# Strategic Positioning in Big Data Utilization: Towards a Conceptual Framework

Milla Wirén and Matti Mäntymäki[(✉)]

Turku School of Economics, University of Turku, 20014 Turku, Finland
milla.wiren@utu.fu, matti.mantymaki@utu.fi

**Abstract.** This paper introduces a conceptual framework for strategic big data utilization. We discuss big data through (i) its origins (where and how does the data accumulate), (ii) its constitution (what is the nature of the data), and (iii) its applications (how and why can the data be processed and utilized). Based on this conceptual analysis, we argue for three continua that can guide the process of making strategic decisions regarding the utilization of big data. We further use these continua as a foundation for proposing a conceptual framework for strategic data use and strategic positioning. The conceptual framework facilitates understanding the firm-specific possibilities that emerge from aligning the overarching business goals with the opportunities emerging from big data.

**Keywords:** Big data · Strategy · Big data utilization

## 1 Introduction

Big data has become a richly researched phenomenon from the perspectives of its accumulation (Brynjolfsson and McAfee 2012), attributes (Chen et al. 2012), implications (Newell and Marabelli 2015; Zuboff 2015; Yoo 2010) and individual applications (Davenport et al. 2012; Varian 2010). At the same time, there are increasing calls for integrating the information systems, or digital strategy with the overarching business strategy (Bharadwaj et al. 2013). However, maybe reflecting the fundamentally amorphous nature of the phenomenon labelled big data (Constantiou and Kallinikos 2015), the discussions under the same label suffer from a level of amorphousness where the origins, constitution, threats, possibilities and utilization are somewhat conflated. As argued by Zuboff (2015), the abundance of definitions for the concept of big data suggests that understanding the phenomenon is still suffering from a lack of sufficient conceptual clarity.

We ground our approach on a threefold conceptualization of big data: the phenomenon can be explored by zooming into its origins, emergence and accumulation; from the viewpoint of addressing its attributes, nature and constitution; or from the perspective of outcomes, emerging both through deliberate utilization and from unintended implications. Our focus lies in the third vantage, explicitly in the part of the deliberate, strategic level utilization of the big data.

Previous discussions about this third vantage of big data, its outcomes, are either focused on specific business applications or use cases such as customer relationship management (McAfee et al. 2012), predictive maintenance (Wang et al. 2017), or game

analytics (Koskenvoima and Mäntymäki 2015). The practitioner-oriented advices pivot primarily on the need of employing capable data scientists and on the need of making decisions based on data driven knowledge (Davenport et al. 2012; McAfee and Brynjolfsson 2012; Schildt 2017). Another avenue within this third vantage is the critically tinged stream of literature assessing the implications of the phenomenon of big data (Zuboff 2015; Newell and Marabelli 2015), however while we acknowledge a set of insights derived from those approaches, our focus is not on the unintended implications. Instead, we address another gap: research providing strategic directions for the use of big data is largely absent in the literature. We argue that to reap the strategic benefits of big data, data-driven decisions need to be underpinned by knowledge of how do the specific choices regarding big data utilization reflect and fit the overall business strategy.

To address this void in the literature, we draw on Newell and Marabelli (2015) and Zuboff (2015) to categorize the sources of big data into five types. We thereafter categorize big data into two types based on its constitution (Constantiou and Kallinikos 2015). In practice, we take the aforementioned threefold approach and identify the phenomenon labelled big data through (i) its origins (where and how does the data accumulate), (ii) its constitution (what is the nature of the data), and (iii) its applications (how and why can the data be processed and utilized). Distinguishing between these three vantages is important for the sake of conceptual clarity, because conflating the three vantages in the descriptions threatens to muddle the conceptual understanding of the complex phenomenon.

As a result of this conceptual analysis, we argue for three questions that can guide the process of making strategic decisions regarding the utilization of big data. We further use these questions as a foundation for proposing a conceptual framework for strategic data use and strategic positioning. In addition, the proposed framework facilitates positioning the data-use related firm-specific choices into a wider context of strategic tradeoffs that allow assessing the level of alignment between the overarching business goals of the firm and its choices related to utilization of big data.

The remainder of the paper is structured as follows: we begin by outlining and continuing the discussions (Constantiou and Kallinikos 2015; Tilson et al. 2010; Zuboff 2015) delineating and crystallizing the phenomenon labelled big data and its impacts on strategizing. In sum, we view big data from the three vantages of its origins, constitution and use, with the emphasis on the last part. Then we continue by identifying three continua along which the firms need to position their strategic approaches in regards to utilizing big data. Subsequently, we propose a framework to support strategic utilization of big data and conclude by recapping the contributions and providing future research avenues.

## 2  Big Data: Origins, Constitution and Utilization

According to Zuboff (2015), as most discussions of big data are accompanied by efforts of defining it, a satisfactory definition of the phenomenon loosely given the label seems to not yet exist. The origins of the concept are often traced to the 2001 paper by Laney at Gartner (Gartner and Laney 2001), where the qualitative changes in data and its

accumulation were defined by three v's, volume, velocity and variance. Subsequent definitions have introduced for example additional v's, like veracity, variability, visualization, value to name a few (Newell and Marabelli 2015), blurring the definition further. For the sake of conceptual clarity, here we identify the phenomenon through three vantages: origins, constitution and use.

With respect to the origins of big data, Newell and Marabelli (2015) propose a category of two (human-digital interaction, e.g. social media and internet searches and signals from embedded sensors), whereas Zuboff (2015) identifies three more sources (computer-mediated digital economic transactions, corporate and government databases, and surveillance systems) to present altogether five source categories. Here we utilize the more comprehensive categorizing by Zuboff to ground our approach.

The human-digital interaction refers to the traces people leave when utilizing the digital devices by e.g. browsing the web, using social networking sites and navigation services, or storing online fitness activities recorded by wearable devices. Especially when using mobile digital devices, we are creating data harnessed also by embedded sensor technology tracking for example our location (Abbas et al. 2014) or even the intensity of our discussions (Greene 2014). The developments in sensor technology are quite notable: it is for example possible to detect from a distance whether the driver of a car driving by is under the influence of alcohol (Hewitt 2014), or even to monitor brainwaves (Sundaresan 2017), each of these applications resulting in creating data in numerous forms. Equally notable are the sensor technology developments underpinning the Industry 4.0 scenarios (Gilchrist 2016; Hermann et al. 2016; Kagermann 2015) in the industrial setting, and in the creation of the so called internetof-everything, the development of smart gadgets and equipment for both industrial and domestic use. These developments in the sensor technology have also an impact on the third data source identified by Zuboff (2015), the surveillance systems (Lyon 2015), taking additional advantage of the increasingly ubiquitous cameras and satellites.

These three types of data sources represent the new forms of data creation, where the data being harvested is far from uniform or categorized at the outset. In contrast, the fourth source, government or corporate databases, is old and as such, mostly reliant on *a priori* categorizing the desired data, including its form. This means that the type of data originating from these more traditional sources is qualitatively different to the data resulting from the more contemporary sources of big data: for example, a governmental health related database contains only such information of the citizens that someone at some point has deemed pertinent to ask and store. The fifth identified source, the traces of digital transactions (Varian 2010) falls in between these categories, as a part of that data is predefined and -categorized, consisting of alphanumerical information, whereas a part of that data is mere imprints of the events and actions accompanying the transaction processes.

In terms of the constitution of big data, Constantiou and Kallinikos (2015) detail the diversity of the types of data being accumulated. Part of the data is pre-categorized and alphanumerical (like the data in governmental databases, for example), but more notable are for example the images, location specific data, audio signals and social network system tokens, such as Facebook "Likes", resulting from the newer types of data creation. In contrast to such data that has long been used in business intelligence (i.e. harvested by firms to facilitate decision-making by seeking out answers to

predefined questions), these new forms of data do not result from *a priori* planning and categorizing the specific type of data needed for specific business purposes but accumulates as a trace of all types of digital activities (Chen et al. 2012). This type of data is not "sorted in the way in" (Weinberger 2007), but instead requires advanced analytical and processing capabilities in distilling the meaningful from the noise.

To recap, the sources of big data can be categorized into five types, and the constitution of data can be reduced into two types based on prior literature. However, similar typologies focusing on the uses of big data are largely absent in the literature. To address this void in the extant literature, we argue for three questions that can guide both data use related theorizing and strategically exploiting the potential of big data.

## 3   Three Continua of Strategic Utilization of Big Data

The first continuum stretches between generalizability and personalization. It pertains to the expected value generation of data: do we want to use the data in seeing patterns and trends, or do we want to be able to predict the actions of an individual, be it human or machine? The second continuum relates to our access to data, and consists of the polar ends of proprietary and networked sources of data: can we source all data through our own efforts, or do we need data from external sources? The third continuum relates to the level of investments we are willing to make in our data processing capabilities: is our business primarily in refining the raw data, or in utilizing the refined data?

Next, each of these continua is explored in more detail, followed by a synthesizing framework that can provide starting points for strategic data use.

### 3.1   Utility: Generalizability vs. Personalization

The first strategic question in thinking about the potential use of the enablers of big data is: does the business benefit more from understanding general trends and patterns, or from being able to understand the behavior of an individual, be it a human, a process or a machine? The answers to this question form a continuum along which the firm can position itself.

A furniture retail firm was struggling with staffing issues in their physical stores – the customers peaked seemingly at random, which meant that the store was constantly either overstaffed or understaffed in regards to the customer needs. The firm hired a consultancy specialized in data analysis, and after some twists and turns, it was found out that the customer peaks correlated with local weather patterns: customers kept pouring in the day after it had rained and poured outside. While the data revealing this insight did not enable analyzing the causes of this, it however helped the firm in its staffing issues: a year later the personnel costs were notably smaller, and the turnover rate of sales on the peak days was notably better.

This example captures nicely the potential of big data in identifying patterns and creating generalizable preferences. The massive amounts of data about seemingly unrelated issues processed with algorithms enables seeking and seeing correlations. In industrial setting, such correlations can be found between error logs and for example

specific type of use or location, thus helping in designing predictive maintenance protocols.

However, while these correlations and patterns reveal a lot about the aggregated tendencies, they do not yet yield information about the individual dispositions. Newell and Marabelli (2015) highlight the issue with an example from the field of insurance. Based on comparing the accident logs with the information about the drivers of those vehicles, young male drivers, as a group, drive more recklessly than other groups. However, this information cannot be used as grounds for charging young men more for their car insurance, because that risks discriminating against such young men who drive safely.

This leads towards the other end of the continuum; the use of big data in personalization, also dubbed as little – or small – data by some authors (Boncheck 2013, Newell and Marabelli 2015). Continuing with the insurance example, it is possible to gather data from the driving behavior of an individual with equipment fitted into the vehicle (e.g. use of gas, brakes, speed, acceleration) or through identifying a passing car by means of surveillance equipment (cameras, road sensors) and a connection to for example license plate registries. This data can then be compared to the general data about what type of driving typically leads to accidents, and from those correlations it is possible to derive a personalized prediction about the likelihood of a specific individual having an accident. That information can then be used in allocating a risk premium or offering a discount on the insurance, as for example some companies already do (Progressive N/A).

In social media marketing, the recent case of Cambridge Analytica (Cadwallar and Graham-Harrison 2018) highlights other applications of personalized insights extracted from big data. By first harvesting vast masses of Facebook "likes" from diverse individual and connecting that data to the results of the online personality quizzes of those same individuals, it has been possible to deduce for example the political opinions of people based on their "liking" Kitkat or Harley Davidson (Kosinski et al. 2013). This means that through analyzing the "likes" of an individual, it is possible to create fairly accurate predictions of the preferences of that individual, which then in turn enables microtargeting of ads, personalized marketing. Additionally, in the industrial side this approach enables monitoring a specific piece of equipment and predicting its maintenance and service needs. Comparing the general error log data with the data about how specific environmental circumstances and use of that equipment correlate with that data, and then analyzing the accumulated data of the environment and use of a specific piece of equipment enables estimating the need for servicing of that specific piece of equipment.

To sum, while from the technical perspective of harvesting and analyzing data this use-related distinction can be to a degree neglected, understanding the business utility of big data is an essential strategic choice.

## 3.2   Access: Proprietary vs Networked Data

The question facilitating the positioning along the second continuum of data access is: to what extent is it possible to ground the strategic decisions of the firm on the

internally generated (operational) data, and to what extent is access to external data sources necessary for strategic purposes?

The second continuum is underpinned by the insight that data required and generated in the process of firm activities (stored often in diverse CRM and ERP solutions), and the data needed in strategic business related decisions should be seen as two distinct types of data as defined by its use. The first type of data is operational, necessary in and generated as a side product of the operations of the firm, but it is the second type of data that is required in strategic decisions. The important distinction to understand is that while the underlying data sets (including generating, storing and processing the data) may be the same, the use of that data is different depending on whether it is used for operational or strategic purposes. In our second dimension, the question is about the tightness of the coupling between these two uses of data.

At the one end of this continuum, it is possible for the firm to exploit the data it generates for the operational purposes also in strategic decisions to the degree where it is not necessary for the firm to source data externally. An example can be found from traditional heavy machinery industries where the main economic logic is that of economies of scale, and the competitive advantage results from superior cost-efficiency. In that context the data generated by the operations, including operative costs like fuel, material, downtime of machinery and personnel can be utilized in identifying factors that increase the costs and thus diminish the bottom line. What follows is that from the business perspective, such decisions that reduce the costs are directly aligned with the low-cost strategy, and as such, strategic. In these contexts, the benefits from big data emerge from the capability of the firm to process its operational data also in its strategic pursuits of low costs.

In this realm, the only external information required is the buying and selling prices of the relevant markets. The required data sources are primarily internal, related to the operations of the firm, however the technological advances in sensor technology and connectivity mean that in order to exploit the potential of big data fully in this context, it may be necessary to retrofit the mechanical pieces of equipment with such systems that enable the more comprehensive monitoring of all operation related processes and resulting costs.

At the opposite end of this continuum this connection between the internally generated data and the data required for strategic decisions is decoupled. The firm may or may not generate data for its operational purposes, but the business decisions are highly dependent on data from external sources. Shipyards serve well to illustrate this position at the other far end of the continuum. Shipyards operate in the nexus of on the one hand the ship owners and carrier lines, and on the other hand the multitude of subcontractors responsible for designing, constructing and implementing diverse systems, ranging from the heavy machinery like the engines to the interior design, electrical systems and various navigating related solutions. In short, the shipyards can be envisioned as a type of platform where the diverse needs of the ship owners and operators meet the diverse offerings in the field of maritime engineering. In order for this platform to create value, it needs access to several external sources of data, related to not only the contemporary technological advances in the maritime engineering, but also to the drivers of competitive value in the shipping industry, the overarching

regulatory and environmental developments in the maritime industries, and the overall market trends and ecosystems in global maritime transportation industries.

In these types of settings, the value from big data emerges from the capabilities geared towards creating appropriate data collaboration relationships grounded on strategic insights about which types of big data have strategic value, where do such types of big data originate from, and how and by whom should the volumes of heterogeneous data be processed. This last type of capability leads towards the third continuum in the conceptual framework of strategic data use.

## 3.3    Investments: Raw vs Refined Data

The third continuum assesses the optimal tradeoff between the investments on the requisite data-processing capabilities and the expected value returns of those invest-ments: can the firm create value from possessing such sophisticated data-processing capabilities that enable dealing with the raw data, or does the value emerge only from wielding refined data?

The third continuum can be approached through looking at the diverse stages of making sense of the masses of data. Each stage is accompanied with specific requirements, in other words investments in specific types of data-processing capa-bilities. This translates into strategic decisions, where the required investments in data-processing capabilities in different stages of data analysis can be weighed against the business goals: in short, where is the firm-specific cut-off point where the benefits of having such capabilities in-house outweigh the costs of the requisite investments.

As previously discussed, the big data consists of both pre-categorized alphanu-merical data, and uncategorized, highly heterogeneous data. This means that the pro-cessing of the latter type of data has to begin with rendering the data uniform enough to be processed together: ultimately this underpins the notion of digitalization, which means that the diverse types of data signals are digitized, made into bits (binary digits) (Tilson et al. 2010) that can then be processed by various digital technologies. So, the first set of data processing capabilities are such technologies that facilitate digitizing of previously heterogeneous data.

However, the subsequent mass of digital data is still amorphous and unorganized. Unlike traditional data already gathered based on predefined categories, this mass of data needs to be categorized. This is the second stage of data processing, where the artificial intelligence, i.e. sophisticated algorithms are necessary, because processing the vast amounts of data isn't a feasible task for human capabilities. However, this has two pre-requirements: first of all, the firm needs to have the resources needed in developing or accessing the requisite algorithmic capabilities, and secondly, the firm needs to have the human capabilities required in delineating the desired categories then filled with the sorted data – even the most sophisticated algorithms function only based on what they are coded to do, based on the categories they are programmed to identify.

In the third stage, the algorithms are equally necessary. When the unorganized noise of big data is first digitized and then categorized, the artificial intelligence can be used in finding patterns and correlations within or in between the diverse categories. It is here that the business benefits begin to emerge. However, what is still needed, especially on the human side, is the fourth stage, the capability of understanding how

and why certain patterns and correlations could have business value, what is the strategic importance of the unearthed insights.

The strategic positioning of the firm in this third continuum emerges through asking the question of whether our competence lies in the first three, or in the fourth stage – or maybe the competitive advantage is a result of having the requisite capabilities in all stages. Amazon is an example of a firm that has invested heavily on capabilities spanning the whole continuum, there are data analyzing firms exploiting the business potential of focusing on the first two stages, whereas for most of the businesses the value of big data derived insights begins to emerge only when the results of the third stage of analysis are combined with the specific strategic needs and ambitions of the firm identified with the fourth stage capabilities.

This third continuum has also another dimension, related to the discussion in the context of the second continuum. Many firms generate masses of operational data that they may or may not use in the strategic decision-making. In the case where the operational data does not carry business value for the focal firm, it might however carry business value to another actor. This means that one of the offerings a firm could consider would be selling (or sharing) the data it generates, either as raw or refined, depending on the emphasis on the data analyzing capabilities in-house. So, ultimately a firm has four choices in regards to data collaboration, depending on its choices in data analysis capabilities, its evaluation of the strategic value of its internally generated data, and its evaluation of its strategic external data needs.

First, if the firm has invested little in the data analyzing capabilities, it can sell (or share) such internally generated raw data it regards of little strategic importance to itself and acquire such refined external data it deems having strategic business value. The other three options presume a level of internal data analyzing capabilities. The firm can still sell (or share) raw data with little value to itself but acquire such raw data it deems valuable and possesses the capabilities to analyze. With adequately sophisticated analyzing capabilities, the firm can sell its data in an already refined form and acquire raw data it then processes to meet its needs – or the needs of its customers, if the firm is in the actual business of data analyzing. Finally, the firm can also acquire refined external data, and in turn offer refined internally generated data. This last option requires a crystallized understanding of the specific data needs in both the strategic actions of the business itself, and in the businesses of the data collaborators.

## 3.4   Conceptual Framework: Starting Points for Strategic Big Data Utilization

Traditional data gathering processes result from predefined data needs answered by gathering pre-categorized data to answer premeditated questions. As a result, the data in traditional corporate and governmental databases is relatively homogeneous (alphanumerical), structured, and it fills a predefined purpose. In contrast, the big data generated from the sources of human-digital interaction, surveillance systems, sensors and partly digital transactions is highly heterogeneous and trans-semiotic (including not only text and numbers, but also image, sound and activity tokens), and un- or semistructured, and its accumulation is not only a result of purposeful data collection but a residue of the actions and interactions within the digital realm.

Tackling these issues emerging from the origins and constitution of big data leads towards our conceptual framework. To begin with, can the firm benefit from making the heavy investments required in having the capabilities spanning the whole continuum or should it focus on a specific set of data processing capabilities? In addition, as the raw data is in itself useless, how does the firm ensure the adequate access to the other stages of data processing capabilities, if its position on this continuum is narrow? Furthermore, what is the level of coupling between the internally generated operational data and the data required for strategic purposes: i.e. is possible to rely of proprietary data alone or is external data needed, and if so, should it be raw or refined – and with whom and how should the data collaboration agreements be sketched?

Ultimately, however, the aforementioned questions need to be aligned with the anticipated business utility of data: in terms of the business goals of the firm, is it necessary to glean personalized insights or can most value be derived from generalizations? To sum, these choices can be envisioned as a three-dimensional matrix, depicted below (Fig. 1).
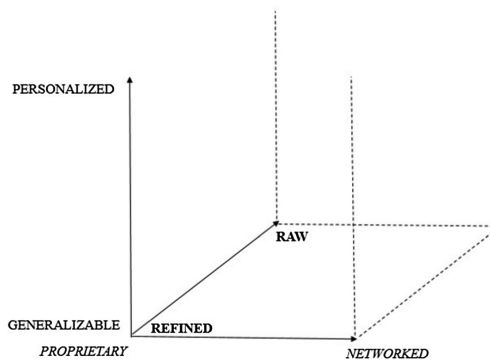


**Fig. 1.** Conceptual framework of strategic data use

In the nexus where the utility of the data is realized from generalizable insights, access to data is proprietary, and the investments on data processing capabilities are primarily based on dealing with refined data, the strategic data use is still quite traditional, merely enhanced by the efficiency enabling facets of data and automation. However, as soon as the strategic aspirations move onwards on any of the dimensions, the changes in the data use become more profound, increasing both the potential benefits of harnessing big data, and the risks, costs, and complexities in handling the big data. This means that moving onwards on any of these dimensions should be firmly grounded on the overarching strategic choices and business goals of the firm, which subsequently implies that the optimum position along each continuum is highly firm specific, dependent on both the endogenous capabilities and dispositions, and the exogenous environmental elements.

## 4    Discussion

### 4.1    Contributions

As its chief contribution to the literature on the strategic potential of big data, the present paper has put forward a framework for strategic utilization of big data. In doing so, the paper responds to the call for research on big data and decision-making (Abbasi et al. 2016). Moreover, this research also continues the discussion initiated by Constantiou and Kallinikos (2015) who highlighted the changes in the strategic contexts resulting from big data: we acknowledge the inevitable changes and go further in exploring the subsequently emerging choices regarding the strategic approaches to utilizing data. Furthermore, we heed the call (Bharadwaj et al. 2013) for the need of digital business strategy where the strategic choices of data use are embedded in the overarching business strategy. As our contribution, we present the conceptual framework identifying the questions underpinning the deliberate data use, which enables assessing the alignment between the overall business goals of the firm and its big datarelated choices.

The framework has both theoretical and practical value. In terms of theorizing, it provides a possible way of categorizing the discussions concerning the third vantage of big data, the deliberate use of big data. For practice, our framework highlights that the further from the nexus of generalizable, proprietary and refined data use the ambitions of the firm in terms of data are, the bigger the complexities, required investments and potential rewards are – and the less the firm can rely on traditional methods of using big data as a subset of business intelligence. Furthermore, understanding these strategic choices of data use can serve as a basis for differentiation through developing analytical capabilities and insights (Abbasi et al. 2016).

### 4.2    Limitations and Future Research

Like any other, the present study is not without limitations. We thus suggest future research addressing the limitations and shortcomings of the present study.

First, due to the conceptual nature of the present study, we cannot provide first-hand empirical evidence of the applicability of our framework. As a result, future research should empirically scrutinize the framework, including also the firm-dependent guidelines for assessing the optimum position along the continuums.

Second, due to space limitations, we have not discussed role of analytics and business intelligence (cf. Chen et al. 2012). Future research should investigate and elaborate on the analytical capabilities, tools and processes needed to obtain the desired business impact of each type of big data use presented in our framework.

Third, in addition to analytical capabilities, tools and processes, the strategic choices presented in and guided by our framework are likely to have ramifications on the firm level but also the network-level business model. We thus suggest future research focusing on how utilizing different types of big data should manifest in the business model and vice versa (cf. Woerner and Wixom 2015), including the network-level effects.

And finally, the big data driven changes to the strategic choices and decision-making processes are by no means limited to the questions of data use. The research on these changes is nascent, with ample room for further scholarly contributions.

# References

Abbas, R., Michael, K., Michael, M.G.: The regulatory considerations and ethical dilemmas of location-based services (LBS) a literature review. Inf. Technol. People **27**(1), 2–20 (2014)

Abbasi, A., Sarker, S., Chiang, R.H.: Big data research in information systems: toward an inclusive research agenda. J. Assoc. Inf. Syst. **17**(2) (2016)

Bharadwaj, A., El Sawy, O.A., Pavlou, P.A., Venkatraman, N.V.: Digital business strategy: toward a next generation of insights **37**(2), 471–482 (2013)

Boncheck, M.: 05/03, 2013-last update, Little Data Makes Big Data More Power-ful [Homepage of Harvard Business Review]. https://hbr.org/2013/05/little-data-makes-big-data-mor. Accessed 21 Feb 2018

Brynjolfsson, E., Mcafee, A.: Race against the machine: how the digital revolution is accelerating innovation, driving productivity, and irreversibly transforming employment and the economy. Brynjolfsson and McAfee (2012)

Cadwallar, C., Graham-Harrison, E.: 03/17, 2018-last update, How Cam-bridge Analytica turned Facebook 'likes' into a lucrative political tool [Homepage of The Guardian]. https://www.theguardian.com/technology/2018/mar/17/facebook-cambridge-analytica-kogan-data-algorithm?CMP=share_btn_tw. Accessed 18 March 2018

Chen, H., Chiang, R.H., Storey, V.C.: Business intelligence and anaytics: from big data to big impact. MIS Q., 1165–1188 (2012)

Constantiou, I.D., Kallinikos, J.: New games, new rules: big data and the changing context of strategy. J. Inf. Technol. **30**(1), 44–57 (2015)

Davenport, T.H., Barth, P., Bean, R.: How big data is different. MIT Sloan Manag. Rev. **54**(1), 43 (2012)

Gartner, Laney, D.: Last update, 3D data management: Controlling data volume, velocity and variety [Homepage of Gartner], [Online] (2001). https://blogs.gartner.com/doug-laney/files/2012/01/ad949–3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf. Accessed 21 Feb 2018

Gilchrist, A.: Introducing Industry 4.0, pp. 195–215. Springer (2016)

Greene, B.: 02/02, 2014-last update, How your boss can keep you on a leash [Homepage of CNN], [Online]. http://edition.cnn.com/2014/02/02/opinion/greene-corporatesurveillance. Accessed 15 Jan 2018

Hermann, M., Pentek, T., Otto, B.: Design principles for industrie 4.0 scenarios, 2016. In: 49th Hawaii International Conference on System Sciences (HICSS), pp. 3928–3937. IEEE (2016)

Hewitt, J.: 06/17, 2014-last update, Engineers create the first laser breathalyzer, for driveby DUI enforcement [Homepage of Extremetech], [Online]. https://www.extremetech.com/extreme/184050-engineers-create-the-first-laser-breathalyzer-for-drive-bydui-enforcement. Accessed 10 Jan 2018

Kagermann, H.: Change Through Digitization—Value Creation in the Age of Industry 4.0. Management of Permanent Change. pp. 23–45. Springer (2015)

Koskenvoima, A., Mäntymäki, M.: Why do small and medium-size freemium game developers use game analytics? In: Janssen, M., et al. (eds.) I3E 2015. LNCS, vol. 9373, pp. 326–337. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-25013-7_26

Kosinski, M., Stillwell, D., Graepel, T.: Private traits and attributes are predictable from digital records of human behavior. Proc. Natl. Acad. Sci. U.S.A. **110**(15), 5802–5805 (2013)

Lyon, D.: Surveillance After Snowden. Wiley (2015)

McAfee, A., Brynjolfsson, E., Davenport, T.H., Patil, D., Barton, D.: Big data: the management revolution. Harvard Bus. Rev. **90**(10), 60–68 (2012)

Newell, S., Marabelli, M.: Strategic opportunities (and challenges) of algorithmic decision-making: a call for action on the long-term societal effects of 'datification'. J. Strateg. Inf. Syst. **24**(1), 3–14 (2015)

Pariser, E.: The Filter Bubble: What the Internet is Hiding from You. Penguin UK (2011)

Progressive, N/A-last update, Snapshot discount system [Homepage of Progres-sive Casualty Insurance Company], [Online]. https://www.progressive.com/auto/discounts/snapshot/. Accessed 21 Feb 2018

Schildt, H.: Big data and organizational design–the brave new world of algorithmic management and computer augmented transparency. Innovation **19**(1), 23–30 (2017)

Sundaresan, S.: 04/05, 2017-last update, Don't speak, just think: The technology that turns thoughts into text [Homepage of World Economic Forum], [Online]. https://www.weforum.org/agenda/2017/04/technology-that-could-turn-your-thoughts-intotext/. Accessed 10 Jan 2018

Tilson, D., Lyytinen, K., Sorensen, C.: Desperately seeking the inrastructure in IS research: conceptualization of "digital convergence" as co-evolution of social and technical infrastructures. In: 2010 43rd Hawaii International Conference on 2010a System Sciences (HICSS), pp. 1–10. IEEE (2010)

Varian, H.R.: Computer mediated transactions. Am. Econ. Rev. **100**(2), 1–10 (2010)

Wang, J., Zhang, L., Duan, L., Gao, R.X.: A new paradigm of cloud-based predictive maintenance for intelligent manufacturing. J. Intell. Manuf. **28**(5), 1125–1137 (2017)

Weinberger, D.: Everything is Miscellaneous: The Power of the New Digital Disorder. Macmillan (2007)

Woerner, S.L., Wixom, B.H.: Big data: extending the business strategy toolbox. J. Inf. Technol. **30**(1), 60–62 (2015)

Yoo, Y.: Computing in everyday life: a call for research on experiential computing. MIS Q. **34**(2), 213–231 (2010)

Zuboff, S.: Big other: surveillance capitalism and the prospects of an information civilization. J. Inf. Technol. **30**(1), 75–89 (2015)