Master Thesis

The integration of Big Data in purchasing, as designed in a new Big Data Purchasing Maturity model

University of Twente

Faculty of Behavioral, Management and Social Sciences

Department of Technology Management and Supply

Submitted by: ***Laura de Haan***

Date: *29-06-2018*

Contact: *l.m.dehaan@student.utwente.nl*

Version: *Final version*

Supervisory committee

Internal supervisors

Prof. Dr. Habil. H. Schiele

Faculty of Behavioral, Management and Social Sciences

Chair of the Department of Technology Management and Supply

University of Twente

Ass. Prof. M. de Visser

Faculty of Behavioral, Management and Social Sciences

Department of Technology Management and Supply

University of Twente

External supervisor

M. Schoenaker

Data Driven Business analyst

Bright Cape

Website: www.brightcape.nl

“Data is the new science. Big Data holds the answers.”
– Pat Gelsinger

Acknowledgements

After a period of hard work, this research has come to an end. I am proud on the result that is in front of you. I learned a lot from my research, since it is an interesting and hot topic at this moment, with a lot of potential for the future. And also because of this, the topic contains a lot of grey areas and undiscovered situations.

The Master Thesis journey that I went through, was not possible without the support of my family, friends, professors, colleagues and many others. Therefore, I would like to take this moment to express my gratitude to these people. My special gratitude goes to Mike, with his endless supporting motivation and feedback, and to all my other colleagues and managers from Bright Cape. Thanks for letting me be part of the team, the answers on all my questions, and of course, the wonderful ski trip to Sankt Anton.

I would also express my gratitude to my first super visor Holger Schiele, for his help, knowledge sharing and feedback during all the master courses and this research. It really has strengthened my knowledge about purchasing in all its aspects. Besides, I would like to thank my second supervisor Matthias de Visser, for his insightful comments and feedback.

What I want to say for now: I hope that you will enjoy reading this Master Thesis.

Yours sincerely,

Laura de Haan

Utrecht, June 2018.

Management summary

This research is conducted with support from Bright Cape, a company for 'Small and Big Data solutions'. Bright Cape is founded in 2014 and developed its own general Bright Cape Maturity model. This model is based on Gartner's Enterprise Information Management (EIM) Maturity model and has five stages. The Bright Cape Maturity model is developed for common use. During the years, they discovered that the more the maturity model is specified to a department, the more accurate the model will work. Since purchasing is one of the topics where Bright Cape find data solutions, a more specified model will improve their data results. There are many maturity models for both Big Data and Purchasing, but they are not combined yet. Therefore, the goal of this research is to develop a maturity model for Bright Cape where Big Data and Purchasing are integrated. To achieve this result, there are main and sub questions designed. The main question is: ***"How are the different steps designed and specified in the new Big Data Purchasing Maturity model?"*** The sub research questions are *"What is the current situation of Big Data in purchasing?"* and *"How does a purchasing maturity level relate to a Big Data maturity level?"* As final product, a scorecard for the new Big Data Purchasing Maturity Model will be developed.

The new model is based on three different subjects; Big Data, Purchasing and Maturity models. Big Data can be described as the 3V's of Doug Laney; large **volumes** of **varied** data that are developed and handled at high **velocity.** The main rule for volume is, when it is too big to handle, then it will be Big Data. In this research is that bigger than excel or a dataset of 1TB. The rule for variety is related to semi-structured and/or unstructured structure. Finally, the rule for velocity is, that the speed is (almost) near real-time. Next, purchasing as function is throughout the years developed from a supplementary to a more strategic function. When it is structured according the six steps of Van Weele, the purchasing function can add several advantages. Finally, maturity models conduct steps in sequence to show the current as-is situation. That can be for a company in general or specified to a subject or department. There are many Big Data Maturity Models and Purchasing Maturity Models developed throughout the years. In most models, there are four of five levels defined.

The new designed model is based on the Bright Cape Maturity Model, the Industry 4.0 Purchasing Maturity model of Torn (2018) and The Purchasing Maturity Model of Schiele (2007). The model is based on eight dimensions: 1) Strategy, 2) Process & Systems, 3)

Physical level, 4) Purchase to Pay (P2P), 5) Controlling & KPI's, 6) Sourcing, 7) Suppliers, and 8) Employees & Uses. There are four levels defined to distinguish the different stages. There are defined as:

- **Stage 1**: "The purchasing processes are well defined following the best practices of Industry 3.0. There are no (Big) Data applications integrated."
- **Stage 2:** "The purchasing processes are standardised and digitalised. There are the first applications of Big Data and there is one person assigned to perform the task."
- **Stage 3**: "The Big Data is fully integrated into the purchasing processes, and are cross-functional integrated through the company."
- **Stage 4:** "The Big Data processes are fully autonomous organised within the strategic purchasing department. The systems and processes are self-learning and continuously improving".

The scorecard is based on the best-case scenario of Big Data integrations in Purchasing, applied and specified to the eight dimensions. **Strategy:** Big Data and Purchasing are on the C-level of the company and the strategic decisions are significantly improved, with the help of a structured roadmap. **Processes & Systems:** There is a totally integrated and autonomous flow in Purchasing, with a continue monitoring system driven on Big Data. **Physical level:** The seamless connection between machine-to-machine communication through the Internet of Things, results in a real-time analyse of the data, for automatically made purchasing decisions. **Purchase to Pay**: There is a total data driven and seamless payment systems, which can help by complex purchasing contracts. **Controlling & KPI's:** There is always and anytime up to date and complete data driven information available, for controlling and with relevance for purchase managers. The process is self-learning and optimizing, which is driven on data and is in favour of purchasing. **Sourcing:** The sourcing process has integrations of e-sourcing and e-procurement with a data driven predictive demand. The real-time analysis of external and internal data can have a predictive influence on purchasing. **Suppliers:** With data sharing between buyers and suppliers, the purchasing process can be more optimized for both parties. There are Early Supplier Integrations (ESI) possible for further improvements and early detection of supply chain disruptions. **Employees & Users:** The user adapt the Big Data integrations and advance purchasing processes easily and understand what they are doing. There is a feedback loop and regularly planned evaluations, with constant improvements.

List of tables

List of figures

List of abbreviations

3D	Three dimensional
AI	Artificial Intelligence
BDMM	Big Data Maturity models
BI	Business Intelligence
BOM	Bill of Materials
CSC	Category Sourcing Cycle
CMM	Capability Maturity Model
CPO	Chief Purchasing Officer
CPS	Cyber-physical systems
DEA	Data Envelopment Analysis
E-procurement	Electronic-procurement
E-sourcing	Electronic-sourcing
ECR	Efficient Customer Response
EDI	Electronic Data Interchange
EIM	Enterprise Information Management
ERP	Enterprise Resource Planning
ESI	Early Supplier Involvement
GB	Gigabyte
GDPR	General Data Protection Regulation
GPS	Global Positioning System
I4.0	Industry 4.0
IDC	International Data Coporation
IoT	Internet of Things
IP	Internet Protocol
IT Systems	Information Technology Systems
KB	Kilobyte
KPI	Key Performance Indicators
M2M	Machine-to-Machine
MB	Megabyte
MRP	Material Requirements Planning
MRP II	Manufacturing Resource Planning
P2C	Purchase-to-Cash

P2P	Purchase-to-Pay
PDC	Purchasing Department Cycle
PDCA	Plan Do Check Act
PMM	Purchasing Maturity models
PQP	Product Quality Project
RDBMS	Relational Database Management Systems
ROP	Reorder Point System
SKU	Stock Keeping Unit
SQL	Structured Query Language
UCD	User Centered Design
UX	User Experience
VMI	Vendor Management Inventory

Table of content

1. Big Data; the next generation in analytics. What is in it for purchasing?

1.1 Current trends in data science and analytics; the rapid development of simple data to extended datasets leading to exponential growth of Big Data.

Searching through large amounts of unstructured data, looking for usable information. Nowadays, it is one of the most rapidly growing businesses in the world. How different was this many years ago? In the last decades, the uses and applications of the produced data have undergone a huge development and the amount of applications is still growing rapidly. Data can be seen everywhere nowadays, from information of the smallest production machines to the tracking data of the biggest companies in the world. Due to the increased amount of volume and the usage of it, the merge of these datasets will grow into Big Data. The applications of Big Data create new ways of working, living, behaving, communicating and cooperating and is the connection between companies, governments, electronic devices and humans. Merging datasets result in enormous data resources[1] and the creation of data is still growing exponentially. In 2009, the digitally created and replicated data was about 0.8 Zettabytes[2] worldwide and this grew to 10 Zettabytes in 2015. According to the International Data Corporation (IDC), it is expected that the global production of data will have grown to 180 Zettabytes by 2025[3]. Table 1 below shows an overview of the different volumes of data.

1 Byte	= 8 Bits	
1 Kilobyte	= 1024 Bytes	= 8.192 bits
1 Megabyte	= 1024 Kilobytes	= 1.048.576 Bytes
1 Gigabyte	= 1024 Megabytes	= 1.073.741.824 Bytes
1 Terabyte	= 1024 Gigabytes	= 1.099.511.627.776 Bytes
1 Petabyte	= 1024 Terabytes	= 1.125.899.906.842.624 Bytes
1 Exabyte	= 1024 Petabytes	= 1.152.921.504.606.846.976 Bytes
1 Zettabyte	= 1024 Exabytes	= 1.180.591.620.717.411.303.424 Bytes
1 Yottabyte	= 1024 Zettabytes	= 1.208.925.819.614.629.174.706.176 Bytes

Table 1 Overview different volumes of data
Source: (Dictionary, 2011)

In 2001, the upcoming concept of Big Data was defined by industrial analyst Doug Laney. He described the characteristics of Big Data as the 3V's; large **volumes** of **varied** data that are developed and handled at high **velocity**[4]. The volume refers to the size of the data files and is measured in scales of Terabytes, Petabytes or Exabytes[5]. The variation in the data is

[1] See Loebbecke & Picot (2015), p. 149
[2] See Gantz & Reinsel (2010), p. 2 ; Manyika et al. (2011), p. 16
[3] See Press (2017), webpage
[4] See Laney (2001), p. 1-2
[5] See Gandomi & Haider (2015), p. 138

twofold. First, there is a dissimilarity between internal data and external data[6]. Second, there is variation in the structure of data and could be described as structured data, semi-structured data and unstructured data[7]. Further, the velocity refers to the speed of processing the data.

Overall, there are many views and opinions about the use of Big Data. That results in an increasing awareness of the benefits Big Data applications. Unfortunately, many companies are currently not able to use the applications optimally and professionally. Throughout the years, the use of Big Data has proven itself as a tool with advantages in all different areas. It can be used as extension of information that already exist within companies[8]. For every company and sector, it will vary on the strategic goals of the company or sector what will be valuable data[9]. The value can be described as both economic and social. Economic value refers to the benefits that can be measured by the increases in profit, competitive advantage, better financial performances and business growth. Social value refers to the improvement in the social well-being in fields such as healthcare, security, education and public safety[10]. For instance, various government agencies can increase citizen engagement in public affairs and enhance transparency by using information found in Big Data. Because of social value, every single or group of users enhances the benefits that arise with the use of Big Data.

Next to the advantages and endless possibilities of Big Data, there are also some barriers that need to be discussed for a successful implementation. Issues as storage problems in the traditional capacity[11], lack of skilled people[12], data security[13], data privacy[14] and data complexity[15] are commonly faced in the Big Data industry.

1.2 Big Data in purchasing: the current situation of two independent strategic functions which can reinforce each

Due to the possible applications of Big Data, it can have an additional value for every part of a company. One of the departments where Big Data can play an important role, is the purchasing department. According to research, purchasing data is even the most frequently

[6] See Halper & Krishnan (2014), p. 4
[7] See Oussous, Benjelloun, Ait Lahcen, & Belfkih (2017), p. 3
[8] See Portela, Lima, & Santos (2016), p. 604
[9] See Günther, Rezazade Mehrizi, Huysman, & Feldberg (2017), p. 191
[10] See Günther et al. (2017), p. 191
[11] See Oussous et al. (2017), p. 2
[12] See Dhanuka (2016), p. 18
[13] See Oussous et al. (2017), p. 4
[14] See El-Darwiche et al. (2014), p. 14 & p. 18
[15] See Zschech, Heinrich, Pfitzner, & Hilbert (2017), p. 2614

collected data type and is a function that each company faces[16]. Overall, purchasing was defined as the acquiring of services and goods to accomplish specific goals. Therefore, purchasing was a supplementary activity with little importance for the company throughout the years[17]. In some companies purchasing is still a supportive and administrative function, but in an increasing amount of organizations the purchasing function is growing into a more strategic function[18]. Companies in a highly competitive market are forced to focus more on product innovation, supplier relationships, lead-times and cost savings. A strategic purchasing function can make the difference in this competitive market and obtain more value with reduced costs[19]. Purchasing as a strategic role can play a crucial role in the long-term goal of the company.

The use of data is not new within the purchasing function and the first best case data influences started from a computer network more than 50 years ago. At that time, the network established a paperless way of doing purchasing, billing and payments. The computer network was the basis for the Electronic Data Interchange (EDI), which can be seen as an inter-organizational information system[20]. The EDI facilitates a link between organizations, especially between buyers and sellers. The intra-organizational link emphasized an automated computer-to-computer data exchange of commercial documents and information. The main purpose of EDI was the exchange and processing of data between organizations with as less human intervention as possible. The implementation of EDI caused increases in the speed and accuracy of the purchasing process. Next to that, in 1978, Charnes, Cooper and Rhodes published a pioneering paper over the Data Envelopment Analysis model (DEA)[21]. That model is a linear programming model that enables the development of an improved evaluation system of purchasing performance[22]. In 1996 the DEA model was applied in the evaluation of a supplier of individual products[23]. This model is shown as the first combinations of data analytics and purchasing. Currently, it is still applied as a mathematical programming approach for evaluating the relative efficiency[24]. Some given examples in the literature are applied for determining the efficiency-based

[16] See Lismont, Vanthienen, Baesens, & Lemahieu (2017), p. 116
[17] See Úbeda, Alsua, & Carrasco (2015), p. 177
[18] See Knoppen & Sáenz (2015), p. 123
[19] See Úbeda et al. (2015), p. 177
[20] See Banerjee & Sriram (1995), p. 29
[21] See Charnes, Cooper, & Rhodes (1978), p. 432
[22] See Easton, Murphy, & Pearson (2002), p. 124
[23] See Liu, Ding, & Lall (2000), p. 143
[24] See Ebrahimnejad, Tavana, Lotfi, Shahverdi, & Yousefpour (2014), p. 308

ranking of energy firms[25] and as efficiency multiple criteria model in the banking industry[26]. Another case of the implementation of a data related system goes back to 1993, when Kurt Salmon and Associates made a statement. They found: "By expediting the quick and accurate flow of information up the supply chain, ECR enables distributors and suppliers to anticipate future demand far more accurately than the current system allows"[27]. ECR is the abbreviation of Efficient Customer Response and refers to the flow of data. Throughout the years there are several other implementations with the use of data. One important system to mention is the electronic-procurement (e-procurement) system, which is still in use nowadays. E-procurement is a technological solution and has the possibilities to transform the purchasing process into a buying system that uses the internet[28].

Hence, it is not surprising that the use of data has an added value for purchasing. In each step of the whole process, there is creation of data. The use of data through the internet helps the purchasing function for improvement in all stages, from the supplier selection process to the final buying process and so on. The data influence can help reducing long throughput times, waiting times in the process and eliminate bottlenecks. It gives more control and insights in the buying process, makes processes compliant and reduces waste of unnecessary activities. In the supplier selection process, the data can support electronic-sourcing (e-sourcing) and in the buying process, it can support, for example, decreasing the product life cycle and reducing time to the market[29]. From the suppliers' side, there can be advantages with the use of data by adjusting the supply of their products and processes to the preference of the buyer[30]. Therefore, there are many data integrations in the purchasing processes. Due to the increasing amount of functional possibilities, data is more and more important for purchasing. With the current applications, data has a more real-time information characteristic. When the data meets the requirements of Big Data, it becomes possible to make forecasts and predictions in trends and behaviour concerning every part of the process[31]. The same as in other parts of a company, the use of Big Data in the purchasing department is also still in its infancy. Even in purchasing, Big Data can produce competitive

[25] See Khalili-Damghani, Tavana, & Haji-Saami (2015), p. 760
[26] See Ebrahimnejad et al. (2014), p. 308
[27] See Lummus & Vokurka (1999), p. 13
[28] See Presutti (2003), p. 221
[29] See Presuttim (2003), p. 221
[30] See Schulz (2017), p. 3
[31] See Schulz (2017), p. 5-6

advantages in, as for example in optimizing the processes. In addition to that, with both purchasing and Big Data as strategic decision makers, they can reinforce each other.

1.3 Bright Cape; a company for Small & Big Data solutions

In every company is the knowledge and implementation level of Big Data different and depends on the existing situation. In countless companies the use of Big Data influence is still in the initial phase and in some companies there is a more developed implementation of Big Data. Since many companies do not have the right knowledge or facilities to indulge in the advantages and possibilities of Big Data, they ignore it and miss the opportunities. Bright Cape, founded in 2014, is a company that performs as the link between the deeper knowledge and insights of Big Data and to support the implementation and knowledge about the company's specific data, in organisations which are not able by themselves to use. Therefore Bright Cape is a consultancy company and focuses on 'Small and Big Data solutions'[32].

Bright Cape uses the data as a competitive advantage, establishes new connections, understands the market and discovers trends in new and existing target groups. All solutions are found in the own data of a company. The solutions are based on mathematical methods, algorithms, the 'User Centered Design (UCD) process' and Process Mining. This last technique is involved with discovering, monitor and improve the real processes, by extracting knowledge from the available systems[33]. Next to this, Bright Cape created strategic partnerships with software companies, which are supporting these solutions: 'Celonis Process Mining' and 'SAS'. As a result Bright Cape makes the data understandable and ready to use for the company. Their first main driver was to evolve the gut feelings into underpinned decissions. That driver was related to the fact that most of the times, decisions within the companies are based on intuition and feelings from the heart. Decisions can be stronger and better in line with the strategy of the company, when it is supported by data. The decisions are then based on intuitions and substantiated with hard and accurate facts. Bright Cape focusses and specialised on a different kind of analytics and combine this with domain experience and knowledge in the following domains: finance, risk, logistics, energy and marketing. Similarly, there are analytical projects involved with UX (User Experience).

[32] See Bright Cape (2017), website
[33] See Sahin, Accorsi, Frey-Luxemburger, Knahl, & Atkinson (2017), p. 151

To analyze how the current situation of the company is, in the degree of data intelligence, Bright Cape designed it's own maturity model. This model is based on Gartner's Enterprise Information Management (EIM) Maturity Model. This model identifies the current stage of the maturity level that the company or specific department has reached. The overall maturity level can be dissimilar than a deparment related maturity level due to department structures and developments. Further, it shows which actions are necessary to grow to the next level[34]. The first steps of the EIM Maturity Model are in the reactive part, where action follows as a reaction. How higher the degree of intelligence, the more competitive advantage for the company and the more the reaction is proactive. The model is shown in figure 1 below.

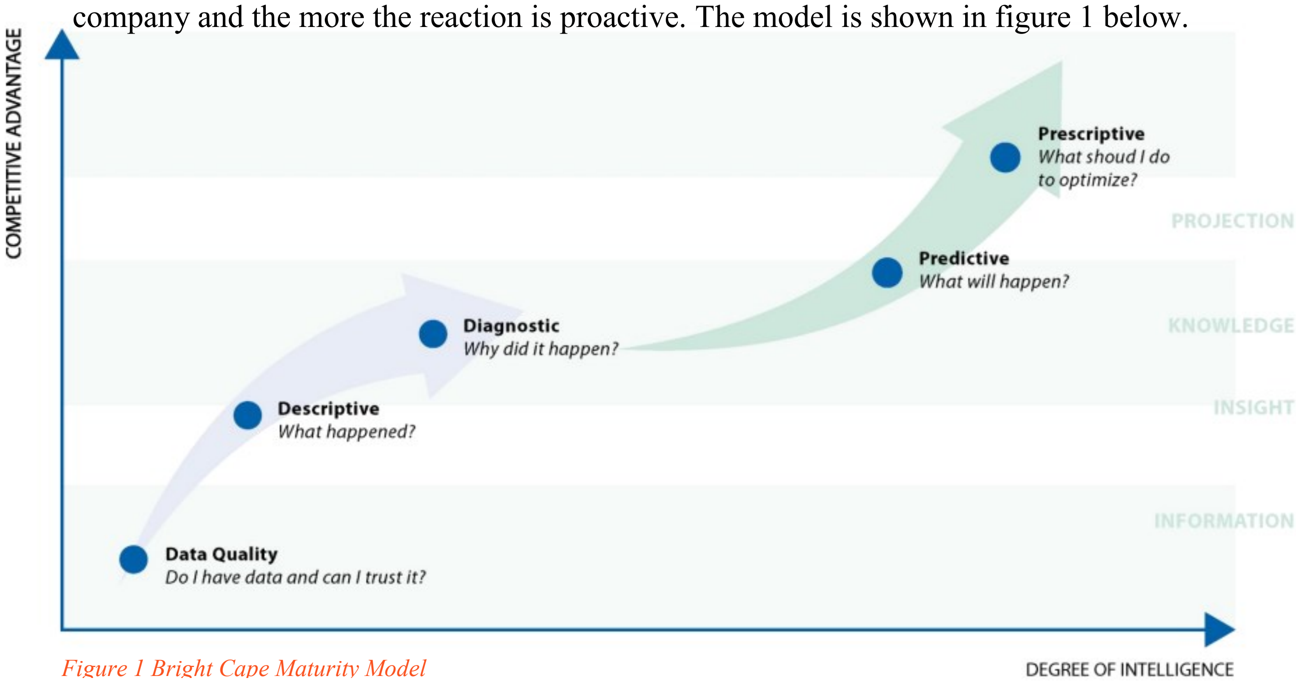


Figure 1 Bright Cape Maturity Model
Source: (Bright Cape, 2017)

For organisations it is necessary to mention that they cannot skip stages or activities during their growth. That makes an organization weaker instead of stronger and leads to dysfunction of the maturity model. The reason for this is that the growth to the next level is building on the current level, what serves as a foundation. When there is not a solid basis, the next step does not fit and can cause weaknesses and unstable situations.

The Bright Cape Maturity Model is a global model that is used to define the current situation of the total company. When the model is more specified to a subject, it gives a more detailed overview of the specific situation. It provides specific scoring criteria to define the level of maturity the company is categorized in. The more the maturity model is specified, the better Bright Cape can find their guidelines and solutions for improvements. One of the domains

[34] See Newman & Logan (2008), p. 3-8

where Bright Cape finds data solutions, is in the purchasing department. Purchasing and Big Data strengthen each other since both can support in making strategic decisions, as stated before. For Bright Cape and the specific customer, it is in both their advantage and saves time when the solution fits easily in the current Big Data environment.

To analyse how developed the degree of intelligence is, a Big Data Purchasing Maturity Model is a good idea to show the combination of Big Data and Purchasing. In the last thirty years, several existing Purchasing Maturity Models were developed as stated in the current literature. These models differ in specific content and in numbers of different steps but have in common that they show the 'as-is' situation of the maturity of the purchasing function. In addition to that, in Big Data literature there are recent Big Data Maturity models which are developed or adapted in the last couple of years. The history of Big Data Maturity Models is shorter than the Purchasing Maturity Models since Big Data is a pretty new phenomenon. Building on this, both in the current literature as within Bright Cape there are yet none designed maturity models for Big Data and Purchasing combined. That results in the fact that Bright Cape can have an advantage of the combined models, but that model does not exist at this moment. In schematic view the current situation leads to the following problem:

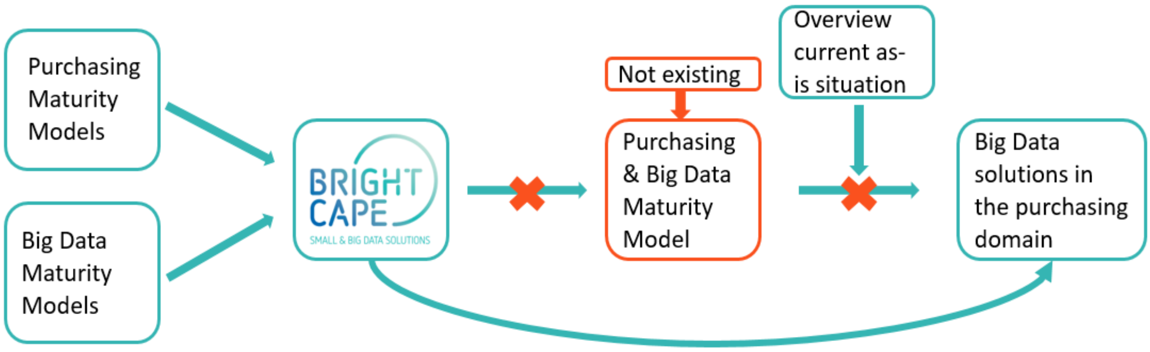


Figure 2 Problem statement

Since there is not an existing Big Data Purchasing Maturity model, the final goal of this thesis is to design a new maturity model that Bright Cape can use. In this research, the following main research question is treated:

"How are the different steps designed and specified in the new Big Data Purchasing Maturity model?"

To explore this research question, the following sub questions are formulated:

- *"What is the current situation of Big Data in purchasing?"*
- *"How does a purchasing maturity level relate to a Big Data maturity level?"*

To answer the (sub) research questions, the remainder of this thesis is structured as follows. The next part explains and explores the available literature in the literature review, to gain deeper knowledge about the different subjects. The literature review is spread out over three chapters. It starts in chapter two with literature about the history of Industry 4.0 and Big Data, followed with purchasing as function and purchasing theories in the third chapter. The fourth chapter is the literature review about maturity in general and maturity models specified to Big Data, purchasing and industry 4.0. Also, the maturity model of Bright Cape further explained. In that chapter, there are also different schematic overviews of the existing and relevant maturity models. In chapter five the methodology for collecting data is explained. This chapter elaborates on the set-up and background of the literature review, interview methods and for design a new model. In the sixth chapter the results of the different interviews are worked out and analysed. There is also an ideal situation created for Big Data in Purchasing. Further, the chapter ends with developing a new Big Data Purchasing Maturity Model for Bright Cape. Supplementary, a discussion and conclusion that answers the main research question and sub research questions is provided. Finally, the limitations and opportunities for further research are included in the last chapter. In figure 3 below there is a schematic overview of this research.

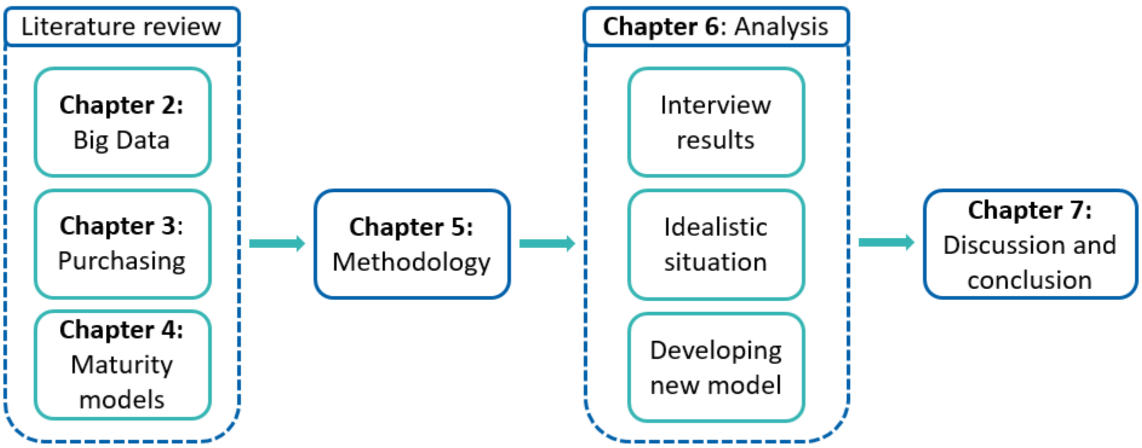


Figure 3 Schematic view of the research

2. The evolution and characteristics of Big Data in the history

2.1 Big Data is part of the current industry 4.0, with its cyber physical systems, machine-to-machine communication and other technological developments

The rise of Big Data is not a stand-alone technical development, but it is part of the industrial revolution. The present stage of the industrial revolution is industry 4.0 and it made its entrance some years ago, at the Hannover Exhibition in 2011[35]. This built on industry 1.0, industry 2.0 and industry 3.0, what went through a transformation from an economy driven by agriculture and handicraft to an economy based on industry and machine factoring. Subsequently, it changed to an industry that uses electronic and Information Technology (IT) systems and results in an industry driven by cyber-physical systems. The main characteristics of the industrial revolution are related to technological, socio-economic and cultural factors.

The start of the first industrial revolution goes back to Great Britain in the late eighteenth and early nineteenth century. This industry is characterised by the benefits of mechanisation and the industrial society that replaced the agriculture[36]. The introduction of the machinery in the manufacturing industry changed the production processes. The machines were driven by steam, water and wind energy. This offered possibilities for starting a comprehensive cooperation between animal/ human labour and machines. In 1784, the first mechanical loom made its entrance and was a sign for other factories to improve. With the new conditions and ways of working, production locations left the local homes and shifted to central locations in large cities. Factories with the same specialization grouped together and created industrial cities close located to their core sources[37]. The production process grew impressively and achieved a space for optimisation. The supply of products could better meet the strongly increasing demand. Due to the new know-how and other major developments, the first blueprints for the factories were formulated as we recognize them nowadays[38].

Nearly a century later, in the late nineteenths, the second revolution made its appearance in the slaughterhouses of Cincinnati. The main factor of the change is the emergence of new energy sources, namely electricity, oil and gas. That resulted in the first production lines and

[35] See Drath & Horch (2014), p. 56-58
[36] See Yin, Stecke, & Li (2017), p. 1
[37] See Rodrigue (2017), webpage
[38] See Sentryo (2017), webpage

increased the output enormously. Due to the use of the new energy sources, the development made big changes to products how we use them nowadays. With the discovery of the telephone and the telegraph, pioneering communication methods were revolutionised[39]. Moreover, breakthrough innovations such as cars, planes, electronic and mechanical devices where included in the developments. The long-distance transport and communication methods result in an expansion of the market reach. The continuous improvement theory and lean manufacturing, as envisioned by Henry Ford, underlies all these productivities and developing explosions[40]. Therefore, the second industrial revolution can be seen as the revolution of mass production[41].

The third industrial revolution started when Modicon presented the first programmable logic controller in 1969. That controller made digital programming of systems possible and used electronic and information technology to automate the production[42]. The machines are able to repeat tasks under minimal supervision and with well-defined parameters. This led to a rise of the era of high-level automatization. At the same time, a new type of energy emerged, nuclear energy. The use of this energy results in an introduction of electronics as the transistor and microprocessor. Moreover, globalisation enables a minimisation of sources, labour hours and transport costs and thus results in a new manufacturing landscape and advanced economies. Because of all the changes and digitalisation, traditional industries changed in character whilst new industries arose.

Since 2011, the fourth revolution has unfolded, and the applications are multiplying at lightning speed. In contradiction to the other three revolutions, the fourth revolution was not related to the need for a new type of energy but rooted in a new technological phenomenon; the digitalisation. The digitalisation leads to build a new virtual world, driven by cyber-physical systems and robotics. In the same line as the industrial revolution, the technical evolution of the Internet is growing parallel. The current industry aims to connect all factories and production lines with technological applications for real-time interaction. The communication is possible through applications as Big Data, Cloud and Internet of Things (IoT). Other additional technologies such as Block chain, three dimensional (3D) printers, Artificial Intelligence (AI), real-time sensor technologies and so on will help to influence

[39] See Sentryo (2017), webpage
[40] See Bhuiyan & Baghel (2005), p. 763
[41] See Tanenbaum & Holstein (2016), webpage; O'Hare (2017), webpage
[42] See Drath & Horch (2014), p. 56

and integrate future production and services[43]. The application possibilities are enormous, as improved decision-making, anticipating of inventory, predictive maintenance, and improved coordination among jobs. The focus shifted to global value chains where supply chain management and global manufacturing became closely embedded.

To explain the way of working of the technologies, some important usages will be further explained. First, Internet of Things is the network of physical devices and appliances embedded with sensors, software, electronics, actuators, that enables ways to make a connection between the devices and measurement appliances, with the exchange of data[44]. Second, the block chain technology is a public ledger where transactions are recorded and confirmed anonymously and is shared between many parties. The network is owned by nobody and can be used by everybody as a public record. Once the transaction is entered, it cannot be changed afterwards. Developments as cryptocurrencies uses that technologies[45]. Additionally, the cloud computing can be seen as outsourcing computer applications or data management to an external service provider. The data is stored across various servers without knowing exactly what the location is. It can be seen as an important part of the internet developments, since everyone can work with it and can use it everywhere[46]. Finally, cyber-physical systems (CPS) relates to systems that have transformative technologies for managing interconnected systems between computational capabilities and its physical assets. Cyber-physical systems used for further developing and managing Big Data. With the implementation and integration of these systems, it enables companies to further develop industry 4.0 applications[47].

2.2 Big Data is characterised by the 3V's; Volume, Variety and Velocity.

As result of the fourth industrial revolution, Big Data is an enormous growing potential. When further analysing Big Data, there is not one clear definition of the statement, since several researchers have different views on Big Data. For instance, Manyika et al. (2011) states Big Data as "Datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyse"[48]. Davis and Patterson (2012) define Big Data

[43] See O'Hare (2017), webpage
[44] See Brown (2016), webpage
[45] See Augur (2015), webpage
[46] See Koops, Leenes, de Hert, & Olislaegers (2012), p. 1
[47] See Lee, Bagheri, & Kao (2015), p. 18
[48] Manyika et al. (2011), p. 1

as "Data too big to be handled and analysed by traditional database protocols such as SQL"[49]. Likewise, Gantz and Reinsel (2011) defines Big Data as "Big data technologies describe a new generation of technologies and architectures, designed to economically extract value from very large volumes of a wide variety of data, by enabling high-velocity capture, discovery, and/or analysis"[50]. Meanwhile Hashem et al. (2015) refers to Big Data as "A set of techniques and technologies that require new forms of integration to uncover large hidden values from large datasets that are diverse, complex, and of a massive scale"[51]. Finally, Wu, Zhu, Wu & Ding (2014) characterised Big Data as "Large-volume, heterogeneous, autonomous sources with distributed and decentralized control, and seeks to explore complex and evolving relationships among data"[52]. These definitions all show one common point, namely that it is all related to the size of the data. Based on these mentioned definitions and on the results of the interviews, in the remainder of this thesis the following definition as Big Data will be used: *"Big Data are an extreme voluminous and complex dataset(s) where the data comes from a variety of structured and un-structured data sources, what can be transformed with new technologies and advanced analytics into valuable structured data and knowledge"*.

The data from Big Data, are not new, but are parts of the small datasets which are used and stored since the first data integrations. Therefore, it could be assumed that Big Data are extensive versions of (different) small datasets. The size of the dataset is not the only leading criteria what can be quantified for defining small or Big Data. Many authors use explicitly the combination of the 3V's to characterise Big Data from small data[53]. The three main mentioned V's are Volume, Variety and Velocity. Some author's added Value and Veracity as additional V's and defined it as the 5V's. Other additional added characteristics to define Big Data are Visualisation, Verification, Validation, Variability, Vision, Immutability and Complexity[54]. Therefore, the terminology Big Data is mostly used as an "umbrella term" that covers all related and involved characteristics[55]. For comparing the differences for each

[49] Davis & Patterson (2012), p. 4
[50] Gantz & Reinsel (2011), p. 6.
[51] Hashem et al. (2015), p. 100.
[52] Wu, Zhu, Wu, & Ding (2014), p. 98.
[53] See Laney (2001), p 1-3; Chen, Chiang, & Storey (2012), p. 1182; D. T. Moore (2014), p1; El-Darwiche et al. (2014), p. 5; Halper & Krishnan (2014), p. 14; Radcliffe (2014), p. 4-5.
[54] See Jagadish et al. (2014), p. 88; Gandomi & Haider (2015), p. 138-140; Özköse, Arı, & Gencer (2015), p. 1043; Emani, Cullot, & Nicolle (2015), p. 78-79; Portela et al. (2016), p. 605; Oussous et al. (2017), p. 3; Günther et al. (2017), p. 10.
[55] See Merino, Caballero, Rivas, Serrano, & Piattini (2016), p. 124.

characteristic between small data and Big Data, the following overview in table 2, summarizes the differences of the 5 most important V's. Each characteristic is further explained in the next paragraph, with some illustrative examples.

	Small data	**Big Data**
Volume	Limited to large datasets	Large to very large datasets
Variety	Limited to wide and is structured	Wide and is semi- or un-structured
Velocity	Slow, freeze-framed/ bundled, controlled and a steady flow	Fast, continuous and can accumulate within very short time
Value	Business Intelligence, analysis and reporting	Complex data mining for prediction, pattern finding etc.
Veracity	Contains less noise as data is collected in a controlled manner	Usually quality of data not guaranteed.

Table 2 Comparing small and Big Data
Sources: (Bista, 2018), webpage; (Kitchin & McArdle, 2016), p. 2.

Volume relates to the magnitude of the data[56]. Volume can be seen as an important V of the 3V's because it defines the size of the dataset and refers to the storage space required to record and store data[57]. As seen in the table, a large dataset can be either small data or Big Data. That results in the fact that a concrete definition of the threshold when small data will be Big Data cannot be clearly stated. This threshold depends on factors related to the time, the type and source of the data, the specific industry and the storage. What today may hold for Big Data, may not meet the thresholds in the coming years, because the environmental elements which transform data into Big Data will also change in the years[58]. Say fifty years ago was a dataset of 50 kilobytes (KB) mentioned as Big Data, around 2012 was that 200 Gigabyte (GB) and nowadays datasets reach more than 1Terrabyte (TB) of storage[59]. The type of the data relates to the involved data that is measured by one data point. For example, sensors for pollution or sound requires little storage of 1 GB annually for all the measurement records, while Facebook processes daily more than 500 TB of data[60]. Each sector handled other types of data, what leads to different volumes for processing. In 2012 for example, econometrics define a dataset of 200 GB as Big Data, while physics see 200 GB still as a small dataset[61]. Another mentioned additional example, is that a 2.5-meter wide-angle optical telescope creates 200 GB per night of data what can be mentioned as Big Data, while

[56] See Gandomi & Haider (2015), p. 138
[57] See Kitchin & McArdle (2016), p. 6
[58] See Gandomi & Haider (2015), p. 138
[59] See Diebold (2012), p. 1
[60] See Kitchin & McArdle (2016), p. 6; Sinha (2016), webpage
[61] See Diebold (2012), p. 2

200GB per night of social media or telecom data is more defined as a small dataset[62]. For volume related to storage, it is mostly mentioned that normal used computers and laptops with programs as Excel are a threshold for defining data as Big Data. When the computers and programs are not able to progress the dataset, then it is Big Data. Concluding could be said, out of the extended literature review, it is hard to define one specific threshold that holds for all Big Data volumes. For the remainder of this master thesis, and in cooperation with professors from the University of Twente[63], manager from Bright Cape[64] and one exceptional literature research[65], the threshold between small data and Big Data is stated into two different thresholds. Hereby one of those thresholds should always be possible to determine data or Big Data. The first threshold is stated in line with the limitations of Excel, where it is not possible to have a dataset of more than 1,048,576 rows and/or 16,384 columns[66]. If the dataset is designed in another format that Excel, the threshold for 1 TB is stated[67].

Variety of data types refer to the heterogeneity in the dataset since there is not one fixed data structure[68]. The data can only be stated as Big Data, when there is a variety in sources among the dataset. The origin of the data can have both internal and external entries for the company. That can come from public or private sources, local or far-away and shared or confidential sources[69]. The formats of the data can be structured (from databases and spreadsheets, and constitutes only 5%), semi-structured (from social media, weblogs, email, sensors, mobile devices etc.) or un-structured (from video's, audio, images etc.)[70]. Therefore, the representation is rarely perfectly ordered nor ready for usage. This variety of formats made it difficult for what concerns the warehousing, processing, and data management. The current warehouses, or Relational Database Management Systems (RDBMS) do not have the capabilities to store all the different data formats.

Velocity of data involves the rate at which the data come in and the speed at which they must be acted upon[71]. The data arrive at high velocity with >500TB each time frame. The importance is situated in the speed of the feedback loop, capture real-time data from

---

62 See Katal, Wazid, & Goudar (2013), p. 406
63 Source: personal communication Dr. E. Constantinides (2018), D. Bucur (2018) and R. Haverkort (2018).
64 Source: personal communication J. Hilberink (2018).
65 See Kitchin (2014), p. 1.
66 See Microsoft (2018), webpage.
67 Based on personal communication respondents of Bright Cape.
68 See Jagadish et al. (2014), p. 88; Emani et al., 2015), p. 71.
69 See Oussous et al. (2017), p. 3.
70 See Gandomi & Haider (2015), p. 138; Emani et al. (2015), p. 72.
71 See Gandomi & Haider (2015), p. 138.

input through to analysis and at high feed rates of speed[72]. Therefore, the data are often non-stationary and should be analysed more quickly. Examples of the unprecedented rate of data creation is the proliferation of digital devices such as sensors and smartphones as well as retailers as Wal-Mart and website YouTube. They process, for instance, more than hundreds of thousands of streaming data sources. With a real-time analysis, it can create real customer value. Traditional data management systems are not able to handle that amount of data feeds, wherefore new Big Data technologies play an important role. For the remainder of this master thesis, the threshold between data and Big Data is hold on an (almost near) real-time processing power of the data.

Next to the 3V's, the additional characteristics will be discussed. Starting with the two additional from the 5V's. First is **value**, which is the benefit as result of the analyses. Big Data is often mentioned as 'low value density'. That means that the wide variety of data as it is received in their original form, has usually a low value[73]. With structuring and analysing the data, a high value can be obtained. The value is both for analytical use of supporting human decisions and needs and is enabling new business models[74]. **Veracity** refers to the accuracy of the data if it is conformed to the truth or facts[75]. Due to the characteristics of veracity, results deducted from Big Data are not able to be verified; but they can be allocated to a probability. The data should come from accurate sources and its security should be provided, because uncertainty and noise can be caused by many sources[76].

Furthermore, **Visualisation** is the transformation of data to a visible application, as a report, dashboard, graph or Big Sheets. It makes it possible to support easy result monitoring and possibilities for formulating recommendations for decision making issues. The partner software tools of Bright Cape are also based on the visualisation of the data into visual graphs, which supports the decision-making process. **Verification** is to conform the processed data to preconceived specifications[77]. **Validation** is the check of the intention is fulfilled with the results[78]. **Variability** relates to the variation in the data flow values. Often the variability is not consistent and has peaks and valleys in the data[79]. **Vision** is the purpose

[72] See Emani et al. (2015), p. 72.
[73] See Gandomi & Haider (2015), p. 139.
[74] See Emani et al. (2015), p.72.
[75] See Emani et al. (2015), p.72.
[76] See Gandomi & Haider (2015), p. 139; Günther et al. (2017), p. 200.
[77] See Benjelloun, Lahcen, & Belfkih (2015), p. 1; Oussous et al. (2017), p. 3.
[78] See Benjelloun et al. (2015), p. 1; Oussous et al. (2017), p. 3.
[79] See Gandomi & Haider (2015), p. 139.

of the use of Big Data[80]. **Immutability** is the variation when the data is gathered and stored accurately managed, then the Big Data can be permanent[81]. **Complexity** of the data refers to the situation that the Big Data are generated from a huge number of sources. Therefore, the connection, cleaning and transformation of the different data sources imposes a critical challenge. It is difficult to organise and hard to analyse[82]. **Privacy** refers to access and privacy of the storage of data sources. That includes data of social media sources, geospatial sources and anonymous tracking of Global Positioning System (GPS) sources[83].

2.3 Data Analytics, with data mining, process mining and machine learning will be used to extract knowledge from Big Data

The characteristics of Big Data do not enable data researchers to use and process the data. There is data management necessary, that describes the different steps of the process from collecting and transforming the data, to providing information and making analyses[84]. Emani, Cullot and Nicolle (2015) define the four different steps of the analysis as the 4A's; Acquisition, Assembly, Analysis and Action. Mountasser, Ouhbi and Frikh (2016) mentioned the same four steps as layers and named them; Acquisition, Organisation, Analysis and Knowledge Capitalisation[85]. Building upon that, the data analytics is part of the steps of data management, and digs further in technology. Analytics can be seen as a sub-process in the overall process and uses techniques to analyse and master intelligence from Big Data and even from smaller data sets. Kwon, Lee and Shin (2014) defined data and Big Data Analytics as "Technologies and techniques that a company can employ to analyse large scale, complex data for various applications intended to augment firm performance in various dimensions"[86]. Since there are many characteristics and variations of Big Data, there are numerous different methods and techniques used for the analytics on the data[87]. These methods are for both structured and unstructured data. The three most important methods are: data mining, machine learning and process mining, and they will be further explained.

Data mining is the process of discovering patterns in large datasets. It is not required that the dataset will meet all the different characteristics of Big Data. The term 'data mining' is

[80] See Benjelloun et al. (2015), p. 1; Oussous et al. (2017), p. 3.
[81] See Benjelloun et al. (2015), p. 1; Oussous et al. (2017), p. 3.
[82] See Benjelloun et al. (2015), p. 1; Oussous et al. (2017), p. 3.
[83] See D. T. Moore (2014), p. 1; Jagadish et al. (2014), p. 88 & p. 92, Hashem et al. (2015), p. 111.
[84] See Zschech et al. (2017), p. 2616.
[85] See Emani et al. (2015), p. 72; Mountasser, Ouhbi, & Frikh (2016), p. 105.
[86] Kwon, Lee, & Shin (2014), p. 387.
[87] See Emani et al. (2015), p. 74.

also related to the process of mining valuable knowledge in the data[88]. The process of exploring and analysing allows extremely large datasets. The main purpose is mostly predictive (to predict a value based on data) or as classification tool (sorting data into groups). The patterns are created by analysing for frequent if-then situations, and to locate the most important relationships within the data[89]. Data mining is applicable in all different sectors, and the advantages may vary depending on it. As overall benefit counts the ability to uncover hidden relations and patterns, which can be used as predictive decision-making tool. As examples for specific goals or industries, data mining can serve as a tool for marketing campaigns, detect historical sales patterns for predictions for the purchasing department, improved supply chain operations for the production companies or as a risk model or a fraud detection tool for the financial or insurance industry. There are several algorithms that process the data for data mining, and it depends on the input and output which algorithm is the best suitable. An overview of these algorithms is in table 3.

Algorithms	**Description**
Classification trees	Classify the dependent categorical variable, based on predictor variable(s). Results in a tree with nodes and links between the nodes that can be read to form if-then rules.
Logistic regression	Statistical technique of a standard regression extends with classification. Results in a prediction formula that the probability of the occasion as a function of the independent variables predict.
Neural networks	Software algorithm that consists of input nodes, hidden layers and output notes, where each part is assigned to a weight. Data is send to the input node, and by a trial and error method, the algorithm adjusts the weights until a defined stopping criterion is reached.
Clustering techniques	Technique of identify groups within similar data. It calculates the distances between the record and points in historical (trained) data.

Table 3 Data mining algorithms
Source: Kaur & Singh (2017), p. 407

Machine learning

Machine learning refers to the discovery of knowledge and making intelligent decisions[90], and is also known as statistical learning from the complex data[91]. The theory behind machine learning is based on studying algorithms for prediction and interference and uses statistical models to predict an output[92]. It is mainly discussed that machine learning is in the same line as data mining, since they have common characteristics of finding patterns in data. However,

[88] See Kaur & Singh (2017), p. 407.
[89] See Rouse (2017a), webpage.
[90] See Oussous et al. (2017), p. 5-6.
[91] See Emani et al. (2015), p. 74.
[92] See Zakir, Seymour, & Berg (2015), p. 82.

machine learning emphases on designing algorithms that can learn from and define forecasts on the data, while data mining have concentrations on determining properties of data sets[93]. The algorithms for machine learning are mostly defined as supervised or unsupervised[94]. Supervised algorithms mean that there is a human interaction within the process. The input and desired output are manually labelled for classification and are used for training or as learning basis. Once the training is successfully completed, the algorithm will apply that knowledge into the new data. Contrarily, the unsupervised algorithm does not involve human participation, but it is more an iterative approach. In unsupervised learning, an artificial intelligence (AI) system groups the data with similarities and differences, even though there are not stated categories beforehand. In this situation, the algorithm can perform more complex datasets and processing tasks than supervised algorithms. While the results may be more unpredictable than supervised algorithms.

Process mining is the final important data analytics method. Process mining is a method that analyses the event data and discovers the underlying processes what results in a process model[95]. Event data is data from the execution of a system and holds for example a time stamp. So, process mining shows what has been done during the process and detect discrepancies in what should have been done. Once there is a process model, the event data can replay on the model to check if there are any bottlenecks or delay in the process[96]. When there is a bottleneck in the process, the underlying reason is important to know. That requires an extended analyse if there is any correlation in the different process characteristics. An overview of the most well-known process characteristics is in table 4 below.

Process characteristics	Description
Control-flow perspective	Control on the flow how the next step is going to be accomplished.
Data-flow perspective	The control of the data flow in sequence or in line is.
Time perspective	Relates to the time stamp of the processes. It shows the duration of the different steps or the remaining time to the next step or to the end of the process.
Resource/ organisation perspective	How the different sources performing a particular activity.
Conformance perspective	The process of skipping a mandatory part of the process or executing some steps in a different order.

Table 4 Process mining characteristics
Source: De Leoni, Van Der Aalst, & Dees (2016), p. 236.

[93] See Amatriain (2015), webpage.
[94] See Rouse (2017b), webpage.
[95] See Dabrowski et al. (2017), p. 199.
[96] See De Leoni, Van Der Aalst, & Dees (2016), p. 236.

2.4 The use of Big Data in the organization results in several advantages and profit for the purchasing department

The influence of Industry 4.0, and especially data science, Big Data and Big Data analytics, results in many application possibilities and advantages for the entire industry. In every part of the production or within the organisation, more than one advantage can be mentioned. Since the main topic of this research, next to Big Data, is about the purchasing function, only the advantages and application possibilities in relation to purchasing will be discussed. Further, for the advantages in relation to Big Data, the threshold as stated in chapter 2.2 will be determined. It must be mentioned that the grey area around the thresholds is not necessary to be excluded since there should be some relevant data available. For understanding is chosen to use only the term Big Data in the advantages, since the result is more important than the specific data characteristic.

First of all, obtaining competitive advantage with the use of Big Data plays a crucial role in the positioning among competitors[97]. With the knowledge gathered from the available data, it can generate new information for competitive and complex decisions. The real-time information makes it possible to react more agile on the current situation[98]. It creates opportunities to understand the customer behaviour, segment the customer base and gives insights on the current trends and market movements[99]. These movements are earlier visible in the data than that it is noticed by the company. The competitors will stay behind, what will lead to an increase in the competitiveness of the market, in their own advantages.

Another mentioned example as competitive advantage is at the inventory management. With datamining of different datasets, Big Data can improve the quality to a full transparent process at the Stock Keeping Unit (SKU) level of the inventory. With a data driven forecasting and linked barcodes in the systems, automated replenishment systems can be integrated[100]. As result, the incidents of running out of stock are reduced and there is always an optimal inventory level available. When the stock level is not well determined, over capacity and under capacity are in the warehouses, which results in damages in the financial results. At the moment when the stock level is as optimal as possible, it saves money what

[97] See Portela et al. (2016), p. 605.
[98] See McAfee & Brynjolfsson (2012), p. 6.
[99] See Assunção, Calheiros, Bianchi, Netto, & Buyya (2015), p. 3.
[100] See Manyika et al. (2011), p. 70.

could be lead to profit for the customer. With that profit, the company can increase in competitiveness.

Next, due to the characteristics, Big Data enables the possibilities to analyse extremely large datasets and combine disparate datasets from different sources[101]. These datasets might include spend data, contract data, performance management, contract terms, invoice data, SKU data, bill of material (BOM) details, warranty/claims data and so on. With combining and integrate these data, the processes could be optimized and reduces mistakes.

Next to that, applications of Big Data are able to solve complex sourcing challenges when looking at the total costs[102]. The ability to increase the optimization of the most sourcing occasions will be one of the first areas where Big Data approaches become predominant. The advantages of this method are threefold. First, better award decisions are made relating to the total costs. Second, Big Data can play a crucial role for lower cost structures by influencing the sourcing process before the real sourcing take place. That is in the design phase, by suggesting changes to specifications and starting the most profitable collaboration in early phase.

The following advantage can be related to the integration of real-time external information, what can help in the decision-making process. Real-time information about (natural) disasters, war, political situations, price changes and availability of raw materials, or other influential factors, can influence when, where and with whom to source[103]. Integrated with data results in optimal knowledge of the external environment.

Further it can be mentioned that the productivity will increase. When the manufacturer has an integrated Big Data system, they can receive early in the process significant and predictive information for what they can expect, so that they are able to prepare. That result in more just in time management with raw materials available and ready to use[104]. Additionally, the manufacturers can drive efficiency in their production processes and optimize their facilities. That enables the company to do more with less products and produce with a higher-quality output.

[101] See Busch (2012), p. 1-2.
[102] See Busch (2012), p. 2.
[103] See Busch (2012), p. 2.
[104] See Manyika et al. (2011), p. 77.

Additionally, the business performance will be improved with the use of Big Data intelligence[105]. With the techniques managers are able to measure all the insights and they know radically more about their organisations. They can directly translate the knowledge from Big Data into the improved decision making. That will lead to an increased business performance due to the better results. As a result, the most internal savings in terms of finance can be made in the contributions of labour, capital, inventory, purchased services and IT investments..

Finally, it can be said that the Big Data revolution has many more advantages. Some mentioned practices are for example that it enables companies to run more and faster queries. In another case it gives the ability to constantly check the massive spend cube. Moreover, it is said that due to Big Data it can better measure, implement, manage and forecast cost avoidances.

2.5 Big Data has also some barriers that needs to be covered for the full benefits of the implementation of the data

Different companies, departments, functions or teams would not reap the full benefits of the implementation and the use of Big Data, unless they are able to work and manage it efficiently. The settles threshold between data and Big Data plays an important role for the barriers that needs to be fully covered. Also, the internal change for the implementation is an important foundation for a successful integration. McAfee and Brynolfsson (2012) defined five important factors which are particularly important in that process[106].

The first mentioned factor is **leadership**. Companies that are successful with the use of their data, are not more successful because of bigger or better datasets, but it is the way they manage it. Leadership teams that set clear goals, ask the right questions and define their success, will succeed to transform data into Big Data in the company. In addition, the leaders also need to see the overview of the market development and can spot opportunities for implementation. Building on that, another mentioned factor is **talent management**. Combining the right data and a clear leadership style is crucial for data scientists or other skilled professionals. They need people who can really work and translate large quantities of data, can clean datasets and can transform these into visualisations and detect the correlation and causation. As well as work with the connection between data and the business

[105] See McAfee & Brynjolfsson (2012), p. 5.
[106] See McAfee & Brynjolfsson (2012), p8-9

language. Since the growing demand, these skilled people are harder to find. Following that, the next factor is **technology**. The tools for handling the 3V's are increasing throughout the years, as well as the integration of Big Data Analytics technologies in the systems. Since the frequently used program Excel is not able to handle Big Data, another technology needs to be implemented. Therefore, it is important that their IT department integrate all the relevant internal and external data in the same system and between the other different departments. Therefore, all the IT people need to have a good understanding of the business processes and its requirements, both for now and in the future. As mentioned before, there is an extreme variety of all the sources, therefore, the technology plays an important role. Next to that, the fourth defining factor is **decision making**. Normally the information and decision rights are in the same place of the company. With Big Data, the information is after receiving transferred to another department for further processing. The expertise is not as it used to be. Therefore, it is important to maximize cross-functional cooperation and a change in the decision-making process. Lastly, the **company culture** plays an important role. First, it is necessary to define what the departments already know. After that, the company culture needs a shifting to be more data-driven instead of looking into data for supporting material to justify already making decisions.

In addition to McAfee and Brynolfsson, Manyika et al. (2011) state that there are another five issues that need to be addressed to capture the full capabilities of Big Data[107]. They start with **data policies**. An increasing amount of digital data travels across the boundaries of the company, the country and sometimes across the big oceans. Since there is more and more personal information involved, including sensitive data, everyone wants to protect their privacy. Hence, in each company or country there are different rules and regulations, so the policy issues from the company become increasingly important. The policies are including the privacy, intellectual property, liability and security. Building upon that, there are many data regulations stated by the local government or European Parliament for protection the policies. Recently, there is a new General Data Protection Regulation (GDPR) established, as stated by the European Parliament[108]. In this regulation, all the companies in the Netherlands are obliged to draw up a document where and how they detect all the privacy sensitive information from each data subject. A data subject is more than just personal data tracking information, it included also Internet Protocol (IP)-addresses, cookies and device

[107] See Manyika et al. (2011) p. 107-109.
[108] See Strik (2016), webpage.

identifiers. In line with that, another mentioned point is **access to data.** Access to the data is necessary to capture the full additional value of Big Data. There is an increasing demand in access to third-party data sources and external sources, for integrating with their internal systems with their own data. So, the barriers for accessing the data must be overcome and integrated into the systems. Further, the **industry structure** plays an important role of capturing value from Big Data. As an example, the public sector with less competitive forces and highly sensitive information, put a higher barrier up against the use of Big Data. In comparison to sectors with production processes, higher productivity and output driven, they have sometimes much lower barriers. Next, **organizational change and talent** is another difficult issue. It is important for an organisation to learn how to leverage and handle Big Data and change the way of working into an integrated system. Since there is a lack of understanding about the final advantages, people are normally resistant to change. Therefore, it is important that the right people help to support the change of implementing Big Data. Moreover, it should connect with the current company culture. Finally, they mention **technology and techniques** as fifth factor which has an overlap with the factors of McAfee and Brynolfsson. Organisations need IT systems with new technologies to integrate with the large variety of datasets to capture value from Big Data.

Furthermore, there are many more barriers to overcome to attract the value of Big Data. **Privacy** is one of the main points since a huge amount of Big Data contains a lot of sensitive personal information[109]. People worry about how that data is used, and especially how it affects them. Techniques such as privacy-preserving health data mining are examples of how to use Big Data, while keeping it anonymous. **Higher cost and less flexibility** are also mentioned as issues[110]. Initially, companies see higher costs for investment and integration as a barrier and refuse implementation. Nevertheless, in the end the implementation leads mostly to cost savings in the company, but that is unseen by the lack of knowledge.

[109] See Maltby (2011), p. 5.
[110] See Kaur & Singh (2017), 408.

3. Purchasing: The strategic decision maker that can make the difference

3.1 Purchasing: The evolution from side function to decision maker at strategic level explained based on the Purchasing Department Cycle

One function which is part of each company, is the purchasing function. It depends on several characteristics of the company, the nature, size and maturity level, how developed, structured and organised the purchasing function is. Purchasing is also a function that developed strongly throughout the years. In the past, purchasing had been more serving other departments with just operative actions as buying goods and obtain maximum price savings[111]. Recently, purchasing has developed from an administrative function to a more and more strategic function[112]. Strategic purchasing is in line with the overall strategy of the company in its value- and market proposition. That is reached, for example, through a pro-active and long-term relationship management with the suppliers. It encompasses all levels in the network of the whole supply chain, from the start with the raw materials to the end with the final customers. Even though purchasing is increasing in importance, but due to the lack of recognition, the tactical realities of implementations are far from optimal. This recognition involves top management decisions and their view on the strategic role of purchasing. The top management must introduce the purchasing as an equal partner in the management team. Knoppen and Sáenz (2015) state that "Purchasing must be acknowledged as a key decision maker that impacts optimization of value creation of the firm while also minimizing costs"[113]. To describe the content of the purchasing function, there are many models published. One of the models is the models of the Purchasing Department Cycle (PDC) of Schiele (in press). See figure 4 below.

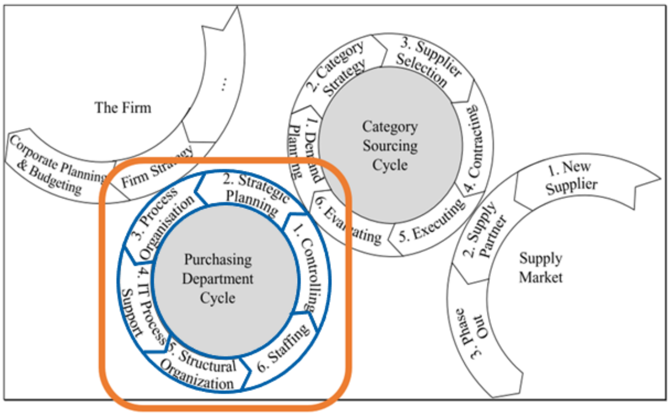


Figure 4 Purchasing Department Cycle
Source: (Schiele, in press)

[111] See Úbeda et al. (2015), p. 177.
[112] See Knoppen & Sáenz (2015), p. 123.
[113] Knoppen & Sáenz (2015), p. 124.

The Purchasing Department Cycle has six different steps, which are continuously following each other. The steps are: 1. Controlling, 2. Strategic Planning, 3. Process Organisation, 4. IT Process Support, 5. Structural Organisation and 6. Staffing. To define the purchasing function further, the steps will be elaborated related to the theory of Schiele[114].

1. Controlling: The controlling part of purchasing has crucial functions in administrating project progress, monitoring and calculating savings, preparing reports for improvements and performance measurements and helps to set the targets and measures the progress. For achieving these targets, the purchasing volume and the size play an important role.

2. Strategic Planning: The strategy of the purchasing department must be in line with the overall company strategy. That is important, since the purchasing department is involved in a major part of the company's turnover. There is not one universal purchasing strategy, but there is a hierarchy of strategies, which are structured among different purchasing categories.

3. Process Organisation: The organisation of the purchasing process includes the design and implementation of the supporting processes. It depends on the type of the industry, the maturity level and the size of the company, whether purchasing is involved with more specific processes. That include that a strategic sourcing process is in line with, for example, a purchase-to-pay process, early supplier inclusion, supply risk management and so on.

4. IT-Process Support: An increasing important IT-tool is e-procurement, what refers to the digital link between supplier and buyer for a quick and easy involvement. Another popular mentioned IT-process is the Electronic Data Interchange (EDI), with a digital exchange between buyer and supplier. These methods are related to the fourth industrial revolution, with machine communication, machine learning and cyber-physical systems.

5. Structural Organisation: The structure of the purchasing department follows the settled strategy, which have to be reviewed periodically. An important point in purchasing is the balance between the level of decentralized and centralized authority. It is in line with the other points that the structure, maturity level and the size of the company have an influence on which authority level is the best suitable.

6. Staffing: When the targets and the structure are defined and the processes are involved in the structure of the company, sufficient personnel must be employed to accomplish the goals.

[114] See Schiele (in press), p. 21-26.

3.2 The sourcing process starts with tactical purchasing followed by operational purchasing, based on the model of Van Weele.

The purchasing department has different functions and are all related to the different steps of the purchasing sourcing process. Schiele mentioned in his model (see figure 4) the Purchasing Department Cycle (PDC), as mentioned before, and the Category Sourcing Cycle (CSC). Whether the PDC focus more on the activities in the purchasing department, the CSC is purchasing executed on a categorical level. That cycle is in line with the often mentioned purchasing model of Van Weele (2010)[115]. The model describes the different steps of the purchasing process and make a distinction between the tactical purchasing and the operational purchasing. In figure 5, see below, is an overview of the model.

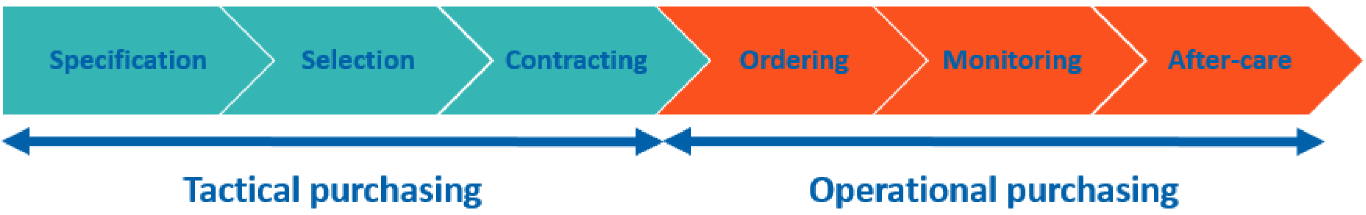


Figure 5 Purchasing process model
Source: Based on van Weele (2010)

Tactical purchasing

1. Specification: At this stage, there is defined what is going to be purchased. Therefore, it is the most determining step of the whole process, regarding to the level of success. It is the basis for everything that will follow in the process. For determining what is necessary, all possible data and information will be collected to make a good decision. This step results in requirements for the goods of services, relationship with the supplier and maintenance at the end. Additionally, this step ends with a list of selection criteria what will be used for evaluation in later stage.

2. Selection: After defining what the purchase order is, the next step is defining who will be the supplier for the purchase. Here, the available information about the suppliers will be evaluated and will lead to a selection of the suppliers. It can be helpful to have comparative research to make sure the best fitted suppliers will be selected. Finally, this will result in selecting the best offer for the specification.

3. Contracting: The following step is entered once the supplier has been selected and the contract is going to be signed. In this phase, there is a decision made in the sourcing strategy, where single, dual, or multiple sourcing will be used. Here you end up with information from

[115] See van Weele (2010), page number undefined.

the contracted parties. The contract contains all information and agreements that are made in relation to the purchase order. That is more than just the goods, it also includes the logistics, payment terms and after sales if they are part of the deal. Once the deal closed, all the involved parties of the company will be involved for further processing. That includes the legal department, accounts payable, users, logistics, technical maintenance etc.

Operational purchasing

4. Ordering: After signing the contract, the critical part of the purchasing process is done, and the operational part will follow. This starts with the ordering process, what is requesting the delivery. That ordering can be a one-off, when it is just for a special occasion. Or it could be a repeating order, and will it be frequently used. It is important to make clear what is ordered, when the delivery will be, to whom and from where it comes from.

5. Monitoring: Monitoring is the part after the specific purchasing order is placed. This includes the processing of the invoices and payments, as well as checking the too late or wrong deliveries. Also, the invoices will be checked to make sure that there is delivered, what is ordered to a price that is agreed upon.

6. After-care: In the last step will the eventual problems solved or when service is necessary during the operational part and needs to be resolved. When things go systematically wrong, it is up to the after-care to talk with the supplier and evaluate things. This step is more related to the supplier management.

The difference between the tactical and operational purchasing are related to the impact the purchasing department has on the total costs of the order. Mostly it is assumed that contracting and negotiating are the most important parts of the purchase. Actually, in this model the specification part is the most influencing on the total costs and continuity of the order, followed by selection and contracting. In the operational purchasing the influence is almost none. The figure on the next page, shows the theory about the influence on the total costs.

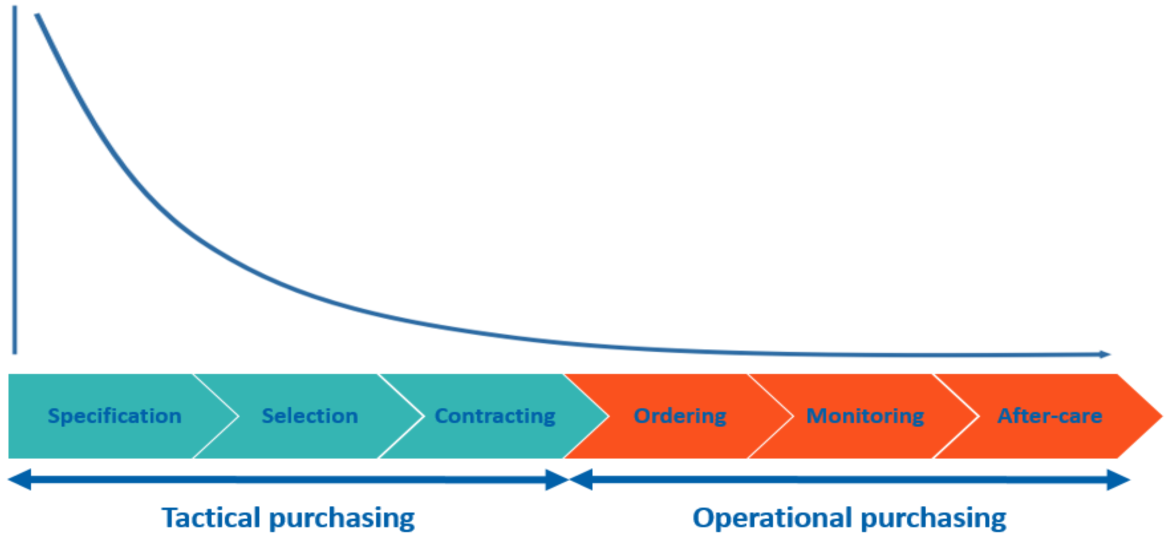


Figure 6 Impact on the total costs of the purchasing order
Source: Based on (van Weele, 2010)

The explanation of this figure builds on several reasons. First of all, and in favour of the cost savings, a difference can be made in the specification phase of defining the specs. In this step, the main cost drivers will be determined. In the negotiation phase that main drivers are already part of the product and therefore, a huge price change during the negotiations is almost impossible. With an influence on the specification, there can be cost savings made with just changing or remove some details. In subsequent stages, the purchasing is already being done, and the biggest change in cost savings can be made in internal processes and the internal operation. In addition, in the specification phase it happens often that organisations find out what they really want or changes their mind. With a great attention to that, in that phase people can change their mind due to the new possibilities and optimize their purchase order, in their opinion. But, for the purchasing department it could look like they are too overenthusiastic for unnecessary additions, which will increase the costs. With early involvement in the specifications, it can be as optimal as possible for the lowest costs.

3.3 Based on the purchasing portfolio of Kraljic, there are different strategies needed for the purchase of leverage, bottleneck, routine and strategic goods/services

The purchasing process is on average in the same order and is a stable and structured way of doing business. Although, recently the stability of purchasing departments has been increasingly jeopardized. Factors as resource scarcity, political instability, intensified competition and technical changes can have an enormous impact. Even the supply and demand can be turned completely within a couple of hours. The purchasing department must respond in the right way to that. It depends on the nature of the products how to react on that changing environment and instability. To define the type of products, Kraljic (1983) made a purchasing portfolio matrix, see figure 7 below, to define the different levels and characteristics of the products[116].

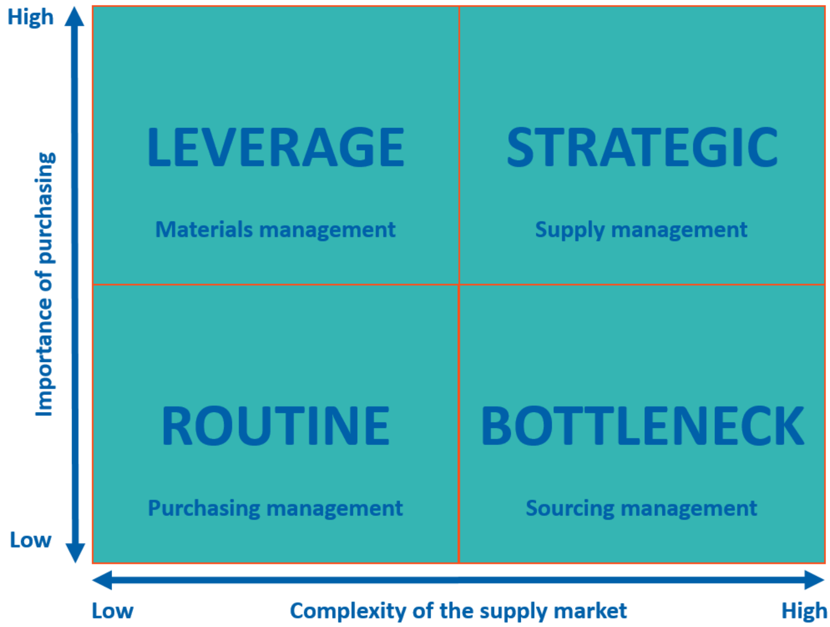


Figure 7 Purchasing portfolio
Source: based on Kraljic (1983)

The matrix is based on two different factors. On the horizontal line, the extent presents from low to high the complexity of the supply market, also supply risk. On the vertical line from low to high, the importance of the purchasing is qualified, also stated as the financial impact. All types of products, goods and services are positioned in one of the four typologies. In line with the four options, another type of management is involved.

For each typology of the purchasing portfolio, there is a diversity in the involvement of data science uses. For *routine* items are standardisation of procedures and use of e-procurement options to include data for optimisation[117]. Routine items have a low risk and low value, but take many times to arrange. With digitisation, there is a reduction in time, effort and administrative costs. For *leverage* items, the influence of data can enhance the continuity and efficiency of measuring and monitoring the supplier performance[118]. When the data detect measurements, which are deviated from the settled standards, they can be immediately adjusted. Therefore, earlier in the process, there is a prevention of possible problems, what

[116] See Kraljic (1983), page number undefined.
[117] See Gelderman & Semeijn (2006), p. 214.
[118] See Gelderman & Semeijn (2006), p. 214.

will lead to added value and better cost performance. *Bottleneck* items are difficult to manage and have low value for the company, while they can be crucial in the process. This situation is not favourable and must be avoided. With data influences in the predictive demand, the order size can increase and with standardisation as a result of data influences, the products will be more related to the routine or leverage characteristics[119]. Additionally, data can contribute in searching for alternative suppliers that can reach the requirements, what might lead to more competition and advantages[120]. Last, in the *strategic* items the supplier is mostly the dominant part of the relationship and has a strong position. For adding data in the process of strategic items, an integrated purchasing system and strategic partnerships can result in added value and partnerships. With an integrated system, it is for both supplier and buyer easier to communicate and to negotiate, what will result in better prices and can gain competitive advantages. Besides the competitive advantages, due to the nature of strategic products, out of the matrix of Kraljic it is the less impacted influenced product by the use of data analytics. The more standardised the processes and products, the highest impact is prospected[121].

3.4 The history of system integration in purchasing, with as most important part the enterprise resource planning (ERP)

In the history of purchasing, there are several digitalisation improvements, which are in the preliminary stages from the development of data applications and the final Big Data integrations. It started in the1960s, with the introduction of the computerized reorder point systems (ROP), to satisfy the needs for planning and control[122]. In the late 1960s, the material requirements planning (MRP) was born, although the initial solutions were expensive, big and abstemious. The systems required support from a large technical staff. The storage capacity was a major technology for that moment. Also, there was not a database and the software tools were limited by the standards of that moment. In 1970s, there was a shift towards marketing related to MRP, what results in an integration with forecasting, procurement, master scheduling and control on the shop floor. The master production schedule has a leading role for the system and every part is connected[123]. In the mid 1970's, major software companies were born, what had a great influence on the systems. As an

[119] See Bäckstrand, Tiedemann, & Hedén (2015), p. 5.
[120] See Gangurde & Chavan (2016), p. 1753.
[121] See Moretto, Ronchi, & Patrucco (2017), p. 12.
[122] See Robert Jacobs & 'Ted' Weston (2007), p. 358.
[123] See Miller & Sprague (1975), page number undefined.

example, in 1979 Oracle launched the first structured query language (SQL) relational database system. In the 1980s, a lower cost alternative was developed for small and medium sized business and more related to manufacturing. The term manufacturing resource planning (MPR II) was born[124]. Finally, in the 1990s, the enterprise resource planning (ERP) was developed. That system was the basis for the current ERP systems that are still used nowadays.

Robert Jacobs and 'Ted' Weston (2007) define ERP as a "framework for organizing, defining, and standardizing the business processes necessary to effectively plan and control an organization so the organization can use its internal knowledge to seek external advantage"[125]. In addition to that, Ragowsky and Somers (2002) define the applications of ERP systems as "These systems help organizations deal with their supply chain: receiving, inventory management, customer order management, production planning and managing, shipping, accounting, human resource management, and all other activities that take place in a modem business"[126]. When the ERP systems are properly selected and implemented well, the purchasing departments can have significant benefits. There could be a reduction of inventory costs, raw materials, lead-times, production time and production costs. The benefits are a result of the software tool and information technology, which is used to manage all the relevant enterprise data and provide knowledge for the department that will use it. It is the precursor of the predictive analyses with Big Data implementations.

Another mentioned method for system integration in the purchasing, is the electronic-procurement system. Kaliannan, Chandran and Hashim (2010) state that e-procurement refers to "The use of electronic methods in every stage of the purchasing process from identification of requirements through payment, and potentially to contract management"[127]. E-procurement uses the internet technology in the purchasing process and can be considered as a part of the process and not as a replacement of it[128]. The benefits of e-procurement are in cost reductions due to less paperwork, more efficient purchasing process and fewer mistakes. The system can be used for finding suppliers, managing tenders, submitting bids and prevents maverick buying. One important part of e-procurement is the e-sourcing. E-

---

124 See Umble, Haft, & Umble (2003), p. 242.
125 Robert Jacobs & 'Ted' Weston (2007), p. 357.
126 Ragowsky & Somers (2002), p. 12.
127 Kaliannan, Chandran, & Hashim (2010), p. 1334.
128 See de Boer, Harink, & Heijboer (2002), p. 25-26.

sourcing refers to the process of using internet technology for identifying new suppliers or purchasing requirements. With the new suppliers, the competitiveness could increase for the purchasing department, and reduce the associated supply risk. With more suitable, new suppliers, the dependency is lower and gives a better position in the competitive market.

Finally, the spend analysis is a commonly used method to evaluate all the purchases of the enterprise. It includes all the buying details as value, supplier, number of contracts, order details and purchasing organizations. In addition to that, it also includes a supplier base of the suppliers. That is an overview of the suppliers by industry, geography, firm, dependency, risk, socioeconomic details, influence on the supplier and percentage of the business that it gets from a single supplier[129]. The analysis is a labour-intensive and time-consuming activity with possibilities of mistakes, and is difficult to identify. Notwithstanding, it is helpful for developing supply strategies, selecting the best suppliers, manage supplier relationships, maximize rewards and minimizes risks. Due to the combination of analytical and benchmarking techniques, it can reveal new targets or opportunities that are unknown or difficult to detect. With the spend analysis, it is possible to analyse the purchase by spend, by value by business volume or by strategic importance for the company.

[129] See N. T. Moore, Cook, Grammich, & Lindenblatt (2004), p. 7-8

4. Maturity levels as measurement method for continuous improvement in the organization, specialized in the disciplines Purchasing and Big Data

4.1 Maturity model as analytical model for describing the current situation based on different maturity levels within the organization.

The development of the maturity models goes back to the introduction of the Capability Maturity Model (CMM) from Paulk, Curtis, Chrissis & Weber in 1993[130]. After that moment, several researchers developed different maturity models, with the CMM model as basis design. In the literature, there is not any universally used definition for Maturity models. Drus and Hassan (2017) concluded that "Maturity is an incremental process from some infancy stage to a more sophisticated stage, whereby an entity has to undergo some transitional stages"[131]. Next to that, Pöppelbuß and Röglinger (2011) add "Maturity models basically represent theories about how organizational capabilities evolve in a stage-by-stage manner along an anticipated, desired, or logical maturation path"[132]. In both definitions, maturity is related to different stages. According to several researchers, there is a difference between immature and mature organizations. Immature organizations have processes and tasks that are improvised by managers and practitioners, even if they have been specified. They are not correctly followed or enforced. In opposite, mature organizations have the ability to manage developments and maintenance. Managers communicate well with staff organization-wide and the work is carried out according to the settled plan.

The Capability Maturity Model (CMM) of Paulk et al. (1993) is based on continuous improvement and aid organizations to prioritize the effort for the improvements[133]. The model settled five different levels of maturity, whereby each level is the basis for the next level. Each level established a different part in the process, resulting in an increase in the process capability of an organization. The five levels are described below:

1. **Level 1, initial:** In the first level, the company does not provide any stable environment for developing and maintaining software and has difficulties in making commitments. The success of the company depends on the competence and capability of their individuals, not to the organization.

[130] See Paulk, Curtis, Chrissis, & Weber (1993), p. 18-27.
[131] Drus & Hassan (2017), p. 614
[132] Pöppelbuß & Röglinger (2011), p. 4.
[133] See Paulk et al. (1993), p. 21-23.

2. **Level 2, repeatable:** At the second level, policies for managing a project and implementing procedures are established. The policies and planning are based on experience and have basic management control. Organizations are disciplined, because they are stable and earlier success can be repeated.
3. **Level 3, defined:** In the third level, the process for developing and maintaining the process is documented and integrated across the organization, including engineering and management processes. Organizations have a standard and consistent process, because it is stable, repeatable and based on a common understanding of the process.
4. **Level 4, managed:** The fourth level sets quantitative quality goals for the products and services. The processes are well-defined and have consistent measurements. An organization-wide database will be used for collecting and analysing data from the projects. The processes can be defined as quantifiable and predictable.
5. **Level 5, optimizing**: In the fifth and last stage, the company has a continued improvement. It identifies weaknesses and strengths of the processes and is able to prevent some defects. There is also a major focus on reducing waste. The processes are continuous striving to improve the capability and process performances.

Trying to skip a level in the maturity process is more counterproductive than optimized. Each following maturity level builds on the foundation of the current maturity level.

For designing a maturity model, Pöppelbuß and Röglinger (2011) defined a set of principles that the maturity model should meet. The foundation of the maturity model is to describe the different stages and the maturity path. The purposes of the maturity models are different in their application-specific purpose and will be first distinguished[134].

- **Descriptive:** Used to describe the as-is situation and for assessing the capabilities according to certain criteria. In this purpose, the model is used as a diagnostic tool.
- **Prescriptive:** Used to indicate how to establish desirable maturity levels. In this purpose, the model is used as describing guidelines for improvements with detailed courses of action.
- **Comparative:** Used for comparison between maturity levels of similar business units and organizations, based on sufficient historical data. In this purpose, the model is used as a comparative tool.

[134] See Pöppelbuß & Röglinger (2011), p. 3-4.

The purpose of the use of the maturity model in this research can be described as ‘prescriptive’. The model will be used as guidelines for improvements and detailed actions with regard to the influences of Big Data on the purchasing function. For further designing of a maturity model, Pöppelbuß and Röglinger (2011) defined some basic principles that could be applied in the design phase. These basic principles are in a sort of checklist that can be used to make sure that all the necessary information is applied. This checklist will also be used in this study for designing the desired model. Nevertheless, they mentioned that “We do not require each maturity model to meet all design principle. Instead, the framework serves as a checklist when designing new maturity models”[135]. With an understanding of this statement, the design principles will be followed broadly. The results of the checklist will be further explained and processed in chapter 7. The checklist is in table 5 below.

	Design principles 1. Basic	
1.1	**Basic information**	
	a) Application domain and prerequisites for applicability	
	b) Purpose of use	
	c) Target group	
	d) Class of entities under investigation	
	e) Differentiation from related maturity models	
	f) Design process and extent of empirical validation	
1.2	**Definition of central constructs related to maturity and maturation**	
	a) Maturity and dimensions of maturity	
	b) Maturity levels and maturation paths	
	c) Available levels of granularity of maturation	
	d) Underpinning theoretical foundations with respect to evolution and change	
1.3	**Definition of central constructs related to the application domain**	
1.4	**Target group-oriented documentation**	
	Design principles 2. Descriptive	
2.1	**Intersubjective verifiable criteria for each maturity level and level of granularity**	
2.2	**Target group-oriented assessment methodology**	
	a) Procedure model	
	b) Advice on the assessment of criteria	
	c) Advice on the adaptation and configuration of criteria	
	d) Expert knowledge from previous application	

Table 5 Checklist for designing a maturity model

4.2 In the literature the first Big Data Maturity Models is from 2013, through the years they are upcoming and are enlarged to the currently ten known models

Big Data, as stated in chapter 2.2, has the potential to improve or transform the organization and can reshape whole departments. The technological aspect is increasing in importance,

[135] See Pöppelbuß & Röglinger (2011), p. 5-8.

but organizations must implement it very carefully. Wrong or rushed implementation can result in serious problems, lack of trust, whereby the data is really sensitive for the organization. Therefore, these organizations need to understand how they operate and where they are in terms of Big Data and the defined thresholds. Farah (2017) defined a maturity model for Big Data as "a way of classification to describe the status of an organization with respect to its effort as it relates to data collection, maintenance, and analysis"[136]. They need to know where they are now, as-is state, and where do we want to go to, to-be state[137]. As said before, also for Big Data maturity, it is important to follow the different levels to become successful. All the improvement models and innovation approaches have to follow the steps of development in sequence[138]. The ultimate goal of Big Data integrations in the company, is to transform the company to be completely data-driven and autonomous[139]. Different aspects Big Data maturity models are related to the environment readiness, the organization's internal capabilities and the different stages of maturity in which Big Data applications can be used. In the literature, the first Big Data related maturity model is designed by Halper and Krishnan in 2013. Recently, there is a total of ten maturity models designed. An overview of these models is in table 6.

No	Model	Source	Year	Levels	Purpose
1	4-Stage Big Data Maturity Model	Tiefenbacher and Olbrich	2015	4	Prescriptive
2	TDWI Big Data Maturity Model	Halper and Krishnan	2013	5	Comparative
3	EMC's Big Data Business Model Maturity Index	Schmarzo	2016	5	Prescriptive
4	IBM's Big Data and Analytics Maturity Model	Nott & Betteridge	2014	5	Descriptive
5	Hortonworks Big Data Maturity Model	Dhanuka	2016	4	Comparative
6	IDC's Big Data and Analytics Maturity Model	Moore	2014	5	Comparative
7	Big Data Maturity Framework	El-Darwiche et al.	2014	4	Prescriptive
8	Big Data Maturity Model for Zakat Institution	Sulaiman, Che Cob and Ali	2015	5	Prescriptive
9	Big Data Maturity Model	Radcliffe	2014	6	Prescriptive
10	Value Based Big Data Maturity Model	Farah	2017	5	

Table 6 Overview Big Data Maturity models

Sources in order of the overview: Tiefenbacher & Olbrich (2015), p 7; Halper & Krishnan (2014), p. 5-13; Schmarzo (2016), webpage; Nott & Betteridge (2014), webpage; Dhanuka (2016), p. 10-15; D. T. Moore (2014), p. 4; El-Darwiche et al. (2014), p. 16, exhibit 3; Sulaiman, Che Cob, & Ali (2015), p 63-6; Radcliffe (2014), p. 2-5; Farah (2017), p. 12-14.

As shown in table 6, the most Big Data Maturity Models (BDMM) have four or five levels in the maturity curve, and just one model has six levels. To give a deeper insight in the different levels of each maturity model, an overview is in table 7 on the next page.

[136] Farah (2017), p. 12.
[137] See Radcliffe (2014), p. 1.
[138] See Günther et al. (2017), p. 198.
[139] See El-Darwiche et al. (2014), p. 3.

No.	Level 0	Level 1	Level 2	Level 3	Level 4	Level 5
1		Analytical Projects	Functional excellence	Value Proposition	Business Model Transformation	
2		Nascent	Pre-adoption	Early adoption	Corporate adoption	Mature/ Visionary
3		Monitoring	Insights	**Optimization**	Monetization	Metamorphosis
4		Ad-hoc	Foundational	Competitive	Differentiating	Breakaway
5		Aware	Exploring	**Optimizing**	Transforming	
6		Ad-hoc	Opportunistic	Repeatable	Managed	**Optimized**
7		Performance management	Functional Area Excellence	Value proposition Enhancement	Business Model Transformation	
8		Ignorance	Coping	Understanding	Managing	Innovating
9	In the dark	Catching up	First pilot(s)	Tactical value	Strategic Leverage	**Optimize &** Extend
10		Initial	Defined	Managed	**Optimized**	Strategic

Table 7 Overview different maturity levels of the Big Data Maturity models

The model of Radcliffe (2014)[140] added as sixth layer, the "In the dark" phase, the place where companies are before they do anything with data and do not know anything about it. When comparing the models, it shows that models with five layers have mostly the first layer as the awareness and using baby steps in the use of data. In models with four layers, the first step is mainly the data already implemented and ready for use. There is not a level defined for the beginning stage, except maturity model 5, where the first level is just 'aware'. The status 'optimized' occurs in many models, but on different levels. It is on level 3 out of 5 (model 3), 3 out of 4 (model 5), 5 out of 5 (models 6 &9) and 4 out of 5 (model 10). That shows that the optimization of Big Data analytics is not in all models the final goal, but sometimes it is just halfway the model. In model 8 the first level 'ignorance' correspondents to the 'in the dark' level 0 of model 9. Actually, in the third level there is just the first understanding of the data influences. In comparison to the other models, the first levels are more extensive and thus broader explained. In total, it shows that models with five or six levels, have in general less distinctiveness between the different levels. Therefore, it is harder to define a specific stage since all levels are close related to each other. In models with four levels, there is more distinctiveness in the different steps. Further, when the middle is split into two distinctive steps, the boundaries are more clearly defined.

The different Maturity models have diverse purposes and a dissimilar amount of levels, they are also analysing other domains. Based on results of these domains, a maturity level will be defined. In table 8 on the next page, there is an overview of these different domains, when they are mentioned by the researchers. The table shows that one half do not have domains

[140] See Radcliffe (2014), p. 2-3.

and the other half do have research domains, which can be used as a framework. These are the basis of measure consequently the same domains for a roadmap[141]. In model 2, the five domains are used as the basis of a benchmark survey, while model 5 uses the domains as capability factors[142] and in model 9, the Big Data Maturity model is based on the 'Eight Building Blocks of Big Data'[143]. Overall, the domains look from different views to the maturity of Big Data.

No.	1	2	3	4	5	6	7	8	9	10
Domains	X	• Organization • Infrastructure • Data Management • Analytics • Governance	X	• Business Strategy • Information • Analytics • Culture & Execution • Architecture • Governance	• Data Sponsorship • Data & Analytics Practices • Technology & Infrastructure • Organization & Skills • Process Management	• Intent • Data • Technology • People • Process	X	X	• Vision • Strategy • Value & Metrics • Governance, Trust & Privacy • People & Organization • Data Sources • Data Management • Analytics & Visualization	X

Table 8 Overview of different domains used in Big Data Maturity models

Based on the different set-ups and comparing the most important features, this thesis will use a four-step approach in the new designed Big Data Purchasing Maturity model. The model is structured with different domains to cover the broad view according Big Data. Additionally, the structures of the current models will be used to define the new levels into a constructive distribution of the different levels.

4.3 The first Purchasing Maturity Model is from 1988 and throughout the decades the models are still expanding and improving

Next to the Big Data Maturity models, there are several models specified for purchasing. According to definition of Úbeda et al (2015), "Purchasing maturity models characterize the evolutionary process of purchasing"[144]. Additionally, Schiele (2007) mentioned that maturity in the purchasing function is related to the level of professionalism of that function[145]. As the level of purchasing maturity will improve to the next level, the organization experienced increased performance. Therefore, there is assumed that a greater maturity is associated with a greater performance. Companies can save around 5 – 10 % of the average spend by moving to the next maturity level[146]. In the literature, there are twelve

[141] See Radcliffe (2014), p. 1-2
[142] See Dhanuka (2016), p. 10-15
[143] See Radcliffe (2013), p. 2-4; Radcliffe (2014), p. 1.
[144] See Úbeda et al. (2015), p. 178.
[145] See Schiele (2007), p. 274.
[146] See Keough (1993), p. 1.

Purchasing Maturity models (PMM) described. These models can be more specified as conceptual, since the models are not empirically tested[147]. An overview of these models is shown in table 9 and will be further discussed. Additionally, this overview contains different information than the overview in table 6, since not all the detailed information about the different models was available for this research. Therefore, the summaries of Úbeda and Alsua (2015) and Schiele (2007) are used as basis of this overview[148].

No	Source	Year	Levels	No. of items for assessment
1	Reck & Long	1988	4	11
2	Cammish and Keough	1991	4	8
3	Keough	1993	5	8
4	Burt and Doyle	1994	4	33
5	Barry, Cavinato, Green & Young	1996	3	20
6	Bhote	1989	4	24
7	Freeman and Cavinato	1990	4	9
8	Chadwick & Rajagopal	1995	4	9
9	Cousins, Lawson & Squire	2006	4	24
10	Paulraj, Chen & Flynn	2006	3	42
11	Schiele	2007	4	111

Table 9 Overview Purchasing Maturity models

Based on the article of Úbeda, Alsua & Carrasco (2015) the Purchasing Maturity models could be divided into four different categories[149]. The different categories are distinguished in table 9 with bold lines. Maturity models 1 to 5 belong to the first category, where the models are based on primary observations and interviews with experts. These suggestions and opinions are summarised into a maturity model with stages. The models 6 to 8 are mentioned as the second category. These models are based on one dominant theory with one differentiating criteria. The stages are determined based on ex-ante models. Unless the models should be more consistent, they run the risk of being biased and do not involve other relevant dominants. Further, models 8 & 9 belong to the third category, what could be specified as models where the first stages are deduced from primary observation. Unless the observations, there are not any empirical tests to check the link between purchasing maturity and firm performance of the models. Finally, the model 11 is related to the last category where the model is based on theory and included ex-ante identification.

The table 9 above shows also the different levels the Purchasing Maturity models (PMM) have defined. The range is between 3 and 5, where the most maturity models have 4 levels

[147] See Schiele (2007), p. 274-275.
[148] See Schiele (2007), p. 275-277; Úbeda et al. (2015), p. 178.
[149] See Úbeda et al. (2015), p. 178.

or stages distinguished. Just as the Big Data Maturity models, the preference is also for four distinctive stages, since the same advantages are suitable. The last column, with the number of items for assessment, is the range after which the maturity level is defined. For example, the model of Schiele (2007) has a questionnaire with 111 different questions, what will lead to a maturity level. Contrarily, the model of Cammish and Keough (1991) has eight questions as suitable foundation. The other models are in the range between 8 and 42, and on average that is 27,18. Further, the different PMM's have also diverse topics included in their measurement criteria. An overview of these addressed topics is in table 10 below.

No	Planning	Structural organisation	Process organisation	Human resources	Controlling	Collaborative supply relation
1	X			X	X	X
2	X	X	X	X		
3	X	X	X	X		
4	X	X		X	X	X
5	X		X		X	
6	X	X	X		X	X
7	X			X	X	
8	X	X	X	X	X	X
9				X	X	X
10					X	X
11	X	X	X	X	X	

Table 10 Overview addressed topics in the different Purchasing Maturity models

The chosen topics are based on several works and stated as conclusion in the article of Schiele (2007)[150]. The result of the research is a five-dimensional profile that covers the purchasing maturity. The sixth dimension, Collaborative supply relation, is added because it cannot be deduced from a management point of view and has a prescriptive character. The other five topics are from a management-oriented profile. As seen in the table, the most models do not shelter the five dimensions and are incomplete. Only one model, model 8 from Chadwick & Rajagopal (1995) include all the different topics. For further analysing, the most complete and relative models will be used.

4.4 Industry 4.0 integrated with purchasing maturity model is developed to describe the level of maturity for the applications of Industry 4.0 on the purchasing function

Based on the nature of the maturity models, they are suitable for a diversity of industries, companies and departments, including Industry 4.0. Up to now, there are several published Industry 4.0 maturity models in organisations, as arranged by Schumacher, Erol and Sihn

[150] See Schiele (2007), p. 276-278

(2016)[151]. For Industry 4.0 are maturity models and readiness models. In a readiness model, the readiness assessment takes place before engaging in the maturity process and a maturity is more related to the as-is situation. In addition to the overview Schumacher, Erol and Sihn (2016), Leyh, Schäffer, Bley and Forstenhäusler (2016)[152] and Kleeman and Glas (2017)[153] designed also two maturity models for Industry 4.0. Based on all these different models, Torn (2017) designed a maturity model for Industry 4.0 and Purchasing, called Purchasing 4.0. The purchasing 4.0 Maturity is building on Hazelaar's Master Thesis[154] and the PwC Industry 4.0 Maturity model[155]. In The work of Hazelaar (2016), there is a roadmap designed that lets the company enter and involve with the Industry 4.0 paradigm. The roadmap has four different steps from standardization through integration and automation into machine-to-machine communication. The PwC Industry 4.0 Maturity model (2017) has four stages, called 1) Digital novice, 2) Vertical integrator, 3) Horizontal collaborator and 4) Digital champion. Based upon that, Torn (2017) proposed four definitions of the stages used in his maturity model for Purchasing with industry 4.0 (I4.0). They are cited and described as follows[156]:

Digital Novice: "Industry 4.0 concepts are in a pre-mature stage within separate departments, pilot-projects and novice use of data analytics. Industry 4.0 processes are not standardised within the firm yet."

Vertical integrator: "Cross-functional collaboration and alignment of digital strategy between departments, advanced use of data analytics. Standardised I4.0 processes are present in the organisation."

Horizontal integrator: "Integration of processes outside the boundaries of the firm, competent use of data analytics, machine-to-machine communication, and/or cyber-physical systems. Standardised I4.0 processes are present, and people in the organisation are responsible to ensure compliance."

[151] See Schumacher, Erol, & Sihn (2016), p. 162-163.
[152] See Leyh, Schäffer, Bley, & Forstenhäusler (2016), p. 1299-1300.
[153] See Kleemann & Glas (2017), p. 35-41.
[154] See Hazelaar (2016), p. 58.
[155] See Griessbauer, Vedso, & Schrauf (2016), p. 28.
[156] Torn (2017), p. 66.

Digital champion: "Industry 4.0 concepts fully aligned with corporate strategy across integrated supply chains, proficient in applying autonomous decision making, machine-to-machine communication, and cyber-physical systems."

This model involves eight dimensions: 1) Strategy, 2) Process & Systems, 3) Physical level, 4) Purchase to Pay (P2P), 5) Controlling & Key Performance Indicators (KPI), 6) Sourcing, 7) Suppliers, and 8) Employees & Uses. The model with the eight dimensions is in figure 8 on the next page.

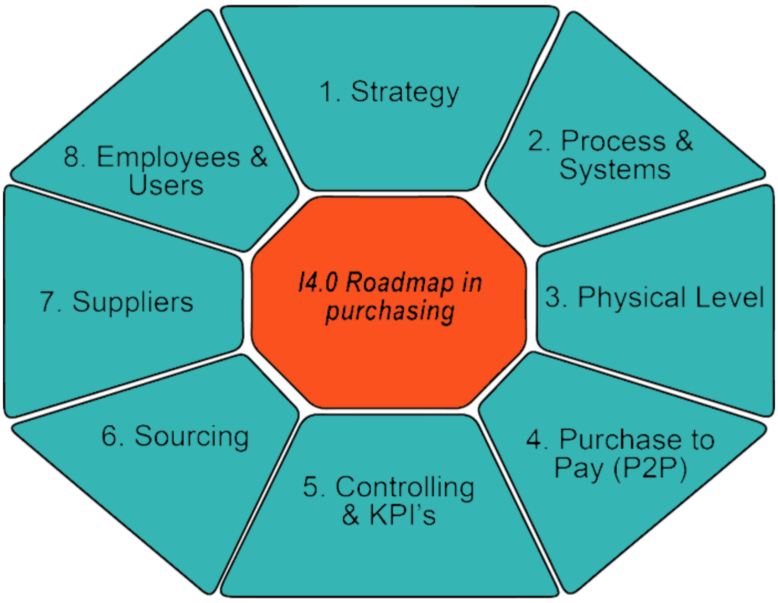


Figure 8 Purchasing 4.0 Maturity model

The eight dimensions together cover the roadmap for Industry 4.0 in purchasing. The content of the different dimensions is explained by the following overview:

1. Strategy: The strategy included the digitisation strategy with the requirements and priorities to focus on. Also, the strategy for Industry 4.0 and for purchasing is determined.

2. Processes & Systems: The way how processes are standardised, digitised or automated. The development level of the supporting systems, e-procurement software and technologies are also involved.

3. Physical level: The establishment of the connection between physical systems, with an integration of services and automatic ordering process between buyers and suppliers.

4. Purchase to Pay (P2P): Included the development of automation in the payment processes and the way of delimited the maverick buying process. The step 'monitoring' of the Purchasing Process model of Van Weele is part of this dimension.

5. Controlling & KPI's: Controlling relates to the completeness, real-time transparency and analytical capabilities of the firm. The predictive analyses, decision support systems, self-learning algorithms and data security are additional tools for controlling KPI's.

6. Sourcing: Parts of the sourcing dimension are the predictive demand and market analysis. The steps 'specification', 'selecting' and 'contracting' of Van Weele are part of sourcing and are evaluated by the involvement of data analytics and e-procurement systems.

7. Suppliers: To which extend the suppliers are prepared and willing to collaborate in the processes. Also, the early involved in the process, supplier capacity and integrating of data exchange are important in the supplier evaluation.

8. Employees & Users: The willingness and capacities of the employees which are involved with Industry 4.0 and purchasing. This included the way how the learning and development are stimulated by the management and the involvement of the employees in the digitisation.

4.5 Bright Cape Maturity model is based on Gartner's Enterprise Information Management maturity model and has five different levels.

As the last specified maturity topic about the different maturity models, Bright Cape developed their own model for measuring the degree of intelligence in data maturity. The model is shown in figure 1, chapter 1 and page 6. This model is based on Gartner's Enterprise Information Management (EIM) Maturity Model. The model of Bright Cape has the same characteristics as the common maturity models, where the stages following each other and are necessary for growing. The model has five different stages, from 1) Data quality 2) Descriptive, 3) Diagnostic, 4) Predictive and 5) Prescriptive. The These five stages could be defined and explained as follows:

1. Data quality *("Do I have data and can I trust it?")*: It is the first or initial stage what is just the beginning and is pretty straight forward. Here it is more basal information about the data, is there data, what data is missing, and what kind of data is available, but there is nothing happening with the data

2. Descriptive *("What happened?")*: In the second stage, the data is more analysed, but it is not able to use. There are more insights in data quantities, for example, with KPI dashboards, spend overviews, overview of orderliness, involved companies, et cetera. Hereby is just the information about what happened when and how.

3. Diagnostic *(“Why did it happen?”)*: The diagnostic stage is more relative to explain the data and what happened, whereby the why question is important. With an increase in parts of the data, sometimes it is easy to explain, but frequently there are some correlations that needs to discover for further processing. Increases can be seen as price fluctuations or change in suppliers or a mutation in the demand.

4. Predictive *(“What will happen?”)*: In the next step the predictions are made about what will happen. Here the company is able to detect trends in the data and extract predictive knowledge about the trends. They make prospects for further processing, based on the history.

5. Prescriptive *(“What should I do to optimize?”)*: The fifth and final stage is related to optimization and autonomous decision making. Based on the data, the company is able to see how it will work in the future, the systems are able to implement trends in autonomous actions. The machine learning tools are an important part of the prescriptive step, since they are able to make autonomous decisions for the future.

5. Methodology review about the different qualitative research methods

5.1 The qualitative literature review gives an in-depth overview of the current already known literature

For each research, the consideration whether a quantitative or qualitative research method is more suitable for the aim of the research, is taken into account. Therefore, both research methods will be briefly discussed. To start with a quantitative research, it is defined as a deductive approach. That implies that there are one or more statements to reach into a conclusion. There is a relationship between different variables, mostly between the theory and research[157]. The researcher deduces one or more hypotheses and test these through fixed variables[158]. Researchers have the possibility to transform data into numbers through mathematical models[159]. In addition to that, a quantitative method is more suitable for large datasets or samples. Within this approach the data collection methods, for example in the literature review and the interview method have a structured character. The data is 'hard', reliable and more static. On the other hand, a qualitative research has an inductive approach, what implies that the findings of a research are fed back to the theory. The emphasis is on the creation of theories[160] and is regarded as a subjective method. In this type of research, the researchers develop questions and guidelines in order to gather insight information, understand the meaning of individuals and is not generalizable[161]. In contrast to quantitative research, the qualitative research is more suitable for small datasets. Furthermore, the data collection method is mostly based on interviews, where the structure is a semi-structured or unstructured design. The data can be more classified as deep and richer.

Both quantitative and qualitative research methods, the most important criteria for evaluation are the reliability and validity. Reliability is the aim towards whether the results of the research are the same when the research will be executed again. That can be both with the same researchers at another place or time, or with other respondents, or that can be with different researchers. For quantitative studies, it is easier to increase the reliability than with the qualitative studies, since results may change to the perception and knowledge of the

157 See Golafshani (2003), p. 597.
158 See Bryman & Bell (2011), p. 38.
159 See Tracy (2013), p. 24.
160 See Bryman & Bell (2011), p. 38.
161 See Sofaer (2002), p. 334.

different researchers[162]. The reliability has two different parts, namely the internal reliability and external reliability. Internal reliability is the extent to which a measure is consistent within itself. That means that there is a consistency or correlation of the results across items in the test. External reliability is the extent to which a measure varies from one use to another. That means that independent researchers can reproduce the research and obtain similar results to the one obtained in the original study. Next to that, the validity is also an important part of the research. Validity is the extent to which a research measures what it is intended to measure. It applies both in the design as well as to the methods. Research validity has also two different parts, internal validity and external validity. Internal validity refers to how the findings match with the reality, while external validity refers to which findings can be replicated to other environments.

For both qualitative and quantitative research studies, the literature is always part of the research. The literature review summarizes and criticize prior results about the chosen research topic. It shows what is already known and discovered by other researchers and what part is still unknown[163]. The goal of a literature review is forming a basis of the current information for further research[164]. They have different similarities as answering a research question, relating the data to the literature and are concerned with variations. Next to that, there are also any differences between both literature reviews, which are in line with the differences between qualitative and quantitative research methods[165].

For this research is chosen for a qualitative research method, with a comprehensive literature review. That is because of the research introduces new insights to the topic and therefore needed in-depth and rich information to build further upon. The review contains two main topics, namely the first topic 'Big Data' and the second topic 'Purchasing'. The searching words for relevant articles about the first topic contains 'Big Data', 'Data-driven', 'Data-based', 'Process mining' and 'Industry 4.0'. The searching words for the second topic contains 'Purchasing', 'Procurement', 'Buying' and the Dutch word 'Inkoop'. In addition to those topics there will be searched for 'Maturity' and 'Maturity model'. Furthermore, there is also searched for articles where all the different topics are combined in one article together. As literature databases the databases from 'Scopus', 'Web of Science' and 'Google Scholar'

[162] See Bryman & Bell (2011), p. 49.
[163] See Denney & Tewksbury (2015), 219-220.
[164] See Cronin, Ryan, & Coughlan (2008), p. 38.
[165] See Bryman & Bell (2011), p. 413.

will be used. For structuring the first outcomes of the literature review, a literature matrix is used. That matrix is a clear overview of the specified content of the articles and an overview of the used models. That leads to a structured and clear literature review. Later on, this matrix is extended with catchwords and relating articles to the subject. That literature matrix is used as a basis for the structure of the research.

5.2 Semi-structured interview method for gathering in-depth information about Big Data and Purchasing

The second part of the research builds on the literature review. Since there are less articles available about Big Data, Purchasing and Maturity together, the second part gathers information in a qualitative way. The main research methods for qualitative research are sampling, ethnography and participant observation, interviews and focus groups[166]. The sampling method is studying a subset of a larger population consistent with the same characteristics and has two variations. One is probability sampling, where participants are chosen randomly, and the other is purposive sampling, what a more strategic form is of sampling of cases and participants. Two examples of purposive sampling are theoretical sampling and snowball sampling. With ethnography and participant observation the researcher is involved in the social life or natural habitat of those who are subject of the research and collect information. Interviewing is a method to gather deeper information of the participant and is possible in both quantitative and qualitative research[167]. A quantitative interview can also be seen as a structured interview, and a qualitative interview is further divided into two main different types, the semi-structured interview and the unstructured interview. Lastly, there is a focus group what is a form of an interview with several participants together, what can be mentioned as a group. A focus group can also be divided into specific forms, namely the group interview, where selected people discuss the list of topics, and the focused interview, where people are selected to discuss a particular situation on behalf of their own involvements. For this research, the chosen method is the interview method, since there are less articles available about Big Data, Purchasing and Maturity together. With the interview method, it is possible to structure the questions and get deeper information. The following definition describes this method; “Interviewing distinguishes itself from other research approaches by engaging participants directly in a conversation with the researcher in order to generate deeply contextual, nuanced and authentic accounts of

[166] See Bryman & Bell (2011), p. xv – xvi.
[167] See Bryman & Bell (2011), p. 441-444.

participants' outer and inner worlds, that is, their experiences and how they interpret them."[168]

As mentioned above, there are two main different types to distinguish, namely the quantitative interview and the qualitative interview. In a quantitative way, the interview has a structured composition. That has the result that each respondent has to answer the same questions as the other respondents and that all questions are exactly in the same order. The goal of this type of interview is to make sure that the answers can be aggregated and compared. The advantage of a structured interview is an outcome that is maximized in validity and reliability and reflects the researchers' concerns. The disadvantage of structured interviews is due to the standardization it is not possible to react to the answers or go deeper into the information. It is not possible to deviate from the interview questions. Therefore, some essential information can be missed.

On the other hand, as mentioned before is a qualitative interview method and it has two main different types, the semi-structured interview and the unstructured interview. In a semi-structured interview the interviewer made an interview guide with a list of questions or topics. It depends on the situation, whether specific questions from the guide or additional questions will follow. Therefore, it is possible to ask in-depth questions. An unstructured interview starts mostly with just one question where both interviewer and respondent react on each other. Since there is no structure at all, it is more the type of a conversation. The advantage of a qualitative interview is that there is a flexibility to ask more questions that relates to the situation to gain more and insight information, and build upon the answers. Especially when the subject is not very extensively in the literature or at the researcher, gathering a large quantity of useful information. As disadvantage can be mentioned that the answers are harder to compare since not every respondent had answered the same questions.

For this research is conducting a qualitative semi-structured interview the best suitable option. Due to the flexibility and the broader insights, the interview is supposed to generate extensive and detailed answers. That gives the opportunity to gain extended and new information about Big Data, Purchasing, Maturity and other related subjects. The interview results will be integrated during the design phase for the new model. Based on the answers from the respondents, the different steps of the model will be influenced. Their opinions and

[168] Schultze & Avital (2011), p. 1.

ideas play an important role in determining the scorecard for each different step and the sub questions. Also, the respondents give insights in the best-case scenarios of integrating Big Data into purchasing, as a result of experiences.

The literature which is used in the literature review is used as a base for the interview guide. It is studied intensively for building further upon this information. The respondents of the interview are internal employees from Bright Cape and external randomly selected employees from different companies. To limit bias in the answers the selected respondents and companies are in different sectors and spread over the entire country. For a reliable result, there will be conducted enough interviews, until the answers show a saturation.

The main characteristic of all the respondents is that they have a function in the Purchasing department or the Big Data department. The respondents are selected and invited by email, which can be found in appendix I. The interviews are conducted in a direct face to face situation, and are all recorded with the agreement of the respondent. The voice records are in the possession of the researcher and can be requested. The anonymity of the respondents is guaranteed and the interviews will only be used confidentially. All the external respondents have signed the ethical approval, what can be found in appendix III. The language of all the interviews is in Dutch, since all respondents are native Dutch-speaking.

The interview guide, as shown in appendix II, was developed as a guide for the questions and makes sure that all planned subjects are covered in detail. The interview questions for the internal colleagues are in appendix IV. These interview questions are standardised and during the interview was decided which questions are suitable, depending on the situation, knowledge level and experiences. Therefore, not all the questions are used in each interview. The duration of the interview with internal employees was between 15 and 30 minutes and took place in the office of Bright Cape. The questions of the external interview are in appendix V and took place at the preferred location of the respondents, what was mostly in their personal workspace. The duration of the interviews was approximately 45 minutes. After 18 internal interviews and 7 external interviews, the results show a saturation. Table 11 shows the overview of the internal respondents and table 12 of the external respondents. Further, besides the scheduled interviews, there where were additional topic related conversations next to the coffee machine, during lunch walks and further casual meetings. These conversations are not involved in the overview of table 12.

Respondent	Function	Start date at the company
I	Data driven business analyst	October 2017
II	Data scientist	March 2016
III	Data scientist	March 2017
IV	User Experience consultant	May 2017
V	Junior consultant	January 2018
VI	Junior consultant/ data scientist	July 2017
VII	Manager commercial analytics	June 2017
VIII	Data scientist	August 2016
IX	Innovation consultant	January 2018
X	User Experience consultant	April 2017
XI	Data scientist	September 2016
XII	Data scientist	October 2017
XIII	Data scientist	August 2017
XIV	Junior data scientist	April 2017
XV	Junior consultant/ data scientist	December 2017
XVI	Manager business analytics	June 2016
XVII	Data driven business analyst	Augustus 2015
XVIII	Chief officer	February 2016

Table 11 Details about the internal respondents for this study

Respondent	Company	Company size
A	Production company	1.001 - 5.000
B	Publisher	1.001 - 5.000
C	Merged municipality	501 - 1.000
D	Production company	201 - 500
E	Governmental company	10.001+
F	Municipality	201 - 500
G	Wholesaler	201 - 500

Table 12 Details about the external respondents for this study

5.3 Methodology about the design process of the new Big Data Purchasing Maturity Model, based on different steps and the ideal situation

For using the results of the interview, the answers are fully transcribed in appendix VI for the internal employees and in appendix VII for external employees. For the internal respondents are only the ones in table 11 transcribed, since the other casual talks are not documented. The transcription is the translation of spoken words into a written text[169]. Unless the transcription is made very carefully, it is possible that it captures not the total interview with the verbal communication nor it is totally error free[170]. Since there was not a necessary reason to validate the transcriptions with the respondents, the transcription is not presented to them. If they prefer to see their answers, they can receive the transcription on request. The translation of the internal and external respondents is mostly separated from

[169] See Halcomb & Davidson (2006), p. 38.
[170] See MacLean, Meyer, & Estable (2004), p. 113.

each other since their knowledge level is different. Because of this, the respondents are separated and organised with different characters. The Roman numerals (I, II, III, etc.) are for the internal employees and the capital signs (A, B, C, etc.) are for the external employees. In order to turn the interview results into analysable content, the following actions were taken. First the results of the respondents are separated and grouped to the same parts of the interview. For the external respondents, the format of the Industry 4.0 Purchasing Maturity model will be followed, as seen in chapter 4.4. After that, the answers are clustered to different answers to the same dimension or subject. The answers are coded to have a clear overview of the results. The clustered overviews of the internal employees can be found in appendix VIII and for external employees in appendix IX. Further, the knowledge level of the respondents, their background and the current maturity level of the organisation are taken into account. In the following step, the grouped answers are reflected to a maturity curve for a position on the maturity. The maturity curve has four different levels, which are in line with the four stages of Schiele (2007)[171]. The positioning creates an overview of where the proposed applications have any influence and possibilities at which maturity level. In the next step, the ideal situation is created out of the interview results. The results of the interviews analyse the situation from two sides. On the one hand, the interview results of the internal employees are based on their extensive knowledge and experience with Big Data. That knowledge applies to purchasing, whether they have a limited view of the possibilities. Therefore, the results are more based on data and involved a lower level of purchasing details. On the other hand, the external respondents do have an extended view on the purchasing function, but unfortunately a limited view on the use of Big Data. So, one of the questions of the external respondents was the creation of their ideal purchasing situation, when the possibilities are unlimited. Since their undeveloped knowledge of Big Data, their ideal situation did mostly not involve any Big Data. Both with different views to Big Data and purchasing, there is a gap in the results that is bridged by making some assumptions. Next to the ideal situation and the corresponding maturity levels, there is also a scorecard developed. This scorecard includes standardized questions for determining the maturity level. Subsequently, the information that was given in addition to the specific topic, but provides valuable information, was analysed and discussed. These results are also in the analysis of the interviews.

[171] See Schiele (2007), p. 278.

6. Analysing the research results and the development of the new Big Data Purchasing Maturity Model

6.1 Analysis of the differences and similarities of Big Data maturity models and Purchasing models

Maturity models are available and applied companywide with possibilities for each different sector or on each department or task separately. The aforementioned different Big Data and Purchasing Maturity models have many characteristics in common and some parts are dissimilar. About the things in common could be said, that all the models have a certain form of constructive levels, where the next level is built upon the current situation. There is a logical order in the levels and are classified in the same way. However, for both Big Data and purchasing maturity models, there is a distinction in the average of levels. For the Big Data, the ten maturity models have on average 4.8 levels. For the purchasing maturity models, the average is on 3.9 levels, resulting in almost one level difference on average. Further, one can say that the year of publication is influencing the results. The Big Data maturity models are published in a period between 2013 till 2017, whereby the average is calculated on 2014.8. While the purchasing maturity models are published in between the years 1988 till 2007, with an average in 1995.9. With a difference of more than 19 years, the time period is an important part to identify. In the current industry, the developments are following up each other in a high pace what makes it harder keep up with the competition. For comparing both parts of the maturity models, the purchasing maturity models are relatively outdated, while the content is still applicable in the current purchasing role. That shows a strong content of the models, since the purchasing function is strongly developed recently. Comparatively, the development in Big Data is growing and extending rapidly, what makes maturity models earlier outdated.

Content wise, the maturity models have also similarities and differences. Although the facts of the different purchasing maturity models are not detailed accessible for the current literature review, it is possible to make analyses concerning the content. It is remarkable that the half of the Big Data maturity models are indulged with measurement dimensions, while all the purchasing maturity models do have mentioned measurement dimensions or addressed topics. That makes the Big Data maturity models harder to define on the dimensions, but the different levels and purposes can add information. The different levels show that the models are structured in in another way, relative to the purpose. A descriptive model has five levels and describes more or less the situation. All the comparative models have dimensions where they can evaluate the situation on the same dimensions as other models. Almost all the prescriptive models do not have any dimensions but are used for describing the guidelines. In the same way, purchasing models do not have an obvious purpose of the model. That results in different ways of using the maturity models and situations where the models can be applied.

Next to the maturity models of Big Data and purchasing, Bright Cape has designed its own model. This model is based on Gartner's Enterprise Information Management (EIM) Maturity Model. The model of Bright Cape has five different stages, from 1) Data quality 2) Descriptive, 3) Diagnostic, 4) Predictive and 5) Prescriptive. For comparing that model with the four stages designed model of Schiele (2007), the following situation is created, see figure 9 below.

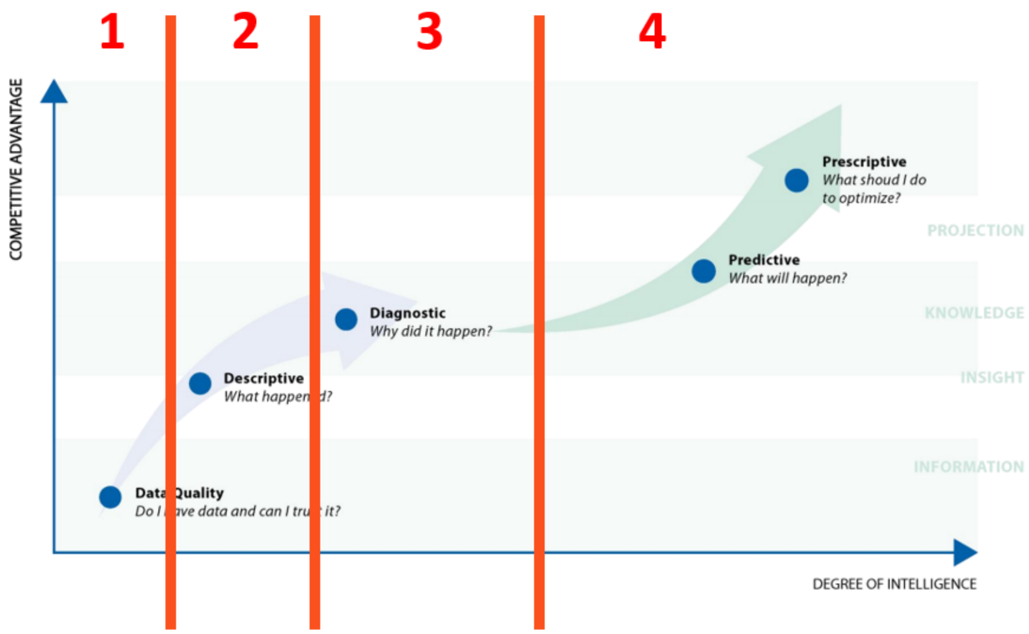


Figure 9 Bright Cape Maturity model compared to the four stages of Schiele

As seen in the model, the first three stages of the Bright Cape maturity model relate to the first three stages of the model of Schiele (2007). The fourth and fifth stage of Bright Cape are together Schiele's fourth stage. In both models, the first stage is the initial phase where all the (right) data available is collected and check which data is generated in the current processes. If this is wrong or incomplete, the first step will erase the wrong data before making analyses, otherwise the analyses will be made on wrong or biased data. The second stage offers similarities between the data collecting of the Bright Cape maturity model, with assigning tasks to people or positions to gather data. The third stage is in both models more related to systems and processes. The final stage has automated systems and predictive assumptions. The two separate stages of Bright Cape differentiate from the final stage of Schiele, which is more related to the details and interpretation. In the end, both models have as a final step an autonomous and predictive stage for further development in the company. Due to the overlapping character features, there can be a smooth integration with both models for the coming designed “Big Data Purchasing Maturity model”.

Finally, the maturity model of Bright Cape is based on data maturity. For analysing the model of Bright Cape with the Big Data models, the models 5 (Hortonworks Big Data Maturity Model from Dhanuka) and 7 (Big Data Maturity Framework from El-Darwiche et al.) have overlapping interfaces. Model 5 is with four defined levels a simpler explained version of the Bright Cape model. The levels are building on each other, starting with an awareness phase, followed by the exploring level, then optimizing and rounding off with transforming. Moreover, model 7, with also four levels, show parallels with both mentioned models, but is expressed with rarely used descriptions. In either event, these models have all in the beginning steps an opening or initial phase. Build upon that, the next level shows points of contacts with dash boarding, exploring, functional excellence, all related to the gathering the data. The third phase is more in line with further optimizing, making predictions and value enhancement. For the Bright Cape model, this is the fourth level, and the third level of models 5 and 7. Finally, all these three models have a transforming and implementation phase, whereby the further integration is involved.

6.2 Analysis of the results of the internal interviews

The interviews of the internal employees were classified in different subjects. The first part of the interview was related to the working experience from the colleagues in the Bright Cape holding and the current and previous projects. The second part was about the way how the employees were collecting the necessary information. The third part was about their opinion about different subjects in relation to Big Data and data science. The last part was related to purchasing. During the semi-structured interviews, the order could have been changed in relation to the given answers. When it was clear there was not enough knowledge, not all questions are further asked. The four different subjects will be discussed separately.

6.2.1 Analysis of the successful and current projects in the recent history of Bright Cape, related to the interviews of the internal employees

Since the founding date in 2014, Bright Cape is very active in different companies and sectors. The main facts about the respondents are related to their history within the company, the projects they have captured and the functions they held. Among the eighteen internal employees, on average they were working for 13.3 months at Bright Cape. In function, there are two data driven business analysts, ten data scientists, two user experience consultants (UX), one innovation consultant and three employees with a managing role. Together, all different and relevant functions of Bright Cape employees are covered among the respondents. Throughout the years, they developed an extensive customer database with referral cases. Bright Cape operates in three different ways. They work on 1) consultancy projects, 2) on an interim (project) basis and 3) at innovation projects with cooperating organizations. One of the main cooperating organizations is EIT-Digital. Last October one of the main specializations of Bright Cape, Advanced Analytics, was dissociated from the holding and continued under a different name, called 'Pipple'. The team of Pipple and Bright Cape are working still together on the same desks, buildings and projects, what results in a strong team connection within the Bright Cape holding. There are a few people who worked earlier for Bright Cape and now for Pipple, on the same projects. Therefore, the colleagues of Pipple are also respondents of the internal interviews. As reference cases of the consulting projects can the companies can be mentioned, as shown in figure 10[172].

Figure 10 Satisfied customers of Bright Cape
Source: Bright Cape

172 See www.brightcape.nl

6.2.2 Analysis of the way how the internal respondents collects the necessary information related to (possible) projects

The beginning of each new project for Bright Cape and Pipple starts with an initial phase whereby the involved employees gather information about the ideas, opinions and goals of the new project. For designing the final maturity model, the way in which the employees collect new information will be used, to fit the model into the working behaviour of Bright Cape. To determine how that processes is evolving, the second part of the interview was related to collecting information. Overall, the main reaction of the employees is that they answer with stating that it depends on the situation, the occupations of the conversation partner and the moment in the negotiations about the projects, how the conversation is going. Mostly, they do not have any organised questionnaire, but structure the conversation more related to feelings. Almost all of them mentioned the factor of, especially, the conversation partner, talking. The more the other person will talk, the more information will be given.

When going deeper into the way how they gather information, then the way of questioning is related to: 'How', 'Why', 'What', 'When' and 'Where'. The most common topics that are highlighted, are related to the company organogram or company structure, expectations, responsibilities, entitlements, process structures, demand, databases, problems, issues, and applications. Further, things that also discussed are the timetable, steps for further processing, visualisations, goals, available data, planning and implementation. The given examples, which are part of the initial phase, are holding interviews and meetings, mapping ideas on the maturity curve, clustering information, testing hypothesis and setting up a business case. As final given information on the collecting of information, is that for an optimal conversation, the goals and information of the questionnaire is applied to the conversation partner and specified to the company. Terminology and word choices has to be changed and applied in each conversation. When all before mentioned subjects extensively are discussed, there will be a good coverage of the related and involved information.

6.2.3 Analysis of their opinion about subjects related to Big Data applied within a company

The third part of the internal interview was applied to the employees' knowledge about Big Data applications and their opinion about the use of data into an organisation. As a result of this part has to be mentioned that there is a difference into their attitude towards the definition of Big Data. In the table on the next page are the results of their opinion about Big Data.

Unfortunately, three respondents were able to answer this question, since they have recently left the company and not able to reach in any other way.

Number	Definition of Big Data
I	Big Data is a combination of the different V's what together define Big Data. It is not possible to set a threshold for the volume.
IV	The concept of Big Data is an overused terminology, where the functionality is more related to the three different V's.
V	The definition of Big Data is a concept and holds different factors which together are Big Data. The volume is not able to state in figures.
VI	Data which is too big for processing on a normal laptop and requires special software and hardware tools. Threshold for volume is more than 32 GB.
VIII	Small or big is just a label and nothing special. More important how the available data, results in a desired solution and/or useful decision. Choice of the right methods, tooling, databases for the individual situation. The volume is not necessary, it is more about the solutions.
IX	A collection of reliable information, which is big and consistent enough for developing statistical conclusions. Big Data is a technical concept and not able to be written in hard and reliable numbers.
X	Enormous amount of web data, meta data etc., which can be processed with a data mining program or Google Analytics for developing statistical analyses. What result in information or can develop algorithms. So, it is many information what can be used.
XI	Data which is different from a traditional source as figures and tables in Excel. Volume is not the most important, but more the diversity and unstructured data as whole text massages. In my opinion, there is not a threshold for the volume of Big Data.
XII	High volume of data which makes it impossible to see individual data points. It is about the velocity and volume. 1000 points should be able to analyse manually, for Big Data there are algorithms necessary for analysing the data. The threshold is on 1 TB.
XIII	Small data and Big Data is just data. There is no threshold when data is Big Data. When the volume is too big, then there are another solutions for processing that volume of data. The infrastructure is changed and depending on the situation. The boundaries of Big Data are defined by the functionalities of programs as Excel.
XIV	Big Data is an enormous data set, or the velocity of the receiving data is extremely fast, or the data can be unstructured. The minimum volume of Big Data is around 1 TB.
XV	The volume of the data is huge and unstructured. It is not defined manually, and it is not hundreds but thousands of data points which are mentioned as Big Data.
XVI	Extremely high volumes of data. The term Big Data is enormous overhyped. The data is not that special, and the infrastructure needs to be changed to handle the data, that is. As boundaries, the limitations of Excel are a good point to determine Big Data.
XVII	Enormous amount of data which is not only generated but also stored and used for further processing. The volume of one dataset will be 1 TB, but it is a combination of more different datasets.
XVIII	Everything that is not able to progress in Excel or a volume of more than 1 TB. It is a hype, but it is not more than connecting different data sources in another program. Different sources include video sources, and data about a CRM or HRM system.

Table 13 Overview definitions of Big Data of the internal respondents

Based on this overview can be summarized that there is a big difference between the different respondents. Out of the fifteen respondents, there some respondents who mentioned the combination of the 3V's. Based on the fact that almost all the definitions and descriptions of Big Data in literature are related on the 3V's, it is surprisingly that the respondents see Big

Data in another way. After some pressure and repeating asked questions about the volume, the half of the respondents have given numbers about the threshold of Big Data. The others mentioned specific reasons why they don't want to and are able to give a specific threshold. Overall, the six mentioned thresholds of the limitations of Excel, or a minimum value of 1 TB are in line with the stated threshold in chapter 2.2. There are just two respondents who mentioned a lower threshold which will be hold for Big Data. That shows a big different with the other volumes and is hard to clarify. While many respondents mentioned something in relation to the volume in their definition, there are still differences among these respondents. Since there is spoken about no difference between small data and Big Data it is just high volume, and the enormous hype of the terminology of Big Data, it is completely different attitude than respondents who mentioned the applied systems for the dataset. Further, there was just one respondent who mentioned the timeframe in Big Data. Where the current situation is not the same for over ten or twenty years. In conclusion can be stated that the results are surprisingly, based on the fact that all the respondents are employees of the same Big Data consultancy. It was assumed that there is a consistency among the respondents, since the knowledge about Big Data is highly developed. Besides that, the results are in line with the fact that Big Data can be stated as an 'Umbrella term", what makes it hard to define. Finally, the stated threshold of chapter 2.2 is in line with the opinions of the respondents and will be held in the remainder of this master thesis. If there is spoken about Big Data, the threshold of 1 TB and/or the maximum values of Excel is involved.

About the application of data into an organisation, the respondents share the same opinion. They agreed that the addition of data is not a reason on itself, but it has to be added to accomplish settled and measurable goals. The fact that data can have an additional value to the goals and is part of strategic choices, is clear for all the respondents. An important part is that the strategy and goals have to be specified, before the data is added to the process. This is because it has to be clear which results have to be achieved, instead of looking into the data and see what might happen. The usage of data without a specified goal is a result of the current hype about data applications among companies and their competitors. Companies want to go ahead with their competitors without knowing exactly what they are doing. As a result, there is most of the time no additional value, but costs of extra money and time.

6.2.4 The involvement of the internal respondents in relation to everything with purchasing, included possible applications.

As a result of the fourth part of the internal interviews, the paragraph is an extended overview of Big Data applications in purchasing. The respondents of Bright Cape and Pipple have a deep understanding about the data possibilities and applications. They are involved in the achieved goals which are a result of Big Data and know in what way data can have an additional value. Unfortunately, they have not an extended knowledge and experience with the purchasing function itself, but after a detailed summary about the most important purchasing details, they see can add their data skills into the purchasing function. The respondents mentioned possibilities based on their data knowledges and experiences. Additionally, they refer to the six steps of the Van Weele purchasing process as basis for their own processes. An extended overview of the common mentioned ideas for data integration is in table 14 below.

<table>
<tr><th colspan="3">Applications of Big Data in purchasing</th></tr>
<tr><td>➢ Process mining</td><td colspan="2">- Following processes within the systems
- Controlling the processes
- Detect bottlenecks</td></tr>
<tr><td>➢ Machine learning</td><td colspan="2">- Information of parts of the processes within machines
- Detect reduced quality during the process
- Clustering data sources</td></tr>
<tr><td>➢ Finance</td><td colspan="2">- Connection with the finance department and systems
- Pay-to-Cash (P2C) & Purchase-to-Pay (P2P)
- Price negotiations
- Process more based on cost-driven
- Spend-analysis about spend, suppliers, invoices et cetera</td></tr>
<tr><td>➢ Process</td><td colspan="2">- Early supplier involvement (ESI)
- Electronic Data Interchange (EDI)
- Detailed order knowledge at the purchasing department
- Interchange and connection between departments
- Digital and automatic digitalisation</td></tr>
<tr><td>➢ Insight</td><td>- People’s working skills
- Logistics
- Demand planning
- Quality standards
- Performance measurements
- Customer satisfaction
- Market research
- Throughput times</td><td>- Product details
- Buyer details
- Supplier details → Preferred suppliers
- Responsibilities
- Partnerships
- Vendor Management Inventory (VMI) systems</td></tr>
<tr><td>➢ Predictive</td><td>- Costs involved
- Duration time
- Maintenance</td><td>- Order quantity
- Production planning
- Occupation degree</td></tr>
</table>

Table 14: Overview mentioned application possibilities for Big Data in purchasing

Based on the mentioned examples for Big Data applications, there are some given results that can be accomplished in purchasing. They see possibilities for efficiency and optimization with the processes. Further, there can be results achieved in reducing throughput time and other time saving opportunities when the departments are smooth connected to each other.

6.2.5 Future ideas for Bright Cape

At the moment and as said before, purchasing analytics is not a core business of Bright Cape. To understand how the employees will see the future of purchasing analytics, the respondents are asked about their point of view for the future. Since Bright Cape is founded in 2014, the strategy is still evolving based on experiences. A big part of the respondents mentioned that Bright Cape needs to determine a focussed strategy first. This is because they are doing projects in every segment, every department and in each company. To become the best in what you are doing, they mentioned to specify the goals first instead of waiting and see which project will be the best valued part of Bright Cape. Therefore, when purchasing analytics will be a core business, then it has to be a prepared strategy. Further, it is said by the respondents, that their data skills could be perfectly applied within the purchasing department, because the purchasing process is somehow structured and is easily to add the techniques into it. Finally, the respondents that have a managing role mentioned that there are no plans for purchasing analytics in a short time, while the other respondents are more enthusiastic to add it into their competences.

In the case that Bright Cape decide to specialise more in purchasing related data projects, then the following options can be possible. Bright Cape has strongly developed techniques in process mining, which is also applicable on the purchasing process. Since the volume of Big Data is not related to a settled threshold, it is variable to add different data sources. In the purchasing process, there are timestamps and other details on the different steps in the process. With process mining, it is possible to detect bottlenecks and can control the process flow. Another possibility in the purchasing process is related to the inventory in the warehouses. By monitoring the receiving and outgoing goods, and implement and automotive re-ordering systems, data science can result in savings by optimisation. Further possibilities can be seen for example in the contract management. The contracts are more and more digital signed, which create important information for data science. With specific algorithms based on earlier successful and unsuccessful contracts, new contracts could be easier, faster and with more protection signed.

6.3 The results of the external interviews

The external interviews were held at different companies, in different sectors and in different cities in the Netherlands. Among the respondents, there were both public organizations (A, B, D and G) and private ones (companies C, E and F). In total seven interviews were held, until the point that there were many similarities among the answers. All the interviews were with two persons, and the last interview was with four persons. A schematic overview with the details of the respondents, company characteristics and duration of the interview is in table 15 below.

	Occupation	**Company characteristic**	**Location interview**	**Length of interview**	**Number of participants**
A	Operational Buyer	Production company	North Brabant	40 min.	2
B	Product Manager Purchasing	Publisher	North Holland	34 min.	2
C	Employee Policy & Real Estate	Merging municipality	North Holland	33 min.	2
D	Strategic Procurement	Production company	North Brabant	51 min.	2
E	Purchasing Advisor ICT-systems	Governmental company	North Holland	46 min.	2
F	Junior Buyer	Municipality	North Holland	38 min.	2
G	CEO, Purchase Manager and Business Intelligence Manager	Wholesaler	North Brabant	69 min.	4

Table 15 Detailed overview of the interviews

The interviews will be further discussed, based on the model with the eight different dimensions from Industry 4.0, as shown before in figure 8 in chapter 4.4. The questionnaire is structured according to the same dimensions, to create consistency in the answers and for further developing of the new designed model. Besides, in the remaining part of external respondents, all the different respondents/businesses/companies will be identified as 'companies', even though public organisations are structured and operate in a different way. Additionally, for analysing the results, it is necessary to mention that the Dutch government plays an important role for the purchase order from public companies. Those companies are obliged according to several thresholds to have all the purchase orders open for everyone, and make a procurement order online on TenderNED[173]. That is an online order where each applicant has the same possibilities to accomplish that order. Since there is public money involved, with procurement orders the government tried with the help or rules to prevent conflicts of interest. While public companies are obliged to, sometimes private companies also use the online TenderNED for their purchase orders, to keep the order process transparent and fair, to prevent any problems and try to get the best deal as possible.

[173] See www.tendered.nl.

6.3.1 Results of the Big Data & purchasing strategy

The interview questions about the strategy is subdivided in a part of purchasing in the strategy and the use of data science in their strategies. For the involvement of purchasing in the company strategy, the answers show a difference among the public and private companies. Where the public companies C, E and F, have in general a 'Purchase- and Procurement policy' (Inkoop- en aanbestedingenbeleid), the private companies are diverse. For private company G is purchasing a core and strategic function, therefore, it is integrated in their corporate strategy. On the opposite, the other companies are just in the growing phase with the purchasing department and are realizing more and more that purchasing can have a strategic role for the company. The importance of purchasing as a strategic decision maker is yet identified by the respondents. Additionally, almost all the respondents mentioned that they have, or created recently, a centralised purchasing department. In the purchasing department are also different purchasing functions, as the tactical, operational and strategic buyers.

The answers related to data science in the strategy, show many differences. First of all, in companies A, B, D and G it is unclear and not defined what exactly Big Data is and how they use it with data science. Those companies see it more related to just the use of data and data applications. The other three respondents react that they don't know, and have to check it. When they don't know, it probably would not have a big influence on the strategy, otherwise they would have heard about it. Company G mentioned that their strategy is based on three guiding principles, where data-driven is one of them. Unless the company is data-driven, the data in purchasing is not really strong developed and is a point of attention. The data security and data privacy is for all the respondents a big issue. As mentioned before, the new GDPR regulation reflects all Dutch companies, both private and public companies, and is a current big issue. Company B has a whole department for the new regulations as well as companies' C and G mentioned it. The more data privacy and data security is under attention, the more the data regulations are involved in the strategy. Most respondents have also specified that in the answers.

6.3.2 Results of Big Data & purchasing in relation to the process & systems

For the purchasing process, the respondents are asked to define their purchasing processes to one of the four characteristics of chaotic, standardised, digitalised or automatized. Each respondent describes their purchasing process in reality and how it should be according the structure. To compare the results to a purchasing maturity level, the respondents are asked to position their own purchasing department to a maturity curve. The purchasing maturity model that is used, is not one as mentioned before, but a simplified version with 5 levels; 1) basic, 2) defined, 3) foundation, 4) advanced and 5) extended[174]. To see how these processes and maturity levels are related to each other, see table 16 below.

Company	How the process feels	How the process is	Maturity level
A	Chaos	Standardised, partly automatized	2) Defined
B	Standardised, work with Van Weele.	Standardised	3) Foundation
C	Digitalised	Digitalised	2) Defined
D	Chaos	Partly digitalised	Between 1) Basic and 2) Defined
E	Standardised, work with Van Weele	Standardised	Between 2) Defined and 3) Foundation
F	Chaos	Partly digitalised	1) Basic
G	Outdated	Partly digitalised	2) Defined

Table 16 Overview structure of the process related to purchasing maturity level

Based on the results could be said that the purchasing process of the companies A, D, F & G, feels for the purchasers as chaotic or outdated. That is also reflected to the maturity curve. They score on 1) basic or at maximum 2) defined. Conversely, companies B and E reflect their purchasing department as standardised and mentioned the model of Van Weele[175]. Company F mentioned in the results also the model, but that model was not implemented yet. When the purchasing department is standardised according the structure of Van Weele, they are higher on the maturity curve in relation the other departments. Noteworthy, it is not related to the differences between public and private company structures, since company B is a private company and company E a public company. Further, the overview shows that the digitalisation is under development within the purchasing department. Each respondent gave examples about their status of digitalisation. The first step is mentioned as saving all the documents on a digital server. The online part is not yet the most important part, the companies are even struggling with the digitalisation of their information. Next to that, in each interview, the respondents have a negative feeling by the huge number of different steps in the process the purchase order has to go through, for there is an approval. They mentioned for example that with less bureaucracy, more structure, more LEAN processes, and the removal of all additional steps, the procedure will lead to a faster and a more efficient

[174] See Irfan Sabir & Irfan (2014), 52-55.
[175] See van Weele (2010), page number undefined.

process. Even company E agreed that there are too many steps, what leads currently to a process of more than one year.

Another part of the second dimension, are the systems that are part of the purchasing process. In the interviews, each respondent could name their systems which are involved, both the company specified systems and as well as the standardized systems. Overall, there were no respondents that reacted that their systems are perfectly working and integrated together. That should probably be related to the low maturity scores. Companies D and G have an ERP system, but they see also points for improvement. Where company D mentioned that they do not know all the possibilities the systems offer, and see many white spots. Company G adheres to that. The use of a solid integrated ERP system is one thing that the other respondents should add for a better use of systems. Besides all the different systems, the way in which all the information and documents are saved, is also diverse. Where company A works totally in the Cloud, companies B, E, F, and G have different systems for the whole purchasing process. Sometimes was mentioned that companies have still a part of paperwork integrated in the purchasing process. Paper in the process could by risky, since they have experienced that papers are lost, damaged or wronged.

Further, the use of Big Data integrations in the process is different for each respondent. To determine where and how the Big Data applications are integrated in the purchasing process, the interview included also a question to determine their own position on a Big Data Maturity model for a comparative view. As model is chosen for the Big Data Maturity Model of Radcliffe, as mentioned before as model 9 in table 6[176]. This model is chosen due to the clear visualisation of the curve in a picture. This model has six different levels; 0) in the dark, 1) Catching up, 2) First pilot(s), 3) Tactical value, 4) Strategic leverage and 5) Optimize & Extend. In table 17 below is an overview of the maturity levels, given by each respondent.

Company	Big Data Maturity level
A	3) Tactical value
B	3) Tactical value
C	Between 3) Tactical value and 4) Strategic leverage
D	Mostly 0) in the dark
E	Between 0) in the dark and 1) Catching up
F	Between 0) in the dark and 1) Catching up
G	1) Catching up

Table 17 Overview Big Data maturity levels

[176] See Radcliffe (2014), p. 2-5.

As seen in the table, there are little differences in the given maturity levels. Company A has a lot of data and uses it in the process. Their use of data could be called Big Data and is related to the third level maturity. Company B and G monitor their processes, but use mainly the historical purchases for further processing. It is all with human decision making, because they make still the decisions. Surprisingly, without the integration of Big Data, company B scores a third level. Company C uses data in predictive decision making, where they make prognoses about the population growth for example. It is not clear for the respondent how big the volume is of the used data, and if they are using normal data or Big Data. The variety and velocity are in line with the characteristics of Big Data. Company C is a public municipality where their purchases are mostly related to the inhabitants, and it should be helpful to know the prognoses. Therefore, it is to clarify why they scored between the third and fourth level. In contradiction, the other two public companies, company E and F mentioned the restrictions as result of the public procurement tender, as mentioned before. Since every supplier is on the same level, historical data makes no sense. The only part where they see options for adding data should be in the specification phase, as mentioned by Van Weele. There is it possible to optimize the procurement order, based on Big Data. Finally, company D mentioned that the use of Big Data is far away, they just grown up from the Stone Age and have still difficulties with even small data. It is not surprisingly that they positioned themselves as level 0) in the dark. Overall could be said that the lower maturity levels are related to the systems which are involved within the purchasing department, and also the use of data in the processes.

6.3.3 Results of Big Data & purchasing in relation to the physical level

The physical level in the interview is related to the involvement of any physical applications in the purchasing process. Based on earlier answers, the expectation about physical applications was relative low. Company A referred as answer to their vendor management inventory (VMI) system. They have an inventory of small pieces where their supplier replenished the inventory automatically when they are warned that the stock is below a settled value. That is the first step to integrate an automated system with predictive actions. The same example is given by company G, but they are warned in their own system by their own settled values. Their process is according the principle of Plan Do Check Act (PDCA), whereby the Check step is a crucial part for moving to a fully automated system. The theory of Deeming behind PDCA, is that it is a wheel where each step is following the previous one, and is a continuous process[177]. Company D mentioned that their system has the features to integrate a physical system, but that their processes are not compliant and have too much human errors. Type of errors that are controllable, when the process is well defined. Further, the other companies B, C, E and F react that each transaction is driven by a human step. There is nothing that goes automatically, based on data. The companies explained this by giving examples about what the possibilities are in physical systems, but agreed that the implementation is a bridge too far.

6.3.4 Results of Big Data & purchasing in relation to the purchase to pay (P2P)

For the fourth dimension about purchase to pay, there was a corresponding answer of all the respondents. They react that the financial systems or departments are strictly separated from each other. That results in a four-eye principle with the creditors department, as mentioned by company B. For the financial part of the purchasing process, they mentioned that in general the systems of the purchasing department create an overview of the outstanding account. Human interaction is involved to transfer the details to the finance department for a second check and an agreement for payment. Only when that step is accomplished, the payment is carried out. At the moment, the processes and systems of these companies are not ready for a seamless and automated purchase to pay principle.

6.3.5 Results of Big Data & purchasing in relation to the controlling & KPI's

In the interview is the fifth dimension split up into three different parts, namely; the rules and regulations in relation to purchasing, as well ass related to Big Data and the way how the controlling is maintained. First of all, the rules and regulations which are applicable on the purchasing department are for the most respondents different. Company A had actually no idea about the rules and regulations, and these are not well communicated within the company. Public companies C, E and F refer to the 'Purchase- and Procurement policy' (Inkoop- en aanbestedingenbeleid), which they are obligated to follow. These are also mentioned at the first dimension, the strategy. Company B added that there are requirements to the board members that are registered by the Chamber of Commerce. Only these registered members are allowed to make purchase orders. Company G added the same situation for procuration regulations. Overall, all the companies have rules that they need to follow.

[177] See Sokovic, Pavletic, & Pipan (2010), p. 477.

The rules and regulations with regard to the use of Big Data in the company are in the same line as the ones for purchasing. For both public and private companies, there are requirements from the government in relation to the use of data and are mandatory to follow. As given example is the GDPR law and data security and privacy issues. Actually, the respondents react with the fact that there is mostly nothing about Big Data, but more about just data and data science influences. Company C give as example, that they have penalties for people who are leaking information or do not handle it carefully. Companies B and F added as example, that they have a specific department, that is just for checking data related situations, to make sure everything is according the rules and regulations. Further, company B mentioned that all their agreements are according their contract standards and conditions. In these conditions are parts of the data influences integrated, and have to be signed for closing a deal. Company G adds that they have a warehouse on their source systems as a security layer. They mentioned also their protected Business Intelligence system and the authorisation in the ERP system is strictly related to the function of the employee. With restricted access, they tried to prevent and secure the internal data and information for unauthorized actions or errors.

Furthermore, the reactions about the controlling part are surprising. Four out of the seven respondents react that they have no idea how the companies maintain the controlling part for both purchasing and Big Data. They did not notice anything in relation to controlling. The other three companies gave examples to explain their situation from the company. Company B mentioned that there is a team where the privacy officer checks if everything goes according agreements. Company F added that in their company, there are many employees just work solely and sometimes they skipped the purchasing department for a purchase. That could happen but there is nobody who noticed that. As a result, this can lead to missing earlier agreements of discounts et cetera. Company G mentioned that they have an accountant who has a frequently check on the figures and procedures.

6.3.6 Results of Big Data & purchasing in relation to the sourcing

The sixth dimension gives a view on the sourcing process for selecting suppliers for a new contract of new suppliers for a current purchase contract, or in some other way. Since the respondents represent both public and private companies, the sourcing process is also depending on the earlier mentioned governmental rules and regulations. Since companies C, E and F are public companies, they have to deal with that restrictions, what is also seen in the answers of these respondents. They show a limited input of suppliers' evaluation and earlier experiences, what is used for a next purchase order. Additionally, these three respondents mentioned, that there is a preference for local suppliers instead of the bigger national or even global companies. Company C added that they do supplier evaluations, but they discussed it only among themselves, and leave it just by words and attitude towards a supplier. For the private respondents, there are different results. Company A states that only the initial purchasers are directly involved in the supplier selection process, and that the biggest parts of the supplier database are prescribed by their buying companies. For that reason, purchasing in company A is more an executing role instead of a strategic role. Company B added, that they follow for the greater part the tender procedure to have an open and transparent purchasing process. While it for public companies is mandatory to do that above settles thresholds, private companies are free to choose for the tender procedures, like company B did. For the lower amount of purchase order, they use their supplier database, which are mostly involved with standard, low valued goods. These goods are similar to the routine goods as mentioned by Kraljic (1983)[178]. Only when their database is not satisfactorily enough or there are suppliers with a negative supplier evaluation, they will search on the internet for new suppliers. In that way, the supplier evaluation is part of the sourcing process. Further, company D mentioned that they have an extended supplier database with all the information with regards to the supplier evaluation and performance measurements. That is surprisingly, since they score relative low on the data involvement in their processes, and are still on the lower maturity levels in both purchasing and Big Data. That results in a higher score for the sourcing process, because their suppliers have an important influence towards the outgoing product. Therefore, the sourcing process of company D is more developed as other departments of the company. Finally, company G describes their sourcing process as different and depending on origins of the supplier. For European companies, there is much historical data, what is used as background information for new orders. Controversial, the suppliers from the Far East are mainly new and without any information. The purchasing process of company G is mentioned as a complex process. Therefore, they tried to gather as much as possible information about the service, complaints and delivery reliability on forehand of making the order to be sure it is an intelligent decision.

[178] See Kraljic (1983), page number undefined.

6.3.7 Results of Big Data & purchasing in relation to the suppliers

For the seventh dimension, there was a part in the interview about the suppliers and how they use Big Data in their processes. For all the companies, it is different where their suppliers are located around the world. For the public companies, it is around the company, sometimes another part of the Netherlands or at least in Europe. The companies A, D, and G import (part of) their products outside Europa and have to deal with other standards and working conditions. In the interviews, the respondents react mainly that they have no idea if or how their suppliers use Big Data. Company A mentioned that it is not his part of the job to visit and see their suppliers, since there is just only contact by orders and confirmations. In addition to that, the respondent can see whether a supplier has an advanced system or not. Sometimes the response of a supplier is just a sign underneath the order confirmation form. But sometimes, there are suppliers who translates the order to their own systems and send a redesigned confirmation back, what should be assumed that they use an order system. The public companies C, E and F mentioned all three that they do not have any idea about that. Due to the procurement tenders, the suppliers react with the necessary information and is mostly for all companies the same structure. Company B added that their buyers have mostly a supplier database for their own usage. The supplier's database of company D has to deal with somewhat bigger suppliers, with an own ERP system. On the experience of company D, the biggest companies do not necessarily have the most advanced systems, it is more in contradiction. Finally, company G makes a difference between their suppliers located in Europe and in the Far East. The ones in Europe are in general more developed. The suppliers in the Far East are more underdeveloped, and the companies that are 'smart' companies, are mostly just production locations. In the current tendency, in China, for example, recently more and more young educated people take over the companies from their uneducated parents, what results in more English spoken and more developed companies. For the improvement of the suppliers, it is a good development.

6.3.8 Results of Big Data & purchasing in relation to the employees & users

The last dimension is related to the employees, and in this research specific towards the involvement of Big Data. That is investigated based on the knowledge level and how willing the employees are to add Big Data into their working environment. As mentioned before in chapter 2.6, skilled employees who can work with data is a strong barrier that needs to be covered for an integration. As result of the interviews, an overarching view is that department wide, there are at least certain people with knowledge about data. Big Data as concept is almost too hard to understand and to apply, since the different V's are wide ranging. Therefore, the usage is definitely a bridge too far. At companies A, B, E, F and G, there is knowledge about the use of Big Data at the IT-department, but at the purchasing department, the knowledge level is still limited. Company C relates to the little applications they already have in predictive prognosis. Company D did not even mention their IT-department; therefore, it is still unknown if there is any Big Data knowledge available. Related to the willingness of the employees, that is not mentioned as a difficult point. Overall, the employees are somehow adapting the current trends and are going on with the applications. Company D mentions that the time and priorities are the most important factors for the willingness of Big Data application. When the employees have the time and see it as a priority to integrate, the involvement will be going on. Actually, the important reason is mostly missing, what results in a postpone in integration.

6.4 The ideal situation for Big Data and Purchasing based on the literature research and interview results

Based on the previous literature research and the results of the internal and external interviews, an ideal situation for the integration of Big Data applied into purchasing will be formulated. The same structure as previous paragraph will be used, as based on the model of industry 4.0 and purchasing. Each part of the eight dimensions will separately be discussed in the optimal way, as defined for the most ideal and last stage.

Strategy

As first stated dimension, strategy is the most important part of the new model, because of the fact that it positioned the importance and strategic power of the purchasing department in relative to the other departments. When the purchasing function is with a Chief Purchasing Officer (CPO) on the C-level of the company, strategic decisions can be significantly improved. With the additional information, the decisions can be made with a long-term perspective. As the final level of the new model, the adopted CPO is part of a strategy where there is a fully systematic and company-wide adoption of Big Data techniques integrated in the purchasing decision making process. The purchasing function is an important part of the strategy of the company. They follow a structured roadmap, that is designed for a long-term and optimal integration and with a clearly defined vision. The strategy is periodically updated with the developing trends and application possibilities to strengthen the purchasing position within the organisation.

Processes & Systems

For now, the process is mostly related to a standardised, digitised, automated or even chaotic flow. In a future perspective, the last level is a fully integrated, autonomous and predictive system, where all the human involved tasked are taken over by systems. The tactical, operational and strategic purchasing processes are integrated in the autonomous flow and there is a continuing monitoring system for controlling. All the human errors are erased out of the processes, which are based on Big Data science. E-sourcing and E-procurement are standardised and involved for leverage, routine and bottleneck purchase orders, the strategic purchase is separated. Since the purchase of strategic goods is more crucial, there is a split in the process to involve a higher rate of control. The decision support systems and measurement values are based on the business analytics, data mining, machine learning and/or process mining. Finally, the most optimal planning for the processes in relation to the demand will be created, as a form of predictive knowledge created out of the produced Big Data.

Physical level

In the current situation, organizations are making the first steps in implementing cyber-physical steps. In the future, it is expected that there is a predictive and automotive integration between the systems. That is a virtual and seamless connection between machine-to-machine (M2M) communication through the Internet of Things (IoT). Due to the seamless connection, there is a real time analysing of the involved purchasing data, where the level of demand is automatically created and executed. The order is placed automatically in the system of the most suitable supplier from the supplier database.

Purchase to Pay (P2P)

Developments in the purchasing process should be in line with a seamless connection between systems and an automated process flow. For now, the current payment processes are involved by humans to have a final check on all liberties the involved employees have in the systems. As last level in the model, the automated systems are arranged so that only ordering according the process is allowed and all sideways or short cuts for buying products out of the procedure is prohibited. That results in effective and more control in payment systems with an election of unnecessary payments and insights in the spend analysis. The same real-time tracking of the processes and physical systems are involved in the payment processes. Further, Big Data integrations help the financial department to handle complex contracts and fulfils it efficiently.

Controlling & KPI's

When the human involvements are erased out of the processes, the human feedback loop for controlling is also erased. Therefore, with all the automated and integrated systems, the controlling part of the process is increasing strengthened in importance. In the last level, there is always, anywhere and anytime up to date and complete information available, which is relevant for purchase managers. There is a complete transparency in the data, with a strong cyber safety. Further, there is a predictive decision support system for adjusting possible errors. The systems are self-learning and optimizing out of monitoring the real time data of Big Data. The controlling part includes also the involvement of data security and data privacy. With a detection of fraud, misuse, errors, destruction or any other modification protection in an early stage, further problems will be drain off. With data as leading part of the processes, the level for preventing certain damage is an important part for growing into the future.

Sourcing

The sourcing process with the specification, selecting and contracting steps, it is a time-consuming process, since there is a high degree of human intelligence involved. In the new model, the ideal situation is sourcing defined as a total integration of e-sourcing and e-procurement with a predictive demand. The real-time analysis of data results in an increasing benefit based on the historical data. With the integration of external information about the current situation in the market, which can be influencing the decision-making process, it will result in the most suitable sourcing decisions. With connected e-sourcing systems, the supplier database can be used as optimal as possible, and allows companies to extend when there is a data-driven demand for new suppliers. Predictive and smart sourcing results in an increasing efficiency in demand-supply and in the total cost of ownership.

Suppliers

Nowadays, the buyers and suppliers work according their own structure, to have the best situation for themselves. In the most ideal situation, the suppliers collaborate closely for creating the best and equally profit for the involved parties. That is directly through digital integrated supply chains and connecting systems. By transferring knowledge in two ways, both buyer and supplier will evolve and develop, even the suppliers with less capabilities. The purchasing company is as strong as the supplier with the lowest capability. With integration possibilities, the most ideal situation results in early supplier involvement (ESI) for improvements in the purchasing process. Driven on Big Data integrations, the predictive demand strengthens the ESI and identifies possible supply chain disruptions and risks.

Employees & Users

The last dimension is related to the employees as users, which are part of the process, work with the systems and have to deal with the Big Data in purchasing. In the ideal situation, the users adapt the new structure of working and understand what they and the systems are doing. The learning and developing skills of the users are coordinated and stimulated by a managing purchasing function. In the most ideal situation, the learning and developments skills go through a further level with feedback loops and evaluations. With continuous improvements, the users can easily adopt the next development to stay updated.

6.5 Building the new Big Data Purchasing Maturity model with four stages

The ideal situation, which is created in chapter 6.4, will be used as basis for the new designed Big Data Purchasing Maturity model. For the new model, it is decided to distinguish four different stages, which are structured and based on the guidelines of Schiele (2007). The four different stages of Schiele are defined as the following stages:

Stage 1: *"A particular best practice activity/tool/method is known within the organisation."*

Stage 2: *"A position or person is assigned to perform the task."*

Stage 3: *"The process for completing the task is defined and documented as well as applied."*

Stage 4: *"Cross-functional integration in the company is assured while basic requirements are met."*

To specify the four stages of the new Big Data Purchasing Maturity model, the four stages of Schiele and the four stages of Torn (2017) will be used as referring stages. As mentioned before, the four stages of Torn are specified as; 1) Digital Novice, 2) Vertical Integrator, 3) Horizontal Integrator and 4) Digital Champion. The extended descriptions of these four stages are explained and defined in chapter 4.4. Building upon that, the four stages of the Big Data Purchasing Maturity model are defined as the following definitions, as stated on the next page:

Stage 1: *"The purchasing processes are well defined following the best practices of Industry 3.0. There are no (Big) Data applications integrated."*

Stage 2: *"The purchasing processes are standardised and digitalised. There are the first applications of Big Data and there is one person assigned to perform the task."*

Stage 3: *"The Big Data is fully integrated into the purchasing processes, and are cross-functional integrated through the company."*

Stage 4: *"The Big Data processes are fully autonomous organised within the strategic purchasing department. The systems and processes are self-learning and continuously improving".*

After defining the four stages, the dimensions of the new Big Data Purchasing maturity model are elaborated in the scorecard. The scorecard, which can be find in appendix X, is structured according the same eight dimensions, and explained in detail. To determine a final maturity score and position on a maturity curve, score points can be given as mentioned in the model.

7. Conclusion

7.1 Conclusion

The aim of this research was to design a new maturity model where the purchasing function or department was combined with integrated Big Data applications. As a result, there was a new Big Data Purchasing Maturity model designed, with an attached scorecard for determining the score on the curve. Since Big Data is an upcoming trend among companies and there are less known applications of Big Data in relation to purchasing, the literature review and designed model had a more explorative background. Therefore, there are less best case scenario's implemented in the research. The stated main research question and sub research questions are based on the problem as detected and mentioned by Bright Cape. The (sub) research questions will be answered as follows:

"What is the current situation of Big Data in purchasing?"

Based on the explorative research, there is a little to no integration of Big Data in purchasing. The literature review shows the initial steps for possible integration options, but there are not any best practice cases written. In addition to that, based on the internal interviews, there are some examples given about the first steps which are taken to look at data by process mining and with purchasing dashboarding. Hence, both Big Data integration into a company and a strategic purchasing function are recently in strong evolvement. Therefore, it is assumed when combining both parts of Big Data and purchasing, and implementing the integration in the corporate strategy, that they can reinforce each other strongly and strengthen the company.

"How does a purchasing maturity level relate to a Big Data maturity level?"

When analysing the given maturity models, related to Big Data and purchasing, there are many similarities. Despite the differences among the maturity models related to Big Data or to purchasing itself, the structure is based on the same guidelines. The first stage is mostly related to an initial or basic phase, followed by the first steps of implementing and structure the processes. The third step is mostly defined to standardised descriptions and goes on the curve to automotive processes and ends on the curve with a predictive and fully integrated process. It is model specific to define a certain number of steps, where four or five are the most common chosen.

"How are the different steps designed and specified in the new Big Data Purchasing Maturity model?"

The different steps are designed in the Big Data Purchasing maturity model, as stated in appendix X. There are four different stages defined, where the first stage is related to an initial and/or pre-mature phase for applications of Big Data and purchasing as a supplementary function. The second stage is based upon a situation where the person or position is skilled to perform a more strategic purchasing task, which is related to Big Data. The third stage is defined to a more standardised process, as organised in a roadmap. There are applications of communication among machines. Finally, the fourth and ideal stage is designed as a fully integrated and autonomous decision-making system, driven on Big Data and is aligned with the corporate strategy. The eight different dimensions are the foundation of the topics which are in the model, for covering all important subjects. These dimensions build on the eight dimensions of the Industry 4.0 Purchasing Maturity model.

7.2 Limitations and further research

This research looks into a wide area of different theories and practices, however, there obviously is a chance that the information is already outdated in the area of data science. In the last couple of years, the development is at such a high rate, that it is hard to keep it updated with the most recent information, what could be mentioned as a limitation for the designing part of the maturity model. Next, the overviews of maturity models show that the models of purchasing are in another time span than the Big Data models. Thus, the purchasing function has to be developed in the last decade, but unfortunately that is not seen in the models. Further, another limitation of this research has to be mentioned that the background of the interviews is too abstract. For a deeper and more complete view into the purchasing department, the variety of respondents should be more extended in the next research. At this moment, a limited number of respondents can give a biased view. Additionally, the respondents should be separated in, for example, private or public companies, or to globally and local companies et cetera. Then the analysing part is more related to the companies operating according the same rules and options. Now the governmental rules are an important issue in the discussion and analysis. Additionally, the respondents of the interview are held with employees working on different levels in the organization. That can be a limitation to the experience and own view on the purchasing department.

Another limitation about the research is the umbrella terminology of Big Data. For the research it is crucial to define the threshold when there is Big Data in the process or it is just normal data. In almost all the literature about Big Data, and there is many literature, none of the researchers give even global guidelines about the volume of data. To make sure the literature is not outdated, the same threshold is tried to detect among professors and researchers at the University of Twente. More than 20 skilled persons react with the fact that there is no threshold and it is not able to have one. Therefore, it was really hard to work further on this point and results in a limitation of the model with Big Data integrations.

For further research should be suggested to update or improved the newly designed Big Data Purchasing Maturity model. Now, the model is primarily based on people with knowledge and experiences in Big Data or in purchasing, but mostly without combining knowledge or experiences. For further processing, it is suggested that the research include respondents with experiences and knowledge of both parts and hopefully with best case scenarios. When more companies involve their best practices, and with more companies in different industrial areas, the model can be more specified and with increasing details.

During this research, the final model is not tested among the respondents or other companies, because of the short time frame. For further research, the model have to be tested among the external respondents, and see if they have the same score as they assumed during the interviews. Additionally, the researcher can also fill in the Big Data Purchasing Maturity model, based on the given answers during the interviews. As a double check, both answers can be compared and see what the similarities and differences are. Some influencing factors of the result could be the different knowledge level about both Purchasing and Big Data, the experience with maturity models and the development of the company. It could be helpful to use the internal and external respondents, since their answers are part of this new Big Data Purchasing Maturity Model. The internal respondents are also able to fill in the model, since they can give insights about the different levels in relation to Big Data integrations.

Finally, in further research it could be possible that there are the coming years barely people with best cases about Big Data and Purchasing together. In that case, some methods as a world café, brown paper session or a brainstorm session with people from both functions together, will create the dialogue underneath. That can also result in some more useful details and developments for the model.

8. Bibliography

Amatriain, X. (2015). Xavier Amatriain's answer to What's the relationship between machine learning and data mining? - Quora. Retrieved 17 May 2018, from https://www.quora.com/Whats-the-relationship-between-machine-learning-and-data-mining/answer/Xavier-Amatriain?srid=cgo&share=4db6d33c

Assunção, M. D., Calheiros, R. N., Bianchi, S., Netto, M. A. S., & Buyya, R. (2015). Big Data computing and clouds: Trends and future directions. *Journal of Parallel and Distributed Computing*, *79–80*, 3–15. https://doi.org/10.1016/j.jpdc.2014.08.003

Augur, H. (2015). WTF Is The Blockchain? A Guide for Total Beginners. Retrieved 13 March 2018, from http://dataconomy.com/2015/10/wtf-is-the-blockchain-a-guide-for-total-beginners/

Bäckstrand, J., Tiedemann, F., & Hedén, E. (2015). Competitive advantage based purchasing matrix -A portfolio-approach to differentiated purchasing strategy, (July).

Banerjee, S., & Sriram, V. (1995). The Impact of electronic data interchange on purchasing: an empirical investigation. *International Journal of Operations & Production Management*, *15*(3), 29–38.

Benjelloun, F. Z., Lahcen, A. A., & Belfkih, S. (2015). An overview of big data opportunities, applications and tools. *2015 Intelligent Systems and Computer Vision, ISCV 2015*, 1–6. https://doi.org/10.1109/ISACV.2015.7105553

Bhuiyan, N., & Baghel, A. (2005). An overview of continuous improvement: from the past to the present. *Management Decision*, *43*(5), 761–771. https://doi.org/10.1108/00251740510597761

Bright Cape. (2017). Bright Cape | Big data als concurrentievoordeel. Retrieved 4 January 2018, from https://brightcape.nl/

Brown, E. (2016). Who Needs the Internet of Things? Retrieved 12 March 2018, from https://www.linux.com/news/who-needs-internet-things

Bryman, A., & Bell, E. (2011). Business research methods. *Oxford University Press*, 806. https://doi.org/10.1017/CBO9781107415324.004

Busch, J. (2012). The Meaning of Big Data for Procurement and Supply Chain : A Fundamental Information Shift. *Spend Matters*, 1–2.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, *2*(6), 429–444.

https://doi.org/10.1016/0377-2217(78)90138-8

Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *Mis Quarterly*, *36*(4), 1165–1188. https://doi.org/10.1145/2463676.2463712

Cronin, P., Ryan, F., & Coughlan, M. (2008). Undertaking a literature review: a step-by-step approach. *38 British Journal of Nursing*, *17*(1), 38–43. https://doi.org/10.1177/107808747000500401

Dabrowski, J., Kifetew, F. M., Muñante, D., Letier, E., Siena, A., & Susi, A. (2017). Discovering requirements through goal-driven process mining. *2017 IEEE 25th International Requirements Engineering Conference Workshops, REW 2017*, (October), 199–203. https://doi.org/10.1109/REW.2017.61

Davis, K., & Patterson, D. (2012). *Ethics of Big Data* (First). Sebastopol,: O'Reilly Media, Inc.

de Boer, L., Harink, J., & Heijboer, G. (2002). A conceptual model for assessing the impact of electronic procurement. *European Journal of Purchasing & Supply Management*, *8*(1), 25–33. https://doi.org/10.1016/S0969-7012(01)00015-6

De Leoni, M., Van Der Aalst, W. M. P., & Dees, M. (2016). A general process mining framework for correlating, predicting and clustering dynamic behavior based on event logs. *Information Systems*, *56*, 235–257. https://doi.org/10.1016/j.is.2015.07.003

Denney, A. S., & Tewksbury, R. (2015). How to Write a Literature Review Paper? *JOURNAL OF CRIMINAL JUSTICE EDUCATION*, *24*(2), 218–234. https://doi.org/http://dx.doi.org/10.1080/10511253.2012.730617 How

Dhanuka, V. (2016). Hortonworks Big Data Maturity Model. *Hortonworks*, (March), 1–24. Retrieved from http://hortonworks.com/wp-content/uploads/2016/04/Hortonworks-Big-Data-Maturity-Assessment.pdf

Diebold, F. X. (2012). 'A Personal Perspective on the Origin(s) and Development of "Big Data": The Phenomenon, the Term, and the Discipline', Second Version. *Social Science Research Network*, *13*(3), 1–8. https://doi.org/10.2139/ssrn.2202843

Drath, R., & Horch, A. (2014). Industrie 4.0: Hit or hype? [Industry Forum]. *IEEE Industrial Electronics Magazine*, *8*(2), 56–58. https://doi.org/10.1109/MIE.2014.2312079

Drus, S. M., & Hassan, N. H. (2017). Big Data Maturity Model – a Preliminary Evaluation. *ICOCI Kuala Lumpur. Universiti Utara Malaysia*, (117), 613–620. Retrieved from http://icoci.cms.net.my/PROCEEDINGS/2017/Pdf_Version_Chap14e/PID117-613-620e.pdf

Easton, L., Murphy, D. J., & Pearson, J. N. (2002). Purchasing performance evaluation: with data envelopment analysis. *European Journal of Purchasing & Supply Management*, *8*(3), 123–134. https://doi.org/10.1016/S0969-7012(02)00002-3

Ebrahimnejad, A., Tavana, M., Lotfi, F. H., Shahverdi, R., & Yousefpour, M. (2014). A three-stage Data Envelopment Analysis model with application to banking industry. *Measurement: Journal of the International Measurement Confederation*, *49*(1), 308–319. https://doi.org/10.1016/j.measurement.2013.11.043

El-Darwiche, B., Koch, V., Meer, D., Tohme, D. W., Deckert, A., & Shehadi, R. T. (2014). Big Data Maturity. An action plan for policymakers and executives. *The Global Information Technology Report 2014: Rewards and Risks of Big Data*, 1–24.

Emani, C. K., Cullot, N., & Nicolle, C. (2015). Understandable Big Data: A survey. *Computer Science Review*, *17*, 70–81. https://doi.org/10.1016/j.cosrev.2015.05.002

Farah, B. (2017). A Value Based Big Data Maturity Model. *Journal of Management Policy and Practice*, *18*(1), 11–18. Retrieved from http://www.na-businesspress.com/JMPP/FarahB_Web18_1_.pdf

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, *35*(2), 137–144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007

Gangurde, S. R., & Chavan, A. A. (2016). Benchmarking of purchasing practices using Kraljic approach. *Benchmarking: An International Journal*, *23*(7), 1751–1779. https://doi.org/10.1108/BIJ-01-2015-0011

Gantz, J., & Reinsel, D. (2010). The Digital Universe Decade – Are You Ready? *IDC iView*, (May), 1–16.

Gantz, J., & Reinsel, D. (2011). Extracting Value from Chaos. *IDC iView*, (June), 1–12. https://doi.org/10.1007/s10916-016-0565-7

Gelderman, C. J., & Semeijn, J. (2006). Managing the global supply base through purchasing portfolio management. *Journal of Purchasing and Supply Management*, *12*(4), 209–217. https://doi.org/10.1016/j.pursup.2006.10.002

Golafshani, N. (2003). Understanding Reliability and Validity in Qualitative Research. *The Qualitative Report*, *8*(4), 597–607. https://doi.org/10.1111/j.1475-4762.2010.00954.x

Griessbauer, R., Vedso, J., & Schrauf, S. (2016). Industry 4.0: Building the digital enterprise. *2016 Global Industry 4.0 Survey*, 1–36. https://doi.org/10.1080/01969722.2015.1007734

Günther, W. A., Rezazade Mehrizi, M. H., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. *Journal of Strategic Information Systems*, *26*(3), 191–209. https://doi.org/10.1016/j.jsis.2017.07.003

Halcomb, E. J., & Davidson, P. M. (2006). Is verbatim transcription of interview data always necessary? *Applied Nursing Research*, *19*(1), 38–42. https://doi.org/10.1016/j.apnr.2005.06.001

Halper, F., & Krishnan, K. (2014). TDWI Big Data Maturity Model Guide. *RDWi Resarch*, *2013–2014*, 1–20.

Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Ullah Khan, S. (2015). The rise of 'big data' on cloud computing: Review and open research issues. *Information Systems*, *47*, 98–115. https://doi.org/10.1016/j.is.2014.07.006

Hazelaar, R. (2016). From standardisation , through integration and automation , into machine-to-machine communication. *Master Theis*.

Irfan Sabir, R., & Irfan, M. (2014). Levels and Barriers to Supply Chain Integration: A conceptual model of Supply Chain Performance. *The International Journal of Management Science and Business Administration*, *1*(1), 52–59. https://doi.org/10.18775/ijmsba.1849-5664-5419.2014.11.1005

Jagadish, H. V., Gehrke, J., Labrinidis, A., Papakonstantinou, Y., Patel, J. M., Ramakrishnan, R., & Shahabi, C. (2014). Big data and its technical challenges. *Communications of the ACM*, *57*(7), 86–94. https://doi.org/10.1145/2611567

Kaliannan, M., Chandran, S. D., & Hashim, R. (2010). Electronic procurement implementation in Malaysia: Suppliers' readiness. *2010 International Conference on Science and Social Research*, (December), 1334–1339. https://doi.org/10.1109/CSSR.2010.5773746

Katal, A., Wazid, M., & Goudar, R. H. (2013). Big data: Issues, challenges, tools and Good practices. *2013 6th International Conference on Contemporary Computing, IC3 2013*, 404–409. https://doi.org/10.1109/IC3.2013.6612229

Kaur, N., & Singh, G. (2017). A Review Paper On Data Mining And Big Data. *International Journal of Advanced Research in Computer Science*, *8*(4), 407–409. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=aci&AN=123430617&lang=es&site=ehost-live&scope=site

Keough, M. (1993). Buying your way to the top. *McKinsey Quarterly*, (3), 41–62.

Khalili-Damghani, K., Tavana, M., & Haji-Saami, E. (2015). A data envelopment analysis model with interval data and undesirable output for combined cycle power plant performance assessment. *Expert Systems with Applications*, *42*(2), 760–773. https://doi.org/10.1016/j.eswa.2014.08.028

Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, *1*(1), 1–12. https://doi.org/10.1177/2053951714528481

Kitchin, R., & McArdle, G. (2016). What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. *Big Data & Society*, *3*(1), 1–10. https://doi.org/10.1177/2053951716631130

Kleemann, F. C., & Glas, A. H. (2017). *Einkauf 4.0 - Digitale Transformation der Beschaffung*. Springer Gabler. https://doi.org/10.1007/978-3-658-17229-9

Knoppen, D., & Sáenz, M. J. (2015). Purchasing: Can we bridge the gap between strategy and daily reality? *Business Horizons*, *58*(1), 123–133. https://doi.org/10.1016/j.bushor.2014.09.006

Koops, B., Leenes, R., de Hert, P., & Olislaegers, S. (2012). Crime and Criminal Investigation in the Clouds Threats and Opportunities of Cloud Computing for Dutch Criminal Investigation. *WODC, Ministerie van Veiligheid En Justitie*, 1–5.

Kraljic, P. (1983). Purchasing must become supply management. *Harvard Business Review*, (83509), 109–117. https://doi.org/10.1225/83509

Kwon, O., Lee, N., & Shin, B. (2014). Data quality management , data usage experience and acquisition intention of big data analytics. *International Journal of Information Management*, *34*(3), 387–394. https://doi.org/10.1016/j.ijinfomgt.2014.02.002

Laney, D. (2001). 3D Data Management: Controlling Data Volume, Velocity and Variety.

Application Delivery Strategies, *949*(February 2001), 1–4. https://doi.org/10.1016/j.infsof.2008.09.005

Lee, J., Bagheri, B., & Kao, H. A. (2015). A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems. *Manufacturing Letters*, *3*(October 2017), 18–23. https://doi.org/10.1016/j.mfglet.2014.12.001

Leyh, C., Schäffer, T., Bley, K., & Forstenhäusler, S. (2016). SIMMI 4.0 – A Maturity Model for Classifying the Enterprise-wide IT and Software Landscape Focusing on Industry 4.0. *Proceedings of the Federated Conference on Computer Science and Information Systems*, *8*, 1297–1302. https://doi.org/10.15439/2016F478

Lismont, J., Vanthienen, J., Baesens, B., & Lemahieu, W. (2017). Defining analytics maturity indicators: A survey approach. *International Journal of Information Management*, *37*(3), 114–124. https://doi.org/10.1016/j.ijinfomgt.2016.12.003

Liu, J., Ding, F., & Lall, V. (2000). Using data envelopment analysis to compare suppliers for supplier selection and performance improvement. *Supply Chain Management: An International Journal*, *5*(3), 143–150. https://doi.org/10.1108/13598540010338893

Loebbecke, C., & Picot, A. (2015). Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. *Journal of Strategic Information Systems*, *24*(3), 149–157. https://doi.org/10.1016/j.jsis.2015.08.002

Lummus, R. R., & Vokurka, R. J. (1999). Defining supply chain management: A historical perspective and practical guidelines. *Industrial Management & Data Systems*, *99*(1), 11–17. https://doi.org/Doi 10.1108/02635579910243851

MacLean, L. M., Meyer, M., & Estable, A. (2004). Improving accuracy of transcripts in qualitative research. *Qualitative Health Research*, *14*(1), 113–123. https://doi.org/10.1177/1049732303259804

Maltby, D. (2011). Big Data Analytics. *ASIST*, 1–6. https://doi.org/10.1002/9781119205005

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). Big data: The next frontier for innovation, competition, and productivity | McKinsey & Company. *McKinsey Global Institute*, (June), 1–143. https://doi.org/10.1080/01443610903114527

McAfee, A., & Brynjolfsson, E. (2012). Big data: The Management Revolution. *Harvard Business Review*, 1–9. https://doi.org/00475394

Merino, J., Caballero, I., Rivas, B., Serrano, M., & Piattini, M. (2016). A Data Quality in Use model for Big Data. *Future Generation Computer Systems*, *63*, 123–130. https://doi.org/10.1016/j.future.2015.11.024

Microsoft. (2018). Excel specifications and limits. Retrieved from https://support.office.com/en-us/article/excel-specifications-and-limits-1672b34d-7043-467e-8e27-269d656771c3

Miller, J. G., & Sprague, L. G. (1975). Behind the Growth in Materials Requirements Planning. *Harvard Business Review*, (September). Retrieved from https://hbr.org/1975/09/behind-the-growth-in-materials-requirements-planning

Moore, D. T. (2014). Roadmaps and Maturity Models : Pathways toward Adopting Big Data. *Proceedings of the Conference for Information Systems Applied Research*, 1–8.

Moore, N. T., Cook, C., Grammich, C., & Lindenblatt, C. (2004). *Using a Spend Analysis to Help Identify Prospective Air Force Purchasing and Supply Management Initiatives*. https://doi.org/RAND

Moretto, A., Ronchi, S., & Patrucco, A. S. (2017). Increasing the effectiveness of procurement decisions: The value of big data in the procurement process. *International Journal of RF Technologies: Research and Applications*, *8*(3), 79–103. https://doi.org/10.3233/RFT-171670

Mountasser, I., Ouhbi, B., & Frikh, B. (2016). From data to wisdom: A new multi-layer prototype for Big Data management process. *International Conference on Intelligent Systems Design and Applications, ISDA*, *2016–June*, 104–109. https://doi.org/10.1109/ISDA.2015.7489209

Newman, D., & Logan, D. (2008). Gartner Introduces the EIM Maturity Model. *Gartner Research ID Number: G00160425*, (December), 1–8. Retrieved from papers3://publication/uuid/87C92356-9333-4565-965C-7AE5C4A4D7CA

Nott, C., & Betteridge, N. (2014). Big Data & Analytics Maturity Model. Retrieved 7 March 2018, from http://www.ibmbigdatahub.com/blog/big-data-analytics-maturity-model

O'Hare, M. (2017). Industry 4.0 and what we learnt from the previous 3 industrial revolutions? Retrieved 12 March 2018, from https://www.linkedin.com/pulse/industry-40-what-we-learnt-from-previous-3-industrial-michael-o-hare

Oussous, A., Benjelloun, F. Z., Ait Lahcen, A., & Belfkih, S. (2017). Big Data technologies: A survey. *Journal of King Saud University - Computer and Information Sciences*, (In Press). https://doi.org/10.1016/j.jksuci.2017.06.001

Özköse, H., Arı, E. S., & Gencer, C. (2015). Yesterday, Today and Tomorrow of Big Data. *Procedia - Social and Behavioral Sciences*, *195*, 1042–1050. https://doi.org/10.1016/j.sbspro.2015.06.147

Paulk, M. C., Curtis, B., Chrissis, M. B., & Weber, C. V. (1993). Capability maturity model. *IEEE Software*, *10*(4), 18–27. https://doi.org/10.1109/52.219617

Pöppelbuß, J., & Röglinger, M. (2011). What makes a useful maturity model? A framework of general design principles for maturity models and its demonstration in business process management. *Ecis*, Paper28. Retrieved from http://aisel.aisnet.org/ecis2011/28/

Portela, F., Lima, L., & Santos, M. F. (2016). Why Big Data? Towards a Project Assessment Framework. *Procedia Computer Science*, *98*(WoTBD), 604–609. https://doi.org/10.1016/j.procs.2016.09.094

Press, G. (2017). 6 Predictions For The $203 Billion Big Data Analytics Market. Retrieved 7 February 2018, from https://www.forbes.com/sites/gilpress/2017/01/20/6-predictions-for-the-203-billion-big-data-analytics-market/#684c31062083

Presutti, W. D. (2003). Supply management and e-procurement: Creating value added in the supply chain. *Industrial Marketing Management*, *32*(3), 219–226. https://doi.org/10.1016/S0019-8501(02)00265-1

Radcliffe, J. (2013). The Eight Building Blocks of Big Data. *Radcliffe Advisory Services Ltd*, (December), 1–4.

Radcliffe, J. (2014). Leverage a Big Data Maturity Model to Build Your Big Data Roadmap. *Radcliffe Advisory Services Ltd*, (January), 1–6. Retrieved from www.radcliffeadvisory.com

Ragowsky, A., & Somers, T. M. (2002). Special section: Enterprise resource planning. *Journal of Management Information Systems*, *19*(1), 11–15.

https://doi.org/10.1080/07421222.2002.11045718

Robert Jacobs, F., & 'Ted' Weston, F. C. (2007). Enterprise resource planning (ERP)-A brief history. *Journal of Operations Management*, *25*(2), 357–363. https://doi.org/10.1016/j.jom.2006.11.005

Rodrigue, J.-P. (2017). The four industrial revolutions. Retrieved from https://transportgeography.org/?page_id=1363

Rouse, M. (2017a). What is data mining? Retrieved 25 February 2018, from http://searchsqlserver.techtarget.com/definition/data-mining

Rouse, M. (2017b). What is machine learning? Retrieved 25 February 2018, from http://whatis.techtarget.com/definition/machine-learning

Sahin, F., Accorsi, R., Frey-Luxemburger, M., Knahl, M., & Atkinson, S. (2017). A Process mining framework to simulate and predict patient pathways and health related costs based on medical claim data. *Collaborative European Research Conference CERC2017*, (September), 1–384.

Schiele, H. (n.d.). Handbook of Operations. *DRAFT*, (Chapter 5), 1–49.

Schiele, H. (2007). Supply-management maturity, cost savings and purchasing absorptive capacity: Testing the procurement-performance link. *Journal of Purchasing and Supply Management*, *13*(4), 274–293. https://doi.org/10.1016/j.pursup.2007.10.002

Schmarzo, B. (2016). De-mystifying the Big Data Business Model Maturity Index. Retrieved 14 March 2018, from https://infocus.dellemc.com/william_schmarzo/de-mystifying-big-data-business-model-maturity-model-index/

Schultze, U., & Avital, M. (2011). Designing interviews to generate rich data for information systems research. *Information and Organization*, *21*(1), 1–16. https://doi.org/10.1016/j.infoandorg.2010.11.001

Schulz, A. A. (2017). Analysis of the current state of Industry 4 . 0 in purchasing : A multiple case study based on the purchasing year cycle. *9th IBA Bachelor Thesis Conference*, 0–14. Retrieved from http://essay.utwente.nl/72702/1/Schulz_BA_BMS.pdf

Schumacher, A., Erol, S., & Sihn, W. (2016). A maturity model for assessing Industry 4.0 readiness and maturity of manufacturing enterprises. *Procedia CIRP*, *52*, 161–166. https://doi.org/10.1016/j.procir.2016.07.040

Sentryo. (2017). Industrial revolutions: the 4 main revolutions in the industrial world. Retrieved 12 March 2018, from https://www.sentryo.net/the-4-industrial-revolutions/

Sinha, K. (2016). Big Data: The Future Has Arrived. Retrieved 17 May 2018, from http://www.dbtalks.com/article/big-data-future-has-come/

Sofaer, S. (2002). Qualitative research methods. *International Journal for Quality in Health Care*, *14*(4), 329–336. https://doi.org/10.1093/intqhc/14.4.329

Sokovic, M., Pavletic, D., & Pipan, K. K. (2010). Quality Improvement Methodologies – PDCA Cycle, RADAR Matrix, DMAIC and DFSS. *Journal of Achievements in Materials and Manufacturing Engineering*, *43*(1), 476–483.

Strik, J. (2016). GDPR compliance in zes stappen. Retrieved 28 March 2018, from https://www.computable.nl/artikel/opinie/security/5851788/1509029/gdpr-compliance-in-zes-stappen.html

Sulaiman, H., Che Cob, Z., & Ali, N. (2015). Big data maturity model for Malaysian zakat institutions to embark on big data initiatives. *4th International Conference on Software Engineering and Computer Systems (ICSECS)*, 61–66. https://doi.org/978-1-4673-6722-6/15/$31.00

Tanenbaum, M., & Holstein, W. K. (2016). Mass production. Retrieved 12 March 2018, from https://www.britannica.com/technology/mass-production#ref17834

Tiefenbacher, K., & Olbrich, S. (2015). Increasing the Level of Customer Orientation - A Big Data Case Study from Insurance Industry. *ECIS 2015 Research-in-Progress Papers*, 0–10. Retrieved from http://aisel.aisnet.org/ecis2015_rip/1

Torn, I. A. R. (2017). THE FUTURE OF PURCHASING AND INDUSTRY 4.0: How Purchasing can Progress and Benefit the Fourth Industrial Revolution. *Master Thesis*, (December).

Tracy, S. J. (2013). *Qualitative Research Methods: Collecting Evidence, Crafting Analysis, Communicating Impact*. *John Wiley & Sons, Ltd.* https://doi.org/10.5613/rzs.43.1.6

Úbeda, R., Alsua, C., & Carrasco, N. (2015). Purchasing models and organizational performance: A study of key strategic tools. *Journal of Business Research*, *68*(2), 177–188. https://doi.org/10.1016/j.jbusres.2014.09.026

Umble, E. J., Haft, R. R., & Umble, M. M. (2003). Enterprise resource planning:

Implementation procedures and critical success factors. *European Journal of Operational Research*, *146*, 241–257. https://doi.org/10.1016/S0377-2217(02)00547-7

van Weele, A. J. (2010). *Purchasing & supply chain management: Analysis, strategy, planning and practice* (5th ed.). Cengage Learning EMEA.

Wu, X., Zhu, X., Wu, G. Q., & Ding, W. (2014). Data mining with big data. *IEEE Transactions on Knowledge and Data Engineering*, *26*(1), 97–107. https://doi.org/10.1109/TKDE.2013.109

Yin, Y., Stecke, K. E., & Li, D. (2017). The evolution of production systems from Industry 2.0 through Industry 4.0. *International Journal of Production Research*, *7543*, 1–14. https://doi.org/10.1080/00207543.2017.1403664

Zakir, J., Seymour, T., & Berg, K. (2015). Big Data Analytics. *Issues in Information Systems*, *16*(2), 81–90.

Zschech, P., Heinrich, K., Pfitzner, M., & Hilbert, A. (2017). Are You Up for The Challenge? Towards the Development of a Big Data Capability Assessment Model. *25th European Conference on Information Systems (ECIS)*, *2017*, 2613–2624.

9. Appendices

I. Appendix 1: Dutch invitation letter for an interview

Geachte heer/mevrouw

Mijn naam is Laura en ik ben een student aan de Universiteit Twente. Daar volg ik de master Business Administration met de richting Purchasing & Supply management. Voor het afronden van mijn master doe ik voor Bright Cape een onderzoek op het gebied van Purchasing en Big Data, waarvoor ik uiteindelijk een gecombineerd maturity model ontwikkel. Een maturity model is een analyserend model waarin verschillende fases worden beschreven. Voor purchasing bestaan er hedendaags verschillende soorten maturity modellen, echter allemaal nog zonder de invloed van Big Data. Door de steeds groter wordende aanwezigheid van Big Data binnen het bedrijfsleven wil ik dit gaan combineren met Purchasing in een maturity model.

Voor dit onderzoek ben ik op zoek naar personen die vanuit verschillende invalshoeken met Purchasing, Big Data of een combinatie daarvan te maken hebben. Daarom zou ik u graag willen interviewen om meer en/of nieuwe inzichten te krijgen om verder te verwerken. Ik hoop dan ook dat er vanuit u interesse is om deel te nemen aan mijn onderzoek en dat u beschikbaar bent in de week van maandag 22 januari 2017 t/m vrijdag 25 januari. Ik heb een aantal vragen op het gebied van Purchasing en Big Data binnen uw organisatie en uw visie daarover in het algemeen. Het interview zal ongeveer 30 á 45 minuten in beslag nemen. De voorkeur gaat uit naar een face-to- face interview, maar uiteraard is het telefonisch ook mogelijk.

Ik hoop dat er interesse is om mee te doen aan het onderzoek en dat we een geschikt moment voor kunnen vinden om het interview af te namen. Als u verder nog vragen heeft, dan kunt u mij gerust een bericht sturen. Op onderstaande gegevens ben ik te bereiken.

Bedankt alvast voor de tijd die u genomen heeft om dit verzoek te lezen en ik kijk uit naar uw reactie.

Met vriendelijke groet,

Laura de Haan
L.m.dehaan@student.utwente.nl

II. Appendix 2: Interview guide

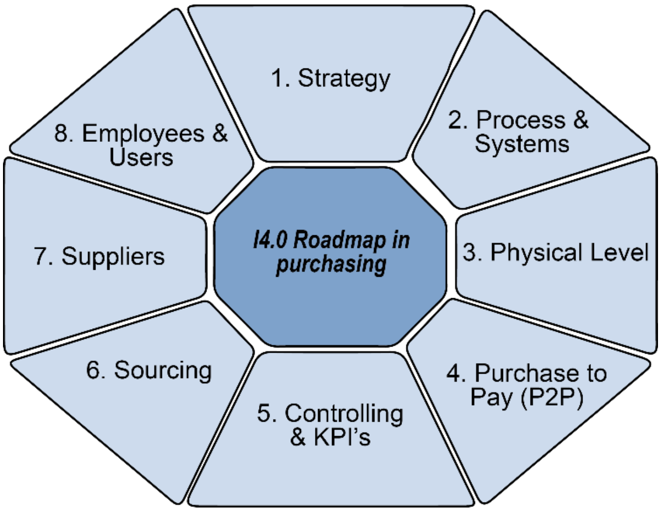
Interview guide

Bij aanvang van het interview

- Goedkeuring van de geïnterviewde voor ethische verantwoording
- De interviews zullen ongeveer 30 minuten in beslag nemen.
- Toestemming vragen voor het opnemen van het interview.
- Anonimiteit van de geïnterviewde wordt beschermd → Aangegeven dat de details anoniem verwerkt worden door middel van codes voor naam.

Afnemen van het interview

- Interview vragen zie onderstaand voor zowel interne medewerkers als externe medewerkers
- Vragen gebaseerd op dit model:



Afronding van het interview

- Controleren of alle vragen zijn behandeld.
- Aanvullende documentatie vragen, wanneer noodzakelijk.
- Bedanken van de geïnterviewde.
- Interview afsluiten.

III. Appendix 3: Ethical approval

Toestemmingsverklaring formulier (informed consent)
Titel onderzoek: Ontwikkeling van een Purchasing Big Data Maturity Model
Verantwoordelijke onderzoeker: Laura de Haan

<u>In te vullen door de deelnemer</u>

Ik verklaar op een voor mij duidelijke wijze te zijn ingelicht over de aard, methode, doel en [indien aanwezig] de risico's en belasting van het onderzoek. Ik weet dat de gegevens en resultaten van het onderzoek alleen anoniem en vertrouwelijk aan derden bekend gemaakt zullen worden. Mijn vragen zijn naar tevredenheid beantwoord.

Ik begrijp dat film-, foto, en videomateriaal en geluidsfragmenten uitsluitend voor analyse en/of wetenschappelijke presentaties zal worden gebruikt.

Ik stem geheel vrijwillig in met deelname aan dit onderzoek. Ik behoud me daarbij het recht voor om op elk moment zonder opgaaf van redenen mijn deelname aan dit onderzoek te beëindigen.

Naam deelnemer:
Datum:
Handtekening deelnemer:

<u>In te vullen door de uitvoerende onderzoeker</u>

Ik heb een mondelinge en schriftelijke toelichting gegeven op het onderzoek. Ik zal resterende vragen over het onderzoek naar vermogen beantwoorden. De deelnemer zal van een eventuele voortijdige beëindiging van deelname aan dit onderzoek geen nadelige gevolgen ondervinden.

Naam onderzoeker: Laura de Haan
Datum:
Handtekening onderzoeker:

IV. Appendix 4: Interview questionnaire internal employees

Algemene informatie over de functie binnen het bedrijf

1) Wat is je functie binnen Bright Cape?
2) Hoelang bent je hier al werkzaam?
3) Heb je nog andere functies gehad binnen Bright Cape?
4) Welke projecten heb je gedraaid en draai je nu binnen Bright Cape?
5) Kan je wat meer vertellen over deze projecten?

Potentiele klanten

6) Welke vragen stel je aan een nieuwe/potentiele klant om te peilen waar de behoefte ligt?
7) Hoe zet je de informatie en behoefte vanuit het bedrijf om naar een projectvoorstel?
8) Hoe bepaal je het plan van aanpak wat je gaat aandragen richting de klant?
9) Maak je gebruik van een maturity curve als je naar een klant gaat en de behoefte bespreekt?
10) Zijn er nog controle vragen of aandachtspunten waar op je specifiek op let bij een nieuwe klant?
11) Loop je nog weleens tegen zaken aan die pas later bij het bedrijf aan het licht komen?
12) Op wat voor een manier evalueer je of het voorstel overeenkomt met de behoefte of wens van de klant?

Big Data

13) Wat is voor jou Big Data?
14) Hoe denk jij over het gebruik van (Big) data in het bedrijfsleven?

Inkoop gerelateerd

15) Heb je binnen één of meerdere projecten te maken gehad met inkoop?
16) Door de informatie van het procurement event, waar zie jij mogelijkheden voor een toepassing van het gebruik van data op inkoop?
17) Wat is jouw visie op inkoop analytics?

Afsluitend

18) Zijn er nog andere zaken die betrekking hebben met inkoop, Big Data of Bright Cape, die nog niet aan bod gekomen zijn, maar die wel nog van toepassing zouden kunnen zijn?
19) Heb je nog vragen of andere aanvullende opmerkingen?

V. Appendix 5: Interview questionnaire external employees

Introductie van onderzoek
Toelichten van onderwerp en focus van mijn studie en onderzoek.

Algemene informatie over het bedrijf

1) Hoe zou u uw bedrijf en haar diensten in het kort kunnen beschrijven?
2) Wat is uw functie binnen het bedrijf en hoelang bent u hier al werkzaam?
3) Heeft u eventueel nog andere functies gehad binnen dit bedrijf?
4) Op welke manier bent u betrokken bij het inkoopproces?

Proces gerelateerde vragen
Kleine uitleg wat Big Data is en wat de toepassingen zijn

1. Strategy

5) In hoeverre is de inkoop opgenomen in de strategie van het bedrijf?
6) Is er iets in de strategie opgenomen over (het gebruik van) Data?

2. Process & Systems

7) Zou u het inkoopproces meer beschrijven als gestandaardiseerd, gedigitaliseerd of geautomatiseerd? Of is het gewoon chaos?
8) Zijn er bepaalde system waar jullie gebruik van maken voor het inkoopproces?
9) Wordt er ergens in het proces informatie opgeslagen en op welke manier?
10) Wordt er in het proces gebruik gemaakt van Big Data?

3. Physical level

11) Zit er in het proces een koppeling tussen de systemen en de fysieke handelingen en processen?

4. Purchase to Pay (P2P)

12) Is er een geautomatiseerd betaalproces of werkt het op een andere manier?

5. Controlling & KPI's

13) Zijn er binnen het bedrijf regels en/of richtlijnen die betrekking op het inkoopproces?
14) Zijn er binnen het bedrijf regels en/of richtlijnen die betrekking op het gebruik van (Big) Data?
15) Op wat voor een manier worden deze eventuele richtlijnen gehandhaafd?

6. Sourcing

16) Wordt er in het leveranciers zoek- en selectieproces gebruik gemaakt van bij het bedrijf bekende data, bijvoorbeeld zoals leveranciers evaluatie?

7. Suppliers

17) Voor zover het bekend is; maken de leveranciers gebruik van Big Data?

8. Employees and Users

18) Is er bij de medewerkers van de inkoopafdeling kennis over het gebruik en de toepassingen van Big Data

19) Is er een bereidheid onder de betrokken medewerkers om het inkoopproces te veranderen (naar het gebruik van Big Data)?

Maturity model

20) Bent u bekend met het gebruik van een maturity model?

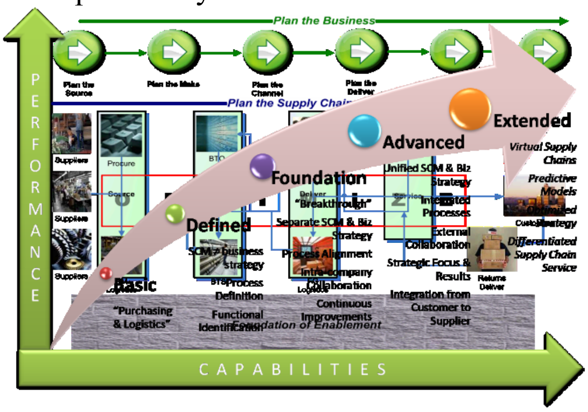
Kleine uitleg wat een Purchasing Maturity model en een Big Data Maturity model is en hoe je het model gebruikt. Van beide modellen een voorbeeld laten zien.

21) Kijkend naar het inkoop maturity model en na het beantwoorden van de voorgaande vragen, waar zou zelf uw organisatie plaatsen?
22) Kijkend naar het Big Data maturity model en na het beantwoorden van de voorgaande vragen, waar zou zelf uw organisatie plaatsen?

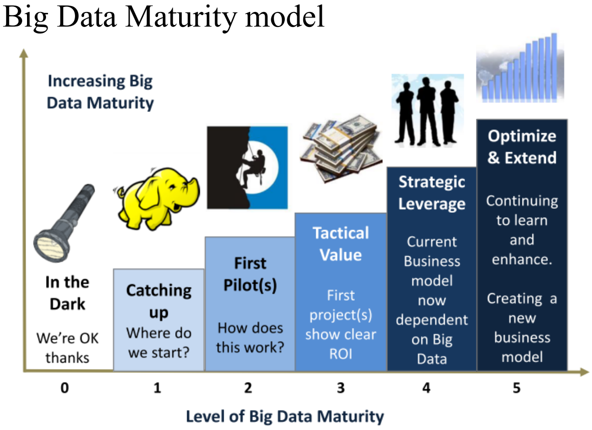
Afsluitend

23) Nu er redelijk wat onderwerpen zijn aangestipt, hoe zou u uw ideale inkoopproces omschrijven als alles mogelijk is?
24) Zijn er nog andere zaken die betrekking hebben met inkoop en Big Data, die nog niet aan bod gekomen zijn, maar die wel nog van toepassing zouden kunnen zijn?
25) Heeft u nog vragen of andere aanvullende opmerkingen?

Inkoop Maturity Model



Big Data Maturity model



VI. Appendix 6: Results interview internal employees

The interview results of the internal employees are confidential, and are not available in this version of the Master Thesis. If there are further questions, please contact l.m.dehaan@student.utwente.nl.

VII. Appendix 7: Results interview external employees

The interview results of the external employees are confidential, and are not available in this version of the Master Thesis. If there are further questions, please contact l.m.dehaan@student.utwente.nl.

VIII. Appendix 8: Clustered overview interview results of the internal employees

The clustered overview of the interview results from the internal employees is confidential, and are not available in this version of the Master Thesis. If there are further questions, please contact l.m.dehaan@student.utwente.nl.

IX. Appendix 9: Clustered overview interview results of the external employees

The clustered overview of the interview results from the external employees is confidential, and are not available in this version of the Master Thesis. If there are further questions, please contact l.m.dehaan@student.utwente.nl.

X. Appendix 10: Scorecard for the Big Data Purchasing maturity model

Big Data Purchasing Maturity model - Scorecard									
						Stage 1	**Stage 2**	**Stage 3**	**Stage 4**
						The purchasing processes are well defined following the best practices of Industry 3.0. There are no (Big) Data applications integrated.	The purchasing processes are standardised and digitalised. There are the first applications of Big Data and there is one person assigned to perform the task.	The Big Data is fully integrated into the purchasing processes, and are cross-functional integrated through the company.	The Big Data processes are fully autonomous organised within the strategic purchasing department. The systems and processes are self-learning and continuously improving.
	Purchasing Element	**Question for analysis**	**% observed**	**Points**		**0-25% = 0- 5 points**	**26-50% = 6-10 points**	**51-75% = 11-15 points**	**76-100% = 16-20 points**
ST	**1. Strategy**								
ST1	Digitisation strategy	Are there Big Data applications involved in the corporate strategy? Is the digitization strategy involved in the corporate strategy?				The corporate strategy does not involve digitisation related integrations. There is low attention to digital techniques.	The first digital solutions and Big Data applications are involved in the different departments. The digitisation strategy is partly described.	Big Data is fully integrated in the cross-functional processes. Persons are responsible for the implementation of the strategy, which is sometimes updated.	There is a clearly defined vision for implementing the digitisation strategy into the company. The strategy is updated, adjusted to the company and aligned with the purchasing function.
ST2	Purchasing strategy	Is the purchasing function involved in the corporate strategy? Is purchasing involved in the C-level of the organisation?				The purchasing strategy is implemented in the purchasing department and does not involve any data related integrations.	The purchasing department makes strategic decisions with the first data driven integrations. Purchasing is not involved in the strategy of the company.	The purchasing strategy has strategic functions within the company and Big Data is fully integrated. People are responsible to align the strategic decisions with the strategy of the company.	There is a clearly defined vision for implementing the purchasing strategy into the C-level of the company. The strategy is updated, adjusted to the company and aligned with the digitisation strategy.
ST3	Corporate strategy	Is there a roadmap for development structured? Is there a long-term vision in the				There is an outdated roadmap, designed with a long-term view. There	There is a roadmap, designed with a long-term view but not any clearly	There is a structured roadmap, designed with a long-term view and	There is a structured roadmap, designed with a long-term view and

		strategy? Is the strategy up-to-date or in which time frame will it be updated?				is not an up to date strategy	defined vision. The strategy is not updated.	clearly defined vision. The strategy is sometimes updated.	clearly defined vision. The strategy is periodically updated.
PS	**2. Processes & Systems**								
PS1	Standardisation	To which degree are the purchasing processes standardised? Which tasks are digitally automated or could be automated in the near future? How are Big Data influencing standardisation the planning and processes?				The purchasing processes are well defined, but not consequently accomplished. There are not any digital automated processes or Big Data integrations.	The processes are digitally connected to each other, and the first applications of Big Data are integrated. There is an exchange of data through the systems. The connection is automated, but are executed with human involvements.	The purchasing processes are digitally connected to each other, and cross-functional through the company. Big Data is integrated within the processes. People are responsible for the quality and quantity of the digital send information.	The tactical, operational and strategic purchasing processes are integrated in the autonomous flow and there is a continuing monitoring system for controlling. The Big Data integration are self-learning and continuous and autonomous improving.
PS2	E-procurement	Is there a distinction in e-procurement for leverage, bottlenecks, routine and strategic goods? Is machine-to-machine communication used in addition to e-procurement?				There is not an e-procurement software integrated, even digitisation is not in the focus of the purchasing department. There are not any connected systems added to e-procurement.	The e-procurement software is involved in the processes and able to manually execute. There is not any distinction between the different goods. There are the first integrations between machines in addition to e-procurement.	E-procurement is standardised for leverage, routine and bottleneck purchase orders. Strategic orders have a higher rate of control. There is an autonomous connection between e-procurement and machine-to-machine communication.	E-procurement is standardised and specialised for the different goods. It is executed automatically. The machine-to-machine communication is fully integrated and automatic in the e-procurement processes.
PS3	Systems	To what degree are the systems totally autonomous and integrated, and is human involvement erased? How are the systems and processes continuing controlled and monitored?				The systems are at least integrated within the own department. The processes are only controlled, monitored and improved by human intervention.	The systems are fully integrated, but have to be accomplished by human interaction. There are the first data driven integrations in the systems, but are not able to execute digitally.	The systems are fully integrated cross-functional in the company and are completely data driven. The human interaction is only as controlling part involved.	The systems are fully integrated, autonomous and have a predictive system, where all the human involved tasked are taken over by systems. The processes are continuously

									monitored, improving and are self-learning.
PS4	Big Data analytics	To which degree results Big Data in a trustful and predictive knowledge? How results Big Data analytics to reduced errors, problems or bottlenecks? Are there any business analytics (e.g. data mining, machine learning or processmining) involved in the decision support system?				There are not any (Big) data related applications, and there is a lack of knowledge about (Big) data analytics. The biggest bottlenecks are detected and reduced by human involvement during the processes.	The first Big Data applications are integrated into the processes and systems. They are manually executed by humans. The decision support system is based on human thinking.	The Big Data applications are fully integrated and are automatically organised into the processes and are able to create predictive knowledge. The Big Data knowledge can make adjustments to bottlenecks and errors.	Due to the self-learning capabilities of Big Data analytics, the predictive demand is trustful and continuously improving in order to reduce errors. The analytics are authorised to make autonomous decisions with a high impact.
PL	**3. Physical level**								
PL1	System integration	To which degree is there a real-time analysis? How far are the systems advanced and able to make predictive decisions?				There is not a data driven analysis, only analysed by humans. There are not any predictive decisions based on data, even there is no data integrated.	There is a partial analysis, driven by the first Big Data integrations. There are not any predictive decisions based on data.	There is a continuous analysis of the data, while it is not real-time driven. The systems are able to translate the knowledge into predictive decisions.	There is a continuous and real-time analysis of the data out of the systems. The systems make autonomous predictive decisions.
PL2	Integration of techniques	Is there an automotive connection between different systems (e.g. machine-to-machine, Internet of Things and/or cyber-physical things)?				There are not any integrations between any systems, only partly within the purchasing department.	The systems are connected with the first integrations of machine-to-machine communication, but there is no connection yet between the cyber and physical components. Stationary sensors recognise demands for goods. Ordering happens still manually.	The systems are fully integrated and the demand is recognised and communicated through machine to-machine communication. There is a connection between the cyber and physical components to exchange data and make a predictive demand.	The demand is generated and supported by both machine-to-machine communication and cyber-physical things. Real and virtual systems are seamlessly integrated. The ordering process is automatically executed.
PL3	Automatic ordering system	Is purchasing and ordering process supported by digital				The workflow for ordering and executing	The purchasing ordering process are involved with	The ordering process is fully integrated with	Automatic ordering takes place autonomously and

		techniques? Are there suppliers automated selected or is it a manual process? To which degree is the ordering process automatically executed after analysis a demand?				the purchase orders is a structured and manual process. The suppliers are selected by human involvements.	the first digital techniques, and are controlled and executed by humans. The suppliers are digitally selected with human involvement.	digital techniques. The orders are executed automatically. The suppliers are selected in the ordering flow.	is based on digital techniques. The demand is forecast driven and the supplier is automated selected based on the settled criteria.
PP	**4. Purchase-to-Pay (P2P)**								
PP1	Connecting systems	Is there a seamless connection between systems, outside the financial department? Is there an automated process flow organised, or possible to implement in the near future? Are there any cyber-physical connections involved to handle complex contracts?				There systems are well working but are not connected with different departments. Possible plans for connections may exist, there is not yet an automated process flow. There are no cyber-physical connections involved.	There is a connection from the finance department with other departments, although they are manually connected. There is not yet a connection between the cyber and the physical involved, and it is all manually executed.	The systems are seamless connected and automated between different departments. There are any autonomous cyber-physical connections integrated.	The process flow is fully autonomous and seamless connected between different departments. In the process flow, there are cyber-physical connections implemented and they are autonomous organised.
PP2	Ordering	Is buying from suppliers outside existing contracts or procedures, so-called maverick buying, delimited in the systems? To what degree is historical or Big Data influencing the ordering process? Is there any real-time tracking of data and how is this influencing further processes?				The systems are not designed to decline outside ordering, therefore it is easily to create maverick buying within the systems. This process is prohibited as a company policy. The historical data is manually involved in the ordering process.	Maverick buying is restricted by the used e-procurement system but employees can still overrule the system. The first Big Data integrations allow to make analysis out of the historical data, and can implement it manually into a new ordering process.	People are responsible to prevent maverick buying and receive a notification when an employee tries to overrule the system. There is an analysis of data, which is not real-time, and is used to influence the new ordering process. The analysis is autonomous integrated.	Due to a seamless and automated machine-to-machine communication it is impossible to order goods outside the current procedures. The ordering process is fully automated without human intervention. Big Data analytics is real-time used and is implemented as a self-learning system to adjust outgoing orders.
PP3	Monitoring	Is the processing of incoming invoices and payments, and the checking whether orders meets their agreed conditions automated? Is the derived				The monitoring processes are not automated within the organisation, or pilot projects have started.	Monitoring processes are (partly) automated but problems have to be solved manually. The	Monitoring processes are (partly) automated but problems have to be solved manually. The derived information is	Monitoring processes are fully automated and the system itself is capable to solve problems, except for very complicated

		information shared internally and with supply partners, and included in future purchases?				The derived information is not shared internally, but stays within the purchasing department.	derived information is manual shared internally.	manual shared both internally and externally.	problems. The derived information is automatically saved and shared both internally and externally.
CK	**5. Controlling/ KPI**								
CK1	Monitoring	How is the Big Data used for monitoring processes? To which extend is the control visible for the employees?				The monitoring process is done by humans and could be visible to the employees.	The monitoring process is integrated with the first Big Data applications. The monitoring process is mostly manual and supported by data.	The monitoring process is advanced by Big Data analytics and is an autonomous process, which is involved with human interaction.	The monitoring process is advanced by Big Data analytics and is an autonomous process without explicit involvement of human interaction.
CK2	Transparency	To which extend is there transparency during the processes and is the information always, anywhere, anytime up to date and complete available?				There is no up-to-date database available, the data is handled manually and is available at the purchasing department, in paper or in digital version.	Due to the first Big Data integrations, the data is digital available at the purchasing department. The data is not up-to-date, because it is all manually digitalised.	The data is always and anywhere available due to the cross functional integrations. There is not a real-time analysis of the data, so it is not always up-to-date.	There is always, anywhere and anytime up to date and complete information available. There is a complete transparency in the data, with a strong cyber safety.
CK3	Big Data analytics	How advanced are the analytical and self-learning capabilities of the firm? To which degree does monitoring and data processing take place? Is a predictive decision support driven on Big Data implemented yet, or ready to do in near future?				The self-learning capabilities are based on the capabilities of the employees. If there is a monitoring process, then it is done manually. There are not any data related integrations.	Due to the first Big Data related integrations, the data is manual extracted with limited analytic capabilities. The monitoring process is done manually and there is not yet a data driven predictive demand.	The Big Data is fully integrated, and predictive analytics are real time assisted by decision support systems. Self-learning algorithms are not yet implemented. The monitoring process is autonomous.	The self-learning and improving capabilities of the Big Data analytics are real-time assisted by decision support systems. The algorithms are data driven. The monitoring process is fully autonomous and automatically executed.
CK4	Data security	How are data and services in digital systems protected against misuse (e.g. fraud, misuse, and/or errors)? Is there an internal				The firm has little knowledge on cyber-security. There is no collaboration with	The organisation sometimes takes part in meetings with partners to discuss cyber security and	The organisation regularly takes part in meetings with partners to discuss cyber security	Data and services in digital systems are strictly protected by the firm and they provide openness on

		evaluation procedure or feedback for the employees to test their security level? How is the privacy of data guaranteed (e.g. with usage analytics or confidential information)?				partners to face this challenge and there is no internal evaluation procedure.	how to ensure protection. There is sometimes an own internal evaluation and feedback session for employees.	and how to ensure protection. People in the organisation are responsible to achieve cyber security targets. There are regularly evaluation and feedback sessions.	how this is done. Cyber safety is regularly checked and guaranteed by independent organisations for both systems and employees.
SO	**6. Sourcing**								
SO1	Big Data integration	To which extend is e-procurement integrated in the purchasing process? To which degree is sourcing involved with a predictive demand? How are the demand-supply predicted and evaluated, and is there Big Data involved?				E-procurement systems are known, but are not adopted by the company. There is not any predictive demand used in the sourcing processes.	The firm has adopted an e-procurement system but the system handles less than 75% of the total purchasing volume. The predictive demand is partly based on data and partly by humans.	E-procurement systems handle more than 75% of the total purchasing volume. Th predictive demand is data-driven.	E-procurement systems handle more than 95% of the total purchasing volume. The sourcing demand predicted and driven on Big Data The data during the process is real-time and up-to-date.
SO2	Specification	Is there in the specification phase an integration of a connected e-sourcing system? To which degree is market analysis (e.g. the identification of new suppliers, goods or services) automated? Is the determination of requirements supported by data analytics?				There is not an e-sourcing system yet. The market analysis and requirements are done manually and are based on own experiences and earlier purchases.	There is an e-sourcing system and is used manually in the specification phase. The market analysis is partly driven on the first Big Data analytics. The requirements are mostly based on earlier purchases.	The e-sourcing is fully integrated and is an autonomous process. The market analysis is data driven and is manual implemented in the purchase order. The requirements are mostly data-driven.	The external market information is autonomous and up-to-date integrated in connecting e-sourcing systems. The requirements are supported by integrated Big Data analytics.
SO3	Selecting	To what degree is knowledge from the supplier database up-to-date and involved in the selecting phase? Is the selecting-phase an autonomous process?				The supplier data base is not up to date or is involved in the selection process. The selecting phase is not an autonomous process and it is done manually.	The supplier database is digitalised and is not totally up to date. The selecting phase is digitalised but needs human involvement for executing.	The supplier database is mostly up to date and involved in the supplier selection phase. It is fully digitalised and it is an autonomous process.	The supplier selection is based on integrated Big Data analyses and the selection is an autonomous process. Due to data integration, the supplier database is always up-to-date with all

									the relevant supplier information.
SO4	Contracting	To which degree is the contracting phase automated? Is the process ready to erase human involvement in the contracting phase, or is it already executed?				The contracting processes are not automated nor there is a connecting between departments. The process is not ready yet to erase human involvement.	The contracting processes are connected but needs to be executed manually. The process is in an early phase to erase human involvement.	The contracting processes are connected and autonomous. The processes are ready to erase human involvement.	The contracting phase is autonomous organised in connecting systems, and the involvement of humans is erased to a more monitoring position.
SU	**7. Suppliers**								
SU1	Data exchange	Which data is shared with the suppliers and is this a seamless data exchange? To which degree feels the buyer the responsibility to develop the knowledge of the supplier?				Sharing of data in relation to purchasing details with the suppliers occurs regularly. There is less responsibility for data exchange in for increasing developments.	Data is regularly shared with suppliers on the basis of mutual benefits, which are agreed beforehand. The responsibility to data exchange for development occurs on a voluntary base.	There is a seamless and extensive sharing of purposeful (Big) data between buyer and suppliers. Both buyer and supplier feels responsible to exchange knowledge for developments.	The exchange of data is extensive, transparent and goes autonomous through the systems. There are many responsibilities for improvements in knowledge as result of data exchange.
SU2	Early supplier involvement	Are the suppliers willing to collaborate? To which degree are the processes designed and suitable for early supplier involvement (ESI)? Does the organisation have a concept for digital integration with suppliers or an integrated supply chain?				The organisation is cautious towards digital integration with suppliers and/or integrated supply chains. The processes are not able to have ESI. There is no concept yet for a digital integration.	The organisation is willing towards digital integration with suppliers and/or integrated supply chains. Meetings to discuss digital cooperation have started. The systems are able to implement ESI. The company is in the early phase for an integration concept.	Close digital collaboration with suppliers. There are no integrated supply chains yet. The first steps for ESI integrations in the systems are designed by the company. The company has a concept for digital integration.	There is a close collaboration with suppliers and are early involved in the processes. This is realised through integrated supply chains where all parties are full involved. The company has an extensive concept for digital integration.
SU3	Predictive demand	Is there a predictive demand communicated with or to a supplier? To which extend is Big				There is not a predictive demand manually created nor	The predictive demand is driven on the first Big Data applications, and is executed manually.	The predictive demand is autonomous created on the Big Data analytics and is	The predictive demand goes autonomous through the systems from buyer to supplier. The demand is

		Data analytics involved in the predictive demand?				communicated to the supplier.	Mostly the predictive demand is used for own predictions.	manually send to the suppliers	based on Big Data integrations and is always up-to-date.
SU4	Supply chain	How do the supply chain disruptions or risks detect and/or prevent?				There is a low involvement in detection or preventing supply chain disruptions.	There is limited view in detecting supply chain disruptions. The first steps for preventing and detecting are taken.	Based on the fully integrated Big Data analytics, it is possible to detect disruptions in an early phase or even prevent them for the future.	The disruptions in the supply chain are detect and prevent by integrated real-time Big Data analytics
EU	**8. Employees/ Users**								
EU1	Involvement	To what extend are the employees involved in the digitalisation of the purchasing function and adopting Big Data techniques?				Employees are sometimes invited to team meetings from the purchasing department, and receive occasionally new information. Employees are not involved in adopting Big Data techniques and receive seldom new information.	Employees are regularly invited to team meetings for purchasing and sometimes for Big Data. Sharing knowledge on digitisation between different departments is stimulated.	Employees are regularly invited to team meetings for both purchasing and Big Data. Employees are responsible for systematically sharing knowledge on digitisation cross-functional through the company.	Employees are structurally invited in team meetings and have direct influence on the development in the digitisation of the purchasing function and integration of Big Data techniques.
EU2	Learning environment	Is there learning and development stimulated by management? To which degree is there a stimulation for continuous improvements and knowledge updating by the employees? Are there feedback and evaluation sessions periodically?				Employees have the opportunity to follow purchasing related courses in their free time. There are occasionally feedback and evaluation sessions.	Employees are stimulated to follow courses in relation to purchasing and Big Data. The feedback and evaluation sessions are sometimes scheduled.	The organisation actively encourages employees to follow courses linked to purchasing and Big Data and the employees are responsible by themselves to reach specific goals. The feedback and evaluation sessions are periodically scheduled.	Learning and development is not only actively stimulated by management, but have also regular assessments which take place to ensure the knowledge of employees keeps up with current developments. This is evaluated in periodically scheduled feedback sessions.

EU3	Capacity	To which extend are the employees capable to understand and execute within the systems in a digital environment?				The employees understand the purchasing processes and are capable to execute it in the systems. There is not any data related knowledge involved.	The employees deeply understand the purchasing function and are capable to understand systems in the digital environment.	The employees deeply understand the purchasing function and Big Data applications.	The employees have all the knowledge and understands the processes behind the systems. They are fully able to execute all purchasing and Big Data related processes.