Smart Industry Adoption: how analytics affect SMEs in the Netherlands

Author: Regina Na-Young Chang
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands

ABSTRACT,
Over the past years, "big data" has received a great deal of attention and was found extremely important for companies as it gave imperative insights that could, for example, help gain better understanding of customer needs. As the need for and recognition of understanding big data grew, several frameworks were developed to detect the maturity level of analytics and indicate improvements of sophisticating current analytic tools. Even though SMEs are important drivers for the economy and technological change, it became a question whether these maturity level of analytics framework aided them in navigating the right analytic direction. This paper aims to cognise "big data" and the maturity level of analytics, to observe and provide arguments on whether the current maturity level analytics frameworks can provide assistance to SMEs through data collection.

Graduation Committee members:
1st Supervisor: Dr. A.B.J.M. Wijnhoven
2nd Supervisor: Dr. R.P.A. Loohuis, MBA

Keywords
Smart Industry, Big data, Analytics, Maturity levels, SME, Industry 4.0

11th IBA Bachelor Thesis Conference, July 10th, 2018, Enschede, The Netherlands.

1. INTRODUCTION

Smart Industry is a development that is influenced by networks, latest manufacturing technologies, information and digitisations that improve and increase several factors such as quality, flexibility, automation, the participation of the value chain and importantly, enhance interaction with customers. It is aided by a network-centric approach, making use of information and appreciate its value, directed by ICT and upcoming manufacturing techniques. Smart Industry Adoption also includes Industry, Internet of Things, Analytics and "Big Data", they are all interrelated. These technological developments introduce what has been called the "smart factory," in which cyber-physical systems monitor the physical processes of the factory and make decentralised decisions (VID/VDE, 2015). The physical processes systems in Industry 4.0 is in direct relation to the Internet of Things which is the concept that is "basically connecting any device with an on and off switch to the Internet (and/or to each other)" (Morgan, 2014). Analytics is a comprehensive and complex field that 'involves statistical analysis, computational linguistics, and machine learning' to find meaningful patterns and knowledge in recorded data (Gandomi & Haider, 2015). It is used to understand "big data" which is a large amount of data that overwhelms business on a daily basis. However, quantity is not of main importance but rather how business interacts with this data.

2. BACKGROUND OF SMART INDUSTRY

As Zheng et al., (2018) point out, "These technologies are permeating the manufacturing industry and make it smart and capable of addressing current challenges" (p. 1). Typically, the characteristics of these challenges in Smart Industry is having flexible production capacity in terms of products (specifications, quality and design), volume, timing and resources and cost efficiencies. Smart industry makes it possible, to enhance the competitive advantage of the organisation. Through research, several frameworks of how Smart Industry affect businesses were found. They seem to have in common that there are three layers (see figure 1, 2 & 3 in the appendix) of the value chain that interacts with each other. These three layers are (1) Cloud systems, (2) Industrial Networks (3) and Physical Resources. There is a recurring theme that proves that these dimensions constitute the impact Smart Industry has on businesses. Through cooperation with each other, the layers align their own behaviours in order to approach a common goal within the system. For example, there is a demand for a customised product. The big data analytics block knows this and transfers its knowledge towards the coordination technologies. The coordination technologies then help the manufacturing assets to interpret this data. The manufacturing assets, in turn, try to achieve (for example) maximum efficiency & quality. With feedback data, the loop transfers back to analytics which tries to optimise the data again. (B. Chen et al., 2018; Lee, Kao, & Yang, 2014; Wang, Wan, Zhang, Li, & Zhang, 2016)

2.1 Layers of Smart Industry

2.1.1 Cloud Systems

Cloud systems are the big data analytics assisting downstream self-organisation and upstream supervisory control and should be capable to analyse the semantics of various data (Wang et al., 2016). Thus, there are great similarities between the cloud systems and analytics, as it is the data forecasting analysis that involves using historical data to research potential future scenarios and help companies gain knowledge to make improvements/changes. In addition, it analyses the effects of a certain decision that could potentially harm or benefit the company. Additionally, data mining could be used to ensure design optimization and active maintenance. This dimension is the catalyst of the other two dimensions.

2.1.2 Industrial Networks

Industrial Networks is the backbone of the systems architecture, providing efficient data exchange and controllability. Technologies in this layer ensure reliable communication and cooperation among equipment (B. Chen et al., 2018). Therefore, this dimension includes all the technologies that assist the manufacturing dimension. Some examples are toolkits, sensors, 3D Printing, drones, advanced robotics, etc. Besides, it is worth mentioning the Intelligent transportation systems. It is an application of "sensing, analysis, control and communication technologies" that can increase transport safety, mobility and efficiency and has a notable effect on transportation in applications such as traveller information systems and traffic signal control (Rouse, n.d.). Another coordinating technology is value chain control towers. They monitor, manage, and control decision-making across several functions and even companies. All these technologies have importance in Smart Industry. We will call these technologies coordination technologies. They improve the value chain and manufacturing process efficiency via support of analytics to help create a more profitable value chain.

2.1.3 Physical Resources

Physical resources are the assets that organisations use in the process of manufacturing. Manufacturing is the process of transforming raw materials and components into an end product that fulfils the specifications and needs of the customer. The process usually consists of many assets in a large-scale operation. With the use of Coordination technologies fed by Analytics, the process could gain in terms of quality, turnover time, cost efficiencies and many more dimensions. Therefore, it is important that the physical equipment gets real-time data and the devices in the other layers have to provide it with 'heterogeneous and high-speed information'. (B. Chen et al., 2018).

These three layers are, as you can see, different, but do constitute one cyber-physical system. Also, these three layers will all be measured in a different way as can be seen with the operationalised variables in the methodology part. In figure 1.1 we see after researching for literature and reviewing the frameworks, it was agreed to create our own framework influenced by the information found:

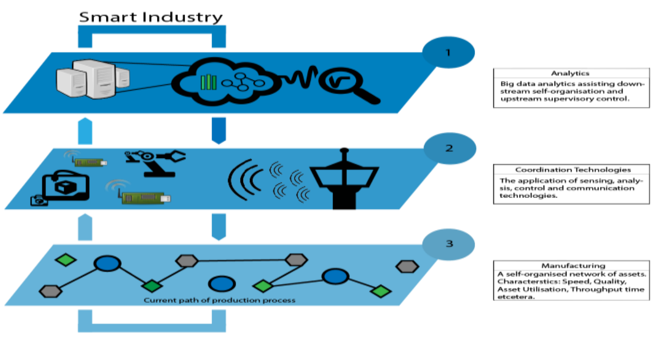

Figure 1: Smart Industry Cyber-Physical System

Large organisations may have the resources to incorporate such a framework and the separate dimensions successfully. However, according to the Chamber of Commerce, only 15% of the SMEs in The Netherlands has heard of the term 'Smart Industry' or 'Industry 4.0' whilst most of the organisations in the Netherlands are labelled as an SME; up until 250 employees (Smetsers, 2016). Moreover, on the technical level, competencies in ICT, automation technology & electronic and software engineering are needed to develop or successfully use cyber-physical systems. On the organisational and business level the development of employee training and new business models are especially demanding for SMEs (Jäger, Schöllhammer, Lickefett, & Bauernhansl, 2016). Therefore, given the fact that SMEs are important drivers for the economy and technological change, the central research question will focus on this group. Moreover, it will encompass three different bachelor theses and will work towards multiple research goals. Firstly, the research aims at broadening the understanding of Smart Industry and its three dimensions. With understanding we mean what kind of impact Smart Industry currently has on organisations. Secondly, the research targets to lessen the gap between expectations and the reality of Smart Industry. Thirdly, it is intended to observe the SMEs industries' current state of adoption Smart Industry.

3. THEORETICAL FRAMEWORK

In this section of the research, it will go into depth of the first layer of Smart Industry Cyber-Physical System. It analyses the relevant literature on the different theories used in this research. Firstly, to further understand this research paper, it will discuss about what big data is and the ongoing importance. Secondly, the literature on the evolution of analytics will be reviewed and highlight how data and analytics go hand-in-hand. Lastly, the maturity of level frameworks by R. Grossman, W. Eckerson and Waston et. al will elaborated along with its main factors to identify the maturity levels.

3.1 Big data

In the recent years, from academics to corporations, big data are the subject of attention and even to the extent, fear. It has quickly evolved into a "hotspot" that entices attention from academia, industry, and even governments (Jin, Wah, Cheng, & Wang, 2015). Many were left unprepared as big data suddenly rose. Before, the developments of any "new" technology appeared first in technical and academic journals. Later on, the knowledge and production were published into other ways of knowledge mobilization (including books). The swift evolution of big data technologies and the acceptance rate by the private and public sectors left little time for the converse to develop and mature in the academic domain (Gandomi & Haider, 2015). Even McKinsey, the renowned consulting firm, claimed that the big data has grasped into every area of the industry and business functions and has become a crucial factor in production (James Manyika, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, 2011). Using and mining data prefigures a new wave of consumer motivation and growth of productivity.

What is "big data"? It seems to entice and scare people in the private and public sectors from every area. There is no universally accepted definition (Jin et al., 2015). For IBM, big data is "a term applied to data sets whose size or type is beyond the ability of traditional relational databases to capture, manage, and process the data with low-latency" (IBM, 2018). Another definition of big data, given by TechAmerica Foundation, it "describes large volumes of high velocity, complex and variable data that require advanced techniques and technologies to enable the capture, storage, distribution, management, and analysis of the information" (Agarwal et al., 2012). To understand how difficult it was and is to pinpoint the definition, please refer to figure 2 as there was an online survey conducted by SAP that showed how different executives had a different understanding of what big data is (definitions ranging to what its functions are to what big data is) (Interactive, 2012).

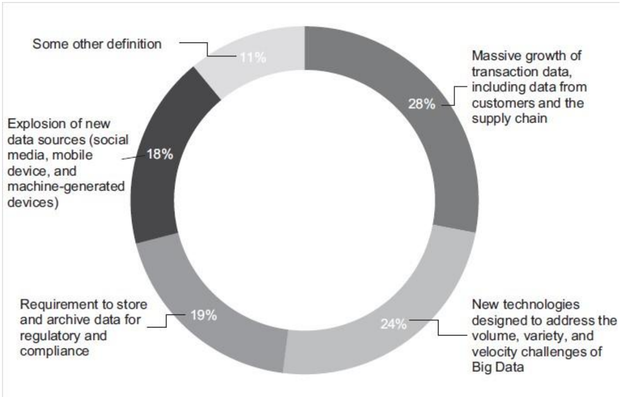

Figure 2: Definitions of big data based on an online survey of 154 global executives in April 2012

Size is a significant characteristic when concerning the question "what big data is?" Over the years, other characteristics of big data have surfaced and these characteristics are Volume, Variety and Velocity (the Three Vs) (Laney, Management, & Volume, 2005) and thee Three V's have been used as a common framework to describe big data. First, Volume refers to the scale of the data and is stated in multiple terabytes and petabytes. The term of big data volumes is comparative and vary due to factors such as time and type of data. Because it is accepted as big data in present day does not mean that it will meet the threshold in the future as the storage capabilities will increase, enabling even bigger data sets to be obtained (Gandomi & Haider, 2015). It is worth mentioning that the type of data and the type of industry influence the definitions thus, it is unrealistic to determine a specific threshold for big data volume. Second, Variety is about the structural heterogeneity in a data set and as technologies becomes advance, it allows firms to use various types of structured, semi-structured and unstructured data (Gandomi & Haider, 2015). Structured data, which establishes only 5% of all existing data (Cukier, 2010) and this data can be found in spreadsheets or relational databases. Examples of unstructured data are images, text, audio and video and unstructured data denotes to the lack of organisation required by machines for analysis. Semi-structured data is the in-between of the two data structures mentioned above, the format does not follow the strict criterions. Organisations have been hoarding unstructured data from internal sources (e.g. historical data) and external sources (e.g. social media). The emergence of new innovative data management technologies and analytics have allowed organisations to control the data in their business processes e.g. facial recognition has allowed retailers to obtain data about the age, gender and movement patterns of their customers which can aid in decisions related to product promotions and placements (Gandomi & Haider, 2015). Lastly, Velocity regards to the rate at which data are generated and the speed it takes for the data to be analysed. As digital devices such smartphones and sensors evolve and are used everywhere, the rate of data creation is unparalleled and the need for real-time analytics and evidence-based planning grows (Laney et al., 2005). It does not have to be the unconventional to conjure data as such, even usual retailers generated high-frequency data, for example, Wal-Mart (an American multinational retail corporation) process more than one million transactions per hour (Cukier, 2010). This type of data, from smartphones and smartphone apps that produces information that can be used to generate real-time, personalised offers, can provides information about the customers that can be analysed to create real customer value.

Besides the three V's, three other dimensions of big data have been discovered and they are Veracity, Variability and Value.

Veracity was pegged as the "fourth" V by IBM (IBM et al., 2018) and this characteristic represents the unreliability in some sources of data e.g. customer sentiments in social media as they are biased due to human judgement. This can be addressed by using tools and analytics developed for management and mining of uncertain data. Variability (and complexity) was the additional characteristics to be discovered by SAS (SAS, 2018b), variability represents the variation in the data flow rates as the velocity of big data is not always consistent and have periodic peaks and dips. The complexity signifies that big data are generated through numerous sources. Thus, this enforces a confrontation: the need to connect, match, cleanse and transform data received from various informants (SAS, 2018b). Value is often used as a defining attribute of big data and that "big data are often characterised by relatively low value density" (Oracle, 2014). This means that the data collected in the original form has a low value relative to its volume but analysing large volumes such data, high value can be gained.

Big data is a by-product of the information era, an era that "has been marked by increasingly pervasive digital technologies that have reconstituted organisational life and action" (Caesarius & Hohenthal, 2018), that is sustained by everyday generation, storage and distribution of ample sets of data, in varied formats, at extreme velocity and with increasing granularity. Due to the abundance of data historically, it can be traced back to the overview of "compatible and interoperable digital mediating technologies" (Kallinikos, 2010). First, the process of digitisation that has been constant for years due to the combination of IT and more advanced database technologies in organisations (Zammuto, Griffith, Majchrzak, Dougherty, & Faraj, 2007). Second, the development and arrival of the Internet when "digitisation entered in a first wave the personal sphere and instigated social and cultural changes" (Kjaerulff, 2010). Third, the introduction of social media "when digitisation entered in a second wave the personal sphere and permeated the everyday lives of people" (Bruns, 2008). Fourth and lastly, the influx of network-connected objects that operate without human involvement (known as Internet of Things) (Caesarius & Hohenthal, 2018). In short, big data is an occurrence that disrupts from traditional ways of dealing with data in organisation and proposes a new way of knowledge production that utilises "computational manipulation of complex data sets with algorithmic accuracy to generate insights that were previously deemed impossible" (Boyd & Crawford, 2012).

3.2 Analytics and its evolution

By definition, according to SAS, analytics is "an encompassing and multidimensional field that uses mathematics, statistics, predictive modelling and machine-learning techniques to find meaningful patterns and knowledge in recorded data" (SAS, 2018a). Analytics is a combination of processes and tools that include statistics, data mining, artificial intelligence and language processing (Russom, 2011). It is also used to large and disperse datasets for gaining invaluable understandings to improve firm decision making (Ertemel, 2015). Analytics is a critical organisational IT competency due to the increased amounts, speed of change and types of data in business over the two decades (Kambatla, Kollias, Kumar, & Grama, 2014). There is a pressure of firms needing to improve their data competency to determine better, more informed and faster decisions. Recent studies show the many firms that invest in data analytics do not take full advantages of using the data analytics tools, even though there is evidence that using data analytics help organisations improve their decision-making performance. A report from Deloitte claims that "only 25% of firms reported that analytics has significantly improved their organisation's outcomes" (Deloitte, 2013) and another report claims that they found "only 27% of firms that invested in data analytics reported their initiatives as successful" (Colas, Nambiar, Finck, Singh, & Buvat, 2014). There are several arguments regarding data analytics and its features. One could argue that organisations need to concentrate on various confrontations to gain the benefits, even if data analytics has immense potentials for improving firms results (Akter & Wamba, 2016) whereas another argues that even if a firm is investing in their data analytics, it does not guarantee that it will improve their decision making and different firm can play critical roles in successfully using these tools (Ghasemaghaei, Hassanein, & Turel, 2017). Furthermore, there is another argument that high failure rate of productively using data analytics because of the obligatory required conditions to generate understandings from analytics are unkempt as most firms solely focus on data aspects such as data volume to understand (Wu, Buyya, & Ramamohanarao, 2016). Therefore, it is necessary to empirically study the characteristics of successful data analytics initiatives.

Over the decades, analytics has evolved tremendously and rapidly. To fully understand analytics and this research paper, this part of the section will explicate about the evolution of analytic technologies and applications that were and is adopted in the industry. There are three groupings of the BI&A (Business Intelligence and Analytics) are BI&A 1.0, BI&A 2.0 and BI&A 3.0 (please refer to figure 3) (Chen, Chiang, & Storey, 2012)

Evolution	Key Characteristics	Period
BI&A 1.0 (DBMS-based, structured content)	• RDBMS & data warehousing • ETL & OLAP • Dashboards & scorecards • Data mining & statistical analysis	Between 1970s until 1990s
BI&A 2.0 (Web-based, unstructured content)	• Information retrieval and extraction • Opinion mining • Question answering • Web analytics and web intelligence • Social media analytics • Social network analysis • Spatial-temporal analysis	Early 2000s
BI&A 3.0 (Mobile and sensor-based content)	• Location-aware analysis • Person-centred analysis • Context-relevant analysis • Mobile visualization & HCI	Decades of 2010s

Figure 3: Evolution of BI&A

Beginning with BI&A 1.0, it is mostly where data are collected by companies through various computer systems and often stored in RDBMS (relational database management systems). The techniques generally used in these systems commercialised during the 1990s were mainly in statistical methods that were established in the 1970s and the data mining was developed in the 1980s. The key foundation of BI&A 1.0 are data management and warehousing. Essential for transfiguring and incorporating enterprise-specific data are data marts and tools for ETL (extraction, transformation and load). Simple but insightful graphics are used to discover important data characters such as database query, OLAP (online analytical processing) and reporting tools. BPM (Business performance management) utilise scorecards and dashboards to analyse and visualise variety of performance metrics. Statistical analysis and data mining techniques are implemented for association analysis, data segmentation and clustering, classification and regression analysis, anomaly detection, and predictive modelling in various business applications (Chen et al., 2012).

BI&A 2.0 began during the early 2000s when the Internet and the Web began to offer different data collection and analytical research and development prospects. Web intelligence, web analytics and the user-generated obtained through Web 2.0-based social and crowd systems have been piloted the BI&A 2.0 research in the 2000s (focusing on text and web analytics for unstructured web contents) (Doan, Ramakrishnan, & Halevy, 2011; O'Reilly, 2007). It was the beginning of organisations were allowed to present their businesses online and interact with their customers directly. For example, the introduction of using delegated IP-specific user search and interaction logs that are collected through cookies and server logs suddenly became an important source in understanding customers' needs and recognising new opportunities e.g. using Google Analytics. The BI&A 2.0 systems, unlike BI&A 1.0, "require the integration of mature and scalable techniques in text mining…, web mining, social network analysis and spatial-temporal analysis with existing DBMS-based BI&A 1.0 systems" (Chen et al., 2012).

While BI&A 2.0 enticed research from the industry and academia, BI&A 3.0 emerged with more research opportunities. As mobile phones, tablets, are becoming more integrated into everyday lives compared to the use of laptops and PCs, it is predicted that "the number of connected devices would reach 10 billion in 2020" (The Economist, 2011). As mobile devices (e.g. smart phones, anything that is capable of downloadable applications) are affecting different parts of society from healthcare to entertainment. The decade of the 2010s foreshadows an era for high impact BI&A research development for the industries and academia (Chen et al., 2012).

Figure 5: Evolution of Analytics

Using figure 4 (Chen et al., 2012) to explain the evolution of analytics. First, data analytics signifies to the BI&A technologies that are founded in data mining and statistical analysis. Most of these methods are "data-driven, relying on various anonymization techniques, while others are process driven, defining how data can be accessed and used" (Gelfand, 2011). Second, text analytics has transpired since the early 1990s when search engines have evolved into mature commercial systems that entailed of fast data from search logs analytics and link-based page ranking. This signifies to the BI&A 1.0 development in the text-based enterprise search and document management systems (Chen et al., 2012). Third, web analytics developed as an active field of research within BI&A and it offers exceptional analytical challenges and opportunities e.g. statistical analysis foundations of data analytics. Associated web search engines and directory systems "helped develop unique Internet based technologies for website crawling/spidering, web page updating, web site ranking and search log analysis" (Chen et al., 2012). Fourth, network analytics is an emerging research area that advanced with contributions from computer scientists and sociologist alike, since the early 2000s, that evolved to include new computational models for online community and social network analysis (Chen et al., 2012). Fifth and final, mobile analytics is another radical research in different areas (e.g. crowd sourcing, personalisation and behavioural modelling for apps) and this method developed due to the increasing number of smart phone and tablet owners globally, as an "an effective channel for reaching many users and as a means of increasing the productivity and efficiency of an organization's workforce" (Chen et al., 2012).

3.3 Maturity level frameworks

3.3.1 The TDWI Maturity Model by W. Eckerson

This maturity model examines the characteristics of mature BI implementations and to give organisations goals and motivation to surmount the obstacles regarding BI implementation. Please refer to figure 5 (Eckerson, 2007) to understand the descriptions of each stage in the model.

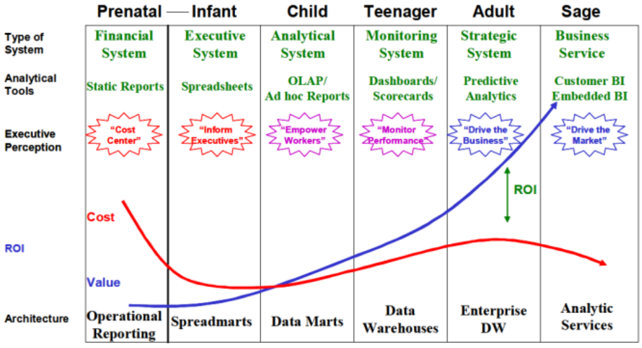

Figure 4 - The TDWI Maturity Model

1) *Prenatal stage*: Established operational reporting system with a standard set of static reports; reports are built into operational systems and limited to that individual system (Eckerson, 2007; Hribar, 2010)
2) *Infant stage*: Faced with numerous partial data sources called "spreadmarts" (spreadsheets/desktop databases used as replacement for regional data) (Eckerson, 2007; Hribar, 2010)
3) *Child stage:* Knowledge workers join community of BI users, companies buy their first interactive tool and capable of analysing trends & past data; it focused on understanding the correlation in data and to gain understanding of past business actions (Eckerson, 2007; Hribar, 2010)
4) *Teenager stage:* Recognising the need and starts to use a standardised set of project and development methodologies (including best practices, learning past experiences and extensive use of external consultants. Software solutions developed on common data model using common consolidation platform and recognises value of consolidating regional DW into centralised DW; use of BI is spread among regular users and enables knowledge workers interactive reporting and analysis (Eckerson, 2007; Hribar, 2010)
5) *Adult stage:* Developing from tactical to strategic business level and BI becomes the central IT system during daily operations IT system driving daily operations of company; there is a generic architecture of data warehouse, fully loaded with data and more (Eckerson, 2007; Hribar, 2010)
6) *Sage stage:* The companies at this level are turning BI system capabilities into technical and business services and are moving development back to basic organisational units through COE (Centres of Excellences) (Eckerson, 2007; Hribar, 2010)

3.3.2 Analytic Processes Maturity Model (APMM) by R. L. Grossman

This maturity model is based off a framework for analytics that divides analytics process into six areas: building analytics models, developing analytic models, managing analytic infra-structure, operating an analytic governance structure, providing security and compliance for analytics assets and lastly, developing an analytic strategy (Grossman, 2018a). This maturity model is divided into five different levels/stages.

1) *AML 1 – Builds report*: "Organization can analyse data, build reports summarizing the data, and make use of the reports to further the goals of the organization" (Grossman, 2018a)

2) *AML2 – Build models*: "Organization can analyse data, build and validate analytic models from the data, and deploy a model" (Grossman, 2018a)
3) *AML 3 – Repeatable analytics*: "Organization follows a repeatable process for building, deploying and updating analytic models. In our experience, a repeatable process usually requires a functioning analytic governance process" (Grossman, 2018a)
4) *AML 4 – Enterprise analytics*: "Organization uses analytics throughout the organization and analytic models in the organization are built with a common infrastructure and process whenever possible, deployed with a common infrastructure and process whenever possible, and the outputs of the analytic models integrated together as required to optimize the goals of the organization as a whole. Analytics across the enterprise are coordinated by an analytic governance structure" (Grossman, 2018a)
5) *AML 5 – Strategy-drive analytics*: "Organization has defined an analytic strategy, has aligned the analytic strategy with the overall strategy of the organization, and uses the analytic strategy to select appropriate analytic opportunities and to develop and implement analytic processes that support the overall vision and mission of the organization" (Grossman, 2018a)

3.3.3 Data Warehouse Stages by Watson et al.

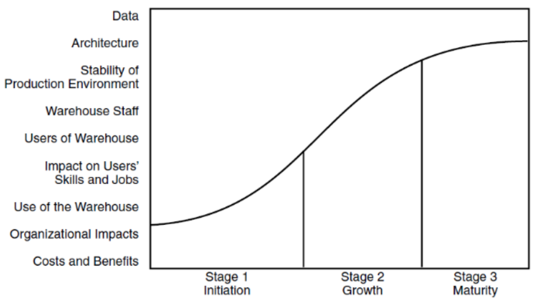

Figure 6: Stages of Growth

This maturity level is built on off nine different variables (data, architecture, stability of production environment, warehouse staff, users of warehouse, impact on users' skills and jobs, use of the warehouse, organisational impacts and cost and benefits) to help describe the different stages.

1) *Initiation*: There is limited amount of data for a single or a few subjects areas; analysts in the business unit are served by the data mart; reports and redefined & ad joc. queries (backward looking to what has occurred) (Watson, Ariyachandra, & Matyska, 2001)
2) *Growth*: There is data for multiple subject areas; users from all of the business are served by the data marts, diverse in the information needs and computer skills; the reports and predefined are more about the analysis of why things occurred and "what if" analyse for future scenarios (Watson et al., 2001)
3) *Maturity*: There is data for enterprise -wide (well-integrated and for multiple time periods); users from the whole organisation can access to the warehouse data; there are reports, predefined queries, DSS and EIS, data mining provides predictive modelling capabilities; integration with operational systems (Watson et al., 2001)

4. RESEARCH QUESTION

4.1 Overall research question

Overall, as a circle group, our research question is:

"What is the impact of Smart Industry technologies on SMEs in the Netherlands?"

4.2 Specific Research question

From the overall research question, we each specified our own question that will help answer the question above:

"To what extent are current frameworks of maturity levels of analytics suitable to SMEs?

And the following sub-questions are:

"How to identify what certain aspects could be improved on for SMEs?"

"How SMEs can have a better understanding on how to improve on certain aspects?"

5. METHODOLOGY

This study presents an exploratory research conducted for the purpose of obtaining an understanding on whether the maturity level frameworks are suitable for SMEs to use and the certain aspects that make the difference. Through this research, understanding of the difference between what is anticipated and what is unanticipated from the maturity level frameworks can be gained. The analysis is based on qualitative and quantitative research focusing on the analytics used in SMEs in the Netherlands. The SMEs examined within this study mostly lies on the industrial and manufacturing (depending how successful the data collection due to time constraints).

5.1 Approach

For this research, my main question is "to what extent are the current frameworks of maturity levels of analytics suitable to SMEs?" What is meant with 'suitable' is to understand how SMEs can identify certain improvements aspects that is relevant to them by using the analytic maturity frameworks and to see how they can actually improve their analytics. Thus, to help understand and collect the correct data, we have to first establish the concept of maturity of analytics and the evolution of analytics. First, what is the maturity of analytics? There are several frameworks found that help identifies where companies current state in regards to how they interact and use their data. These frameworks are made by companies such as SAS and researchers like Watson et. al. (Watson, Ariyachandra, & Matyska, 2001, p. 42) and Robert L. Grossman (Grossman, 2018, p. 45) alike. The maturity means about how much a company embraces analytics and how sophisticated the analytics is to gain better revenue. In short, across many literature papers and research, the main idea correlated throughout is that that the more 'mature' the company is in regards to their analytics, the "financial and general support for BI management and business functions have a positive impact on the overall organizational performance" (Lahrmann, Marx, Winter, & Wortmann, 2011, p. 7) and "it can have a transformative effect on the organization of work in the firm" (Caesarius & Hohenthal, 2018, p. 10). Second, what is the evolution of analytics? It is the change in the characteristics of analytics over the last five decades thus, BI&A 1.0 (DBMS-based and structured content), BI&A 2.0 (web-based and unstructured content) and lastly, BI&A 3.0 (mobile and sensor-based content) (Chen et al., 2012). Thus, it is of the essence to find out whether the evolution stages of analytics held accountable for SMEs in the Netherlands.

5.2 Subjects of Study

Concerning the subjects of the study, there are two main group of subjects that can be recognised: the SMEs in the region of Twente and SMEs in the other regions of the Netherlands. The delegates of the firms, such as directors and employees, are anticipated in providing their knowledge of how the firm practices their data, the outcome of their findings and the certain type of analytics used. The possibility is given to collect data as SMEs play an important role in Overijssel's economy and also Twente, as 25% of employment is situated in companies with fewer than ten employees and 60% in companies with fewer than 100 employees (The European Commission, 2018). Due to the lack of time and access, it is not feasible to directly approach all SMEs in the region. Directors and managers were chosen the basis of their relevance to my research subject and their position within the firm. This is because of their years of experience in their position and direct knowledge of analytics and data. It must be noted that their assumption and beliefs can influence the outcomes that can lead to biased results. Other regions of the Netherlands were chosen as the second subject of the study as it is known that the Netherlands has a strong SBA profile (Small Business Act for Europe) compared to other EU countries as it ranks above the average (The European Commission, 2017). It must be noted that the Netherlands in particularly strong in the "entrepreneurship, 'second chance', 'responsive administration' and skills & innovation…in entrepreneurship the country topped all other EU Member States" (The European Commission, 2017).

5.3 Data collection

As we want to see what the impact of Smart Industry is throughout the Netherlands, we have agreed together to send a questionnaire to a variety of companies. This may not be enough and thus; we can still do two in-depth interviews. However, the focus now lies on gathering enough data through the questionnaire in order to make 'Smart Industry in the Netherlands' representative and reliable. The questionnaire is sent to SMEs in the area of Twente and in the Netherlands that have a manufacturing process of some kind. The contacts are gathered by calling the organisations and searching for their e-mail address on their website. To ensure that the data is kept confidential, we only share the results connected with the company names between the three individual researchers. In the results, no names will be used, just the outcomes of the measurements. The questionnaire will specify this and respondents always have the choice throughout the questionnaire not to join the research. Moreover, if the respondents have any questions and remarks, they have the choice to e-mail the researchers and/or the supervisor F. Wijnhoven at any time. To ensure that our questionnaire is representative, we will have to have a certain number of respondents. The number of individuals in the sample size is all that matters. According to de Veaux (2015), "for a questionnaire that tries to find the proportion of the population falling into a category, you'll usually need several hundred respondents to say anything precise enough to be useful (p. 313)". Thus, a larger sample will make our results more precise. Therefore, we target for 100 organisations. As for our questions that we will ask in the questionnaire, "results depend crucially on the questionnaire that scripts this conversation (irrespective of how the conversation is mediated, e.g., by an interviewer or a computer). To minimize response errors, questionnaires should be crafted in accordance with best practices (Krosnick & Presser, 2010)." Especially within the field of Smart Industry with specific terms, it quite difficult to make the questionnaire one hundred per cent understandable. However, to make it as good as possible, when making our questionnaire, we included the following best practices: 1) Avoiding double-barrelled questions. 2) Using simple jargon for words that are difficult. 3) Avoiding questions with single or double negations. 4) Short questions. 5) Avoiding leading or loaded questions that push respondents toward an answer. 6) Ensuring that everyone interprets the words/sentence in the same way. In order to test our methodology in the questionnaire**,** we have created an operationalisation framework (see figure 7) which makes clear how our subjects and the concepts within the subjects align with each other.

In order to measure the two concepts for my research question, the maturity of analytics and evolution of analytics, we need to operationalise the necessary variables. The maturity of levels is from the following features and overall research (Watson, Ariyachandra, & Matyska, 2001; Grossman, 2018; Eckerson, 2007):

1. Data
2. User
3. Application
4. Output
5. Operation

These variables are found as the general, overlapping indicators that help identify which stage each an organisation is at and what certain analytics is used, for example. Using these variable, it can be operationalised to make it easier to be measured and obtained through the questionnaire (please refer to the operationalisation framework to see the operation to see how the variables were operationalised). The evolution of the analytics will be distinguished based on the features mentioned above and literature review. If an organisation is using analytics such as RDBMS, data warehousing and statistical analysis etc., it will be categorised as "initiation stage" (Watson, Ariyachandra, & Matyska, 2001). If an organisation is using analytics such as web analytics, web intelligence and social media analytics etc., it will be categorised as "growth stage" (Watson, Ariyachandra, & Matyska, 2001). Lastly, if an organisation is using location-aware analysis, person-centred analysis etc., it will be

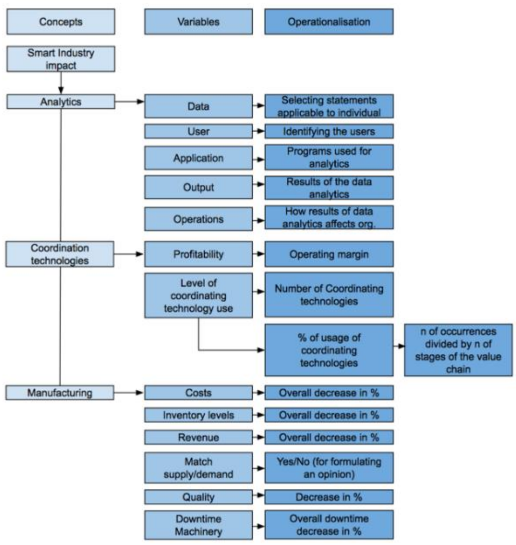

Figure 7: Operationalisation Framework

categorised as "maturity stage" (Watson, Ariyachandra, & Matyska, 2001).

6. DATA ANALYSIS

The following section will describe the outcomes of the data collected from the questionnaires. It must be noted that Microsoft Excel, IBM SPSS Statistic version 25 and Qualtrics were used to analyse the data set (the results can be found in the appendix). Please refer to the appendix for the results of the questionnaires for the exact numbers as they were not referred due to page restraints. Table 1 shows the number of SMEs that participated in this research study and it must be reminded the number of participants is relatively small compared to be what is expected due to various issues (that will be mentioned in another section).

Table 1 - Participants of the questionnaire (N=18)

Company size	Sector	Location Twente	Other province	Occupation
2-50	Industry	No	Friesland	Account manager
2-50	Automotive	Yes		Director Owner
2-50	Industry	No	Friesland	Assistant Director
2-50	Catering	Yes		Project employee
2-50	Industry	No	Zeeland	Office Manager
2-50	Construction	Yes		Owner
2-50	Automotive	Yes		Managing director
2-50	Industry	Yes		Director
2-50	Industry	No	Drenthe	Plant Manager
10 - 50	Industry	Yes		Managing director
10 - 50	Industry	Yes		R&D Manager
10 - 50	Industry	Yes		R&D Manager
10 -50	Industry	Yes		Company director
10 -50	Industry	Yes		Company director
10 -50	Industry	No	Gerland	Director of Operations
101-150	Industry	No	Gerland	Jr. Business Controller
101-150	Industry	No	Gerland	CTO
151-200	Industry	Yes		Director Programs

In order to gain an understanding of the result and to indicate where the SMEs in the Netherlands are places in terms of the three maturity level frameworks, this section will be divided into the following: overall outcome of SMEs and outcome from size of SME.

6.1 Overall outcome of SMEs

Beginning with question 1, this question was about allowing the SME representative to choose the following statements that felt most relevant to them (refer to table 2) and this question was constructed from Grossman's Analytic Processes Maturity Model, with the intent of being able to indicating where the SME stood in this spectrum. Table 2 shows the statements that most of the SMEs selected and this indicates the most of the SMEs fall into the category of AML 1 and AML 2 (Grossman, 2018a).

Table 2 – Results of Question 1 (related to Grossman)

Statements	Percent
Our company can analyse data (1)	20.5%
Our company can build reports summarizing the data (2)	18.2%
Our company can make use of the reports to further the goals of the organisation (3)	13.6%
Our company can use an analytic model (5)	9.1%
Our company uses analytics throughout the company (6)	9.1%

Moving on with question 2 until question 6, these questions were constructed from Eckerson's TDWI Maturity Model and the following questions are stated below (please refer to table 3). Using these questions helped indicate where the overall SMEs belonged on this maturity level framework.

Table 3 - Questions related to Eckerson

Questions	
Question 2	Who uses the analytics in your company?
Question 3	What is your focus on analytics?
Question 4	What is the focus of using analytics?
Question 5	What is the outcome of using analytics?
Question 6	What are the tools used for analytics?

Overall, majority of the results from these questions prove that most of the SMEs fall in the category of Infant and Teenager depending on the certain question and functionalities (Eckerson, 2007).

Table 4 - Result of Question 2

Question 2

		Responses		Percent of Cases
		N	Percent	
Who[a]	1	4	12.1%	23.5%
	2	2	6.1%	11.8%
	3	13	39.4%	76.5%
	4	10	30.3%	58.8%
	6	4	12.1%	23.5%
Total		33	100.0%	194.1%

a. Group

Table 5 – Result of Question 3

Question 3

		Responses		Percent of Cases
		N	Percent	
Your focus[a]	1	12	29.3%	66.7%
	2	2	4.9%	11.1%
	3	9	22.0%	50.0%
	4	8	19.5%	44.4%
	5	4	9.8%	22.2%
	6	6	14.6%	33.3%
Total		41	100.0%	227.8%

a. Group

Table 6 – Result of Question 4

Question 4

		Responses		Percent of Cases
		N	Percent	
Company focus[a]	1	8	18.2%	44.4%
	2	9	20.5%	50.0%
	3	14	31.8%	77.8%
	4	9	20.5%	50.0%
	5	4	9.1%	22.2%
Total		44	100.0%	244.4%

a. Group

Table 7 – Result of Question 5

Question 5

		Responses		Percent of Cases
		N	Percent	
Outcome[a]	1	13	24.5%	72.2%
	2	7	13.2%	38.9%
	3	7	13.2%	38.9%
	4	15	28.3%	83.3%
	5	10	18.9%	55.6%
	6	1	1.9%	5.6%
Total		53	100.0%	294.4%

a. Group

Table 8 - Result of Question 6

Question 6

		Responses N	Responses Percent	Percent of Cases
Tools[a]	1	16	36.4%	88.9%
	2	3	6.8%	16.7%
	3	11	25.0%	61.1%
	4	8	18.2%	44.4%
	5	5	11.4%	27.8%
	6	1	2.3%	5.6%
Total		44	100.0%	244.4%

a. Group

Table 9 - Results of Question 2-6

Questions	Answers	Percent	Result
Question 2	3	39.4%	**Teenager**
Question 3	1	29.3%	**Infant**
Question 4	3	31.8%	**Teenager**
Question 5	4	28.3%	Adult
Question 6	1	36.4%	**Infant**

Ending with the final part with question 7 until 10, these questions, like the other questions, were constructed from Watson et al.'s Data Warehouse stages (refer to table 5. From the results, it can derived that the majority of SMEs are in the category of Initiation and Growth stage (Watson et al., 2001) and from this, please refer to table 6 for the results (for the results of question 10, please refer to the appendix).

Table 10 - Questions related to Watson et al.

Questions	
Question 7	Do you use internal data to improve the processes within the company?
Question 8	Do you use external data to create products specified to the need of one group of people?
Question 9	Do you use data analytics that can help generate useful data information catered towards needs of company?
Question 10	In general, what kind of analytics do you use in your company?

Table 11 - Results of Question 7-9

Questions	Answers	Percent	Result
Question 7	Yes	39.4%	**Teen**
Question 8	No	29.3%	**Infant**
Question 9	Yes	31.8%	**Teen**

Table 12 - Result of Question 7

Question 7

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	15	83.3	83.3	83.3
	No	3	16.7	16.7	100.0
	Total	18	100.0	100.0	

Table 13 - Results of Question 8

Question 8

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	8	44.4	44.4	44.4
	No	10	55.6	55.6	100.0
	Total	18	100.0	100.0	

Table 14 - Result of Question 9

Question 9

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	11	61.1	61.1	61.1
	No	7	38.9	38.9	100.0
	Total	18	100.0	100.0	

6.2 Outcome from size of SME

Interestingly, using SPSS to make custom tables with the size as the independent variable and the other results from the questions as the dependent variable helped understand the bigger picture of the data set. From this data analysis, it showed how size has an impact on the use of analytics. Due to page constraints, please refer to the appendix to see the results of the questionnaire from SPSS for the results compared with size.

First of all, using the results from question 1 and the sizes of the SMEs, table 7 shows an overview of the size of SME and the indication of Grossman's Analytic Process Maturity Model.

Table 12 – Result of Question 1 with Size

Size of SME	Result
2-50	AML 1
10 -50	AML 2
101-150	AML 4
151-200	AML 4

Second, using the results from question 2 until 6 and the sizes of the SMEs, table 8 show an overview sizes of the SMEs and the indication of the Eckerson's TDWI Maturity Model.

Table 13 - Results of Question 2-6 with Size

Size of SME	Result
2-50	Infant/Teenager
10 -50	Teenager
101-150	Teenager/Adult
151-200	Teenager/Adult

Lastly, using the results from question 2 until 6 and the sizes of the SMEs, table 8 show an overview sizes of the SMEs and the indication of the Eckerson's TDWI Maturity Model.

Table 14 - Results of Question 7-9 with Size

Size of SME	Result
2-50	Initiation stage
10 -50	Growth
101-150	Maturity
151-200	Maturity

7. DISCUSSION

As mentioned before in the theoretical framework, there is no universal accepted definition of "big data", nonetheless, the definition established by NIST is relevant for this section: "Big Data refers to digital data volume, velocity and/or variety [veracity] that: enable novel approaches to frontier questions previously inaccessible or impractical using current or conventional methods; and/or exceed the capacity or capability of current or conventional methods and systems" (NIST, 2017). As time goes by, analytic maturity is important for numerous factors when it comes to big data. First, as the variety, velocity and volume of the data develop, the importance of having suitable analytic infrastructure expands (Grossman, 2018b). Second, more multiple models will be probable to be used concerning big data and this is because the number of models will continuously grow and it will become more essential to "have an analytic infrastructure that can build, manage and deploy these models" (Grossman, 2018b). Lastly, once big data becomes recognised to organisation as valuable, models must be built from the big data and can be implemented into "products, services and operations to increase revenues, decrease costs, reduce risk and optimise operations" (Grossman, 2018b). Hence, it is imperative that the more maturity the analytics are, the more likely these entities will happen.

Comparing the results for the overall outcome of question 1 and the results with the size as an independent variable of question 1, there are points worth noting. It must be noted that as the data set is very small, it is infeasible to generalise the whole population thus, it is critically discussed of what could have been the case had the sample size been larger. First, the results of the overall outcome show that most SMEs in the Netherlands could be either on AML 1 and/or AML2. There is subtle difference with AML 1 and AML2, these AML 2 organisations can build models that make predictions about future event rather than summarising past event and comprehend the difference between "(business) rules and analytics and integrate both of them into deployed systems" (Grossman, 2018a). It could be recommended for most SMEs to focus on their skills on setting up a model that can make predictions about the future. Regarding the results with the overall outcome with the size, SMEs that have a size of 2-50 are assumed in the AML 1, SMEs that have a size of 10-50 are assumed in the AML 2 and lastly, there is an exception for SMEs with the sizes of 101-150 and 151-200 as it can be assumed that they are in the AML 4. As AML 1 and 2 were discussed before, AML organisations have defined their processes and structures (including a governance structure); there is an uniform process used across the organisation when choosing analytics (Grossman, 2018a).

Next, examining the results for the overall outcome and with the size as an independent variable for Eckerson's framework (from question 2 until 6). The results of the overall outcome help assume that most SMEs in the Netherlands may be in the Infant and Teenager stage. The SMEs on the Infant stage mainly may use spreadsheets and desktop databases used as replacement for regional data whereas the SMEs on the Teenager stage may recognise the needs and starts using a standardised set of project and development entities such as learning past experiences; also using more sophisticated software compared to the software used in the Infant stage. However, the end results are different when it comes to size because most of the SMEs that have a size of 10-50, 101-150 and 151-200 are assumed to be in the stage of Teenager and Adult unlike the SMEs that have a size of 2-50 that is on the Infant/Teen stage. In the Adult stage, the SMEs can do everything from the Teenager stage and moreover, "BI becomes the central IT system during daily operations IT system driving daily operations of company" (Eckerson, 2007).

Lastly, the results from the overall outcome and with the size an independent variable from the questions derived from Watson et al. can be inspected. The results of the overall outcome assume that most SMEs in the Netherlands may be in the Initiation stage and Growth stage. This means that for the SMEs in the Initiation stage, the analytics may be straightforward as there might be limited amount of data available and rather than sophisticated functions, it may focus on examining what has occurred (Watson et al., 2001). While the SMEs in the Growth stage, the analytics might be more sophisticated but accessible as there might be more users that use it and more analysis on why things occurred and "what if" scenarios for the future (Watson et al., 2001). On the other hand, the results with the size as an independent variable again show a different end result because SMEs with the size of 2-50 may be in the Initiation stage, SMEs with the size of 10-50 may be the Growth stage and lastly, SMEs with the size of 101-150 and 151-200 may be both in the Maturity stage.

8. CONCLUSION

From the several findings, there are improvements and recommendations that could be made for the maturity level frameworks (even when creating a new framework solely for SMEs). Though the sample size is small and it cannot be used to generalise the whole population, it can be deducted that size plays a big influence as this was the case for all three frameworks when the results were compared. In conclusion, the bigger the SME, the more the SME use analytics to their advantage and can utilise their data to the fullest. The size also has an effect on how up-to-date the analytics are, for example, SMEs (that have the size of 2-50) in the Infant stage may spreadsheets compared to the SMEs (that have the size of 151-200) in the Adult stage that may use cascading scorecards. Some improvements that could be made that would make these frameworks more suitable and be used as an improvement tool are 1) being clearer with what tools are relevant for certain aspects, 2) making sure size is incorporated, 3) using the size as the main constant, 4) rather than generalising SMEs, categorising them into their industry as a construction may not have any use of being recommended web analytics compared to the food catering industry. Please refer to the table as an example of this ideal framework, using the other three maturity level of analytic frameworks.

Table 15 -Potential maturity model for SMEs (size)

Size of SME	Levels	Description
2-50	Birth	Build reports
10 -50	Infancy	Build models
101-150	Childhood	Repeatable analytics
151-200	Adolescence	Enterprise analytics
201-250	Adulthood	Strategy-drive analytics

9. LIMITATIONS OF RESEARCH

The purpose of the research was to determine whether the maturity level frameworks were of relevance to the SMEs in the Netherlands and there are limitations of this research that should be noted. First, this research requires additional empirical evidence as it was not sufficient. It was, during the start, limited to a specific region (Twente) which was later on expanded to the whole of the Netherlands. This was because it was difficult to collect even an appropriate amount needed for this research and thus, it is essential to collect further research to classify a considerable sample to generate a sounder framework and most importantly, safeguard a greater reliability and validity of the data. Second, this research requires a longer timeframe to be able to collected the necessary amount of data needed and this limitation goes hand in hand with the first limitation mentioned. Third, the data collected were relied on the subjective viewpoints of the SMEs representative and the representative may not have had the knowledge required to answer certain questions regarding, for example, knowing all the analytics used at the company. Also, the representative may not have answered truthfully in order to have better representation and indication on the maturity level frameworks. Therefore, these limitations may have led to biased results and conclusions. Lastly, the administered questionnaire was not immaculate. This is because there were a few technical difficulties and as most of the questions in the questionnaire was in English, there may have be some misunderstanding, for example, in the terminology of certain analytics and may have been clear for the SME representative had the questionnaire been in Dutch.

10. ACKNOWLEDGEMENTS

First of all, I would like to thank my supervisor, Dr. A.B.J.M. Wijnhoven, for his guidance, supervision, and valuable feedback while writing the bachelor thesis. Second, I would like to thank my bachelor thesis circle, Robbin Baars and Johannes Diederichs, for their feedback, support and teamwork throughout the whole process of writing the thesis. Lastly, I would like to thank my family and friends, from the bottom of my heart, for their encouragement, love and endless support during my studies and without whom I would have never been able to attain this bachelor degree.

11. REFERENCES

Agarwal, M., Sevag, A., Globanet, A., Microstrategy, T. B., Campbell, R., Coggeshall, S., … Kerber, J. (2012). Demystifying Big Data: A Practical Guide To Transforming The Business of Government Listing of Leadership and Commissioners Global Executive Vice President and General Manager, Database and Technology SAP Federal Big Data Commission Staff TechAmerica Foun, 1–40. Retrieved from https://bigdatawg.nist.gov/_uploadfiles/M0068_v1_3903747095.pdf

Akter, S., & Wamba, S. F. (2016). Big data analytics in E-commerce: a systematic review and agenda for future research. *Electronic Markets*, *26*(2), 173–194. https://doi.org/10.1007/s12525-016-0219-0

Boyd, D., & Crawford, K. (2012). CRITICAL QUESTIONS FOR BIG DATA. *Information, Communication & Society*, *15*(5), 662–679. https://doi.org/10.1080/1369118X.2012.678878

Bruns, A. (2008). *Blogs, Wikipedia, Second life, and Beyond : from production to produsage*. Peter Lang. Retrieved from https://books.google.nl/books/about/Blogs_Wikipedia_Second_Life_and_Beyond.html?id=oj2A68UIHpkC&redir_esc=y

Caesarius, L. M., & Hohenthal, J. (2018). Searching for big data: How incumbents explore a possible adoption of big data technologies. *Scandinavian Journal of Management*, *34*(2), 129–140. https://doi.org/10.1016/j.scaman.2017.12.002

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

Colas, M., Nambiar, R., Finck, I., Singh, R. R., & Buvat, J. (2014). Cracking the Data Conundrum: How Successful Companies Make Big Data Operational. *Capgemini Consulting*, 1–18. Retrieved from https://www.capgemini.com/consulting/wp-content/uploads/sites/30/2017/07/big_data_pov_03-02-15.pdf

Cukier, K. (2010). Data, data everywhere | The Economist. Retrieved June 14, 2018, from https://www.economist.com/node/15557443

Deloitte. (2013). The Analytics Advantage - We're just getting started. *Deloitte Network*, 1–21. Retrieved from https://www2.deloitte.com/content/dam/Deloitte/global/Documents/Deloitte-Analytics/dttl-analytics-analytics-advantage-report-061913.pdf

Doan, A., Ramakrishnan, R., & Halevy, A. Y. (2011). Crowdsourcing systems on the World-Wide Web. *Communications of the ACM*, *54*(4), 86. https://doi.org/10.1145/1924421.1924442

Eckerson, W. (2007). Beyond the Basics : Accelerating BI Maturity.

Ertemel, A. V. (2015). *International journal of commerce and finance. International Journal of Commerce and Finance* (Vol. 1). Retrieved from http://ijcf.ticaret.edu.tr/index.php/ijcf/article/view/9

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

Gelfand, A. (2011). Building a Trust Infrastructure Privacy Fears Drive Innovation. *Simbios, The NIH National Center for Physics-Based Simulation of Biological Structures* , (Winter), 22–28. Retrieved from http://biomedicalcomputationreview.org/sites/default/files/w12_f2-privacy_and_research.pdf

Ghasemaghaei, M., Hassanein, K., & Turel, O. (2017). Increasing firm agility through the use of data analytics: The role of fit. *Decision Support Systems*, *101*, 95–105. https://doi.org/10.1016/J.DSS.2017.06.004

Grossman, R. L. (2018). A framework for evaluating the analytic maturity of an organization. *International Journal of Information Management*, *38*(1), 45–51. https://doi.org/10.1016/j.ijinfomgt.2017.08.005

Hribar, I. (2010). OVERVIEW OF BUSINESS INTELLIGENCE MATURITY.

IBM. (2018). Big Data Analytics | IBM Analytics. Retrieved June 14, 2018, from https://www.ibm.com/analytics/hadoop/big-data-analytics

IBM, McKinsey Global Institute, Twitter, Cisco, Gartner, EMC, … QAS. (2018). Infographic: The Four V's of Big Data | IBM Big Data & Analytics Hub. Retrieved June 15, 2018, from http://www.ibmbigdatahub.com/infographic/four-vs-big-data

Interactive, H. S. (2012). Small and Midsize Companies Look to Make Big Gains With "Big Data," According to Recent Poll Conducted on Behalf of SAP. Retrieved June 14, 2018, from http://global.sap.com/news-reader/index.epx?PressID=19188

Jäger, J., Schöllhammer, O., Lickefett, M., & Bauernhansl, T. (2016). Advanced Complexity Management Strategic Recommendations of Handling the "industrie 4.0" Complexity for Small and Medium Enterprises. In *Procedia CIRP* (Vol. 57, pp. 116–121). https://doi.org/10.1016/j.procir.2016.11.021

James Manyika, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, A. H. B. (2011). Big data : The next frontier for innovation , competition , and productivity, (May).

Jin, X., Wah, B. W., Cheng, X., & Wang, Y. (2015). Significance and Challenges of Big Data Research ☆. *Big Data Research*, *2*(2), 59–64. https://doi.org/10.1016/j.bdr.2015.01.006

Kallinikos, J. (2010). *Governing through technology: Information artefacts and social practice*. Springer.

Kambatla, K., Kollias, G., Kumar, V., & Grama, A. (2014). Trends in big data analytics. *J. Parallel Distrib. Comput*, *74*, 2561–2573. https://doi.org/10.1016/j.jpdc.2014.01.003

Kjaerulff, J. (2010). *Internet and change : an anthropology of knowledge and flexible work*. Intervention Press.

Krosnick, J. a., & Presser, S. (2010). *Question and Questionnaire Design*. *Handbook of Survey Research*. https://doi.org/10.1111/j.1432-1033.1976.tb10115.x

Lahrmann, G., Marx, F., Winter, R., & Wortmann, F. (2011). Business intelligence maturity: Development and evaluation of a theoretical model. In *Proceedings of the Annual Hawaii International Conference on System Sciences*. https://doi.org/10.1109/HICSS.2011.90

Laney, D., Management, D., & Volume, C. D. (2005). META Delta, (February 2001).

Morgan, J. (2014). A Simple Explanation Of The "Internet Of Things." *Forbes BrandVoice*, 3–6.

O'Reilly, T. (2007). M P RA What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. Retrieved from http://mpra.ub.uni-muenchen.de/4578/

Oracle. (2014). *Oracle: Big Data for the Enterprise*. *Oracle White Paper*. Retrieved from http://www.oracle.com/technetwork/database/bigdata-appliance/overview/wp-bigdatawithoracle-1453236.pdf

Russom, P. (2011). Big data analytics. *TDWI Best Practices Report, Fourth Quarter*, 6. Retrieved from ftp://129.35.224.12/software/tw/Defining_Big_Data_through_3V_v.pdf

SAS. (2018a). Analytics - What it is and why it matters. Retrieved from https://www.sas.com/en_us/insights/analytics/what-is-analytics.html

SAS. (2018b). Big Data - What it is and why it matters. Retrieved from https://www.sas.com/en_us/insights/big-data/what-is-big-data.html

Smetsers, D. (2016). *Smart Industry onderzoek 2016*.

The Economist. (2011). Beyond the PC | The Economist. Retrieved June 16, 2018, from https://www.economist.com/node/21531109

The European Commission. (2017). 2017 SBA Fact Sheet Netherlands. *Annual Report on European SMEs*, (2017), 1–19.

The European Commission. (2018). Overijssel - Internal Market, Industry, Entrepreneurship And Smes - European Commission. Retrieved June 14, 2018, from https://ec.europa.eu/growth/tools-databases/regional-innovation-monitor/base-profile/overijssel

Veaux, R. de, Bock, D. E., & Velleman, P. F. (2015). *Stats* (4th ed.). Pearson Education limited.

VID/VDE. (2015). Reference Architecture Model Industrie 4.0 (RAMI4.0). *Igarss 2014*, *0*(1), 28. https://doi.org/10.1007/s13398-014-0173-7.2

Watson, H., Ariyachandra, T., & Matyska, R. J. (2001). Data warehousing stages of growth. *Information Systems Management*, *18*(3), 42–50. https://doi.org/10.1201/1078/43196.18.3.20010601/31289.6

Wu, C., Buyya, R., & Ramamohanarao, K. (2016). Big Data Analytics = Machine Learning + Cloud Computing. Retrieved from http://arxiv.org/abs/1601.03115

Zammuto, R. F., Griffith, T. L., Majchrzak, A., Dougherty, D. J., & Faraj, S. (2007). Information Technology and the Changing Fabric of Organization. *Organization Science*, *18*(5). https://doi.org/10.1287/orsc.1070.0307

Zheng, P., wang, H., Sang, Z., Zhong, R. Y., Liu, Y., Liu, C., … Xu, X. (2018). Smart manufacturing systems for Industry 4.0: Conceptual framework, scenarios, and future perspectives. *Frontiers of Mechanical Engineering*, *13*(2), 137–150. https://doi.org/10.1007/s11465-018-0499-5

12. APPENDIX

12.1 SPSS outcome (Overall)

Question 1

		Responses N	Responses Percent	Percent of Cases
Statements[a]	1	18	20.5%	100.0%
	2	16	18.2%	88.9%
	3	12	13.6%	66.7%
	4	5	5.7%	27.8%
	5	8	9.1%	44.4%
	6	3	3.4%	16.7%
	7	3	3.4%	16.7%
	8	2	2.3%	11.1%
	9	8	9.1%	44.4%
	10	2	2.3%	11.1%
	11	4	4.5%	22.2%
	12	1	1.1%	5.6%
	13	1	1.1%	5.6%
	14	1	1.1%	5.6%
	15	1	1.1%	5.6%
	16	1	1.1%	5.6%
	18	2	2.3%	11.1%
Total		88	100.0%	488.9%

a. Group

Question 2

		Responses N	Responses Percent	Percent of Cases
Who[a]	1	4	12.1%	23.5%
	2	2	6.1%	11.8%
	3	13	39.4%	76.5%
	4	10	30.3%	58.8%
	6	4	12.1%	23.5%
Total		33	100.0%	194.1%

a. Group

Question 3

		Responses		Percent of Cases
		N	Percent	
Your focus[a]	1	12	29.3%	66.7%
	2	2	4.9%	11.1%
	3	9	22.0%	50.0%
	4	8	19.5%	44.4%
	5	4	9.8%	22.2%
	6	6	14.6%	33.3%
Total		41	100.0%	227.8%

a. Group

Question 4

		Responses		Percent of Cases
		N	Percent	
Company focus[a]	1	8	18.2%	44.4%
	2	9	20.5%	50.0%
	3	14	31.8%	77.8%
	4	9	20.5%	50.0%
	5	4	9.1%	22.2%
Total		44	100.0%	244.4%

a. Group

Question 5

		Responses		Percent of Cases
		N	Percent	
Outcome[a]	1	13	24.5%	72.2%
	2	7	13.2%	38.9%
	3	7	13.2%	38.9%
	4	15	28.3%	83.3%
	5	10	18.9%	55.6%
	6	1	1.9%	5.6%
Total		53	100.0%	294.4%

a. Group

Question 6

		Responses N	Responses Percent	Percent of Cases
Tools[a]	1	16	36.4%	88.9%
	2	3	6.8%	16.7%
	3	11	25.0%	61.1%
	4	8	18.2%	44.4%
	5	5	11.4%	27.8%
	6	1	2.3%	5.6%
Total		44	100.0%	244.4%

a. Group

Question 7

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	15	83.3	83.3	83.3
	No	3	16.7	16.7	100.0
	Total	18	100.0	100.0	

Question 8

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	8	44.4	44.4	44.4
	No	10	55.6	55.6	100.0
	Total	18	100.0	100.0	

Question 9

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	11	61.1	61.1	61.1
	No	7	38.9	38.9	100.0
	Total	18	100.0	100.0	

Question 10

		Responses		Percent of Cases
		N	Percent	
Analytics[a]	1	7	12.5%	38.9%
	2	5	8.9%	27.8%
	4	13	23.2%	72.2%
	5	7	12.5%	38.9%
	6	7	12.5%	38.9%
	7	1	1.8%	5.6%
	8	4	7.1%	22.2%
	9	4	7.1%	22.2%
	10	2	3.6%	11.1%
	11	1	1.8%	5.6%
	14	3	5.4%	16.7%
	18	2	3.6%	11.1%
Total		56	100.0%	311.1%

a. Group

12.2 SPSS outcome (with size)

12.2.1 Question 1

		Statements																	
		Our company can analyse data	Our company can build reports summarsing the data	Our company can make use of the reports to further the goals of the organisation	Our company can use the analysed data to build analytic mode	Our company can use an analytic model	Our company follows a repeatable process for building analytic models	Our company follows a repeatable process for deploying analytic models	Our company follows a repeatable process for updating analytic models	Our company uses analytics throughout the company	Analytics models are built with a common infrastructure and process whenever possible	Analytics models are used with a common infrastructure and process whenever possible	The outputs of the analytic models integrate together to optimise goals of the company	Analytics across the enterprise are coordinated by an analytic governance structure	Our company has defined an analytic strategy	The analytic strategy aligns with the overall strategy of the company	Uses the analytic strategy to select appropriate analytic opportunities	Our company use the analytic strategy to select appropriate analytic opportunities	Use the analytic strategy to develop and implement analytic processes that support the overall vision and mission
		Count	Count	Count	Count	Count	Count	Count	Count	Count	Count	Count	Count	Count	Count	Count	Count	Count	Count
Size	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	2-50	9	8	7	0	1	0	0	0	3	0	1	0	0	0	0	1	0	1
	10-50	6	5	2	2	4	0	0	0	2	0	1	0	0	0	0	0	0	0
	101-150	2	2	2	2	2	2	2	1	2	1	1	0	0	1	1	0	0	1
	151-200	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0
	201-250	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	More than 250	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

12.2.2 Question 2

		Who					
		Analyst	Knowledge worker	Manager	Executive	Customer	Other
		Count	Count	Count	Count	Count	Count
Size	1	0	0	0	0	0	0
	2-50	0	0	5	6	0	2
	10-50	2	1	5	2	0	1
	101-150	1	1	2	1	0	1
	151-200	1	0	1	1	0	0
	201-250	0	0	0	0	0	0
	More than 250	0	0	0	0	0	0

12.2.3 Question 3

		Your focus					
		To make plans	To assign	To process	To make a strategy	To make services	Other
		Count	Count	Count	Count	Count	Count
Size	1	0	0	0	0	0	0
	2-50	7	1	4	5	4	2
	10-50	2	1	3	1	0	2
	101-150	2	0	2	2	0	1
	151-200	1	0	0	0	0	1
	201-250	0	0	0	0	0	0
	More than 250	0	0	0	0	0	0

12.2.3.1 Question 4

		Company focus					
		To know what will happen	To know why it happen	To know what is happening	To know what to do	To know what we can offer	Other
		Count	Count	Count	Count	Count	Count
Size	1	0	0	0	0	0	0
	2-50	3	3	8	6	3	0
	10-50	4	3	3	1	0	0
	101-150	0	2	2	1	0	0
	151-200	1	1	1	1	1	0
	201-250	0	0	0	0	0	0
	More than 250	0	0	0	0	0	0

12.2.4 Question 5

		Outcome					
		Plans	Procedures and policies	Caution	Action	Recommendations	Other
		Count	Count	Count	Count	Count	Count
Size	1	0	0	0	0	0	0
	2-50	6	2	2	7	5	0
	10-50	4	3	2	5	2	0
	101-150	2	2	2	2	2	1
	151-200	1	0	1	1	1	0
	201-250	0	0	0	0	0	0
	More than 250	0	0	0	0	0	0

12.2.5 Question 6

		Tools					
		Spreadsheets	OAP	Dashboard	Scorecards	Statistical models	Other
		Count	Count	Count	Count	Count	Count
Size	1	0	0	0	0	0	0
	2-50	7	2	4	4	2	0
	10-50	6	0	4	2	1	0
	101-150	2	1	2	1	1	1
	151-200	1	0	1	1	1	0
	201-250	0	0	0	0	0	0
	More than 250	0	0	0	0	0	0

12.2.6 Question 7, 8, 9

		Q7		Q8		Q9	
		Yes Count	No Count	Yes Count	No Count	Yes Count	No Count
Size	1	0	0	0	0	0	0
	2-50	7	2	3	6	4	5
	10-50	5	1	4	2	4	2
	101-150	2	0	1	1	2	0
	151-200	1	0	0	1	1	0
	201-250	0	0	0	0	0	0
	More than 250	0	0	0	0	0	0

12.2.7 Question 10

		Analytics																	
		RDBMS & data warehousing Count	ETL Count	OLAP Count	Dashboards & scorecards Count	Data mining & statistical analysis Count	Information retrieval and extraction Count	Opinion mining Count	Question answering Count	Web analytics and web intelligence Count	Social media analytics Count	Social network analysis Count	Spatial-temporal analysis Count	Location-aware analysis Count	Person-centered analysis Count	Context-relevant analysis Count	Mobile visualisation Count	HCI Count	Other Count
Size	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	2-50	6	3	0	5	2	1	1	2	3	1	0	0	0	1	0	0	0	1
	10-50	0	0	0	5	3	3	0	1	0	0	0	0	0	0	0	0	0	1
	101-150	1	2	0	2	1	2	0	1	1	1	1	0	0	2	0	0	0	0
	151-200	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
	201-250	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	More than 250	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

12.3 Questionnaire (with coding)

Q2 Hoeveel werknemers heeft uw organisatie?

- ◯ 1 (1)
- ◯ 2-50 (2)
- ◯ 10-50 (3)
- ◯ 101-150 (4)
- ◯ 151-200 (5)
- ◯ 201-250 (6)
- ◯ Meer dan 250 (7)

Q3 In welke sector bevindt uw organisatie zich?

- Landbouw, bosbouw en visserij (1)
- Winning van delfstoffen (2)
- Industrie (3)
- Productie en distributie van en handel in elektriciteit, aardgas, stoom en gekoelde lucht (4)
- Winning en distributie van water, afval- en afvalwaterbeheer en sanering (5)
- Bouwnijverheid (6)
- Groot- en detailhandel, reparatie van auto's (7)
- Vervoer en opslag (8)
- Logies-, maaltijd- en drankverstrekking (9)
- Informatie en communicatie (10)
- Financiële instellingen (11)
- Verhuur van en handel in onroerend goed (12)
- Advisering, onderzoek en overige specialistische zakelijke dienstverlening (13)
- Verhuur van roerende goederen en overige zakelijke dienstverlening (14)
- Openbaar bestuur, overheidsdiensten en verplichte sociale verzekeringen (15)
- Onderwijs (16)
- Gezondheids- en welzijnszorg (17)
- Cultuur, sport en recreatie (18)
- Overige dienstverlening (19)

Q4 De organisatie bevindt zich in de omgeving van Twente

- Ja (1)

- Nee, in de provincie... (2) ____

Q7 Please select the following statements that have relevance to your company

- ☐ Our company can analyse data (1)
- ☐ Our company can build reports summarising the data (2)
- ☐ Our company can make use of the reports to further the goals of the organisation (3)
- ☐ Our company can use the analysed data to build analytic models (4)
- ☐ Our company can use an analytic model (5)
- ☐ Our company follows a repeatable process for building analytic models (6)
- ☐ Our company follows a repeatable process for deploying analytic models (7)
- ☐ Our company follows a repeatable process for updating analytic models (8)
- ☐ Our company uses analytics throughout the company (9)
- ☐ Analytics models are built with a common infrastructure and process whenever possible (10)
- ☐ Analytics models are used with a common infrastructure and process whenever possible (11)
- ☐ The outputs of the analytic models integrate together to optimise goals of the company (12)
- ☐ Analytics across the enterprise are coordinated by an analytic governance structure (13)
- ☐ Our company has defined an analytic strategy (14)
- ☐ The analytic strategy aligns with the overall strategy of the company (15)
- ☐ Uses the analytic strategy to select appropriate analytic opportunities (16)
- ☐ Our company use the analytic strategy to select appropriate analytic opportunities (17)
- ☐ Our company use the analytic strategy to develop and implement analytic processes that support the overall vision and mission of the company (18)

Q8 Who uses the analytics in your company?

- ☐ Analyst (1)
- ☐ Knowledge worker (2)
- ☐ Manager (3)
- ☐ Executive (4)
- ☐ Customer (5)
- ☐ Other (6) ____

Q9 What is your focus on using analytics?

- ☐ To make plans (1)
- ☐ To assign (2)
- ☐ To process (3)
- ☐ To make a strategy (4)
- ☐ To make services (5)
- ☐ Other (6) ____

Q10 What is the focus of using analytics?

- ☐ To know what will happen (1)
- ☐ To know why it happen (2)
- ☐ To know what is happening (3)
- ☐ To know what to do (4)
- ☐ To know what we can offer (5)
- ☐ Other (6) ____

Q11 What is the outcome of using analytics?

- ☐ Plans (1)
- ☐ Procedures and policies (2)
- ☐ Caution (3)
- ☐ Action (4)
- ☐ o Recommendation (5)
- ☐ Other (6) ____

Q12 What are the tools used for analytics?

- ☐ Spreadsheets (1)
- ☐ Online analytical processing (2)
- ☐ Dashboard (3)
- ☐ Scorecards (4)
- ☐ o Statistical models (5)
- ☐ Other (6) ____

Q13 Do you use internal data to improve the processes within the company?

- ○ Yes (1)
- ○ No (2)

Q14 Do you use external data to create products specified to the need of one group of people?

- ○ Yes (1)
- ○ No (2)

Q15 Do you use data analytics that can help generate useful data information catered towards needs of company?

- ○ Yes (1)
- ○ No (2)

Q16 In general, what kind of analytics do you use in your company?

- ☐ RDBMS & data warehousing (1)
- ☐ ETL (process in data warehousing responsible for pulling data out of the source systems and placing it into a data warehouse) (2)
- ☐ OLAP (approach to answering multi-dimensional analytical queries swiftly in computing) (3)
- ☐ Dashboards & scorecards (4)
- ☐ Data mining & statistical analysis (5)
- ☐ Information retrieval and extraction (6)
- ☐ Opinion mining (7)
- ☐ Question answering (8)
- ☐ Web analytics and web intelligence (9)
- ☐ Social media analytics (10)
- ☐ Social network analysis (11)
- ☐ Spatial-temporal analysis (analysis of large spatiotemporal, space and time, databases) (12)
- ☐ Location-aware analysis (services are tied to mobile networks, mobile analytics) (13)
- ☐ Person-centered analysis (proposes focusing on the individual (14)
- ☐ Context-relevant analysis (15)
- ☐ Mobile visualization (16)
- ☐ HCI (focused on the interfaces between people (users) and computers) (17)
- ☐ Other (18) ____