterprises according to the Austrian Federal Economic Chamber (WKO, 2019). Eleven participants had a turnover between EUR 10 and 50 million and were classified as middle-size companies. Only three companies had a turnover of under EUR 5 million.

A chi-square test was conducted to test whether the implementation directly depends on a company's turnover, and the p-value was calculated (Table 1 and Table 2).

Table 1. Observed frequency data management/turnover

Observed frequency		
	Company-wide data management	No Company-wide data management
High turnover	30	81
Low turnover	14	16

Source: calculated and compiled by the authors.

The H_0 hypothesis is stated as no dependency between turnover and company-wide data management.

The calculated p-value is 0.039. That value is smaller than the set significance level of 5% (0.05); therefore, set hypothesis H_0 is rejected, and a dependency between the

turnover and the implementation of a company-wide data management can be proven.

Table 2. Expected frequency data management/turnover

Expected frequency		
	Company-wide Data management	No Company-wide data management
High Turnover	34.6	76.4
Low Turnover	9.4	20.6
Grand Total	44.0	97.0

Source: calculated and compiled by the authors.

However, it cannot be ascertained if the data management leads to a high turnover or if it allows allocating resources to establish a company-wide data management. Future works should correlate turnover history to the company's progress with data management to find a possible correlation.

To what degree are evaluation perspectives and Key Performance Indicators (KPIs) standardized?
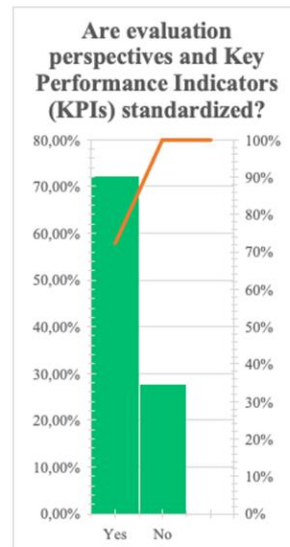
Source: Compiled by the authors.
Figure 6. Questionnaire – Standardization of KPIs, %

72.32% of the participants have stated that evaluations and Key Performance Indicators are standardized (Figure 6). This procedure would allow an inter-divisional comparison of data. In the Bima study 2017/2018, only 29% stated that evaluations are standardized throughout the company. The big difference could arise from a slightly different formulation of the questions between the two studies, or those industrial companies are used to standardize their data compared to other sectors.

To what extent are Big Data projects not started because of not standardized KPIs?

Forty-five participants have stated that their company has already conducted Big Data projects. Sixty-one companies have not conducted such projects. Out of these, 38 have standardized KPIs throughout the company (62%). Of the companies with experience in Big Data projects, 38 out of 45 participants have standardized KPIs. This corresponds to 84% of the participants and shows that standardized KPIs are a vital component in preparing data sets that are comparable and useful for complex evaluations.

A chi-square test was conducted to test whether the standardization of KPIs directly depends on the implementation of Big Data projects, and the p-value was calculated (Table 3 and Table 4). The H_0 hypothesis that was investigated is stated as no dependency between the standardization of KPIs and the implementation of Big Data projects.

Table 3. Observed frequency data Big Data projects/ KPIs

Observed frequency		
	Big Data projects	No Big Data projects
Standardized KPIs	38	38
Not standardized KPIs	7	23

Source: calculated and compiled by the authors.

**Table 4.** Expected frequency data Big Data projects/ KPIs

Expected frequency		
	Big Data projects	No Big Data projects
Standardized KPIs	32.3	43.7
Not standardized KPIs	12.7	17.3

Source: calculated and compiled by the authors.

The calculated p-value is 0.0122. That value is smaller than the set significance level of 5% (0.05); therefore, set hypothesis H0 is rejected, and a dependency between the standardization of KPIs and the implementation of Big Data projects can be proven.

Which technology is used for data storage?

According to the conducted study, under 10% of the Austrian industrial companies use modern data storage technologies such as data lakes, cloud solutions, NoSQL, or data warehouses with data marts (Figure 7). In the BARC study 2014, conducted in the DACH region, 38% of the participants used data warehouses with data marts, and 3% already used data lake architectures.
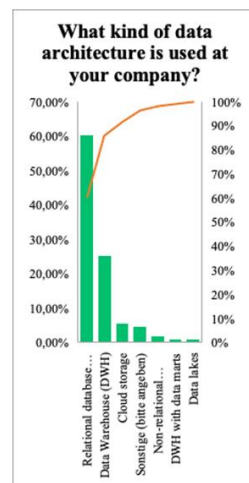
Source: Compiled by the authors.

Figure 7. Questionnaire – Data architecture used, %

How strong is the correlation between companies with a modern data infrastructure and the implementation of Big Data projects by those companies?'

Ten study participants have stated to use modern data architectures such as data lakes, cloud solutions, NoSQL, or data warehouses with data marts. 5 of those companies also have conducted Big Data projects. Out of the companies that use traditional data architectures (93 participants), 37 companies have conducted Big Data projects.

To test whether a company's usage of modern data architectures has a direct dependency on the implementation of Big Data projects, a chi-square test was conducted, and the p-value was calculated (Table 5 and Table 6). The H_0 hypothesis that was investigated is stated as no dependency between a company's usage of modern data architectures and the implementation of Big Data projects.

Table 5. Observed frequency data Big Data projects/ Data architecture

Observed frequency		
	Big Data projects	No Big Data projects
Modern data architecture	5	5
No modern data architecture	37	56

Source: calculated and compiled by the authors.

Table 6. Expected frequency data Big Data projects/ Data architecture

Expected frequency		
	Big Data projects	No Big Data projects
Modern data architecture	4.1	5.9
No modern data architecture	37.9	55.1

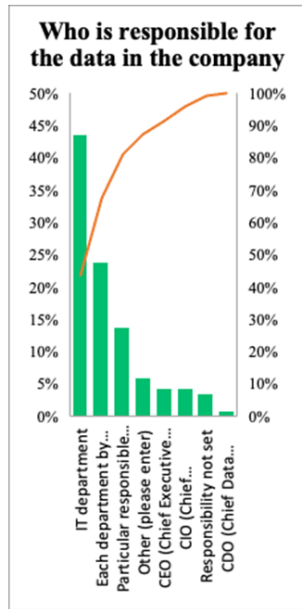Source: calculated and compiled by the authors.

The calculated p-value is 0.5322. That value is higher than the set significance level of 5% (0.05); therefore, the hypothesis H_0 is confirmed, and there is no dependency between the usage of modern data architectures by a company and the implementation of Big Data projects.

In general, it must be stated that modern data architectures and Big Data projects are implemented by a relatively small amount of Austrian Industrial companies. It is presumed that the number of companies using modern data architectures will go up when Big Data projects with more complex data analyses are conducted since these analyses are hard to conduct using traditional data architectures with a limited scope of modification. This does not mean that architecture such as data warehouses will be displaced. It is expected that hybrid solutions will be used that bring the best advantages in each aspect to the companies.

Which department is mostly responsible for data in Austrian industrial companies?

The top answer about the data competence within the company was the Information Technology (IT) department with 43.59%. This answer was followed by 23.39% of the companies who said each department is responsible for the data, while 13.68% said there was a defined data manager. 5% (6 participants) stated that a designated Chief Data Officer or Chief Information Officer was assigned to the company. The other options combined were selected by 6% of the participants (Figure 8). There was also the possibility of entering a different choice in a text field (e.g., CFO).

According to the 2018 Global Management Benchmark Report, 10% of the companies in the US already had a CDO, 20% a CIO, and 14% a data manager assigned. In 51% of the companies, the IT department was responsible for the data (Experian, 2018). That percentages prove that data is more critical in the US than in Austria since C-level positions designated for data management are more established.

Source: Compiled by the authors.

Figure 8. Questionnaire - Responsibility for data, %

To what degree do companies with a dedicated responsibility for data also conduct Big Data projects?

4% have stated that a company chief information officer is responsible for the data. 1% of the participants have a Chief Data Officer, and 14% that there is a responsible data representative without a C-level. Those 22 answers were counted as a dedicated responsibility for data in the company. Of these companies, 11 have conducted Big Data projects, and 11 have not.

To test whether a company's assigned responsibility for data directly depends on the implementation of Big Data projects, a chi-square test was conducted, and the p-value was calculated (Table 7 and Table 8).

The H_0 hypothesis that was investigated is stated as no dependency between a company's assigned dedicated responsibility for data and the implementation of Big Data projects.

Table 7. Observed frequency data Big Data projects/ Data representative

Observed frequency		
	Big Data projects	No Big Data projects
Co. with data representative	11	11
Co. with no data representative	34	85

Source: calculated and compiled by the authors.

Table 8. Expected frequency data Big Data projects/ Data representative

Expected frequency		
	Big Data projects	No Big Data projects
Co. with data representative	7.0	15.0
Co. with no data representative	38.0	81.0

Source: calculated and compiled by the authors.

The calculated p-value is 0.0476. That value is lower than the set significance level of 5% (0.05); therefore, the set hypothesis H_0 is rejected, and there is a direct dependency between an assigned responsibility for data and the implementation of Big Data projects.

However, it must be noted that the calculated p-value is only a little under the set significance level, and therefore the dependency is low, especially considering the limited number of participants. Of the 22 companies with a designated data responsibility, 14 have over 250 employees and are considered large enterprises in Austria (WKO, 2019), and nine have already conducted Big Data projects (64%).

Which industrial sectors in Austria invest the most resources in data management?

The questionnaire was sent to companies from all industrial sectors. The most participants came from the metal sector (34 participants), the chemical sector (13 participants), and the paper sector (11 participants). The response rate corresponds to the industrial landscape in Austria, where the metal, chemical, and paper sectors rank among the biggest industries in terms of

production value in Austria (WKO, 2019). However, just participation in the study is not necessarily an indication that the topic of data management is essential in that company. The implementation of a company-wide data management and the jurisdiction of data in the company were used to answer this research question. Small companies seldom have a Chief Data Officer or Chief Information Officer. However, it can be assumed that they have a determined responsibility for data if the organization invests many resources in data. Only 12 participants answered both questions positively. The metal industry was the only sector with multiple participants who answered both questions positively. 5 of the 34 metal industry participants said they had company-wide data management and designated responsibility for data.

Which is the biggest problem regarding data that Austrian industrial companies have?

For this question, it was possible to select multiple answers. Therefore, the overall number of responses exceeds the number of participants. The dispersion was high since all answers were chosen at least 15 times. More than 55% of the participants have selected communication problems between different departments as the main problems with data management (Figure 9). Data quality was chosen by 25.23% of the participants.

Communication problems between different departments and insufficient resources were the most common problems companies face. In the BARC study 2014, the top answer to the most significant challenge regarding data management was inadequate resources (BARC, 2014).

In the objective survey, 47 participants indicated that lousy data quality is why no big data projects are started. That was the second most selected answer from the technical side. So, the awareness that data quality is essential for further data evaluation is present.

It can be concluded that most departments in the company recognize the importance of data. However, too few resources are allocated for data in Austrian companies.
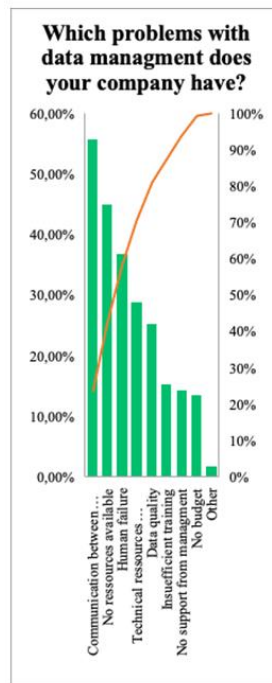
Source: Compiled by the authors.
Figure 9. Questionnaire - Problems with data management, %

To what extent does a company-wide defined data management protect companies from human errors regarding data?

Thirteen companies with a company-wide defined data management have stated that human errors are the main problem regarding data management (29%). Those companies' main reasons were communication problems and a lack of resources. Out of the 60 companies that had not implemented a company-wide data management, 24 (40%) of the participants have stated that human errors are the main reason for problems with data management. The most selected problem for those companies was communication between company members. So, in both cases, human

errors were stated as one of the biggest problems regarding data management.

A chi-square test was conducted to test whether a defined company-wide data management directly depends on human errors, and the p-value was calculated (Table 9 and Table 10). The H_0 hypothesis that was investigated is stated as no dependency between a company-wide data management and human errors regarding data in a company.

Table 9. Observed frequency data Human errors main data problem/ Company-wide data management

Observed frequency		
	Human errors are the main problem	Human errors are not a primary problem
Co. with data management	13	32
Co. without data management	24	36

Source: calculated and compiled by the authors.

Table 10. Expected frequency data Human errors main data problem/ Company-wide data management

| Expected frequency		
	Human errors are the main problem	Human errors are not the main problem
Co. with data management	15.9	29.1
Co. without data management	21.1	38.9

Source: calculated and compiled by the authors.

The calculated p-value is 0.2382. That value is much higher than the set significance level of 5% (0.05); therefore, the set hypothesis H_0 is confirmed, and a direct dependency between human errors as the primary data problem source and an implemented company-wide data management could not be proven.

However, companies with and without a company-wide data management have indicated human errors as one of the primary sources of error regarding data. That leads to the conclusion that processes and responsibilities must be clearly defined to prevent human errors.

To what degree strategic decisions of Austrian industrial companies are based on data?

Eighteen companies (16.07%) stated that strategic decisions are not based on actual data (Figure 10). Hence in most Austrian industrial companies, strategic decisions are based on data.

In the BARC study 2014, 41% of the participants reported that information had a highly significant value in the decision-making process, and 61% stated that this value would be much higher in the upcoming years (BARC, 2014). The current study can confirm this statement, where over 83% of the participants confirmed data-based decision-making.
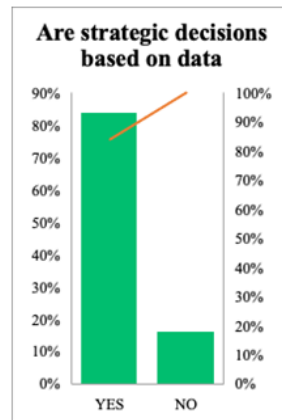
Source: Compiled by the authors.
Figure 10. Questionnaire - Strategic decisions

How strong is the dependency between the implementation of strategic decisions based on data and the turnover of a company?

Only 18 industrial companies in Austria (16.07%) have stated that strategic decisions are not based on data. 13 of those 18

companies had a turnover of over EUR 50 million. That relates to 72.2%. Ninety-four participants have stated that strategic decisions were based on data; 72 had a turnover of over EUR 50 million (76.6%). According to the Austrian Economy Chamber, the threshold of EUR 50 million corresponds to the definition of a large enterprise in Austria (WKO, 2019).

To test whether strategic decisions based on data directly depend on a company's turnover, the p-value was calculated (Table 11 and Table 12). The H_0 hypothesis that was investigated is stated as no dependency between implementing strategic decisions based on data and a company's turnover.

Table 11. Observed frequency data Strategic decisions implemented based on data/ Turnover

Observed frequency	Strategic decisions based on data	Strategic decisions not based on data
Turnover over 50 Million Euro	72	13
Turnover under 50 Million Euro	22	5

Source: calculated and compiled by the authors.

Table 12. Expected frequency data Strategic decisions implemented based on data/ Turnover

Expected frequency	Human errors are the main problem	Human errors are not the main problem
Co. with data management	71.3	13.7
Co. without data management	22.7	4.3

Source: calculated and compiled by the authors.

The calculated p-value is 0.6911. That value is much higher than the set significance level of 5% (0.05); therefore, the set hypothesis H_0 is confirmed, and no dependency between the implementation of strategic decisions based on data and a company's turnover could be proven.

Strategic decisions are mainly based on data in all Austrian industrial companies regardless of turnover. The companies that stated that strategic decisions are implemented without the corresponding data could also have bad communication at the management levels where the reasons for decisions are not well communicated to employees. However, they could also be taken based on data.

5. Conclusion

'Knowledge is power'; this quote by Francis Bacon from 1597 seems more relevant than ever (Bacon, 1597). While many articles and studies on this topic, the concrete work focuses on a particular region and sector. The article intends to overview the current state of data management in Austrian industrial companies?'. Aimed at the general audience and data management professionals, it should start a discussion in the community if Austrian companies lack behind their international competitors in company-wide data management and give an overview of the responsibilities, Big Data projects, infrastructure, problems, and decision-making process.

Approximately 40% of the participants stated that a company-wide data management was implemented in their company. This was not the case for 60% of the companies. Austrian companies have improved compared to other studies but are still behind their international competitors. Of the companies that implemented company-wide data management, only 3 of 44 had a turnover of under EUR 5 million. Hence, a clear correlation between an implemented company-wide data management and the company's turnover could be shown. However, there is also the possibility that the high turnover allowed the investment in data management and not the other way round. Future works should

correlate the turnover growth with the implementation of data management.

72,32% of the participants have stated that evaluations and Key Performance Indicators are standardized. In the Bima study 2017/2018 conducted throughout European companies, only 29% stated that evaluations are standardized throughout the company. This significant difference between the results could either mean that the question was understood differently in both studies or that companies take that issue seriously and prepare data for future complex analysis. A clear correlation could be proven between companies with standardized KPIs and those conducting Big Data projects. No correlation could be proven between the usage of modern storage technologies and experience with Big Data projects. However, it must be noted that not many industrial companies in Austria have experience with Big Data projects and use modern storage technologies.

The IT department was responsible for the data in almost half of the companies participating in the study. In contrast, very few companies in Austria had established C-level positions for data, which is already very common for US companies. There is a direct dependency between an assigned responsibility for data and the implementation of Big Data projects. That could mean that companies that have recognized the importance of data are also the ones who get the most value out of the data and perform Big Data projects. However, since it was confirmed that most companies with a high turnover have experience in Big Data projects, it can also be the case that those mostly big companies have established data responsibilities and implemented Big Data projects, independent of a direct influence between those two aspects.

There was no clear correlation regarding whether they are specific sectors from the industrial companies that have invested more into data management than others. The metal industry had the most participants and companies that answered that a company-wide data management is established.

It was reported that communication problems between different departments and insufficient resources were the most common problems companies face. This statement corresponds to the results of the BARC study in 2014.

Only 16% of the companies said that strategic decisions were not made based on data, and there is no proven dependency that companies with a low turnover more often make decisions that are not based on data.

European companies of all sizes need to recognize the need for company-wide data management and clear data management responsibilities to be prepared for the increasing data volume and more complex data analysis. While most Austrian industrial companies make strategic decisions based on data, they seem to lack behind their international competitors concerning C-level data officials and the percentage of companies that already have a company-wide data management implemented.

References:

Bacon, F. (1597). *Meditationes sacrae*. Montana: Kessinger Publishing, 1. ISBN: 9783772818233

BARC GmbH (2014). *Big Data Analytics.* Retrieved https://www.sas.com/content/dam/SAS/bp_de/doc/whitepaper1/ba-wp-barc-big-data-analytics-2014-2298353.pdf (25.05.2021).

BARC GmbH (2014). *Datenmanagement im Wandel*. Retrieved from https://barc.de/docs/datenmanagement-im-wandel (25.05.2021).

BARC GmbH (2018). *Zeit für eine Neue Kultur durch Business Intelligence & Advanced Analytics.* Retrieved from https://sdv-dialogmarketing.ch/wp/wp-content/uploads/2019/01/bima-studie-2018-abstract.pdf (25.05.2021).

Brüggemann, H., & Bremer, P. (2015). *Grundlagen Qualitätsmanagement*. Berlin, Heidelberg, New York: Springer Verlag, 3. ISBN: 978-3-658-09221-4

DAMA (2009). The DAMA Guide to The Data Management of Body of Knowledge. New Jersey: Technics Publications LLC, 4-83. ISBN 978-0-9771400-8-4

Experian (2018). The 2018 global data management benchmark report. Retrieved from https://www.experian.com.vn/wp-content/uploads/2018/02/2018-global-data-management-benchmark-report.pdf (25.05.2021).

Fayyad, U., Piatetsky-Sharirp, G., & Smyth, P (1996). From Data Mining to Knowledge Discovery in Databases, *AI Magazine, 3*, 37–54. https://doi.org/10.1609/aimag.v17i3.1230

Freiknecht, J., & Papp, S. (2018): *Big Data in der Praxis.* Second edition. Hanser Verlag. 23. ISBN: 978-3-446-45601-3.

Gandomi, A., & Haider, M. (2014). Beyond the hype. Big data concepts, methods, and analytics. *International Journal of Information Management, 35*, p.138-139.

Gartner (2019). *It-Glossary*. URL: https://www.gartner.com/it-glossary/big-data/ (25.05.2021).

Ghavami, P. (2020). *Big Data Management: Data Governance Principles for Big Data Analytics*. Berlin, Boston: De Gryter, 2020, 22. https://doi.org/10.1515/9783110664065

Gordon, K (2013). *The Principles of Data Management.* Swindon: The Chartered Institute for IT, 57-66. ISBN: 978-1-78017-185-2

IBM (2016). *The Rise of the Data Economy: Driving Value through Internet of Things Data Monetization*. New York: IBM Global Services, 2.

IBM (2018). *2,5 quintillion bytes of data created every day how does CPG retail manage it.* URL: https://www.ibm.com/blogs/insights-on-business/consumer-products/2-5-quintillion-bytes-of-data-created-every-day-how-does-cpg-retail-manage-it/ (25.05.2022).

IOT Analytics (2016). *Industrial Analytics 2016/2017.* Retrieved fromhttps://digital-analytics-association.de/wp-content/uploads/2016/03/Industrial-Analytics-Report-2016-2017-vp-singlepage.pdf (25.05.2022).

Krcmar, H. (2015): *Informationsmanagement.* Berlin, Heidelberg, New York: Springer Verlag. 11-59. ISBN: 978-3-662-45862-4

Madhushree, L. M., Revathi, R., & Aithal, P. S. (2019). A Review on Impact of Information Communication Computation Technology (ICCT) on Selected Primary, Secondary, and Tertiary Industrial Sectors. *Saudi Journal of Business and Management Studies*, *4*(1), 108.

Metalltechnische Industrie (2021). *Wer wir sind – Daten & Fakten zur Branche.* Retrieved from https://www.metalltechnischeindustrie.at/news-presse/presse/wer-wir-sind-daten-fakten-zur-branche/ (25.05.2022).

Oracle (2018). *Mining Structured Data and Unstructured Data using Oracle Advanced Analytics*. Retrieved from https://blogs.oracle.com/datamining/mining-structured-data-and-unstructured-data-using-oracle-advanced-analytics-12c (30.05.2022).

Otto, B. (2015). Quality and Value of the Data Resource in Large Enterprises. *Information Systems Management, 32*(3), 234–251. https://doi.org/10.1080/10580530.2015.1044344

Oussous, A., Benjelloun, F., Ait Lahcen, A., & Belfkih, S. (2017). Big Data technologies: A survey. *Journal of King Saud University – Computer and Information Sciences, 30*, 433.

Prakash, S. S. (2020). Evolution of Data Warehouses to Data Lakes for Enterprise Business Intelligence. *International Journal of Innovative Research in Computer and Communication Engineering, 8*, 1038-1042.

Ramasamy, A., & Chowdhury, S. (2020). Big Data Quality Dimensions: A Systematic Literature Review. *Journal of Information Systems and Technology Management, 17*, 11. https://doi.org/10.4301/S1807-1775202017003

Romeike, F., & Hager, P. (2016). Predictive Analytics: Looking into the future. *FIRM Jahrbuch*, Bd. 2016. Frankfurter Institut für Risikomanagement und Regulierung, Frankfurt am Main. p. 168-171.

Schroeck, M.; Shockley, R.; Smart, J.; Romero-Morales, D.; Tufano, P. (2012). *Analytics: Big Data in der Praxis*. Retrieved from https://www-935.ibm.com/services/de/gbs/thoughtleadership/GBE03519-DEDE-00.pdf (05.06.2022).

Techjury (2021). *How Much Data Is Created Every Day in 2021?*. URL: https://techjury.net/blog/how-much-data-is-created-every-day/#gref (30.05.2022).

Wang, R. Y., & Strong, D. M. (1996). Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Information Systems Management*, 12 (4), 5-34. https://doi.org/10.1080/07421222.1996.11518099

WKO (2019). *Industrie aktuell*. URL: https://www.wko.at/branchen/industrie/industrie-aktuell-04-2020.pdf (25.05.2022).

WKO (2019). *Zahlen Daten Fakten*. URL: https://www.wko.at/service/zahlen-daten-fakten/KMU-definition.htm (25.05.2022).

Wallner Martin
Comenius University Bratislava
Faculty of Management
Department of Information Systems
Bratislava
Slovak Republic
wallner4@uniba.sk
ORCID 0000-0003-4367-8711

Peráček Tomáš
Comenius University Bratislava
Faculty of Management
Department of Information Systems Bratislava
Slovak Republic.
tomas.peracek@fm.uniba.sk
ORCID 0000-0002-3498-5640