Data-Driven Decision Making: An Adoption Framework

by
Thilo Weigert

Master in Management, HEC Paris, 2016
B.Sc. Business Administration, University of Mannheim, 2013

SUBMITTED TO THE MIT SLOAN SCHOOL OF MANAGEMENT IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN MANAGEMENT STUDIES
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2017

Signature of Author: Signature redacted
MIT Sloan School of Management
May 12, 2017

Certified by: Signature redacted
Erik Brynjolfsson
Schussel Family Professor of Management Science
Professor of Information Technology
Director, The MIT Initiative on the Digital Economy
Thesis Supervisor

Accepted by: Signature redacted
Rodrigo S. Verdi
Associate Professor of Accounting
Program Director, M.S. in Management Studies Program
MIT Sloan School of Management

Data-Driven Decision Making: An Adoption Framework

by

Thilo Weigert

Submitted to MIT Sloan School of Management
on May 12, 2017 in partial fulfillment of the requirements
for the Degree of Master of Science in Management Studies.

Abstract

The ever increasing abundance of data and advancement of new technologies opens up new possibilities for companies in all segments and causes entire industries to rethink their business models. While there are a multitude of ways for companies to capture these new data-enabled opportunities, an obligatory first step is to make decisions more data-driven, and less guided by intuition. While the positive effects of data-driven decision making on a range of business performance metrics have been proven empirically, the adoption of corresponding practices is rapid but uneven across industries.

Based on examples of the manufacturing and healthcare industries, the rate, speed and effectiveness of a company-wide adoption of data-driven decision making is impacted by factors that include leadership commitment, organization and culture, selection of data, skill depth of both analytics users and consumers, and a company's ability to go beyond the mere collection and analysis of data. While in manufacturing, the main use cases revolve around incremental increases in efficiency, safety and performance, data-driven decision making in healthcare is still in its infancy and starting to uncover exciting use cases with the potential to impact millions of lives.

The more a company embraces data-driven decision making, the more its locus of decision making tends to become centralized. However, this is also largely dependent on the type of decision, the type of data used, as well as the decision's complexity, impact and idiosyncrasy. While there are decisions that can and will be fully centralized and automatized, there will also always be tacit decisions that will fully remain within humans, and decisions that are highly data-driven, but still allow for significant human value contribution.

Data powers insights, decision and actions, and we are only scratching the surface of the value that can be created, captured and redistributed through data-driven decision making.

Thesis supervisor: Erik Brynjolfsson

Title: Schussel Family Professor of Management Science,
Professor of Information Technology,
Director, The MIT Initiative on the Digital Economy

Acknowledgements

To my mentors and teachers.

Thank you for showing me that everything is possible.

To my friends, near and far.

Thank you for never ceasing to support, challenge and inspire me.

To my parents and grandparents.

Your love and dedication will always be the foundation of my accomplishments.

Table of content

1. Introduction

"We are drowning in information, while starving for wisdom."

– E. O. Wilson

The speed of data creation in the world is staggering: With 2.5 quintillion bytes (equivalent to 2.3 trillion gigabytes) of data being generated every day, the volume of data doubles every three years (Henke et al., 2016). This data comes from a rich variety of traditional and increasingly new sources, such as digital platforms, online content, transaction records, sensors, or mobile devices, leading to a complex and ever faster growing pool of structured and unstructured information. At the same time, technological progress has pushed the boundaries of what can be done with that data. By Moore's Law, for decades all elements of computing (storage, memory, processing, ...) have become more powerful and affordable, which may have now may reached a culmination where some elements are virtually free. In fact, some data scientists argue that it may now actually be "cheaper to keep data than to delete it" (Rose, 2016). On top of the abundance of data and the unprecedented computing power, data engineers have developed sophisticated software and algorithms that are often free to use, share and adapt for everybody. In sum, the value that is inherent in data has become accessible as never before.

Evidently, these developments have not gone by unnoticed by business decision makers. On the contrary, 85% of them believe that Big Data will revolutionize business the way the internet did (Accenture Analytics, 2014), 96% of global businesses surveyed by KPMG believe an effective data analytics strategy is important to the growth and future of their organizations (Toon & Collins, 2015), and 73% of respondents of a Gartner survey have invested or planned to invest in data capabilities in the next 24 months (Heudecker & Kart, 2014).

Analytics seems to have become a top-of-mind issue for managers, who recognize the impact of data and even have displayed or announced an intent to act upon it. However, capturing value from data is not straightforward and requires far more than a customary investment in IT technology. As I will analyze in detail, to go from data to "wisdom", companies need to embrace a culture of Data-Driven Decision Making.

2. Data-Driven Decision Making

2.1 Definition

In the following, I will follow Davenport's definition of Data-Driven Decision Making (DDD) as the "use of data and analysis to understand and manage a business more effectively" (Davenport, 2010). This includes the availability and the use of data to support decision making, e.g. for the creation of new products (Brynjolfsson, Hitt, & Kim, 2011).

2.2 Previous research

There is abundant anecdotal evidence where companies across all industries achieve improvements through the use of data or analytical methods (e.g. Glass & Callahan, 2014; Simon, 2013; Walker, 2015). However, research that reliably quantifies the effect of DDD on measurable business indicators on a broader and more general scope was hard to come by until the MIT Initiative on the Digital Economy carried out two widely recognized pioneering studies. The first one analyses survey data on the business practices and IT investments of 179 publicly traded firms in the U.S., and suggests a 5% higher output and productivity through the use of DDD (Brynjolfsson et al., 2011).

The second one is a large-scale study using data of roughly 50,000 U.S. manufacturing establishments collected by the U.S. Census Bureau for 2005 and 2010 (Brynjolfsson & McElheran, 2016a). This representative survey reveals an almost threefold increase in the use of DDD practices in U.S. manufacturing between 2005 and 2010, from 11% to 30%, again indicating a rapidly increasing awareness of DDD among decision makers. The likelihood of DDD adoption among plants surveyed seems to be a positively correlated with plant size, affiliation to a multi-unit firm and investment in IT. Notably, the authors establish a causal relationship of a 3% increase in productivity for plants that have successfully adopted DDD practices. This substantial performance increase can be compared to the expected effect of an investment in IT of $60,000 per employee over a five-year period.

The significant impact of DDD on business metrics is also supported by other accounts. According to PwC's "Global Data and Analytics Survey" of over 2,000 senior executives across a wide range of industries, executives in data driven organizations are three times more likely to report significant improvements in their decisions (Blase, DiFilippo, Feindt, & Yager, 2016). In Accenture's "Big Success with Big Data" survey of over 1,000 companies that have completed at least one big data implementation, 92% are satisfied with the results (Accenture Analytics, 2014), which again validates the value of data-driven methods.

2.3 Motivation and research question

Despite being a relatively young field of research, the previously highlighted findings leave no doubt: Adoption of DDD has a significant impact on business performance on plant and company level. Surveys also indicate that the value of analytics has been recognized by business leaders across a wide range of geographies and industries. Correspondingly, the rate of adoption is staggering, as noted above. Still, around two thirds of U.S. companies have not yet adopted DDD (Blase et al., 2016), and there is significant heterogeneity among those who did (Brynjolfsson & McElheran, 2016a).

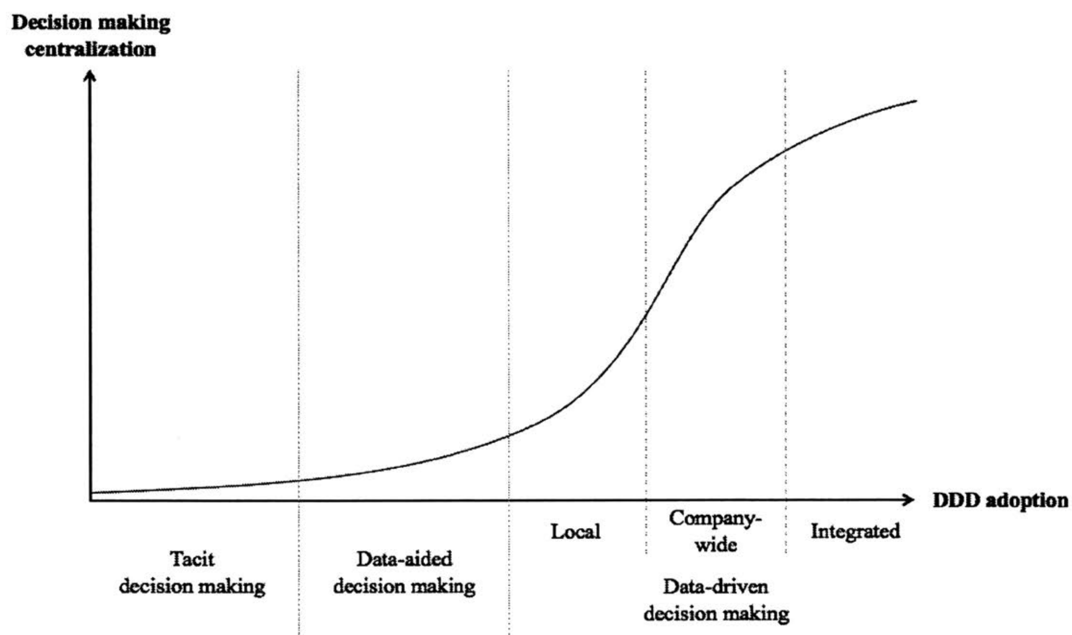
This prompts the following questions: Do industries differ in their stage of adoption? What are the enabling factors and best practices of DDD adoption, and how does DDD adoption change the locus of decision making within a company?

2.4 Structure of the paper

The paper is structured as follows: In Chapter 3, I will introduce a framework to differentiate the different stages of DDD adoption and their impact on the locus of decision making. Applying this framework, Chapter 4 then explores recent examples and cases studies of DDD adoption across a selection of industries. Chapter 5 analyses enabling factors, best practices and challenges of DDD adoption, looking at both the examples studied before as well as observations of other companies. Chapter 6 summarizes my findings and analyses, and provides recommendations for action and further research.

3. Adoption framework

Extrapolating empirical evidence and previous research (Hayek, 1945; Jensen & Meckling, 1976; Kanamori & Motohashi, 2006; Wyner & Malone, 1996; Brynjolfsson & Mendelson, 1993), I will introduce a simple framework to link a company's level of DDD adoption to the degree of decision making centralization.



Conceptualized framework: Impact of DDD adoption on decision making centralization

A high degree of decision making centralization is often correlated with a standardization of solutions and tools across functions and business units, and indicates potential for synergies by pooling expertise of a variety of functions. Low decision making centralization on the other hand can often be found in contexts that require differentiated approaches, where one-dimensional but deep domain knowledge is highly valuable.

In this framework, I propose to differentiate between three broader levels of DDD adoption: Tacit, data-aided and data-driven decision making. As the analysis of selected case studies will confirm, there are a number of factors that are highly correlated with a company's revealed adoption of data-driven decision making practices. For instance, companies that have a low level of decision idiosyncrasy, high decision making frequency and work with machine-readable data are more likely to make use of data-driven decision making practices.

Tacit decision making

Some business models rely on intuition- and experience based decision making to be effective, regardless of a company's intentions to introduce a more data-drive approach. The data and information needed are often "stored" exclusively in humans' brains, muscle memory or subconscious mind, and are thus hard to quantify or codify, or even describe and share. In other cases, tacit decisions involve creativity or a very large number of input factors that would require enormous computational power to optimize. This is where humans beat computers by using their instinct, which "stores" millions of years of comparable situations. What these decisions have in common is that they are inherently difficult to standardize, scale, not to mention centralize.

Data-aided decision making

The second cluster of decisions as categorized by their level of DDD adoption are data-aided decisions. These are decisions that are largely based on insights generated through data, but critically depend on a human input factor. Evidence for this category can be derived from Brynjolfsson and McElheran's study of the U.S. census bureau data on manufacturing (2016b). The reasons to include human contribution in the decision making process are varied. In some cases, adding tacit knowledge to decisions can a lot of value to their outcome (e.g. including customer interaction in the store level repurchasing process of 7/11 in Japan, (Ross, Beath, & Quaadgras, 2013)), in others, a human layer is needed to mitigate a potentially outsized risk or impact of misclassification (e.g. the case of real time in-flight analytics as discussed in the case of Rolls-Royce below). These decisions can be centralized to some degree, for instance by providing dashboards, templates or protocols, and naturally the share of value contribution between agents and centralized data function can vary significantly.

Data-driven decision making - Local

Having introduced the umbrella concept of data-driven decision making (DDD) already above, it is necessary to look one level deeper into the different stages of adoption within a company. The first level for a company is often to carry out selected pioneering data-driven projects in order to gauge impact and return on investment. If the outcomes are satisfactory, often the successful project gets institutionalized and eventually DDD gets adopted on an isolated, functional level.

Frequently, pioneering project are within the area of marketing, since the impact and effect of marketing campaigns can be supposedly easily be analyzed as standalone project. Especially within marketing, there have been significant developments in the automated and standardized use of analytics and data e.g. for a sophisticated personalization of communication and pricing. On an overall basis however, the level of centralization in the case of local DDD adoption is still rather moderate.

Data-driven decision making - Company-wide

As described in more detail below, one of the main ways DDD creates value in an organization is through the enabling of cross-functional collaboration. At the same time, many concrete measures of collaborative value creation are even more effective when they are accompanied by efforts to centralize some of the decision making and standardize best practices. For instance, by aligning teams on the same set of performance metrics and integrating legacy systems on one common platform across all groups and business units (which a company as large as GE managed to implement), significant synergies and cost-savings can be levered beyond the benefits of mere collaboration and idea sharing. For these reasons, centralization of decision making really takes off at the stage of company-wide DDD.

Data-driven decision making - Integration of supply or value chain

Some companies that have successfully adopted DDD across their organization have come to realize that, depending on the industry, there may still be significant value to be captured through the integration of suppliers or customers onto their analytics platform. The example of Walmart shows that this has the potential to significantly reduce friction and transaction costs, increase visibility on replenishment cycles and inventory, and enables faster decision-making at a significantly higher level of certainty. Evidently, in these scenarios, the sheer amount of data - albeit structured and mostly first hand - and decision variables creates an analytical complexity that requires advanced optimization algorithms. Again the example of GE demonstrates that problems of this scale can be most efficiently tackled through the highest degree of centralization.

Interestingly, through computational and analytical power, these centralized decision making facilities sometimes even have more insight into events at the local level than local agents themselves, e.g. the effectiveness of promotions on a SKU level at Walmart. However, many industries are not suitable for such a deep integration, and only few companies have the necessary bargaining power over their suppliers to impose, govern and extract value from centralized decision making at the supply chain level.

4. Case studies of Data-Driven Decision Making

4.1 Manufacturing

4.1.1 Rolls-Royce: Predictive maintenance

Situation

Rolls-Royce is an aircraft and ship engine manufacturer, that had sold engines as standalone product, and later coupled with maintenance contracts. However, due to increased pressure on prices, this traditional business model was threatened by commoditization.

DDD adoption

Rolls-Royce realized there was significant of value to be uncovered through the use of analytics. Placing sensors in the engines that send real-time data to monitoring stations on the ground, the company was able to collect and analyze in real time large amounts of operational data. This allowed Rolls-Royce to get a better visibility of the factors that influenced the performance, lifetime, efficiency and reliability of their engines. That new understanding enabled the users of Rolls-Royce's engines to make a broad range of data-driven decisions with material impact, such as to anticipate performance issues, forecast maintenance, predict engine lifecycle costs, optimize consumption, prevent safety failures, and classify defects in real time.

The data generated is also used to enable data-driven decisions in the design and manufacturing process, where it reduces the number of simulations and experiments needed in product development and quality assurance.

Outcome

To capture some of the enormous value that these data-driven decisions deliver to their customers, Rolls-Royce decided to adapt their business model. Introducing Total Care, Rolls-Royce doesn't sell engine and maintenance contracts anymore, but essentially leases out and services engines, and bills by the hour of uptime. Reducing the risk for their customers and better aligning incentives, this offering proved to be popular with customers and developed into a substantial competitive advantage for Rolls-Royce.

Type of data and decision-making locus

For Rolls-Royce, on the one hand, the cross-functional use of operational engine data unleashes significant synergies, leading to cost and time savings. On the other hand, this new "single source of truth" clearly consolidates and centralizes decision-making. Thereby, scores of roles are rendered redundant, as the 2,600 jobs that were cut "due to increased efficiency" (Marr, 2017).

The wealth of new data enabled also Rolls-Royce's customers to change their decision making process. Key decisions, as listed above, can now be grounded on structured data and become standardized and gradually automated. However, given the complexity and critical nature of real time in-flight decisions, that can have catastrophic consequences for the safety of hundreds of people, these are decisions that are highly data driven but cannot be completely automatized or standardized. Due to the high levels of impact, complexity, and case idiosyncrasy, human judgment will remain necessary, and despite the centralized elements the decision-making process as a whole should be categorized as data-aided.

Correspondingly, the format of information used is two-folded. The first hand data generated by the many sensors in the engines is optimized for real time use and analysis, while the ultimate decisions made by analysts, crew and pilots are influenced by experience, personal preferences and intuition, and sometimes even extend to considerations involving values and ethical components.

4.1.2 GE: The industrial internet

Situation

GE is one of the largest and oldest companies in the US and has traditionally sold "big iron" industrial goods that are used in aviation, manufacturing, energy production and distribution, healthcare and transportation. The company faced similar challenges to Rolls-Royce: a slow but steady commoditization of its core business.

Given the scale of their operations (GE equipment is used in the generation of a quarter of the world's electricity supply), the company is always trying to eek out incremental small improvements in efficiency that can yield big effects. Again similarly to Rolls-Royce, an untapped source of opportunity was the use of data: 63% of GE's customers had indicated that their machines were connected to networks, yet the data remained unused thus far (Lakhani, Iansiti, & Herman, 2014).

DDD adoption

Introducing their "Industrial Internet" initiative, GE envisioned an open, global network that connected machines, data and people to increase operational productivity and efficiency e.g. in manufacturing, transport, finance and aviation. GE thus imagined the Industrial Internet to be a subset of the Internet of Things for industrial equipment. Given their large installed base of equipment, GE controls of the biggest industrial data sets. Generating large amounts of data with sensors in machines, GE would be able to closely monitor the effects and correlations of even minor changes. GE's strategy to capture some of that value was to sell data-enhanced machines on one hand, and using their customers' data to provide outcomes-based services aimed at improving operational performance on the other hand.

Outcome

Early results confirmed the "power of + 1%" (Lakhani et al., 2014), where small increases in operational efficiency and safety had a remarkable monetary impact, leading to millions of annual savings for some of GE's bigger customers. While 1% is already contributing significant monetary value, there is still potential for more. According to the pioneering study on the effects of DDD adoption in U.S. manufacturing firms mentioned earlier (Brynjolfsson & McElheran, 2016b), performance and efficiency increases of 3% have been observed on average through the adoption of data driven decision making techniques. Enhanced by GE technology that has been developed for the purpose of DDD, this number is only expected to go up. Similarly to Rolls-Royce, GE's CEO Immelt therefore foresees a complete overhaul of the traditional business model, where machines are given away for free and companies monetize on services and data coming out of these machines. An essential intermediate step for that is to establish outcome guaranteed maintenance agreements, where GE can leverage the power of their data and insights to carry out data-driven maintenance action, instead of the currently predominating "break-fix" contracts (Lakhani et al., 2014).

The main impact however of GE's advocating of DDD is the development of a platform of cross-industry analytics solutions. Customers of GE in industries as diverse as aviation, energy and healthcare are offered access to the same set of shared tools to make data-driven decisions about monitoring and maintenance for all of GE's industrial technologies.

Type of data and decision-making locus

Prior to rolling out the cross-industry analytics platform, GE has undergone a rigorous restructuring of their internal decision making processes. As a century old company, GE had developed into a dispersed company with siloed and independent functions and teams, leading to substantial heterogeneity in their often incompatible IT and technical products. With the launch of the Industrial Internet initiative, GE built a central global software center to ensure software uniformity across all businesses. This standardization is crucial to reap the benefits and synergies of the cross-industry analytics platform (Evans & Annunziata, 2012).

Another evidence of centralization of data-driven decision making is the pooling of virtually all data science capabilities in a new $1 billion analytics headquarter in Silicon Valley. All functions and teams are encouraged to make frequent use of their services, which effectively puts most decision-making across the company in one place (Fitzgerald, 2015).

Clearly very beneficial for the smooth internal transition and adoption of DDD practices are the characteristics of the data that GE generates and uses. As opposed to the healthcare companies analyzed in the next section, GE can rely on first hand data that is enriched by industry standard external sources. Given the nature of the industries GE operates in, the data is very structured and lends itself well to large scale analyses and centralization.

GE also encourages their customers to monitor data more centrally, e.g. an Oil and Gas customer was advised to develop central analytics capabilities instead of local monitoring on each rig. Despite these efforts and progress towards decision-making centralization, due to the large size, complexity and business impact of many of their accounts, GE still relies on intimate knowledge of a customer and its specific industry sector - thus on a not substitutable human contribution in decision-making.

4.3 Healthcare

According to Russel Glass (2014), “Healthcare may provide the most promising opportunity for big data's transformative powers”. This is due to an array of factors: first, many of the initially described technological advancements of capturing, storing and analyzing data are even accelerated in a healthcare context. For instance, the cost of sequencing a genome has dramatically outpaced Moore’s Law and decreased by a factor of 10,000 from 2008 to 2015. Second, the healthcare industry is experiencing rapid change on a global scale through changing demographics, longer lifespans and shifts in lifestyle and habits. Third, there is seemingly unlimited capital pouring into countless innovations emerging in this industry, of which many are related to data, such as wearable technologies. The resulting rapidly growing volume of healthcare related data presents unprecedented opportunities to understand and analyze some of mankind’s biggest remaining riddles.

4.3.1 Apixio: Digitizing health records

Situation

However, this vast amount of medical data also presents significant challenges: Much of the data about patients’ health is caught in diverse systems with different formats and limited compatibility, and essential information is often buried in hand-written notes or scanned documents. Thus, the problem in healthcare is not lack of data, but the unstructured nature of its data sources, which to some estimates amount to as much as 80% of medical and clinical information (Gough & Rogers, 2016).

Data-Driven Decision Making

By consolidating the vast amount of unstructured data via text recognition methods and other sophisticated machine learning applications, and integrating it with other related data sets such as billing and administrative records, Apixio enables healthcare providers to access the large pool of clinical information and make more data-driven decisions.

Furthermore, by consolidating and applying analytical methods to clinical data on a larger scale, Apixio enables healthcare providers to identify seasonal or geographical patterns in the medical history, and detect correlations between medical conditions and other factors.

Outcome

The impact of providing the healthcare industry with 80% more data is enormous. Two immediate effects can be highlighted. First, healthcare providers are now able to understand a patient's risk profile better. While previously, predicting patients' conditions and cost of treatment involved a lot of manual labor of searching patient charts for documentation, through Apixio this process is now truly data based. This reduces the time and cost per patient classification and improves accuracy by up to 20% (Apixio, 2015).

Second, Apixio's analytics are an important contribution to the healthcare system as a whole: On the one hand, they boost the shift from reactive to predictive healthcare, on the other hand, they enable a much more personalized treatment instead of traditional medicine based on studies and randomized clinical trials.

Naturally, as patient health data is highly sensitive, healthcare and insurance providers are reluctant to share and contribute relevant and current datasets. Hence, an early emphasis on security and data encryption is key for any DDD adopter in healthcare. Another key component of driving DDD adoption in the healthcare industry is understanding the human component.

Concerned of becoming redundant, many practitioners tend to be rather skeptical towards the use of data and analytics at first. As Apixio's CEO Darren Schulte points out (Marr, 2016), it is essential to focus on how data will solve problems and achieve actual results, instead of showing "a lot of slick dashboards which are not very helpful to them".

Type of data and decision-making locus

Transforming unstructured into structured data, Apixio presents an interesting case for the impact of data type on DDD adoption and the resulting decision-making locus. The differentiation between implicit and explicit decision making is often made through the criterion of data machine readability. This happens to be closely related to the boundary that the company is trying to push for the entire healthcare industry by making previously inaccessible sources of data (e.g. handwritten notes) available for use and analysis.

The decision-making locus of healthcare providers using Apixio's data contributions is twofold. On the one hand, when quantifying a patient's estimated cost of treatment based on previous health records and other medical information, decisions are made in a centralized function. On the other hand, when carrying out the arguably much more meaningful decisions concerning actual health intervention and prevention measures on a patient-individual level, physician are empowered and encouraged to use their wealth of experience in addition to the data provided. Therefore, this represents a good example of a shared and data-aided decision making process, drawing on both explicit and tacit information.

4.3.2 Aetna: Personalized care

Aetna is a healthcare insurance provider that has successfully employed DDD methods to offer predictive analytics for personalized care.

Situation

As insurer, Aetna is immediately impacted by the rising healthcare costs. Two of their main sources of expenses are cancer and metabolic syndrome. That is a group of risk factors that raise the risk for heart disease and other health conditions, such as diabetes and stroke, and together account for ~20% of healthcare costs in the United States (Steinberg, Church, McCall, Scott, & Kalis, 2014). Although one in three Americans suffers from metabolic syndrome, treatment of syndromes so far has proven ineffective: according to Michael Palmer, former Chief Innovation & Digital Officer at Aetna, "only 102 patients out of thousands" would see improvements from the rather generic advice of "eating right and exercise" (company information).

Data-Driven Decision Making

Drawing on their rich dataset of member health information, Aetna built a model for physicians to use that allows to predict which new condition a specific member will likely exhibit next and when. Contrasting many other healthcare analytic models that only reveal data associations, Aetna was able to highlight underlying causal relationships (Baldenko, 2014). Based on a member's individual profile, Aetna's approach was to personalize both a patient's risk of suffering a detrimental health event as well as the most effective treatment to prevent that event (Kolodziej, 2015).

Outcome

As a data-rich competitor, Aetna is exploiting the fact that most other medical data is locked up unstructured or in inaccessible formats. The company thus leverages its sophisticated predictive analytics capabilities that saves costs and protects members' health as a competitive advantage. Benchmarked against the base case of providing the same treatment to all at-risk patients, the one-year probability of having metabolic syndrome could be reduced in nearly 90% of people in a 37,000 pilot test (Steinberg et al., 2014).

Type of data and decision-making locus

As opposed to the case of Apixio, where the actual decision-making on the basis of patient data still falls within the physicians' competence, in the Aetna model doctors merely execute based on recommendations of Aetna's centralized predictive analytics function. This high degree of centralization is also heavily supported by the structured, quantified and explicit characteristics of the data that Aetna uses and provides its recommendations in.

4.3.3 Industry summary and outlook

There have been promising initiatives in DDD adoption in healthcare, two of which I briefly outlines above. However, there are still major challenges ahead to capture all of the massive value that can be realized through the use of DDD in healthcare. Some of them include a lack of incentives, regulatory restrictions and the mentioned reluctancy to share data. Also, the industry being characterized by decade-long development cycles, many of the players in the healthcare industry are known for their slow and careful adoption of innovative processes and technologies. While these factors are causal that to date, according to the McKinsey Global Institute, only an estimated 10 to 20 percent of the data-related value in the U.S. healthcare sector has been realized (Henke et al., 2016), the big potential of DDD in healthcare is undeniable, as exemplified by the two case studies above.

5. Analysis

5.1 Capturing value through DDD

Building on the previously explored case studies and including generalizable insights from other sources, I will briefly discuss some of the most common levers that companies have identified to capture value through DDD.

Routinely, decision makers across a broad range of industries state that the top two motivators of their Analytics campaigns are to improve customer experience and process efficiencies (Heudecker & Kart, 2014). As discussed above, the use of Analytics opens up unprecedented opportunities for companies to better get to know their customers' preferences, willingness to pay, consumption habits, and even personality traits. Hence, DDD enables companies to segment customers on a very granular level, and tailor their product offering, marketing and pricing correspondingly.

In industries where product offering and value proposition are more and more converging, business processes are among the last remaining areas that allow for innovation and differentiation (Davenport, 2006). The need to operate at maximum efficiency is often further reinforced by competitive pressure or labor costs disadvantages. The intensified (and often times, first time) use of operations data allows for a new take on a company's internal processes, often uncovering significant inefficiencies. This is sometimes referred to as "process mining" and not by chance an industry that is home to some of the highest growth companies (e.g. Celonis, Fluxicon, …).

Previous research also suggests that a main factor of DDD value creation is the facilitation of "replicating good ideas across the organization" (Brynjolfsson & McAfee, 2008). Still, because of the previously discussed challenges of transforming a company towards a truly data-driven organization, most examples of DDD can still be observed in divisions and functions rather than across companies (Ross et al., 2013). However, the advancements of tools for internal collaboration as well as the rising awareness at all organizational levels to view cross-functional collaboration as "mission-critical activity" (Davenport, 2010) clearly benefits capturing the value of DDD through collaboration.

Evidently, in a company that embraces DDD in the entire organization, different functions have their own specific DDD use cases to capture value. For instance, the use of Data and Analytics enables HR to make better hiring, compensation and promotion decisions, supports supply chain and inventory optimization, and provides R&D with a host of new tools to improve quality, safety and customer fit (e.g. through small-scale experimentation).

5.2 Company-level factors enabling value capturing

After looking at the industry-level implications on a company's ability to capture value through DDD, I will briefly discuss the ramifications of a company's stage of maturity. It becomes evident from the analyzed examples as well as through further research that growth and mature stage both have advantages and disadvantages for companies.

Growth-stage companies are said to have a "clean slate" relative to Data and Analytics (Davenport, 2014): They do not need to worry about integrating big data with traditional sources of information, cater for compatibility of different versions, or devise solutions to merge new technologies with traditional IT infrastructures and architectures. Also, the barriers to adoption, such as a required initial investment or training, are a lot lower. GE has found that the quickest adopters of their cross-industry analytics platform are those companies with fewest existing software-based services in their portfolio. Evidently, also the cultural barrier of implementing DDD is often nonexistent, as many growth stage companies are "born digital".

This is supported by the research of Brynjolfsson, Hitt, and Kim (2011), who find empirical evidence for a negative correlation between DDD and firm age, and conclude that "younger firms are more likely able to adopt new innovations such as business analytics".

On the other hand, practitioners seem to prefer to work with data that has "some history" (Almquist, Senior, & Springer, 2015), as context enables data scientist to better recognize patterns and trends in the data, and as a consequence build better models. Also, some data will only unfold its value after some years, which is why Jeff Bezos famously claims, "We never throw away data" (Davenport, 2014).

5.3 Best practices of effective DDD adoption

While the perceived DDD affinity of many companies is high (almost 75% of companies surveyed by the U.S. Census Bureau report being "relatively intensive both in collection and use of data", Brynjolfsson & McElheran, 2016b), many analytics investments still fail to yield a return (Ross et al., 2013). In the following, I will present some best practices how companies effectively capture value through DDD, again as identified by companies previously studied and others.

Data to insights, insights to action

Some companies seem to see DDD as a miraculous black box that will create value just by itself. Typically, they either accumulate large amounts of data without using it, or apply complex models to historical data and get the "what" without "why" or "what next". As Davenport observed, "the link between analytics and decision making needs to be relearned" (2010): companies need to go from data to insights, and from insights to action.

Especially in light of the ever-growing amount of data, Analytics always needs to follow a hypothesis to avoid losing time, resources and momentum by erratically digging in data. Thus, even companies that have adopted company-wide DDD narrow down their focus when deciding on their resource-intensive analytics efforts (Davenport, 2006) and decide on a single source of truth (Dalle Mulle & Davenport, 2017).

Selection of Data

Now that collecting, storing and processing data is almost free, aspiring DDD adopters are tempted to start by measuring, quantifying and collecting data of everything. Yet, as a proverb among data scientists says, "Data is like garbage - you'd better know what you are going to do with it before you collect it." According to Ross et al. (2013), the single most important reason for analytics initiatives not to be successful is that companies fail to handle well the information that they already have –often is a sizeable trove of unused legacy data, such as decade-old ERP or CRM systems. However, the selection, procuration and integration of data to use in decision-making is as much science as the analytics processes itself. Successful DDD adopters typically enrich internally generated data, such as transaction or supply chain, with well-curated external sources of insights.

Leadership

A transition to DDD is a company-wide effort and affects the culture, processes, incentives and behaviors at all levels of the organization. However, as any major shift in paradigm, DDD adoption is most likely to succeed with senior executive buy-in and backing. As Davenport (2006) observed, "lower-level people [than the CEO] lack the clout, perspective and cross-functional scope" to exert a lasting influence on the transition.

However, before championing a company-wide DDD movement, leaders first need to make sure that they are embracing a data-driven leadership style themselves. In the traditional approach, senior leadership is expected to employ their experience, judgment and intuition to make decisions, and typically request data only to back up their pre-conceived notions. Alarmingly, this confirmation bias can also be easily carried forward to and be reinforced within a data-embracing organization: the more abundant and sophisticated data and analytics become, the easier it is to support multiple, sometimes contradictory, perspectives with the same set of data. As Roland Coase, Nobel Prize laureate in Economics, quipped: "Torture the data, and it will confess to anything." Hence, it is essential that the decision-making process involves the use data from the beginning.

Nonetheless, in a data-driven organization, the value of a leader's experience, judgment and intuition has all but vanished: as discussed above, it is vital to formulate, test and refine a hypothesis, which is "an intuition about what's going on in the data you have about the world" (Davenport, 2013b), until it is validated or refuted. It is exactly this openness towards unprejudiced, data-driven experimentation that presents the biggest challenge for leaders in a data-driven organization, but also their most important element to instill in a DDD culture.

Organization and culture

While leadership has a crucial role to champion a data-driven culture, a company's transition can only succeed if existing people, workflows, processes and incentives also receive adequate training and transformational treatment. Right from the offset, an organization must be structured in a way that efficiently aligns information flows with access and decision rights (Lavalle, Hopkins, Lesser, Shockley, & Kruschwitz, 2010). It has proven effective to set up cross-disciplinary teams, where data scientists collaborate with a variety of other players, benefit from their industry and function expertise, and "ensure that big data is matched by big analytics" (Davenport, 2013a). This also facilitates and enhances the vital cross-functional sharing of information, tools, data and insights, that has been identified to be at the core of value creation through DDD.

As seen in the cases analyzed, the centralization of decision making through the adoption of DDD often leads to substantial overhaul of existing decision-

5.4 Challenges for effective DDD adoption

Companies that have committed to adopting DDD and following the best practices often face a number of common challenges. These preclude organizations from generating a return on their often significant analytics infrastructure investments and include data, people, mindset and regulatory framework.

Data

Data is not easy to handle well. Despite evolving technologies, alone handling the sheer volume can cause problems for data-inexperienced organization. On top of that, data will come most likely from different and incompatible sources, and can be unstructured and messy. These data related challenges may be a reason why today less than 0.5% of all data is ever analyzed and used (Marr, 2017).

Skill gap: domain expertise and data literacy

Data scientists, famously referred to as the "sexiest job of the 21st century" (Davenport & Patil, 2012), come in different variations and will be in great demand: according to an estimation of the McKinsey Global Institute, the United States alone will face a shortage of 140,000 to 190,000 people with "deep analytical skills" by the end of the decade (Manyika et al., 2011).

However, as Garry King of Harvard University stated, “Big data is not about the data”: When data and analytics will eventually become mainstream capabilities and functions, it is domain expertise and other complements to data that will become valuable and allow companies to differentiate themselves by spotting the signal in the noise (McAfee & Brynjolfsson, 2012).

What might appear unassuming next to data prowess and domain expertise but potentially present an even greater skill gap are data literate consumers and users of analytics on all levels. Companies need to remove all barriers and friction between the users and an effective use of analytics. GE leads this effort by introducing a search bar to their cross-industry analytics platform where users can ask questions in natural language, and are thereby able to access the treasure chest of GE’s data more easily and intuitively. For managers to be able to tap the value created through DDD and make effective decisions, they need to be fluent the languages of both data and business. From the viewpoint of 200 of the world’s leading data scientists, there is also a significant translational gap to be closed between analytics and business, as representatives from both sides feel that the respective other side is not focusing on the key questions and issues (Manyika et al., 2011).

Trust in Data-Driven Decision Making

This implicit disregard touches on an important point: DDD adopters need to foster trust in data among all internal customers and users of analytics across their organizations. On the one hand, this includes trust in the actual decision-making capabilities of analytics. On the other hand, a consequent transition towards DDD will provide companies with unprecedented visibility into the activities of their employees and customers. Countering this “creepiness factor” (Davenport, 2013a) and making sure that also those are comfortable who are affected by analytics without having the corresponding data skills will be one of the main barriers towards effective DDD adoption. Certainly a part of the solution will be to offer training and display maximum transparency into the use of behavioral data.

Another aspect of trust in DDD are cultural components: people with different backgrounds often have very different perceptions and attitudes towards privacy and the use of their data.

Security and regulations

This also becomes apparent when looking at the divergence in law and regulations e.g. regarding the storage and use of third person data for some geographies (e.g. U.S. vs E.U.). As this is a significant source of uncertainty and still a very new field of jurisdiction, more convergence can be expected.

6. Conclusion

6.1 Summary of findings

The analyzed case studies, other examples and previous research have revealed a number of interesting insights about the adoption of data-driven decision making in various industries as well as its impact on the locus of decision-making within companies. First off, the impact of DDD is widely recognized among many companies across all industries, not anymore only among tech companies and the early adopters of which I highlighted some. However, there still is a discrepancy between decision makers' enthusiasm about DDD, the de facto adoption, and the use of best practices that make an effective adoption more likely. Yet, the vast majority of DDD adopters yield positive results and are satisfied with the outcomes of their first DDD initiative, which is a strong indicator for the far-reaching potential of DDD.

The bottleneck for a more widespread adoption are not technological barriers anymore, but the mindset, legacy decision-making habits and data literacy of makers and recipients of decisions. Both the studied examples as well as a vast body of other research point to the enormous potential of DDD in literally all industries. While traditional industries like manufacturing generally lag behind in adoption as compared to "digital native" industries, they already benefit from DDD in an earlier adoption stage and can expect to reap even larger benefits the more they transition. Sometimes DDD adoption even pushes and blurs the boundaries of industries. Generally, the more a company adopts DDD, the more its internal locus of decision-making shifts towards more centralization. However, there are and will always be decisions that cannot be automated or codified and thus centralized through DDD, and also decisions that are heavily data-driven, but still materially depend on human involvement or judgment. In general, it is apparent that value capturing DDD is not trivial, and for some adopters it "is often better to have a lot of little data done right than big data done wrong."

6.2 Critique of the methods used in the study

In my analysis of the case studies, I did not get to use previously unpublished or proprietary data, but mostly had to rely on managers' accounts and synthesis of previously conducted analyses in this field. However, I did manage to have conversations with protagonists in the respective companies as well as industry experts, and included these insights were adequate.

Also, the paper only explores cases of successful DDD adoption, while an analysis of failed cases would be also very insightful. Unfortunately, however, they are harder to come by: on the one hand, failures are much more unlikely to publish, especially in a field that enjoys so much attention and coverage as data and analytics currently does. On the other hand, failures in this novel field are hard to identify, as "success" depends a lot on target setting and timeframe.

6.3 Suggestions for further research

Keeping in mind the tripling of DDD adoption in manufacturing between 2005 and 2010 (Brynjolfsson & McElheran, 2016b), or how much has changed between two studies of the McKinsey Global Institute in 2010 and 2015 (Henke et al., 2016; Manyika et al., 2011), an update of my study, assumptions and conclusions will be required soon. Also, as this paper aimed more at providing a general overview of DDD adoption, the scope of industries and depth of differentiation can still be extended notably.

Further research can also focus on an analysis of the initially mentioned discrepancy between enthusiasm in surveys and actual adoption rates. Are the shifts in actual decision-making temporary or sustainable? Have managers embraced DDD and implementation simply lags behind, or, as Dan Ariely of Duke University puts it, is DDD in fact more a like phenomenon where "everyone talks about it, nobody really knows how to do it, everyone thinks everyone else is doing it, so everyone claims they are doing it"?

6.4 Final recommendation

In my opinion, of the best practices examined above, the one with the most impactful long term effects is the openness to share data and resources. Legacy companies coming from traditional industries might not feel comfortable at first with making key information about their business processes available and thus putting them at reach of competitors, partners and customers. However, there are many examples that these concerns are mostly unjustified.

An example for the benefits of sharing key technology and data with the public is the online platform Kaggle. It hosts many interesting datasets across a wide range of ranges for everybody to analyze and experiment with (e.g. Hillary Clinton's published emails, World Bank country data, ...), but also allows companies to host competitions. One of the most popular ones was the Netflix Challenge, where Netflix promised $1m to the team that could improve its recommendation algorithm by 10%. This shows that Analytics can be both open source and a strong and sustainable source of competitive advantage.

The data and analytics community is growing fast and is very diverse. It is home to researchers, individual contributors (developing widely used algorithms and open source packages), corporate contributors (IBM, Twitter, Airbnb, ...), and new forms of education (bootcamps, micro-degrees, online certificates). Its culture of active participation and sharing is good news for all aspiring individual or corporate adopters of DDD, and can serve as a platform for successful DDD adopters who are looking to give back to the community and at the same time continue to evolve.

References

Accenture Analytics. (2014, April). Accenture Big Success with Big Data Survey.

Almquist, E., Senior, J., & Springer, T. (2015). Three Promises and Perils of Big Data.

Apixio. (2015). *Big Data Reveals Crucial Hidden Information in Healthcare Systems* (White Paper – Health Technology Use Case).

Baldenko, A. (2014). *Aetna's Big Data Journey*.

Blase, P., DiFilippo, D., Feindt, M., & Yager, F. (2016). Data-driven: Big Decisions in the Intelligence Age. *PwC's Global Data and Analytics Survey: Big Decisions*.

Brynjolfsson, E., Hitt, L. M., & Kim, H. (2011). Strength in Numbers: How Does Data-Driven Decision-making Affect Firm Performance?

Brynjolfsson, E., & McAfee, A. (2008). Investing in the It That Makes a Competitive Difference. *Harvard Business Review*, (July 2008).

Brynjolfsson, E., & McElheran, K. (2016a). The Rapid Adoption of Data-Driven Decision-Making. *American Economic Review*, (106(5)).

Brynjolfsson, E., & McElheran, K. (2016b, January 1). Data in Action: Data-Driven Decision Making in U.S. Manufacturing, (US Census Bureau Center for Economic Studies Paper).

Brynjolfsson, E., & Mendelson, H. (1993). Information Systems and the Organization of Modern Enterprise. *Journal of Organizational Computing*.

Dalle Mulle, L., & Davenport, T. H. (2017). What's Your Data Strategy? *Harvard Business Review*, (May 2017).

Davenport, T. H. (2006). Competing on Analytics. *Harvard Business Review*, (January 2006).

Davenport, T. H. (2010). Are You Ready to Reengineer Your Decision Making? *MIT Sloan Management Review*, (July 2010).

Davenport, T. H. (2013a). Analytics 3.0. *Harvard Business Review*, (December 2013).

Davenport, T. H. (2013b). Big Data and the Role of Intuition. *Harvard Business Review*, (December 2013).

Davenport, T. H. (2014). *Big Data at Work: Dispelling the Myths, Uncovering the Opportunities*. Boston: Harvard Business Review Press.

Davenport, T. H., & Dyché, J. (2013). Big Data in Big Companies. *International Institute for Analytics*, (May 2013).

Davenport, T. H., & Patil, D. J. (2012). Data Scientist: The Sexiest Job of the 21st Century. *Harvard Business Review*, (October 2012).

Evans, P. C., & Annunziata, M. (2012). Industrial Internet - Pushing the Boundaries of Minds and Machine. *GE Vision Paper*.

Fitzgerald, M. (2015). Gone Fishing — For Data. *MIT Sloan Management Review*, (June 2015).

Glass, R., & Callahan, S. (2014). *The Big Data-Driven Business: How to Use Big Data to Win Customers, Beat Competitors, and Boost Profits*. Hoboken, N.J.: Wiley.

Gough, C., & Rogers, B. (2016). *Solving Practical Problems with Healthcare Analytics*.

Hayek, F. A. (1945). The Use of Knowledge in Society. *The American Economic Review*, *35*(4).

Henke, N., Bughin, J., Chui, M., Manyika, J., Wiseman, B., & Sethupathy, G. (2016). *The Age of Analytics: Competing in a Data-Driven World.* McKinsey Global Institute.

Heudecker, N., & Kart, L. (2014, September). Survey Analysis: Big Data Investment Grows but Deployments Remain Scarce in 2014.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, *3*(4).

Kanamori, T., & Motohashi, K. (2006). Centralization or Decentralization of Decision Rights? Impact on IT Performance of Firms. *RIETI Discussion Paper Series*, (6-NaN-32).

Kolodziej, M. (2015). *Can (big) data improve cancer care?*

Lakhani, K. R., Iansiti, M., & Herman, K. (2014). GE and the Industrial Internet. *Harvard Business School Publishing*, (9-614–32).

Lavalle, S., Hopkins, M. S., Lesser, E., Shockley, R., & Kruschwitz, N. (2010). Analytics: The New Path to Value. *MIT Sloan Management Review*, (Fall 2010).

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). *Big Data: The Next Frontier for Innovation, Competition, and Productivity.* McKinsey Global Institute.

Marr, B. (2016). *Big Data in Practice: How 45 Successful Companies Used Big Data Analytics to Deliver Extraordinary Results*. Hoboken, N.J.: Wiley.

Marr, B. (2017). *Data Strategy: How to Profit from a World of Big Data, Analytics and the Internet of Things* (1st Edition). New York: Kogan Page Ltd.

McAfee, A., & Brynjolfsson, E. (2012). Big Data: The Management Revolution. *Harvard Business Review*, (October 2012).

Ohlhorst, F. (2013). *Big Data Analytics - Turning Big Data into Big Money*. Hoboken, N.J.: Wiley.

Rose, D. (2016). *Data Science: Create Teams That Ask the Right Questions and Deliver Real Value*. New York: Apress.

Ross, J. W., Beath, C. M., & Quaadgras, A. (2013). You May Not Need Big Data After All. *Harvard Business Review*, (October 2013).

Simon, P. (2013). *Too Big to Ignore: The Business Case for Big Data*. Hoboken, N.J.: Wiley.

Steinberg, G., Church, B. W., McCall, C. J., Scott, A. B., & Kalis, B. P. (2014, June). Novel Predictive Models for Metabolic Syndrome Risk: A "Big Data" Analytic Approach. *American Journal of Managed Care*.

Toon, M., & Collins, J. (2015). *Going beyond the data: Achieving actionable insights with data and analytics*. KPMG Capital.

Walker, R. (2015). *From Big Data to Big Profits: Success with Data and Analytics*. New York: Oxford University Press.

Wyner, G. M., & Malone, T. W. (1996). Cowboys or Commanders: Does Information Technology Lead to Decentralization?