Introduction to Big Data in Small Business

Carsten Lund Pedersen, Adam Lindgreen, Thomas Ritter and Torsten Ringberg

This is a book about big data in small businesses. How do small and medium-sized enterprises (SMEs), with limited resources, thrive in a context abundant with data? To address this central question from multiple viewpoints, we introduce a collection of experiences, insights, and guidelines from a variety of researchers, each of whom provides a piece to solve this puzzle.

Today, data often are described as “big,” and they also are increasing in size with each day that passes. Accordingly, the promise of success linked to using these data is presumed to be similarly big, and expectations are still rising. Yet for many firms, especially SMEs, each of the steps they take is quite small. Even if firms can only take small steps, designing a pathway that comprises many, well-placed small steps can create movement and progression. Thus, there is no contradiction between small steps and big data success. It “merely” requires patience, cadence, and an ability to maintain a balance.

In our ongoing research on data commercialization, we often encounter firms that profess to feeling overwhelmed by the amount of data available and the expectations set, both by themselves and partners in their business ecosystem. The feeling of being overwhelmed is intensified by the lack of resources, which is especially pressing for SMEs. Despite these challenges, we also have often encountered intriguing business development initiatives by SMEs that circumvented the obstacles.

These experiences triggered our interest in crafting a book that would be full of ideas to inspire big data-driven success by SMEs that make smaller investments. We are not talking about million-dollar bets and unicorns; rather, our focus is on solid work and experiments that guide the data utilization journey for SMEs. It is not that we are against million-dollar projects and moon-is-the-limit ambitions, but there must be a place for “ordinary” data-driven journeys that firms of any size can take. Before we introduce the chapters of the book, we thus present some general issues and challenges for SMEs seeking to leverage big data today.

BIG DATA

There are more and more digital data, pertaining to basically everything. Digitization is “the technical process of converting streams of analog information into digital bits of 1s and 0s with discrete and discontinuous values” (Brennen & Kreiss, 2016, p. 1), which leads to the “increased availability of digital data enabled by advances in creating, transferring, storing, and analyzing digital data” (Ritter & Pedersen, 2020, p. 181). This staggering development has been well documented. For example, Internet users generate approximately 2.5 quintillion bytes of data each day, and every person is estimated to generate about 1.7 megabytes of data...
each second. In the third century BCE, the sum of human knowledge was believed to be housed in the Library of Alexandria; today, every person alive could obtain 320 times as much information as historians estimate was stored there (Mayer-Schönberger & Cukier, 2013).

Beyond statistics though, just consider what you likely have done today. In your pocket, you have a smartphone that tracked your geolocation on your way to work. The smartphone also gathered meta-data about the people you called or texted and the apps you checked, logging for all of these interactions how long you spent. When you bought this book, information about where you bought the book became available to your credit card company; if you purchased it online, the retailer also collected information about your purchase (e.g., how long you browsed, what other books you clicked on), which it will use to suggest other books of potential interest to you. Everything that we do leaves digital breadcrumbs.

To deal with this wealth of information, data have been described and classified in a multitude of ways. We suggest their massiveness should be framed along three dimensions (similar to the 3Vs of data; see Chapter 13 for a related discussion). First, diversity (partially related to variety) refers to the many different digital data points pertaining to a subject or object. For every person with a device in her or his pocket, data about location position and speed of movement are available, which then can be combined with data about weather conditions and traffic. In homes, various devices register water, heat, and energy consumption, as well as music, radio, and television usage. Each piece of industrial equipment also produces data about sound, vibration, temperature, production efficiency, and so forth. This massive diversity of data is only likely to continue increasing as digitization continues to spread.

Second, comparability (partially related to volume) relates to the data being collected about specific subjects and objects. Knowing the condition of one industrial machine is helpful; being able to compare these data with data gathered from other machines likely offers even more insights. Accordingly, industrial maintenance programs are built on data gathered from many machines that get compared to eliminate potential error sources. As consumers, we are accustomed to such comparisons: Amazon, Spotify, and other platforms suggest what other users, seemingly similar to us, have chosen.

Third, alteration (partially related to velocity) reflects changes in individual measures, including how often and how quickly data points are updated. We have experienced huge shifts toward the real-time, so currently updated data are often available. Mobile devices provide an exact current location, banking apps provide a precise balance in that moment, and the oil temperature gauged by a control center reflects the machine’s condition just seconds ago. Driven by always-on measurement devices and high-speed connectivity (e.g., fiber cable networks, 5G mobile networks), data are produced infinitely; that is, data are available about many aspects pertaining to many objects, and those data also are available many times.

Figure I.1 illustrates this explosion of data, which we attribute to developments in three dimensions: more aspects, more objects, and more often. For a firm, a good exercise is to visualize its own data situation along these three dimensions, relative to its competitors:

- How many different data points do we have access to? How many do our competitors access?
- How many different objects do we have data about? How many do our competitors have?
- How often do our data get updated? How often do our competitors update their data?
The larger the square in Figure I.1, the more extensive are the data available to the organization. The size of the quadrant along these three dimensions illustrates how much data are available for the organization—stated differently, the size of the quadrant indicates just how big the data are. This assessment thus can reveal the competitive situation of an organization in terms of its digital data development. This is not to say that a firm must lead on all three dimensions simultaneously to be successful; as Chapter 5 argues, there is beauty in smallness and appropriate winning strategies. But the mapping helps illustrate where an organization stands and which strategic options are available to it.

**Figure I.1** *Data bigness on three dimensions*

**BIG SUCCESS**

Data alone do not equate to business success, despite a general assumption that they do, as manifested in claims like “data is the new oil” that can do virtually anything for a business. But data must be used before they can contribute to firm performance or market success.

In the meantime, digitization continues to change business practices, business models, and marketplaces (Ritter & Pedersen, 2020), and not all of these changes are value-creating, nor is their potential being fully exploited. Rather, businesses’ vast opportunities are wide, unpre-
dictable, and ever-changing, without any visible horizon. Consider several stylized facts that support this claim:

- 95 percent of businesses identify the need to manage unstructured data as a problem for their operations; some estimates posit that only 0.5 percent of all data are actually analyzed.\(^5\)
- Predictions describe the “datafication” of all aspects of life and its commercialization, yet few of these predictions come to being. Google Glass may have datafied the gaze (Mayer-Schönberger & Cukier, 2013), but little commercial value has originated from these data.\(^6\)

As the chapters in this book argue, there are many different business areas that can use data to realize opportunities: human capital (Chapter 8), learning and training (Chapter 9), sales (Chapter 11), and business development (Chapter 12), to name just a few. A common denominator is that firms need to specify concrete projects based on these data, whether it is a marketing project, a sales project, or an operations project. With such a project in place, firms can describe the aim (e.g., cost efficiency, revenue, both). In turn, the measure of success, achieved with data, becomes more tangible for companies.

**BIG LIMITATIONS**

Most companies facing the contemporary challenge of preempting threats and pursuing opportunities through digitization are SMEs with limited budgets and resource constraints, yet most practitioner-oriented research and advice focuses on Silicon Valley mega-companies that were born digital (e.g., Google, Facebook, Uber) or iconic behemoths undergoing a digital transformation (e.g., Maersk, Siemens, GE). Unfortunately, SMEs get overlooked in research into business applications of digital data, so the guidelines available are either unrealistic (e.g., require more resources than SMEs have available) or not relevant (e.g., they are already agile).

Being small thus can be a limitation; to deal with big data, firms may require substantial resource investments and time, two dimensions that SMEs rarely have in abundance. But even if SMEs face big limitations in their work in the data-related domain, they can overcome them by being aware of the limitations, working within them, and using the related restrictions to think creatively. When SMEs find their own way, it can lead to novel, alternative solutions to pressing challenges.

**BIG PARADOXES**

Finally, like any organization, SMEs confront paradoxes that hinder easy solutions but also enable unique successes when effectively solved.

**Paradox 1: Lack of Data about Data**

Along with the massive increase in data, we also acknowledge the lack of realization of business opportunities, particularly by SMEs. This problem can be attributed, at least partially, to the counterintuitive situation in which data exist about nearly everything, but not enough data reveal the success of data-driven initiatives. That is, decision makers often lack specific data
to support valid assessments of business cases related to big data. Choosing the right initiative then becomes a gamble, without input for making a reliable business case. This situation is particularly challenging for firms with small budgets that cannot initiate a large portfolio of projects and hope that some succeed, while funding a plethora of unsuccessful ones.

**Paradox 2: Voices from the Top vs. Solutions from the Bottom**

Digitization and data-driven business development are discussed everywhere; virtually every top management team has added some digital component to their strategy. Every day, we read about a CEO or politician who argues for the importance of increased levels and paces of digitization and digitalization, and those calls grew even louder in response to the social distancing measures imposed by the coronavirus crisis. However, these discussions are solely agendas. The true development of data-driven business comes from the frontline: engineers with insights into machines and operations, IT experts with capabilities in data and analytics, customer-facing employees who can account for customer needs and value creation—and sometimes, teams of representatives from all these groups who come together around a data project. A stealth innovation can be blocked by an organization that is not ready to promote autonomous projects to official strategy status. So data-driven success requires resolving this issue, and SMEs might stand a better chance, due to the shorter distance between top and bottom, stronger cross-functional insights, and cooperation that tend to appear in their organizations.

**Paradox 3: The More Personal Data about the Customer, the Farther Away the Customer Wants to Be**

Firms know a lot about their customers, reflecting the diverse data they collect. Applying such data brings firms closer to their customers, which allows for a better understanding of their needs and more appropriate value propositions. But the use of data also reveals just how up-close-and-personal suppliers can get, which may alienate customers and cause them to limit further access to their data. Some data utilization clearly appears creepy at times, like big brother or surveillance capitalism (Zuboff & Schwandt, 2019), and it evokes the so-called personalization paradox (Aguirre et al., 2015). There is a fine balance between establishing close, supportive links with customers and respecting their privacy. Similar arguments hold for employees, suppliers, and nearly every other actor in a firm’s business ecosystem.

No solutions to these paradoxes are easily available or obvious; decision makers also must actively address them on an ongoing basis. In doing so, they can make better informed decisions for how to deal with the intricate dilemmas encountered with big data.

**REMAINDER OF THE BOOK**

This book is about the lessons, successes, and failures that SMEs encounter on their quest to obtain data-driven growth. These firms must take small steps toward their goals, because of their limited resources to invest and need to realize immediate success. In describing the journey that SMEs take, one step at a time, we have tried to produce a text that addresses the
Big data in small business

concerns and challenges faced by practitioners. Accordingly, it consists of four general parts, each comprised of related chapters:

- Part I: Foundations: Getting the Basics Right
- Part II: Capabilities: Getting Digitization Right
- Part III: Functions: Getting All Business Areas into Big Data Mode
- Part IV: Transformations: Getting There the Right Way.

Part I: Foundations: Getting the Basics Right

The first part of the book consists of three chapters. Starting us off, Karen Mills and Annie V. Dang propose that the world of small business can be transformed by “Building Small Business Utopia: How Artificial Intelligence and Big Data Can Increase Small Business Success.” They note the frictions, barriers, and stagnation that small businesses confront when seeking access to capital from conventional lenders but also the great promise of financial technology, or “fintech.” Emerging fintech firms leverage big data and artificial intelligence, then build automated underwriting processes and dashboards that streamline the customer experience, make lending applications more efficient, and even help small firms track their cash flows. The success of this market prompted conventional banks and big tech platforms to develop new products and services to appeal to small business customers, expanding their financial access even further. The authors make a convincing case that the resulting product–service ecosystem, or small business utopia, by providing both capital and insights to small businesses, will enable them to prosper, innovate, and grow.

Rather than embracing a utopian ideal, Jan Trzaskowski establishes some practical recommendations for how small businesses should be using big data today, so that they can comply with existing regulations. In “GDPR Compliant Processing of Big Data in Small Business,” he proposes six principles that summarize the 99 articles and 173 recitals of the European Union’s General Data Protection Regulation, which SMEs then can integrate into their operations.

In addition to ensuring compliance, SMEs must address other demands, such as overcoming their size liabilities, so in “Big Data and SMEs,” Vania Sena and Sena Ozdemir propose practical ways they can leverage big data to reduce coordination costs and enhance team-level collaboration. With an organizational perspective, this chapter outlines how business performance can depend on big data investments, with both theoretical and empirical insights for how SMEs can exploit the big data they gather.

Part II: Capabilities: Getting Digitization Right

The second part of the book consists of four chapters, each of which covers a distinct practice or strategy related to effective digitization. For example, in their effort to define data-enabled value-creation strategies that SMEs can use, Bieke Struyf, Wouter van Bockhaven, and Paul MatthysSENS detail effective resource configurations. Their in-depth case study, as reported in “Value-Creation for Industry 4.0 and SMEs’ Data-Driven Growth: Strategies and Resource Alignment,” illustrates how an SME can evolve to a digital platform instigator—and create radically new value-creation opportunities—by overcoming challenges, designing effective organizational processes, and ultimately adopting Industry 4.0 technology. In particular, the case firm’s network capabilities, innovative nature, and incremental digital capability
development efforts supported its shift away from digital servitization and toward digital platform strategies. The authors use these insights to propose a framework of data-enabled value-creating strategies, detailing the capabilities that underlie successful versions, along with a self-evaluation tool that SME managers can use to develop their own digital transformations.

Capabilities also are front and center, along with a project view, in the framework proposed by Carsten Lund Pedersen and Thomas Ritter in the next chapter. To inform data-driven growth in SMEs, they argue that a digitization capability consists of three dimensions (data, analytics, and permission), and data-driven projects involve seven key scenarios. By “Analyzing and Developing Digitization Capabilities for Data-driven Projects in SMEs,” they detail both the steps involved in firms’ data-utilization journeys and the ways firms could apply their framework to develop a digitization capability.

Another case study, describing “How a Glass-Processing SME Developed Its Big Data Competence,” proposes that an SME called Glaston actually used three dynamic capabilities to overcome its limited funds and technical know-how. In defining the firm’s big data competence, Joel Mero, Heikki Karjaluoto, and Tanja Tammisalo describe how Glaston combined its market listening and business imagination skills to monitor industry-level technological developments, then apply them in big data use cases relevant for its operations. With open innovation, Glaston also was able to complement its internal resources and partner with external actors, and then its culture of experimentation enabled it to develop a portfolio of big data initiatives, which it pilot tests carefully before undertaking any full-scale implementation.

Shirley Y. Coleman agrees with the benefits of partnering; she proposes the promise of “Data Excellence in SMEs through Engagement in University Partnerships.” If SMEs work with universities to expand their data capabilities, they might overcome their relatively small pools of in-house expertise, while still leveraging their flexibility and agility, to take advantage of the proliferation of data and digital communication associated with Industry 4.0. The author uses her experience of working with SMEs and developing university partnerships to outline both benefits and risks, as well as to summarize some key lessons learned, to inform other SMEs as they develop plans for their own partnerships.

Part III: Functions: Getting All Business Areas Into Big Data Mode

The four chapters in this third part acknowledge the need to get all departments and operations to embrace big data, even in SMEs. For example, SMEs might ignore the benefits of human capital analytics (HCA), according to Frederikke Amalie la Cour Nygaard and Dana Minbaeva, because most studies of HCA address large, complex organization. But in “Capitalizing on Human Capital Analytics in Small and Medium-Sized Enterprises,” they challenge this misconception by detailing how and why HCA is highly relevant for SMEs, and then offer examples of how they can capitalize on HCA to ensure value creation and competitive advantages.

Next, Camilla Nellemann and Torben Pedersen consider “How Experimental Data Can Optimize E-Learning,” offering the novel prediction that big data can help clarify user behaviors in not just e-learning but also other consumption contexts. When firms can identify behavioral patterns exhibited by their customers, then adapt their offerings to them, using both experimentation and data analytics, they can perform better. To provide evidence for their argument, the authors describe how a software company experimented with ways to teach
customers how to use their products efficiently and thereby discovered that moving physical teaching into a virtual space provided notable benefits, for both e-learners and the firm.

Beyond just efficiency though, SMEs often rely on informal, personal relationships to establish and maintain their competitive advantage. Thus, Poul Houman Andersen wonders and investigates “How Do Big Data Impact Business Market Relationships?” The impact of big data on this competitive advantage appears to reflect the specific interaction patterns between buyers and sellers, which themselves emerge in continuous, interactive market processes. Accordingly, big data constitute meaningful resources that offer potential opportunities, because small businesses can develop data-supported market relationships. The challenges, tasks, and issues facing small business owners even align notably with the services that can be rendered by big data.

Henrik Andersen and Thomas Ritter also consider business relationships, namely, the links of suppliers with their various customers, which differ according to the extent to which suppliers can provide market offerings that match the customers’ demands, sizes, and procurement policies. In “Revenue Blueprinting: Identifying Growth Potential Using Customer Data and Customer Insights,” they propose that suppliers can achieve revenue growth potential if they leverage their best implemented practices, using carefully calculated and verified revenue blueprints that reflect the revenues achieved through each practice. These revenue blueprints highlight the firm’s own best implemented practice, while also accounting for current market offerings and customers. In turn, SMEs can seek revenue growth without having to make expensive new investments or undertake substantial training.

Part IV: Transformations: Getting There the Right Way

This fourth and final part of the book includes two chapters. First, in “Transforming Small and Medium-Sized Enterprises (SMEs) to Digitally Enabled Landscapes,” Bård Tronvoll, Christian Kowalkowski, and David Sörhammar detail the challenges of digital servitization for SMEs, then propose three transformational shifts (dematerialization, identity, and collaboration) that can help them overcome and even embrace these challenges. Such shifts require SMEs to be close to customers, with a strong understanding of their needs. In particular, the collaboration shift should be relatively easier for SMEs, because these small firms already likely coordinate their activities with other actors. But the dematerialization shift (i.e., separate information from the physical world) may be more difficult, because it requires sophisticated technical competencies and financing strength. Finally, the identity shift might be hard too, if the SMEs’ culture remains resistant to change.

Second, Pernille Rydén and Helle Rootzén present the real-world experience of 12 Danish SMEs that have engaged in digital transformation learning processes as the first steps in their big data journeys. Specifically, with their chapter, “Facilitating Big Data Transformation in Danish SMEs: Insights for Managers,” these authors detail the purpose, principles, processes, potential, barriers, and methods that emerged from “KomDigital,” a digital learning concept for SMEs develop by the Technical University of Denmark (DTU). The authors thus suggest practical steps for other SMEs, describing not just managerial concerns but also key drivers of digital learning and transformation. In particular, if they can increase their big data technology competence, SMEs can better exploit both data and analytics to identify consumer insights,
expand their business intelligence, enhance their decision-making power, and evoke new ways of thinking and interacting with their markets.

CLOSING REMARKS

We extend a special thanks to Edward Elgar and its staff, who have been most helpful throughout this entire process. Equally, we warmly thank our contributors, who exhibited a desire to share their knowledge and experience with the book’s readers—and a willingness to put forward their views for possible challenge by their peers. We hope that this compendium of chapters and themes contributes to our colleagues’ and practitioners’ own research and practice. The chapters in this book can help fill some knowledge gaps related to important aspects of capabilities, functions, and transformations of big data that drive business growth; we hope they also stimulate further thought and action pertaining to these topics.

NOTES


REFERENCES


