

1. Introduction

1.1 Statement of the Problem

01101000 01100101 01101100 01101100 01101111 00100000 01110111 01101111
01110010 01101100 01100100

In today's world we are surrounded by a vast amount of zeros and ones. Those numbers, called binaries and based on the Boolean logic, are the common way for computers to communicate with each other (Weyrich et al. 2014). Any digital information can be depicted in binary code. Machines and humans communicate in entirely different fashions since machines' communication is rooted within a computational logic (Kowalski 2011). Much like different languages, the communication between human and machine requires translation. The lines of numbers opening this chapter, for example, translated into writing spell "hello world", a familiar term for anyone studying a programming language. Both worlds (human world and machine world) seem to be separated with only few interconnections, but in recent years, the phenomenon of "big data" has led to an ongoing fusion of the two. Big data are here and big data are here to stay (Davenport 2014).

In an interplay of human and machine, both constantly generate new zeros and new ones. Be it me writing these words, a person talking on her/his phone, someone just walking the streets, be it a traffic light communicating with the control center, a smartphone logging into a Wi-Fi network, or a robot-arm in a factory. Every inhabitant of the modern world now inevitably leaves a digital trail, in addition, to her/his normal trail. The amount of data we generate on a daily basis is staggering, and increasing at an exponential rate. Everybody is participating in this digital world, contributing ever greater amounts of data. Most interestingly, nobody can hide from this digital trail. Even the decision of not participating at least generates the information of someone who does not wish to participate (Hartley & Chatterton 2001). In the future, the digital trail will become more precise, singular, and granular (Kucklick 2014) as there is an increase in the number of digital devices in the world as well as the frequency at which digital devices come to use (Economist 2014). Furthermore, the number of sensors that are capable of tracking people is constantly increasing and some sensors generate information about people as a by-product. These data only add to the already massive pile of existing data.

Although the information deluge (Hoenkamp 2012) is coupled with people leaving a more and more detailed digital footprint, the current development is predominantly driven by technology (Boyd & Crawford 2012). Big data pierce social life and society extensively. Society's solutions, however, are far from being precise or adequate. There is, in fact, a lack of social and ethical solutions (Barabási 2013). The need for a dispute of the social and ethical impact of big data is currently underestimated (e.g. Booch 2014), due to an inherent and precarious misconception. Big data are not the philosopher's stone (Crail 2015). Contrary to

the opinion of some researchers (e.g. Anderson 2008), they are not ever going to reveal a certain and objective truth (Van Dijck 2014). Data are subjective, contextualized, heterogenic, and incomplete (Dalton & Thatcher 2014), while at the same time emitting an “aura of truth, objectivity, and accuracy” (Boyd & Crawford 2012: 664). Partly misled, humans overestimate the preciseness of big data due to the seeming objectiveness and become overconfident on the basis of data (Miller, C. C. 2015). However, shaped by such a narrative (Kosslyn 2015), big data narrow down the image of human behavior excessively and focus on standardized archetypes. Big data contribute to a “demystification of the world” (Weber 1919: 9). Interestingly, the technology behind big data, however, is currently being placed inside of a black box (Pasquale 2015), itself becoming something inscrutable (LaFrance 2015), something mystical (in analogy to Clarke 1977). At the very least, Drucker’s statement (1967) that the computer is the moron, is no longer valid (Dewhurst & Willmott 2014).

There are reasons for outsourcing work and decisions to big data. In a complex world like the one we are living in, decisions need to be made in real time and under the pressure of a fluctuating and volatile environment which, therefore, makes constant change the new “stable” condition (Farjoun 2010). No human is capable of handling such massive complexity without the support of other humans and/or technological augmentations (Anderson & Rainie 2012). Big data are seemingly a technological enabler. Big data are a mixed blessing, supposedly capable of solving nearly any problem, but also the source of a staggering amount of new problems. Consequently, the mere use of big data will not suffice.

Deficit 1: Big data are not researched from a human perspective (Ekbja et al. 2015) and without a focus on the human factor (Zuboff 2014).

It is stated (Chen et al. 2012) that the usage of big data makes people’s behavior more calculable and predictable. On the one hand, there is always the danger of employees believing that they are watched, much like a post-panopticon (Bauman 2000, Bakir 2015) or the electronic whip (West & Bowman 2014). This causes them to adapt their behavior. At first, big data may resemble Taylorism and could possibly lead to Taylorism 2.0 (deWinter et al. 2014), both with negative connotations, although at a second glance Taylorism has the benefit of being comprehensible. On the other hand, the algorithms behind big data are becoming increasingly unintelligible and potentially inaccurate (Kleinberg & Mullainathan 2015). From a technological perspective, we occupy a land of milk and honey where we can “gather data first, produce hypotheses later” (Servick 2015: 493). But, as Davis states, “Big data is pushing us to consider serious ethical issues including whether certain uses of big data violate fundamental civil, social, political, and legal rights” (2012: viii). This discourse is currently lagging behind the technological progress (Kitchin 2014a), despite the increasing significance of big data (Shaw 2014) and the “need of deeper critical engagement” (Crawford et al. 2014: 1664). Given the undeniable potential of big data to solve major problems, such discussion is of utmost importance.

Deficit 2: Big data are not purely technologically driven; they are a social phenomenon. However, the relation between big data and society is highly underresearched.

Big data are closely entangled with humans, as they only unfold their potential when utilized. Big data do not magically develop solutions and do not work independently from humans. It is, therefore, impossible to separate big data from human interaction. Big data may act as a black box. People may not understand the way big data work and may suspect they have a life of their own. Big data have a strong impact at the human level and will influence people drastically (Mayer-Schönberger & Cukier 2013). An interdisciplinary approach to this upcoming discussion is essential since the context of implications will vary. Situational environments determine the use of big data. Differences become obvious in the relationships between government and citizen (Kim et al. 2014), supplier and customer (Strong 2015), and employer and employee (Davenport 2014). Transferring big data strategy to another relationship without making contextual adjustments bears the danger of being inappropriate and even harmful. One field of human interaction is economic organization and, in particular, the usage of big data concerning employees. Employees as an integral part of an organization are neither enemies nor mere resources to be exploited, but rather an employers' partner with shared interests. This makes finding a potential competitive advantage for the company by adapting big data appropriately a delicate process. It might burden the trusting relationship between employer and employee. Marketing methods applying the shotgun principle are promising as they could lead to an increase in sales (Mayer-Schönberger & Cukier 2013), but using such methods with employees may disrupt the employer-employee relationship and harm commitment, performance, and retention.

Deficit 3: Big data are researched in a general way and not from a contextual viewpoint. In particular, the effects of big data in economic organizations are underresearched.

Within an organization, and especially within corporations, every use of big data will influence human relations (Harvard Business Review 2013). Even big data use in apparently nonadjacent fields will have an effect. The use of big data in research and development, for example, will lead to the creation of new products, and new products will impose different requirements of knowledge and skills onto employees. Big data are, therefore, bound to change work within organizations. One point of intersection of big data and humans to be considered is the human resource (HR) department. As a result of electronic human resource management (HRM), HRM have a long history of collecting and applying data. Using data in the analysis of employee relations is not a new turn, but the vastness of available data will represent a challenge to HRM. There is already a lot of information about the employees available to use (Kull 2016). It seems logical that not every individual member of an organization will handle big data but big data require steering by

some entity within. The interests of both employers and employees will be incorporated into the use of big data. Consequently, big data as a technology will be driven by the IT department, however, as a social and human phenomenon will be designed and implemented by the HR department. At the moment, this discussion is predominately driven by practitioners and focuses on operational implementation. Big data will be a transformative power, but they are shaped by the people in the organization. The HR department can use big data to transform the organization proactively and adapt a new role, or leave this emergent but critical field to other departments. HRM will need to reinvent itself in order to deal with big data and use them for their purposes.

Deficit 4: Big data will force HRM to change and assume a new role in the organization. However, it is unclear what this role will look like.

1.2 State of Research

Big data is the buzzword today and many are willingly jumping onto the bandwagon. Big data are new, ubiquitous, and pervasive. However, big data and their effects on organizations are under-researched. Statements claiming that big data would lead to enhanced objectivism are not entirely true since big data are subjective, never neutral, but contextualized (Johnson 2015). Big data are not capable of knowing everything everywhere and anytime. However, this means that the explanatory power of big data is limited and that there is an inherent data bias within big data that leads to distortion between data and *reality*. Big data may potentially lead to a massive paradigm shift in society (Mayer-Schönberger & Cukier 2013) and especially in research (Puschmann & Burgess 2014). Human interaction and its embeddedness within a social network (or organization), in particular, will be shaped differently through this datafication (Lycett 2013). Data are already everywhere and will increasingly become the general mode of communication. Everything can be transformed into a representation of data (Frankel, & Reid, 2008). Big data will impact social life enduringly. This effect is relatively opaque, however, and differs from context to context (Manovich 2011). There are fragmented discussions about the subjective influence of big data (e.g. Boyd & Crawford 2012, Dalton & Thatcher 2014, Kitchin 2014a, Scholz 2015a, Metcalf & Crawford 2016).

Deficit 5: Big data may be subjective; this subjectivity is discussed in a certain context, but not in organization theory or HRM.

Organizations will be transformed by big data, therefore, becoming complex systems (Scholz 2015b). Furthermore, there is an abundance of influences on an organization, which brings about additional obstacles. An emergent trend is analyzing organizations from a complex systems theory perspective (e.g. Amaral & Uzzi 2007), and from the perspective of dynamization (e.g. Stein & Müller 2012). Organizations are already forced into transformation by external pressure from globalization, but big

data will further increase the velocity of such a transformation. Research is currently dealing with several aspects concerning the topic of such transformation, but the digital perspective and, therefore, big data, are still widely neglected. Digitization is undeniably a driver of change (e.g. Castells 2010, Stein 2015), but research is cagey about the topic (Knop 2014).

Deficit 6: Organizations are becoming more complex and dynamic; big data will act as a catalyst for complexity, but research is neglecting it in this context.

Big data will influence the organization, and underestimating their impact will probably be more harmful than dealing with the subject of big data. There are many aspects that make big data interesting. But the most important aspect is that there are big data within an organization that lie fallow. Smart factories and digitization leave a rising pile of data unexploited. No organization that is profit-oriented can look the other way. Big data's role as technological game changer is observable and undeniable. But big data will also change the way we work, although it is quite unclear in what way. It is foreseeable that the amount of data collected will massively increase in the future. Improvements in automation and the development of sensors as well as the gathering of human information will pile up the amount of data collected. In addition to that, data that already exist are normally not forgotten (Rosen 2012), as the capacity of storing data is constantly increasing (Hilbert & López 2011). But what does that mean for people within organizations? Postman singles out two distinct dystopian futures of information (processed data): "Orwell feared those who would deprive us of information. Huxley feared those who would give us so much that we would be reduced to passivity and egoism" (2006: xix). As a result, we are now moving towards a brave new world of data (Scholz, T. M. 2014) in analogy to the title of Huxley (1932). At the moment, both Orwell's and Huxley's predictions appear to be coming true.

Deficit 7: Big data will have an impact on people and organizations, but the potential outcome is still pending and needs further research.

Researchers will deal with big data from a theoretical perspective. Many disciplines discuss big data in very different ways, but thus far lacking a concise theoretical framework. Various existing theories (especially organizational theory) approach big data. Neither, however, is capable of understanding big data entirely. Even whether or not the term big data is precise enough, or whether big data are merely old wine in a new bottle may be debatable. The phenomenon itself, however, will not simply be rationalized away. There is a need to understand big data and the hefty influence of big data on today's world and to utilize this knowledge.

Deficit 8: Big data is not theory-less, however, there are no fitting theories available. This is especially true for organizational theory and HRM.

Up to this point, few publications have dealt with big data in HRM and those few are dominated by practitioners (e.g. Bersin 2012, CIPD 2013, Cornerstone OnDemand 2013, eQuest, 2013, Evolv 2013). In academic research, authors consider looking at individual aspects of big data in HRM, such as the management process (McAfee & Brynjolfsson 2012), analytics (Galagan 2014; Shah et al. 2012), performance (Levenson 2014), talent (Russell & Bennett 2014), workforce management (Miller 2013), and the new employment fields of data scientist (Davenport & Patil 2012, Davenport 2013, 2014), and chief data officer (Lee et al. 2014). There is a gap in the literature with respect to grasping the scope of big data in HRM, scientific discourse is lagging behind practical application (George et al. 2014).

Deficit 9: Big data in HRM are currently driven by practitioners, researchers are already behind them. However, it will be necessary to deal with big data in HRM from a research perspective.

1.3 Terminological Clarification

The term *data* will be omnipresent in the course of this thesis. The plural form will be employed in accordance with the conversation proposed by Kitchin (2014a). He quotes the Oxford English Dictionary:

In Latin, *data* is the plural of *datum* and, historically and in specialized scientific fields, it is also treated as a plural in English, taking a plural verb as in the *data were* collected and classified.

In modern non-scientific use, however, it is generally not treated as a plural. Instead, it is treated as a mass noun, similar to a word like information which takes a singular verb. Sentences such as *data was* collected over a number of years are now widely accepted in standard English.

While this thesis will refer to data as a plural term, the original version will be retained in quotations. Furthermore, 'big data' when labeling a theoretical concept, will be used in the singular form.

1.4 Objective of the Thesis

Concerning the theoretical foundation of big data, the relationship between human and big data, and the role big data play within an organization, research is currently relatively scarce. Although there are thousands of papers on the subject, many are purely technologically driven and neglect the human aspect of big data. But big data are bound to become an integral part of society and organizations. The human factor (Zuboff 2014) as well as the big data lens on humans (Aiden & Michel 2013) require research and a concise theoretical understanding

before actual effects can be analyzed. Big data are not theory-less but, as of yet, lack theory (West 2013, Monroe et al. 2014, Boellstorff 2015). While the obvious perspective on big data is a technological one, big data deeply penetrate the social environment, which is why social knowledge about big data is of utmost importance. Boyd and Crawford accurately describe the current state of research concerning big data as follows: “The era of Big Data has only just begun, but it is already important that we start questioning the assumptions, values, and biases of this new wave of research. As scholars who are invested in the production of knowledge, such interrogations are an essential component of what we do” (Boyd & Crawford 2012: 675).

Nevertheless, the classification of big data within the philosophy of science lags behind. This thesis will be rooted within three distinct philosophies of science. Firstly, in order to capture the impact of big data on society, organizations, and individuals, the thesis attends to the field of science and technology studies to which the relation between society and technology is the object of research. The research stream of organizational theory represents the second philosophy of science, focusing complex systems theory as well as systems theory, cybernetics, and population ecology theory. The third stream is human resource management research. Especially in the context of economic organizations, HRM research focuses on both closing the gap between research and practice, and transforming the organization adequately towards new innovations that emphasize the human factor.

In order to better understand big data and their interrelation with people and, consequently, the role of big data within an organization, a number of goals need to be met:

- **Deriving a theoretical understanding of big data:** It seems that we have a rough understanding of big data. They are vast, however, and there are many definitions.
- **Understanding the impact of big data on the socio-technological, organizational, and human resource-related level:** The technological aspect of big data alone is complex. But the topic of big data becomes even more complex when including society and the individual. Therefore, big data require analysis on different levels. As its main contribution, this thesis elaborates on a theoretical lens on big data from different theoretical perspectives, and constitutes big data as a social construct rather than a technological one.
- **Understanding the effect of big data on any organization, as well as their ability to transform it:** Big data will transform any social system. They will change any economic organization, thus transforming its very structure.
- **Describing the nature of this transformation:** The changes provoked by big data will fundamentally change the role of the human resource department. Therefore, this thesis is not concerned with the possibilities of big data with regards to an organization’s employees, but emphasizes the way in which big

data will be employed within the organization, and how the human resource department will perform the task of supervising big data.

Those goals aim towards developing a theoretical model for a data-augmented homeodynamic organization. This model introduces big data into the organization and describes their impact. It will transform the organization in a comprehensive way, enabling it to deal with big data in an efficient way and utilize them to generate a competitive advantage. That is why this thesis will theorize the impact of big data on an organization, on the HR department, and the people within. In the words of Huxley: "I mean, what I feel very strongly is that we mustn't be caught by surprise by our own advancing technology" (1958).