DESK STUDY - FUTURES BY DESIGN

INCREASING THE DATA MATURITY OF SMEs

WP4 - TOOLS

MATTHIJS BOOKELMANN, RUUD SNEEP
JHERONIMUS ACADEMY OF DATA SCIENCE
m.t.l.bookelmann@tue.nl
r.sneep@tue.nl

Table of Contents

Introduction

This report is written as part of the European Project Futures by Design, which strives to aid SMEs which struggle to adopt the latest techniques in the field of data science. The aim is to help these SMEs to get access to particular technologies and tools which help them innovate, grow and / or increase productivity. The project consists of a collaboration with SMEs in the partnering regions in the Netherlands, Belgium, Germany, Sweden and the United Kingdom. Data science offers an opportunity for value creation that could benefit small and medium enterprises (SMEs) within the European economy. Within any city or village there exist many examples of SMEs that could benefit from working more data driven. They might or might not be aware of the increasing possibilities of existing technologies which can create value from data. Futures by Design is here to help them become aware and provide the tools to help them further.

During the Future by Design project the local partners will share knowledge, ideas and local experiences to support SMEs. In doing so, they help SMEs to become more data driven and provide more information about the relevant developments in the field of data science and its business applications. Futures by Design has created a virtual environment in which the 6 partner regions are connected to support sustainable growth, innovation and productivity of SMEs.

This report focusses on the core of Futures by Design, namely “Understanding how local support hubs could help SMEs become more Data-Driven”. As such, the report is split into two parts. The first part consists of a general view on understanding SMEs and their relationship vis-á-vis data science. In particular, this first part provides a more elaborate view on 1) the importance of SMEs for the EU economy, 2) the relevance of Data Science regarding firm performance, 3) pre-requisites for creating value from data, and 4) reasons why SMEs are generally lagging behind.

The second part of this report addresses the process of creating value from data and ways of how SMEs can be supported in overcoming the barriers they encounter. We start by explaining the importance of SMEs for the EU in general.

Part I: Understanding SMEs & Data Science

SMEs as the backbone of EU society and economy

Small and medium enterprises (SMEs), companies employing less than 250 employees, are a vital part of the European Union’s economy (European Commission, 2018). The importance of SMEs for the EU economy becomes clear when we look at some key statistics about the EU economy. Over 99 % of all companies in the EU are SME. Together, SMEs amount on average for two-thirds of the total European employment and this number is only increasing. Furthermore, SMEs account for over half of the total turnover in the EU (European Commission, 2018).

It is safe to conclude that to sustain competitive the EU relies for a large part on the performance of its SMEs. Therefore, it is of utmost importance that SMEs in the EU innovate for the EU to remain competitive in the global market. Furthermore, Rosenbusch, Brinckmann and Bausch (2011) found that both an innovation orientation and innovation activities generate value for new and established SMEs. One of the new relevant developments on the rise in the area of innovation is the application of data-centric innovations.

90 % of SMEs feel they are lagging in digital innovation (European Commission, 2018). SMEs move more slowly into the field of data science compared to the larger corporations and born digital start-ups, which puts a threat on the sustainability of many SMEs both from the larger and smaller companies.

Why Data Science?

Data Science is on a rise, as an academic field as well as a manner to improve the way business is conducted. Data Science is often projected as the "new oil" and Data Scientists have received the honour of having "the sexiest job of the twenty-first century" (Davenport & Patil, 2012). What has caused this attention for Data Science as a field?

There are three main reasons why it is now that companies can become more and more data-driven: There is 1) more data; 2) more computational power; 3) more opportunities and software for gathering, processing and analyzing data.

Firstly, already since the 90s, more business processes are increasingly digitized, accumulating the amount of digitally stored data. Think of electronic cash registers, websites as a digital sales channel, social media to communicate with customers, and business operation systems (Enterprise Resource Planning (ERP), Customer Relationship (CRM) and Supply Chain Management (SCM), etc. (Aral, Brynjolfsson & Wu, 2006; McAfee, 2002)).

Furthermore, an increasing amount of electronical equipment is connected to the internet, a development often referred to as the Internet of Things (IoT): Turning any connected device into a sensor able to collect and send data. The volume of data being collected is increasing exponentially, in 2012 more data crossed the internet every second than the total amount of data saved in the internet 25 years ago (McAfee et al., 2012). The velocity, the speed in which data is created, is also very important for companies, as it becomes more and more feasible to act upon real time or almost real time information (McAfee et al, 2012).

Moreover, McAfee et al. (2012) find the vast variety in data sources has increased. These data sources such as images, text messages, social media posts, GPS signals from mobile phones and more can be processed and digitized with the latest techniques and developments. As a result, new insights can be derived from relatively novel sources of information and by combining information from these different data sources. In turn, these insights ought to create more value for consumers, organisations and policy makers.

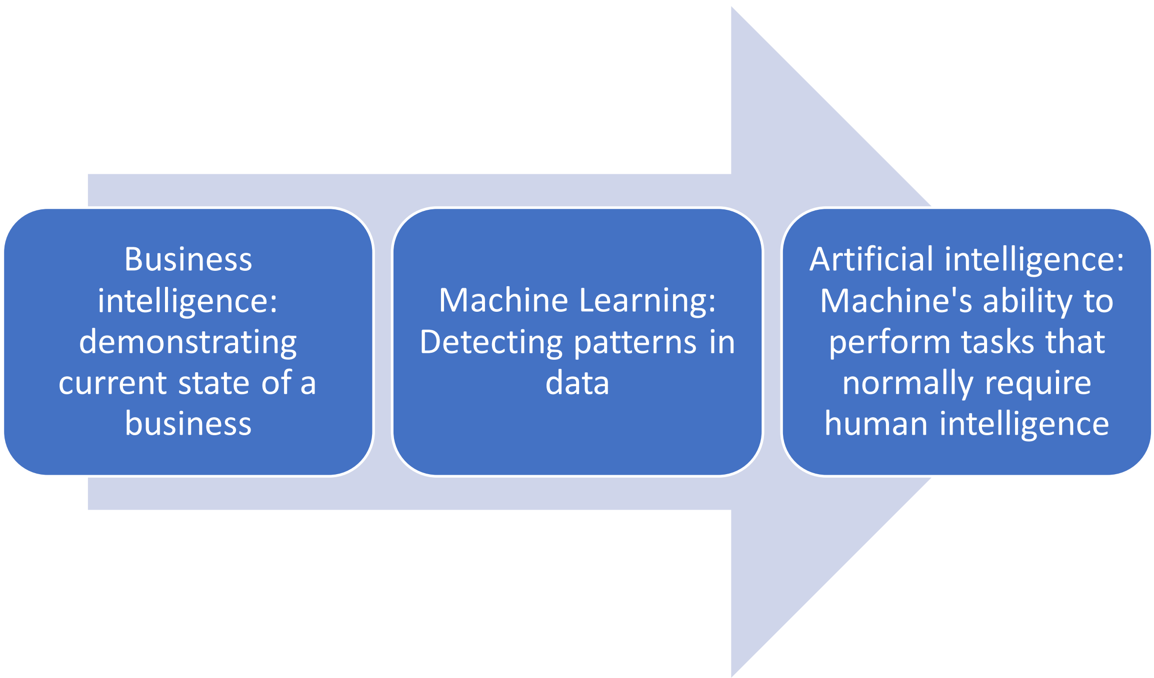
Secondly, our computers have been and are still increasing their computational power and storage capacity, on top of that our network becomes faster (Jaakkola, Henno, Mäkelä, & Thalheim, 2019). Due to the rise in cloud computing solutions offered by the big tech companies like Google, Microsoft and Amazon we can use this computational power from any location in an affordable manner – without owning a supercomputer ourselves.

Finally, as a result, the increased computational power has allowed for the rapid development of new and more mature data analysis techniques, models and algorithms (Venturini, Jensen, & Latour; 2015).

The vast and variate bundle of new data analysis techniques, models and algorithms is encompassed by the field of ‘Data Science’ (Hey, Tansley & Tolle, 2009).

Terminology used in the field of Data Science

The field of Data Science encompasses a large variety of specialist terms and jargon. Below, we highlighted and provided our understanding of three popularly mentioned terms in business and their meaning:



Business intelligence (BI) incorporates analytical tools to present complex information to decision makers (Negash & Gray, 2008). An example of BI is a dashboard, which provides the board of a company with an overview of the organisation’s operations. The decision makers within the organisation use the information displayed to them in the dashboard to make strategic or operational decisions that have an impact on the objectives of the organisation. One of the main aims of BI tools is to help decision makers understand what is going on in their organisation by visualizing the data and information gathered in the company in an understandable manner. Business intelligence includes a presentation of Key Performance Indicators (KPIs) from the organisation (Wimmer & Powell, 2016). Business intelligence therefore is more commonly used to represent visual presentation of a company’s performance in a dashboard.

Some more advanced Data Science techniques are:

Machine learning (ML) as a data science application is a method of programming machines to learn from past experiences (Mitchell, 1997). Machine Learning means as much as computers learning from past experiences with limited or no human interaction. Machine learning is going beyond a step beyond following a set of pre-coded rules, ML dynamically generates 'rules' based on input data (Wimmer & Powell, 2016).

Artificial intelligence (AI) is generally seen as the ability of computers to perform tasks like observing, analysing, learning and making decisions that normally would require human intelligence (e.g. speech recognition, image recognition, decision-making, and language translation). As the AI-related techniques evolve rapidly, so does the general understanding of what AI entails (Ransbotham, Kiron, Gerbert, & Reeves, 2017).

The combination of the increased amount of data as a result of digitization, the increased computing power and accessibility thereof, and the development of Data Science as a discipline brings us into a new era in which a substantive amount of digital information about essentially any topic of interest in business exists (McAfee et al., 2012). McAfee et al. (2012) note that, as we can measure more precisely than ever before we have better information at hands and can manage better than ever before. Therefore, data science applications are even claimed to be "the next frontier for innovation, competition, and productivity" (Manyika, J., 2011), and as "a major differentiator between high-performing and low-performing organisations", as it allows companies to conduct business in a more proactive and forward-looking manner (Liu, 2014).
Brynjolfsson, Hitt, & Kim (2011) found that the application of data science to business practices is associated with higher productivity and market value. Wegener and Sinha (2013) found that across sectors companies that have committed themselves to creating value form their data and investing in analytical capabilities are the top performers in their respective sectors. McAfee et al. (2012) found that companies in the top third of their industry in the use of data-driven decision making, were on average 6 % more profitable and 5 % more productive than their competitors.

The fact that so many more aspects of our society have gone digital means it has become much easier to gather, clean and analyse data. To understand whether a company's budget can be spend more efficiently, to understand whether personnel could work more productively, to understand whether clients are satisfied with the product or service – never in time has it been easier to measure these inputs, learn from them and act upon these insights. The collection of data in IT systems is a key part for the creation of knowledge within an organisation (Lopez-Nicolas, & Soto-Acosta, 2010). When a company is able to transform data into information, this company can make better data-driven decisions and reorganise, learn and innovate (Bean & Kiron, 2013; Yiu, 2012) and therefore strengthen customer relationship management, decrease operational risk, improve operational efficiency and overall firm performance (Bean & Kiron, 2013).
Data-driven decision making is not the answer to all business challenges, furthermore there are different types of challenges that could be tackled using data science techniques. The application of data science techniques fits better with, for example, the optimisation of a factory process then determining a business strategy.

Pre-requisites of a data-driven organisation

There are three key domains, which play an important part in the overall capability of a company to create firm value by the use of data science: data analytics Infrastructure flexibility, data analytics management capabilities, and most importantly, data analytics personnel expertise (Kim, et al., 2012; Wamba, et al., 2017). The competitive advantage a company can extract from developing IT capability is derived both from material components (information infrastructure, connectivity or shareability of data and insights, compatibility of applications, modularity of software) and social/human components (management skills, technical skills, business knowledge and relational knowledge). It is important to understand that to capture the real value that lies in the application of data science in business you need both humans and machines and therefore one should invest in both assets. Moreover, Davenport et al. (2012) emphasise that in the world in which there is an increasing amount of data available, people, management and technology are interconnected and help each other to enhance overall firm performance. Furthermore Barton and Court (2012) support these different pillars of data analytics and their relationships, they specify that management capability is important for the optimisation of decision making processes; technology capability is key to analyse and manage different kinds of data; and lastly data science capability is essential to understand, develop and apply analytics models.

Wamba et al. (2017) also found that having well-developed process-oriented dynamic capabilities has a strengthening impact on translation to value form data analytics capabilities of a company. Process-oriented dynamic capabilities refer to the ability to use information and insights in the improvement of the business processes in an efficient way.

Wamba et al. (2017) found that the capability to work data-driven leads to higher firm performance, Kim et al. (2012) find the IT capability is significantly related to firm performance and Lu & Ramamurthy (2011) found a significant positive relationship between IT capability and a firms agility (market capitalization agility and operational adjustment agility).

Why SMEs are lagging in the move towards a more data-driven organisation

These developments in the field of data science go together with great business opportunities. However, SMEs generally lag behind in the adoption of new technologies which could enable data-centric innovations, such as Customer Relations Management systems (Harrigan, Ramsey, & Ibbotson, 2011), e-commerce (Stockdale & Standing, 2004) and Enterprise Resource Planning software (Chen, Sun, Helms, & Jih, 2008). Madrid-Guijarro, Garcia and Van Auken (2009) identified the main barriers to innovate as an SME as the following: a lack of access to financial capital, having a difficulty attracting or developing the right human capital/skills for innovation, difficulties in accessing external knowledge and reluctance to take risks. Where larger companies have their own data science department (accessing external knowledge, human & financial capital), due to financial constrains SMEs generally do not have access to data science professionals, difficulties in attracting talent and/or difficulties in defining the projects to be taken on by these data science professionals.

The ability of an enterprise to assimilate and use new knowledge depends on whether a company already has prior related knowledge – also referred to as its absorptive capacity (Cohen & Levinthal, 1990), which is generally not the case for SMEs. SMEs might be unaware about the developments in the field of data science or unable to translate the developments in the field of data science towards business strategies for their own business (Van der Veen, Van der Born, Smetsers, Bosma, 2017). To increase the absorptive capacity, an SME could encounter the choice between further specializing the technological personnel or freeing up their time to broaden their knowledge and increase the

absorptive capacity. Due to the intangible nature of absorptive capacity, an enterprise might be hesitant to sacrifice current output as well as gains from specializations to allow its technological personnel to acquire a broader knowledge base, which would lead to an increased absorptive capacity (Cohen & Levinthal, 1990). This effect is expected to be larger in smaller organisations as these organisations depend more heavily on fewer technical personnel and therefore might even be more reluctant to sacrifice current output to increase the absorptive capacity of the company.
From a data-centric innovation perspective specifically there are more concrete obstacles for SMEs to get started with the data-centric innovation beyond the absorptive capacity. To apply data science techniques there might be a lack of access to the required data and analytical skills, the inability to blend in new technologies in the current business model and the hesitation to invest in high risk data-centric innovations (Choudhury, Starr, and Agarwal, 2018).

In a global, cross industry research LaValle et al. (2010) found that the top three impediments for the company-wide adoption and use of information and analytics are:

1) Lack of understanding how to use analytics to improve the business
2) Lack of management bandwidth due to competing priorities
3) Lack of skills internally in the line of business

Conclusion

SMEs are a vital part the EU economy and competitive power, it is important that these businesses keep innovating to ensure the innovative power of the EU. Moreover, innovation has proven to be beneficial to the SMEs themselves as well. Currently the field of data science offers a lot of potential to generate value from data-centric innovations, nevertheless SMEs seem to be lagging both on digital innovations and the implementation of data-centric innovations on a large skill. SMEs seem to not contain the main three ingredients to create value from data, namely data-oriented management, data analytical skills of employees, nor the right software infrastructure. In the next part of this report, the process of value creation through data is further explained to work on an approach in supporting SMEs to move towards value creation from data.

1. SMEs are a vital part of the EU economy;
2. Data Science offers a lot of potential for businesses to innovate;
3. However, SMEs lack data-oriented management, analytical skills and software infrastructure

Part II: Supporting SMEs to become Data-Driven

Introduction

To understand how SMEs can be best supported in creating value from data and tapping into the opportunity that data offers, it is important to understand how data can generate value. In the second part of this report, we will firstly go into the process of data-driven decision-making. Secondly, the importance of data maturity levels in helping companies effectively on their way of becoming more data driven is elaborated upon and considerations for working towards a more data driven company are being discussed. Finally, the approach and learnings of the JADS SME Datalab, a 'good practice' in helping SMEs becoming more data driven (European Commission, 2019) is discussed. To conclude with general recommendations in the exploratory support of SMEs in capturing value from data.

Towards a Data-Driven Business

Before we go into the adaptation of a data-driven mindset, it is important to understand what this means. A data-driven business is a type of business that gathers and processes data in order to get to actionable insights.

In this light, it is important to understand what we mean with terms as 'data', 'information' and 'knowledge'. In his paper on defining data, information and knowledge, Zins (2007) tried to find the most common manner to define the terms 'data', 'information' and 'knowledge' in the field of information science. We have taken his conclusions as a starting point of formulating the meaning of these terms in a business context: *Data* is the input that is measurable (i.e. items sold, cost of production, time spend on performing a given task); *Information* is the explanation of what the data has measured (i.e. more items are sold yesterday than the day before; cost of production is increasing; relatively more time is spend on task A than task B); and *Knowledge* concerns the thoughts in the individual's mind, which is characterized by the individual's justifiable belief that it is true (i.e. we have sold more items yesterday because we invested more in sales). As such, knowledge can be empirical or non-empirical – and it can be fed by data, but is not necessarily so.

In the process of data-driven decision-making, the decision-maker chooses between deciding based on information derived from the data or solely based on previous knowledge. The increased amount of collected data and the increasingly powerful computers enable managers to measure data, translate it in information. As a result, they know radically more about their business, and these managers can directly implement the newly acquired knowledge to make better business decisions and improve performance (McAfee, Brynjolfsson, Davenport, Patil, & Barton, 2012).

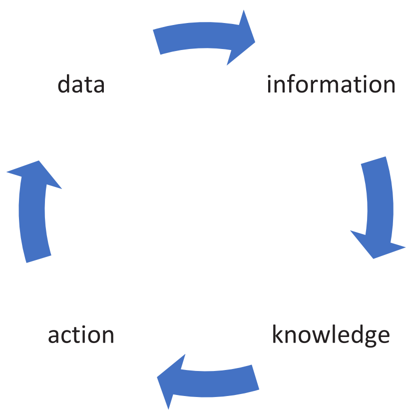


Figure 1 Value chain of data-driven decision-making

Data-driven decision-making means that there is a constant circular flow in the decision-making process form new data to information, knowledge and action inspired on the data-based knowledge. The data resulting from this action leads to new information, new knowledge and new action to be taken. Hence, action is taken based on data analytics and not on the 'gut feeling' or in other words the experience and opinion, of the decision maker. This set of actions is expected to yield higher returns and therefore it is seen as the value chain of data-driven decision-making (fig.1 Value Chain of Data-Driven Decision Making).

This process can be applied to answer a wide range of business questions by applying different analytical skills on the available data. Davenport and Harris (2007) used a graphic produced by SAS to illustrate that to answer a variety of business questions increasing in complexity, one must apply a range of different analytics increasing in level of intelligence and competitive advantage (fig. 2 Business Intelligence and Analytics).

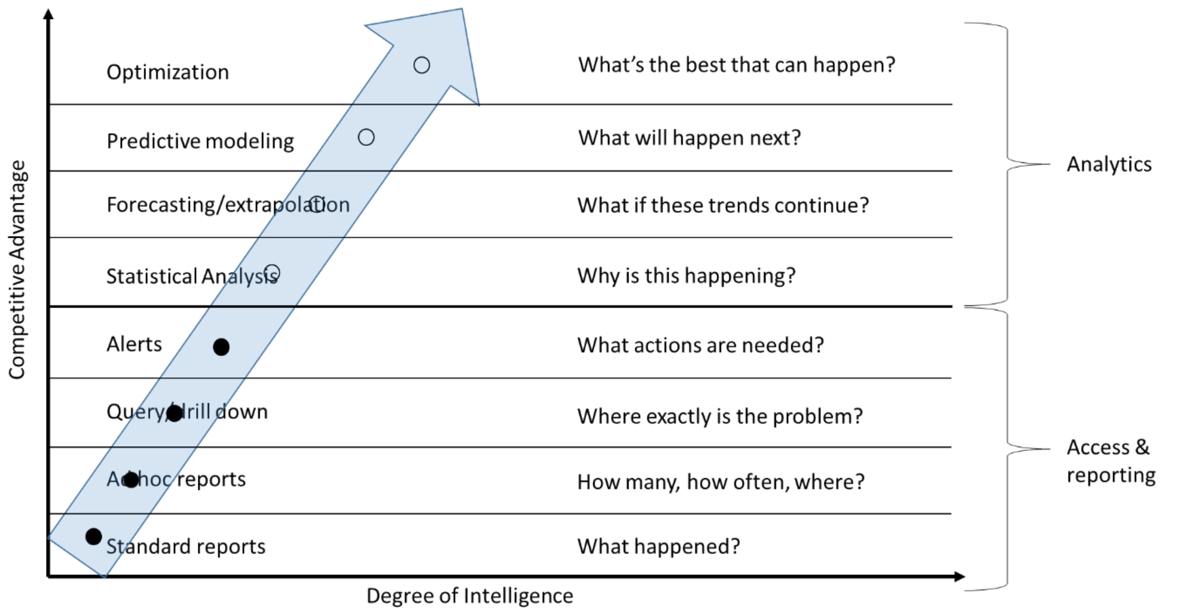


Figure 2: Business Intelligence and Analytics (Adapted from Davenport & Harris, 2007)

Businesses use analytics to answer a wide range of questions. The most basic analytics could be yearly financial reports that show the owners of the business how much they have earned the past year. A slightly more difficult topic could be to have an algorithm determine the value of a newly incoming lead, for the salesperson to get notified by the system when an 'interesting lead' comes in. A more predictive application of data analytics could, for example, help to predict churn of a company that offers a service on subscription basis. The top of the application of data science is to let the algorithm predict the future and act based on its prediction.

Figure 3 used by Kart (2015) demonstrates in a very comprehensible manner that when the level of intelligence of the analytics increases, the human input in the data-driven decision-making loop becomes smaller while the influence of the data and information resulting from analyzing the data increases up until the point that the decision-making on a certain decision process is fully automated. This frees up a lot of time of the human responsible for the decision to be taken. However, to reach this level of data-driven decision-making, a company needs to have a decent data-driven management mindset, enabling infrastructure and the analytical capabilities to retrieve valuable information from the data.

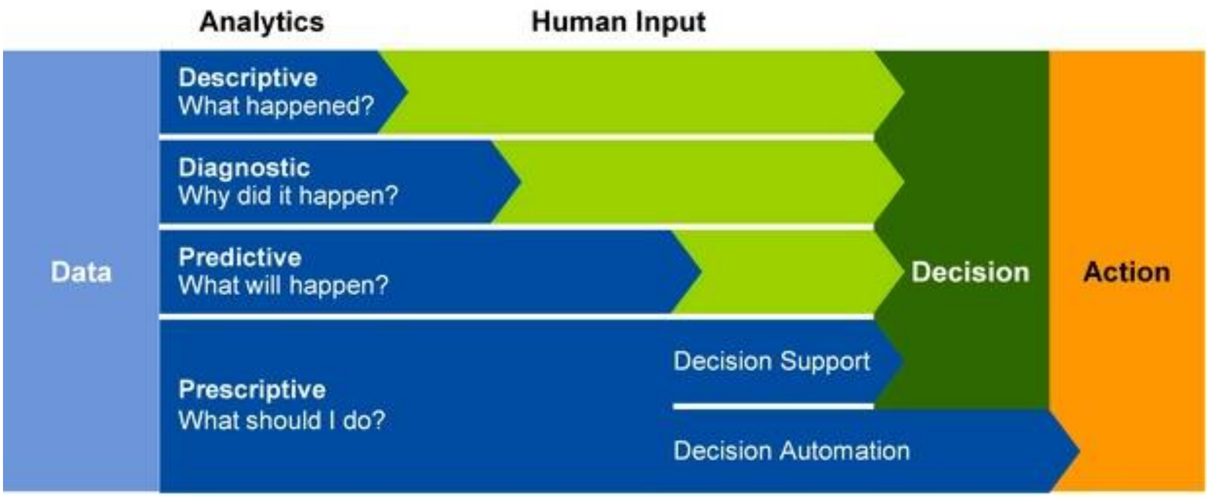


Figure 3 Data-Driven Decision-Making (Kart, 2015)

As most SMEs are still in a relatively early stage when it comes to becoming more data driven, it is important to know that the way towards data-driven decision-making is a step-by-step process and companies first need to lay the right foundations. A company that is on its journey towards becoming a data driven organisation will pass through different levels of data maturity. There are many different models that can be used to estimate the data maturity of a company. It goes beyond the scope of this report to make a comparison of the different existing data maturity models, however it is important to understand the value a data maturity model when aiding a company in improving their data-driven decision-making.

In principle, it is important to know the data maturity of a company to understand what the next steps for the specific company are to reach its goals in relation to their data-driven decision-making processes. Moreover, the data maturity level of the specific company compared to the data maturity level of other companies in their market determines whether they are lagging on the competition or have a competitive edge on their competition in the field of data-driven decision-making. As stated earlier there is empirical evidence that companies that use data-driven decision-making processes outperform competitors that lag on data-driven decision-making, simply because data-driven companies are assumed to have better information at hands to base their decision upon.

How can SMEs be supported to become more Data Driven?

For the value chain of data-driven decision-making to create value for the company, first the right data needs to be collected in the right way. In other words, there must be a data vision and strategy. Becoming more data driven as a company is always a means to an end, therefore it is important to identify the goals you want to reach first. These questions can help a company in establishing a data strategy:

- What does the company want to achieve?
- How can analysing data be used to reach that goal?
- Which data should be collected and how should it be collected?
- How will the data influence future decision-making and what value will it create for the company?

Secondly, a good analysis of the data must take place in order to derive information from the data. The SME should look at their human, technological and management capabilities and check if they are set up correctly to create a more data driven organisation. The SME should answer the following questions about their capabilities to identify their most important area to invest in:

- Do we have the right tools and skills available to make relevant and high-quality analyses of the data?
- Do we need to train, hire or contract human capital to help us apply the right kind of analyses?
- Do we need to invest in the improvement of our IT infrastructure to be able to make the right kind of analyses?
- Do we have the right culture at place to work with data-based knowledge in the decision-making processes?

When the company has the right data and the resulting relevant information, we get to the most important part of the process of data-driven decision making: taking the information into account when making a decision and take action upon the decision to create actual value from the new insights for the company. Furthermore, McAfee et al. (2012) see data-driven decision-making as a real management revolution, as the mindset of decision-makers must change. In the sense that decision makers should ask themselves the question “what does the data say?” when confronted with a challenge instead of relying on their expertise or intuition. Domain-specific information is used to identify specific challenges and opportunities in the domain which than can be tackled with data-driven decisions as result of newly created cycle of the value chain of data-driven decision-making.

Empirical evidence of SMEs benefitting from data-driven innovation

Despite the theoretical reasoning on the advantages of becoming a data driven business, the empirical evidence on the extent to which the implementation of a data-driven mindset is aiding SMEs specifically (i.e. in productivity, innovative capacity or financial growth) is not abundant. Two reasons underlie this observation: 1) in practice, not many SMEs have adopted a data-driven business model; 2) the SMEs that are a data-driven business would entail “frontrunners” (European Commission, 2019) and studying their performance is rather quickly affected by concerns of endogeneity and spurious relationships.

Of course, one could argue that there is ample evidence on the implementation of IT and the effect on SME performance, for example (Lopez-Nicolas & Soto-Acosta, 2010; Harrigan Ramsey & Ibbotson, 2011). Since data science, similar to cyber security and Internet of Things, appears to be seen by SMEs as an operational IT function (European Commission, 2019), these examples can be, albeit in a highly indirect fashion, argued in favor of the effects of becoming a data-driven business. In this report, nonetheless, we prefer to call for more empirical research on practical benefits SMEs enjoy from becoming a data-driven business. This research should consider the baseline data maturity level of the SME and the appropriate accompanying actions undertaken in order to grow towards a data-driven business.

To provide an inspiration for this further empirical research, this report focuses on one of the 'good practices’ put forth by the European Commission (2019), namely the case of the JADS SME Datalab.

Case study – JADS SME Datalab

The JADS SME Datalab is a local example in the Netherlands. It is affiliated with the Jheronimus Academy of Data Science (JADS) in Den Bosch, a collaboration between Tilburg University and The Eindhoven University of Technology. JADS as an institution focusses on the education, research and valorisation of data science as an academic field. The goal of the JADS SME Datalab is to help SMEs to create value with data.

In aiming to achieve this goal, the JADS SME Datalab offers tools and solutions to SMEs that allow them to integrate data science into their business and create value from data. Their workflow entails short projects, so-called “sprints”, of 60-80 hours, executed by a data science student at JADS and supervised by a data-science professional. The approach towards SMEs is 1) tailor-made; 2) accessible; and 3) practically feasible. The approach is tailor-made to the extent that any project the SME engages in with JADS SME Datalab is driven by the data maturity level and organisational challenges of the SME. There is no particular focus on industry, sector, type of project or data maturity level. The approach is accessible in the sense that SMEs pay a (relatively) low fee of €2.750, - and the team running the JADS SME Datalab ensures the SMEs can share any of their doubts, questions and challenges before starting a project. The approach is practically feasible with regards to the achievability of the projects. The JADS SME Datalab does not intend to start a sprint of which it is not certain the intended results can be achieved – unless explicitly and openly discussed with the SME. In Appendix 1 the practical workflow of the JADS SME Datalab is described, including the specific Data Maturity levels.

Hitherto, the JADS SME Datalab has helped over 100 SMEs at different Data Maturity levels with the implementation of data science in their daily business. The (initial) results the JADS SME Datalab is having with SMEs look promising, yet warrant a more sound and rigorous approach. 45% of the SMEs reported that the outcomes of the sprint they took with the JADS SME Datalab is valuable on a daily basis for the firm – and 90% of the SMEs reported they are continuing to work with and integrate data science in their organisation, both outcomes implying an increased data-driven mindset post-sprint[1]. Moreover, 94% of the SMEs working with the JADS SME Datalab would recommend the way of working and starting to work with Data Science to other SMEs[2].

With the approach of the JADS SME Datalab seeming to be successful from an SMEs perspective, the main question remains: how *much* has it really helped the SME? Has the sprint really increased their productivity, their financial position, their innovative capacity? Without a proper ex ante measurement and control group, this remains uncertain. Moreover, given the highly tailor-made approach of the JADS SME Datalab, it begs the question whether one can expect the sprints as "Data Science Interventions" to have a similar impact on any firm with a particular Data Maturity level.

Conclusion

Data science offers a lot of potential, however SMEs are lagging behind in adoption. By providing support to SMEs by experts complemented with the development of tools, guides and checklists, SMEs can be supported to get the value creation form data more in their reach. However, we need to keep in mind that becoming a data driven business is not a panacea, rather a supportive means to reach the end of better firm performance. An artisan bakery does not need to start using data science to improve its recipes. Moreover, automating administrative tasks would be a more obvious application of data science within an artisan bakery. In other words, there are many inspiring stories about how AI has helped a business, however there are a variety of simpler methods that also could lead to better firm performance, and in the end the entrepreneur must lead the business.

[1] N=11 SMEs out of 30 surveyed March 2019
[2] N=30 out of three evaluations rounds in the JADS SME Datalab

References

Aral, S., Brynjolfsson, E., and Wu, D. J (2006) "Which Came First, IT or Productivity? The Virtuous Cycle of Investment and Use in Enterprise Systems," in Proceedings of the 27th International Conference on Information Systems , Milwaukee, WI, pp. 1819-1840.

Barton, D., & Court, D. (2012). Making advanced analytics work for you. *Harvard business review*, *90*(10), 78-83.

Bean, R. A. N. D. Y., & Kiron, D. (2013). Organisational alignment is key to big data success. *MIT Sloan Management Review*, *54*(3), 1-6.

Brynjolfsson, E., Hitt, L. M., & Kim, H. H. (2011). Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance? SSRN. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1819486

Chen, R.-S., Sun, C.-M., Helms, M. M., & Jih, W.-J. (2008). Role negotiation and interaction: an exploratory case study of the impact of management consultants on erp system implementation in smes in taiwan. Information Systems Management, 25(2), 159–173.

Choudhury, P., Starr, E., & Agarwal, R. (2018). *Machine learning and human capital: experimental evidence on productivity complementarities*. Harvard Business School.

Cohen, W.M. and Levinthal, D.A. (1990). Absorptive capacity: a new perspective on learning and innovation*, Administrative Science Quarterly*, *35 (1)*, 128-52.

Davenport, T. H., Barth, P., & Bean, R. (2012). *How'big data'is different*. MIT Sloan Management Review.

Davenport, T., & Harris, J. (2017). *Competing on analytics: Updated, with a new introduction: The new science of winning*. Harvard Business Press.

Davenport, T. H., & Patil, D. J. (2012). Data scientist. *Harvard business review*, *90*(5), 70-76.

European Commission (2018) *Capitalising on the benefits of the 4th industrial revolution* Available online at https://op.europa.eu/en/publication-detail/-/publication/cf1793da-184c-11e8-ac73-01aa75ed71a1 [accessed on 04-06-2020]

European Commission (2019), *Skills for SMEs.* Cybersecurity, Internet of Things, and Big Data for Small and medium-sized enterprises. Available online at https://op.europa.eu/en/publication-detail/-/publication/82aa7f66-67fd-11ea-b735-01aa75ed71a1/language-en [accessed on 04-06-2020]

European Commission (2018), *Statistics on small and medium-sized enterprises.* Eurostat statistics explained. Available online at https://ec.europa.eu/eurostat/web/structural-business-statistics/structural-business-statistics/sme [accessed on 04-06-2020]

Jaakkola, H., Henno, J., Mäkelä, J., & Thalheim, B. (2019, May). Artificial Intelligence Yesterday, Today and Tomorrow. In *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)* (pp. 860-867). IEEE.

Gray, C. (2006). Absorptive capacity, knowledge management and innovation in entrepreneurial small firms. *International Journal of Entrepreneurial Behavior & Research*, *12*(6), 345-360.

Harrigan, P., Ramsey, E., & Ibbotson, P. (2011). Critical factors underpinning the e-crm activities of smes. Journal of Marketing Management, 27(5-6), 503–529

Hey A.J.G., Tansley, S., & Tolle, K. M. (2009). *The fourth paradigm: data-intensive scientific discovery* (Vol. 1). A. J. Hey (Ed.). Redmond, WA: Microsoft research.

Kim, G., Shin, B., & Kwon, O. (2012). Investigating the value of sociomaterialism in conceptualizing IT capability of a firm. *Journal of Management Information Systems*, *29*(3), 327-362.

Kart, L. (2015) How to get started with prescriptive analytics. *Gartner*

LaValle, S., Hopkins, M. S., Lesser, E., Shockley, R., & Kruschwitz, N. (2010). Analytics: The new path to value. *MIT Sloan Management Review*, *52*(1), 1-25.

Lopez-Nicolas, C., & Soto-Acosta, P. (2010). Analyzing ICT adoption and use effects on knowledge creation: An empirical investigation in SMEs. *International Journal of Information Management*, *30*(6), 521-528.
Liu, Y. (2014). Big data and predictive business analytics. *The Journal of Business Forecasting*, *33*(4), 40.
Lu, Y., & K.(Ram) Ramamurthy. (2011). Understanding the link between information technology capability and organisational agility: An empirical examination. *MIS quarterly*, 931-954.
Madrid-Guijarro, A., Garcia, D., & Van Auken, H. (2009). Barriers to innovation among Spanish manufacturing SMEs. *Journal of Small Business Management*, *47*(4), 465-488.
McAfee, A. (2002). The impact of enterprise information technology adoption on operational performance: An empirical investigation. *Production and operations management*, *11*(1), 33-53.
McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big data: the management revolution. *Harvard business review*, *90*(10), 60-68.
Manyika, J. (2011). Big data: The next frontier for innovation, competition, and productivity. Online available at *http://www. mckinsey. com/Insights/MGI/Research/Technology_and_Innovation/Big_data_The_next_frontier_for_innovation* [accessed on 04-06-2020]
Mitchell, T. M. (1997). Machine learning. 1997. *Burr Ridge, IL: McGraw Hill*, *45*(37), 870-877.
Najafabadi, M. M., Villanustre, F., Khoshgoftaar, T. M., Seliya, N., Wald, R., & Muharemagic, E. (2015). Deep learning applications and challenges in big data analytics. *Journal of Big Data*, *2*(1), 1.
Negash, S., & Gray, P. (2008). Business intelligence. In *Handbook on decision support systems 2* (pp. 175-193). Springer, Berlin, Heidelberg.
Ransbotham, S., Kiron, D., Gerbert, P., & Reeves, M. (2017). Reshaping business with artificial intelligence: Closing the gap between ambition and action. MIT Sloan Management Review, 59(1).
Rosenbusch, N., Brinckmann, J., & Bausch, A. (2011). Is innovation always beneficial? A meta-analysis of the relationship between innovation and performance in SMEs. Journal of business Venturing, 26(4), 441-457.
Stockdale, R., & Standing, C. (2004). Benefits and barriers of electronic marketplace participation: an sme perspective. Journal of Enterprise Information Management, 17(4), 301–311.
Van der Veen, M., van der Born, J., Smetsers, D., & Bosma, B. (2017). Ondernemen met (big) data door het mkb
Venturini, T., Jensen, P., & Latour, B. (2015). Fill in the gap: A new alliance for social and natural sciences.
Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J. F., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, *70*, 356-365.
Wegener, R., & Sinha, V. (2013). The value of Big Data: How analytics differentiates winners. *Bain and Company*.
Wimmer, H., & Powell, L. M. (2016). A comparison of open source tools for data science. *Journal of Information Systems Applied Research*, *9*(2), 4.
Yiu, C. (2012). The Big Data Opportunity: Making Government faster, smarter and more personal. Policy Exchange. London: 36. *About the Guest Editors*.
Zins, C. (2007). Conceptual approaches for defining data, information, and knowledge. *Journal of the American society for information science and technology*, *58*(4), 479-493.

Appendix 1 – JADS SME Datalab workflow and data maturity levels

The JADS SME Datalab uses the following steps to guide SMEs to become a data-driven business:

1. Data Maturity: Determine the organisations' data maturity at this moment;
2. Organisational Challenges: Determine the organisational challenges;
3. Check the available data: Reflect with regards to the organisational challenges the extent to which the organisation has a) the relevant data gathered in the right way; b) cleaned the relevant data to provide information able insights; c) whether this information reaches key decision makers; d) whether these decision makers use these insights to undertake the corresponding actions.
4. Organise a project: Determine a project which the JADS SME Datalab deems feasible and which the organisation would like to execute.

Step 1: Determine Data Maturity

Integrating a data-driven mindset in the organisation involves crossing five boundaries – and each of these boundaries relate to a particular data maturity level. The Data Maturity Level can differ between departments of an organisation (if the organisation is of a decent size), but in general SMEs have a pretty good insight when it comes to their biggest challenge vis-á-vis data at different aspects of the organisation. Within the Data Maturity Level of the organisation we refer to the following challenges:

1. **Digitizing data:** Data Science starts with ensuring the data is digital. Companies that are working with analogue or paper sheets and receipts not only have a less efficient way of storing and sharing data, in practice they also often lack the infrastructure and culture to integrate the latest technologies in say Machine Learning immediately.
2. **Cleaning data:** After the data is digitized, it is important that the data is put in a format which with it can easily be accessed for later analyses. Moreover, the data ought to be "clean" – with which we mean it is consistent, accurate and up to date.
3. **From data to information – Descriptive Analysis:** After the data is cleaned, organisations can use the data to gather information on their own operations and workflow. How many items have I sold in week 6 of this year? How many clients are from my neighborhood? Descriptive analyses help an organisation in getting a grip on the operational activities and enable an organisation to more actively take on a direction towards the future.
4. **From information to knowledge – Predictive Analysis:** Whereas descriptive analysis aids an SME in gathering information on the business processes, predictive analysis uses the historic data and insights in aiding the knowledge of an SME beyond the gut feeling. Predictive analysis allows an organisation to receive insights on what is expected to happen in the future with a particular probability and thus in turn helps the decision maker to take a particular course of action, as the uncertainty about future events will be further reduced.
5. **From knowledge to action – Prescriptive Analysis/Algorithms:** Once a company has reached a level of predictive analyses and trusts in the algorithms created, it can turn to an automated decision-making process on the basis of the available data. One can think of an automated algorithm that discovers on the basis of the data that a particular client has a higher probability of cancelling their contract soon – and therefore automatically send this client an offer or particular reminder of why he/she has to keep using the product/service.

The Data Maturity level refers to the readiness of an organisation of working with data. As mentioned, any organisation has data, but it depends on the data maturity to what extent this data can be put to insightful use. To put it simple: if an organisation has most data in an analog format (i.e. paper) it

requires a lot more effort to ensure an actionable insight is derived from the data than when the data is put digital, is cleaned and put in an informative and accessible format. To digitizing a business requires some investment, however if done in an adequate manner, this will unlock value in the long run by allowing fast, accurate and valuable insights to guide business decisions.
There are many Data Maturity Scans available for free. We recommend this scan which considers an organisation's culture, infrastructure, skills, processes and resources.

Step 2: Determine the organisational challenges

After the corona crisis, many organisations will face a general challenge of "becoming more efficient" or more specifically their focus is on "how to cut costs"? We recommend to reframe these challenges, focused on a concrete area of business (e.g. marketing or financial administration) and evaluate the status of the data within your organisation in light of the above data-driven decision-making model.
The below table 1. provides a framework for prioritizing the organisational challenges and mapping these on the data maturity of the organisation for each area of business. It provides an illustration of an organisation for which the financial administration and marketing is the main focus of the organisation in integrating a data-driven mindset. We can observe that the organisation in question is primarily helped by a) gaining insights from the available financial data; b) cleaning the marketing data. Of course, one can also imagine that this organisation is helped by organizing resources to ensure a predictive maintenance analysis, yet since there are other elements this organisation deems as more important, the general recommendation is to prioritize the data science integration accordingly.

Step 3: Check the available data

Even if the available data is digital or used for descriptive analysis in the organisation, it can be the case that the cleaning process of the data could use improvement. Moreover, many organisations are locked in to particular software via which it is difficult to extract the data in a format that it can be used for other analyses that the software is used for.
The data that you want to use to base your decisions on should be digital and 'clean', as you need to be able to trust the data if you want to unlock the value that lays in the potential of having the data make predictions and decisions for your business.

Step 4: Discuss the insights with a data science professional

One of the main issues SMEs have in talking with professional data scientists is that the average hourly rate a data scientist charges is relatively high. Having a conversation with a data scientist without completing the above three steps results in an expensive endeavor which you as an organisation can save already upon.
Nevertheless, at one point, you will have to discuss how to move forward with a professional. Preferably they can guide you in the direction of how to prioritize the next steps further. One may want to use the canvas here as the foundation for the conversation and ensuring that a) the data scientist has a thorough image of the organisation and potential challenges; b) the organisation has an image of the areas where the data scientist can help.