USAGE OF ADVANCED DATA ANALYSIS IN AUSTRIAN INDUSTRIAL COMPANIES

Martin Wallner[1], Tomáš Peráček[2]

Abstract: Data has become one of the most valuable resources for companies. The large data volumes of Big Data projects allow institutions the application of various data analysis methods. Compared to older analysis methods, which mostly have an informative function, predictive and prescriptive analysis methods allow foresight and the prevention of future problems and errors. This paper evaluates the current state of advanced data analysis in Austrian industrial companies. Furthermore, it investigates if the advantages of complex data analyses can be monetarized and if cooperate figures such as the turnover or company size influence the answers of the survey. For that reason, a survey among industrial companies in Austria was performed to assess the usage of complex data analysis methods and Big Data. It is shown that small companies use descriptive and diagnostic analysis methods, while big companies use more advanced analytical methods. Companies with a high turnover are also more likely to perform Big Data projects. On an international comparison for most Austrian industrial companies, Big Data is not the main focus of their IT department. Also, modern data architectures are not as extensively implemented as in other countries of the DACH region. However, there is a clear perception by Austrian industrial companies that forward-looking data analysis methods will be predominant in five years.

JEL Classification Numbers: L60, O32, O52, **DOI:** https://doi.org/10.12955/peb.v2.265

Keywords: Big Data, data analytics, predictive analysis, prescriptive analysis, industrial companies

Introduction

Data has become one of the most valuable resources of our time, and innovations lead to immense amounts of data. The daily data volume was as high as 2.5 quintillion bytes per day in 2018 (Forbes, 2018), and 90% of all existing data has been created in the last two years (Bradshaw, 2013).

Oracle, one of the biggest manufacturers of databases, estimates that this trend will be further boosted in the upcoming years because of the Internet of Things (Andelfinger & Hänisch, 2015).

The linkage between physical and virtual items could lead to a 40% growth in data volume in the next two years (Statista, 2020). These huge amounts of generated data allow institutions almost infinite possibilities of data analytics in various fields (Paganoni, 2019). Besides obvious applications such as the analysis of machine data that allow predictive maintenance to minimize machine downtimes or the analysis of buying behavior to allow more personalized advertisements, there are also governmental applications such as the prediction of citizen behavior (Oussous et al., 2017).

Analytical methods use statistical models to extract information out of a very big pool of data. In the past, explanations of events that already took place could be found with these analyses, however, today, the main purpose is to make predictions in order to enable or avoid certain incidents (Sappelli et al., 2017).

This article is aimed at a general audience as well as data management and data analytics professionals and should point out the shortcomings and qualities of data analysis in industrial companies in Austria compared to other regions. Moreover, the article should induce a discussion in the community, how the current state can be improved, especially in regard to what analytical methods are currently used in Austrian industrial companies and what methods are planned to be used in the future. Furthermore, factors should be pointed out that influence the willingness of companies to utilize these methods and if the usage has a direct influence on the turnover. This article contains parts that had been employed by the author for his Master thesis (Wallner, 2019).

Literature review

One very important aspect in the context of data analytics is Big Data. Gartner, one of the leading IT research and consulting agencies, defines Big Data as "Big data is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making" (Laney, 2001).

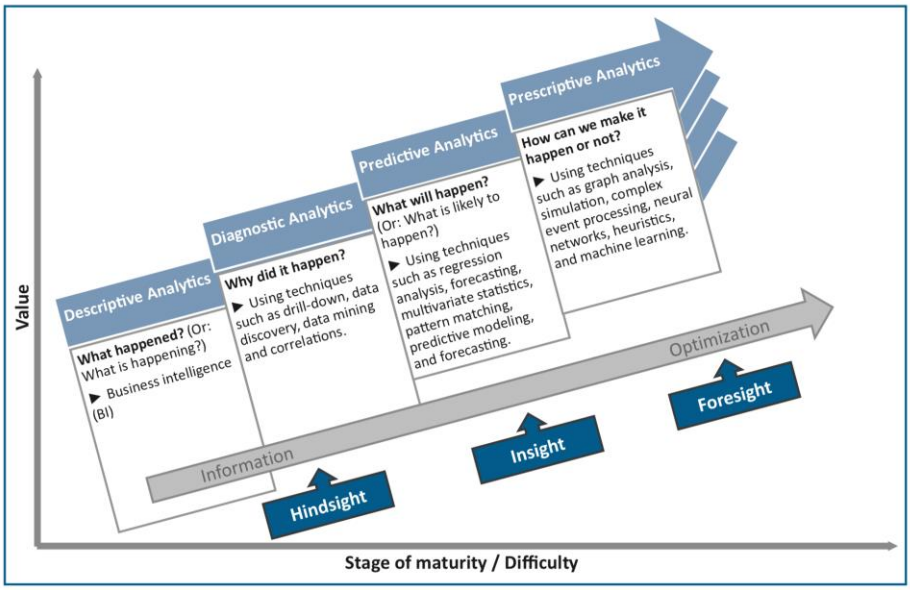
Romeike and Eicher (2016) subsumed that having a broader database as provided by Big Data projects allows institutions more complex analytical methods that can answer more challenging questions.

[1] Comenius University Bratislava, Faculty of Management, Bratislava, Slovak Republic, wallner4@uniba.sk

[2] Comenius University Bratislava, Faculty of Management, Bratislava, Slovak Republic, peracek2@uniba.sk

A classification of the analytical methods based on their value and difficultness is shown in Figure 1.

Figure 1: Classification of analytical methods



Source: Romeike, Eicher (2016)

The simplest form of an analytical analysis is the descriptive analysis, which is used to analyze past data sets in order to be able to answer the basic question 'What happened?'. Correlations between variables are detected to understand past events and to be able to apply this knowledge to present issues (Romeike & Eicher, 2016). A possible problem from the industry could be 'Which parts have failed?' (Gartner, 2014).

The diagnostic analysis is closely related to the descriptive analysis and also evaluates data sets from the past. However, whilst the descriptive analysis only answers the question 'What happened?', the diagnostic analysis investigates why a certain event has happened (Sampaio & Saraiva, 2016).

The descriptive and diagnostic analyses only analyze past data to answer why a certain event happened. The predictive analysis creates models which are able to forecast future events (Romeike & Eicher, 2016).

Tools from statistical methods, data modeling, data mining and machine learning are combined for predictive analysis where historical and present data are processed. This method is more complex, but it can provide predictions that answer the question 'What will happen in the future?' (Romeike & Eicher, 2016).

Prescriptive analysis methods are the newest and most complex addition in terms of data evaluation. Complex algorithms answer the question 'How can we make a certain event happen?' (Romeike & Eicher, 2016). Similar to the predictive analysis, forecasts are made using present and historical data, but in addition, also correlations are used that did not happen but have a defined probability of occurrence (Sappelli et al., 2017).

These complex analysis methods require modern storage architectures like Data Warehouses with Data Marts, Data Lakes, NoSQL or Cloud storages. These technologies offer the possibility to store and access large data volumes very quickly and are therefore ideally suited for complex data analyses (Gorelik, 2019).

There are some detailed surveys on industrial analytics, which focus on the transition to new technologies and applications of these, e.g., BARC Big Data Analytics Survey (BARC GmbH, 2014) or BARC Study 2014 (BARC GmbH, 2014). However, in the fast-paced IT sector, a survey from 2014 could already be outdated.

In the BARC Data Analytics study, only 12% of the companies have stated that they already have experience with Big Data projects, and 22% said that no Big Data initiative is planned. In the BARC Study from 2014 (BARC GmbH, 2014), 6% have stated that Big Data is an important issue for the

company. 41% of the companies have also declared that they use modern data architectures such as Data Warehouses with Data Marts or Data Lakes.

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) did not restrict the geographical location or industrial sector of the participants. This approach is good to reach a high number of participants from different industries, however, regional differences or differences due to the industrial sector are very hard to outline. In the Industrial Analytics Report 2016/2017 (IOT Analytics, 2016), only 30% of the participants said that their companies had performed predictive or prescriptive data analytics, however, 45% have stated that the predictive maintenance of machines should be the main IT concern for industrial companies. This article tries to provide an up-to-date and regionally limited view on the topic of advanced data analytics in Austrian industrial companies.

Data and methodology

With the definition of the research question and the hypotheses, keywords were specified, and a literature review was performed. After the research, which included books, articles and studies, the most important concepts were outlined and information from other studies was extracted. In order to be able to answer the following research question, a survey was conducted and statistically evaluated. The first step of this evaluation included an extinction of invalid answers and a visual presentation of the valid answers, and later certain answers were filtered to answer the hypotheses. To validate the answers and allow a comparison with different regions and countries, the answers of the study were compared with the results of other data analytics and data management studies.

Research question:

'What is the current state regarding advanced data analytical methods in Austrian industrial companies?'

The subsequent hypotheses were established and investigated to answer the research question:

- H1: 'Small companies use descriptive and diagnostic analytical methods, while big companies use more advanced analytical methods'
- H2: 'Best-in-class companies have experience in Big Data projects and recognize the positive effects for different applications and departments'
- H3: 'The current focus of industrial companies is to implement Big Data technologies'
- H4: 'Companies which use complex data analysis methods, such as predictive or prescriptive analysis, are more likely to invest in modern storage technologies'
- H5: 'The future of data analytics are predictive and prescriptive data analysis'

The survey was conducted online and only aimed at industrial companies in Austria. Only companies of the primary and secondary economic sectors were classified as industrial companies (Madhushree et al., 2019).

Therefore, companies from the tertiary and quaternary sectors or commercial enterprises were not considered. To ensure compelling findings, a minimum of 40 feedbacks was defined. In order to achieve a low drop-out rate, the survey could be answered in under 10 minutes, and no open and mandatory questions have been asked. The survey was created with SurveyMonkey, and the answers that could be chosen were either single or multiple choice. Furthermore, the questions and answers were selected in a way to minimize the influence of the respondent. The link to the questionnaire was sent to relevant professionals on 23.09.2019, and a reminder was sent one week later. On 07.10.2019, the survey was closed because 141 valid responses were available. This number was sufficient to answer the research questions statistically robustly.

The introductory questions allowed conclusions on how certain cooperate figures, such as the number of employees or branches, affect the answers to the following questions. In the part 'Organization of data management,' the responsibilities and data quality measures were retrieved. The technical section dealt with the current focus of the IT and data architecture. The last section addressed the usage of Big Data, data analysis methods and the applications of these technologies.

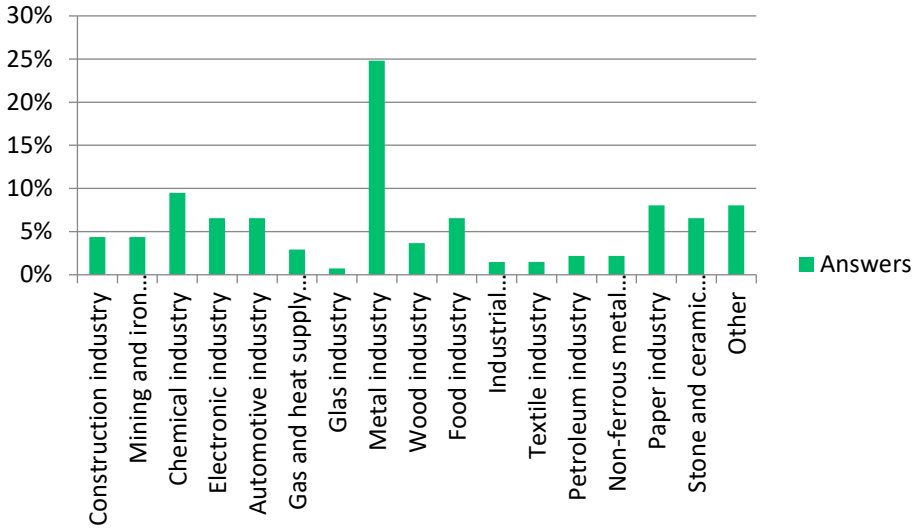
Data collection:

The survey was sent to industrial companies from all industrial branches in Austria. It was conducted in German and translated to English for this publication. 24.82% of the participants were companies from

the metal industry. This sector had by far the most participants. However, this is explainable since this sector contributed a quarter of the overall industrial yield in terms of turnover, production value, and gross value added (Statistia, 2020).

The other sectors had an average participation rate of approximately 6% and were almost evenly distributed between different branches (see Figure 2).

Figure 2: INTERVIEW DATA - Graphical overview of industrial sector of participants



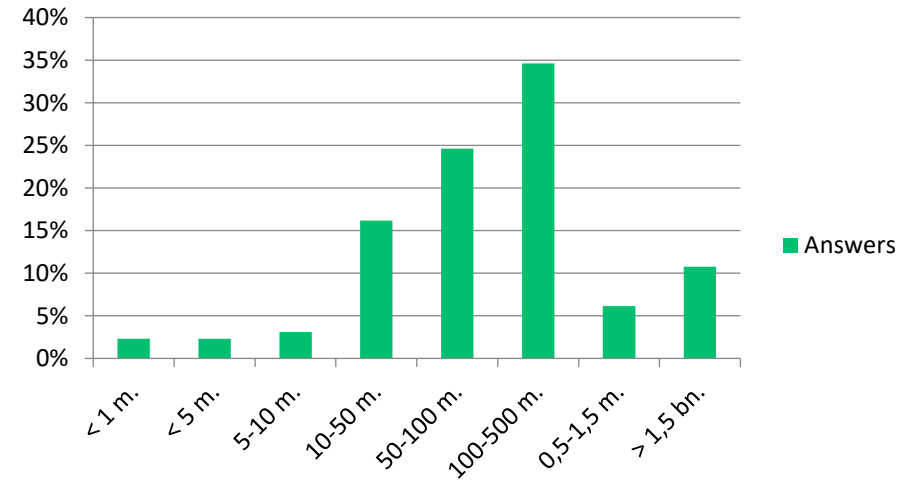
Source: Author (2019)

Most of the participants worked for companies with a headcount of 201-1000 employees (49.26%), followed by companies with a headcount of 1001-5000 employees (18.38%). This question was intentionally asked in a way to only consider employees in Austrian branches. Hence, big cooperations could not falsify the results.

The business divisions of the participants are shown in Figure 3. Most participants worked in management (20.44%), IT (18,98%) and production (16.79%).

The question about the recent yearly turnover was related to the turnover which was achieved with business in Austria. Most of the companies which participated in the survey had a turnover of over 50 million Euro (see figure 3).

Figure 3: INTERVIEW DATA - Turnover of a company in Austria



Source: Author (2019)

Results and Discussion

The evaluation of the survey led to the following results in regard to the set hypotheses:

H1: 'Small companies use descriptive and diagnostic analytical methods, while big companies use more advanced analytical methods'

Of the companies with a headcount of 1 to 50 employees, 50% used modern data analysis methods such as predictive or prescriptive methods, and 50% used traditional methods like descriptive or diagnostic methods. If a company has stated that they used traditional and modern data analytical methods, the company was counted as a company that used modern data analytical methods. 60% of the companies with a headcount of 51-200 employees used traditional data analytics (17 out of 28 participants who answered this question). 200 employees were considered as the threshold between small and big companies. If only those companies are taken into account, which had under 200 employees, 70% of those used classical data analytics. Of the companies with over 200 employees, approximately 50% used modern data analytical methods.

The hypothesis H1 was confirmed.

H2: 'Best-in-class companies have experience in Big Data projects and recognize the positive effects for different applications and departments'

As an indicator for a best-in-class company, a yearly turnover of over 500 million Euro has been defined. 23 participants have stated that they fulfil this turnover threshold, 56% of those had experience with Big Data projects. Of the companies with a turnover below that threshold, only 35% had experience with Big Data projects. Of the companies with a turnover of under 100 million Euro, only 33% have stated to have conducted Big Data projects, while 57% of the companies with a yearly turnover of over 100 million Euro had stated to have experience with such projects.

The hypothesis H2 has been confirmed.

H3: 'The current focus of industrial companies is to implement Big Data technologies'

The question about the current IT focus of the company allowed multiple answers. The number of responses was well-balanced between the possible options. The top answers in descending order were:

- 39.09% for 'Integration of heterogenic data sources' and 'Reduce complexity of the data architecture'
- 31.82% for the answer 'Simple integration of new data sources'
- 28.18% for the answer option 'Big Data'

So, although Big Data seemed to be very important and to be one of the focus areas for industrial companies, there was no apparent trend that would confirm H3.

H4: 'Companies which use complex data analysis methods, such as predictive or prescriptive analysis, are more likely to invest in modern storage technologies'

Only ten participants of the study have said that they use modern data architectures, such as Data Warehouses with Data Marts, NoSQL, Cloud storages or Data Lakes. Of these companies, only five (50%) have already conducted Big Data projects. 93 participants have stated that they use classical data architectures. 37 of those companies (40%) had conducted Big Data projects. A weak overall link between the usage of modern data analysis methods and modern data architectures could be proven, however, the difference in the percentage is too small to establish a general hypothesis.

The hypothesis H4 was not confirmed.

H5: 'The future of data analytics are predictive and prescriptive data analysis'

28.30% of the survey participants used predictive data analysis methods, and 25.47% used prescriptive data analysis methods. The companies have stated in the survey that they think these values will change in five years to 68.87% respectively 63.21%.

It could be shown that there is a clear perception by Austrian industrial companies that forward-looking data analysis methods will be predominant in five years. H5 was confirmed.

Discussion and comparison with other studies:

The results to hypothesis H1 showed a trend suggesting that complex data analytics, which are expensive in the application, were at first used by big companies. However, no correlation can be made if complex

analysis led to a higher turnover or if the high turnover of big companies is the enabler for the usage of these technologies. If the company size is neglected and the total answers are counted, Austrian industrial companies even had an innovative lead in the international comparison with other studies, since 42.65% had already realized complex data analysis projects. In the Industrial Analytics Report 2016/2017, only 30% of the companies had said that they realized complex data analysis projects (IOT Analytics, 2016). However, the Industrial Analytics Report was created in the years 2016 and 2017, while the present study was conducted in 2019. In the meantime, the values in other countries may also have increased. Future research should conduct studies in other countries to assess if Austrian companies actually have an advantage over other geographical regions.

In the BARC Study 2014 (BARC GmbH, 2014), only 6% of the companies had indicated that Big Data is an important issue in their company. In the BARC Big Data Analytics Survey, 22% of the companies have said that no Big Data initiative is planned and 12% that no Big Data project has been realized yet (BARC GmbH, 2014). In contrast, 57% of the companies in Austria with a yearly turnover of over 100 million Euro had stated to have experience with such projects.

What can be assumed is that big companies have realized that Big Data projects are essential to maintain their competitive advantage. In the Industrial Analytics Report, 45% of the participants had stated that the predictive and prescriptive maintenance of machines is the main IT focus right now (IOT Analytics, 2016). This leads to the assumption that these companies use Big Data to be able to run complex data analytics since these methods require a broad database.

The results from the Industrial Analytics Report 2016/2017 study could not be confirmed by the current study, where the top answers to the question were ‘better control of operative processes’ and the ‘optimization of the production planning’. ‘Big Data projects’ was only the third most given answer to this question.

Overall, it can be noted that modern data analysis methods and modern data architectures are used by very few Austrian industrial companies compared to the BARC Study, which was performed in the DACH region in 2014, where 41% of the participants had stated that they use modern data architectures (BARC GmbH, 2014).

These architectures are designed in a way to simplify the implementation of these analysis methods; however, it cannot be assumed that classical architectures will be completely displaced in the industry. In the future hybrid solutions will run in parallel, so the architecture can be chosen that fits the best for every single application.

Although it could be shown that Big Data projects are predominantly used by companies with a high turnover, no correlation can be presumed in the current study, if complex data analysis led to high turnover, or if the high turnover enables the usage of more complex, hence more expensive data analysis methods. The questionnaire revealed that complex data analytical methods had mainly been used by companies with over 200 employees. In general, these companies also have a very high turnover. Future works should investigate how the turnover has developed over time from the moment Big Data projects have been introduced in a company.

Conclusion

Analytical methods offer companies various opportunities that can have a substantial business impact. However, complex analytical methods are expensive and are therefore mostly used by very big companies. The intention of this research work was to close the research gap and answer the research question: ‘What is the current state regarding advanced data analytical methods in Austrian industrial companies?’. For that reason, a survey has been conducted with 141 valid participants.

The study showed that complex data analytics were mostly used by large enterprises. 70% of the companies with over 200 employees had already performed predictive or prescriptive data analytical projects, while only 30% of the companies with under 200 employees have conducted such analytics.

In most cases, these companies were also best-in-class companies in regard to turnover. 56% of the companies with a turnover of over 500 million Euro had experience with Big Data projects. It can be assumed that only companies which have a high investment power can finance such projects. In the long term, this strengthens their position even more.

The investigation clearly showed that although Big Data and complex data analysis were at this time not the main focus of Austrian companies, there was a clear conception that these methods will become crucially important in the future, so machine errors or productions stops do not occur in the future. However, only 28.18% of the companies have stated that Big Data projects were the current focus of the IT department. These complex analysis and Big Data projects require, in most cases, modern data architectures since they are very hard to conduct with classical data architectures. These classical architectures are not intended for this use and a very hard to modify. Very few industrial companies in Austria have adapted to modern data architectures.

It is also apparent that industrial companies in Austria have a clear understanding that these technologies will be very important in the upcoming years and think that predictive and prescriptive analytical methods will be predominantly used in five years. 28.30% of the survey participants currently used predictive data analysis methods, and 68.87% think that this value will go up in five years. 25.47% of the participants currently use prescriptive analysis methods, and 63.21% think that this value will increase in five years. Recognizing this is a very important step, however, Austrian industrial companies of all sizes need to recognize the positive long-term effects on their own business and start making the necessary investments now in order to keep up with their international competitors. In some areas, such as the present focus on Big Data projects and the current usage of modern data architectures, they currently seem to be behind their international competitors.

References

Andelfinger, V.P., Hänisch, T. (2015). *Internet der Dinge: Technik, Trends und Geschäftsmodelle* [*Internet of Things: Technology, trends and business models*]. SpringerGabler, dx.doi.org/10.1007/978-3-658-06729-8.

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

BARC GmbH (2014). *Datenmanagement im Wandel* [*Data management in the course of time*]. https://barc.de/docs/datenmanagement-im-wandel

BARC GmbH (2018). *Zeit für eine neue Kultur durch Business Intelligence & Advcanced Analytics* [*Time for a new culture using business intelligence & advanced analytics*]. https://sdv-dialogmarketing.ch/wp/wp-content/uploads/2019/01/bima-studie-2018-abstract.pdf

Bradshaw, L. (2013). Big Data and what it means. *Business Horizon Quarterly*, *7*(1), 18.

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

Forbes (2018, May 21). *How much data do we create every day?*. Forbes. https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/#5654930d60ba

Gartner (2014). Industrial Analytics. Powered by the Internet of Things. The next wave of business transformation. *Datawatch*, (2), 6.

Gorelik, A. (2019): *The Enterprise Big Data Lake*. O'Reilly.

IBM (2018, May 11). Diagnostic analytics 101: Why did it happen?. IBM. https://www.ibm.com/blogs/business-analytics/diagnostic-analytics-101-why-did-it-happen/

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

Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety. *Meta Group Research Note*, *949*(1), 1-3.

Madhushree L. M., Revathi Radhakrishnan, P.S. Aithal (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.

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*(4), 433. https://doi.org/10.1016/j.jksuci.2017.06.001

Paganoni, M. C. (2019). *Framing Big Data: A Linguistic and Discursive Approach*. Palgrave Pivot.

Romeike, F., Eicher, A (2016). Predictive Analytics: Looking into the future. FIRM Jahresbuch, 168-171.

Sampaio, P., Saraiva, P. (2016). *Quality in the 21st Century*. Springer.

Sappelli, M., Boer, M.D., Smit, S., & Bomhof, F. (2017). A vision on Prescriptive Analytics. *Big Data 2017*, 45.

Statista (2020, Dec 3). Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2024. Statista. https://www.statista.com/statistics/871513/worldwide-data-created/

Wallner, M (2019). *Datenmanagement in österreichischen Industrieunternehmen – Umsetzung, Trends und Hindernisse* [*Data management in Austrian industrial companies – Implementation, trends and obstacles*] [Unpublished master's thesis]. Mining University of Leoben.