

3. Research Framework

3.1 Mental Model

After introducing the term ‘big data’ and defining it on the basis of several theories, the logical next step is the analysis of this reasoning with regards to organizations, as well as an estimation of the HR department’s potential future role. In a first step, I will propose my mental model for this new form of organization and the new role of the HR department. This model serves the purpose of outlining the reasoning behind my framework. Wittgenstein describes the need for such a model as follows:

“We make to ourselves pictures of facts. The picture presents the facts in logical space, the existence and non-existence of atomic facts. The picture is a model of reality” (Wittgenstein 1922: 28).

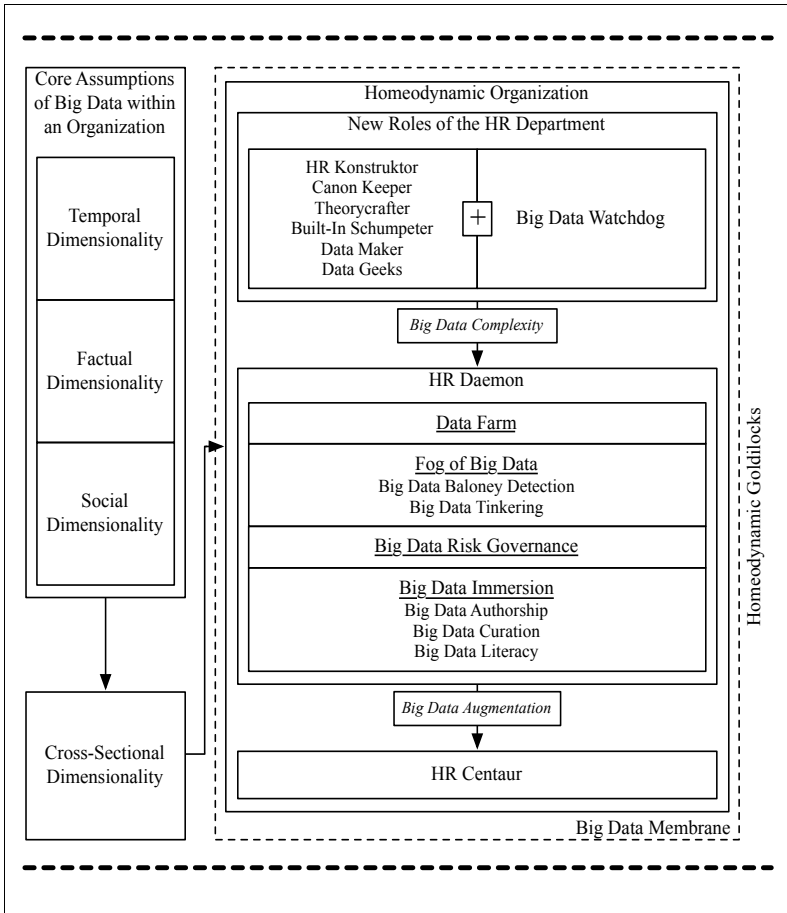
The term ‘mental model’, in this context, was first used by Craik (1943) who described it as the ability of an individual to use external input from alternative models and derive the best alternative, resulting in one concise model that can be presented to other people. Johnson-Laird (2004) describes the cognitive map (Tolman 1948) and the work of Peirce (1958) as precursors to the mental model. However, the term ‘mental model’ gained popularity through Forrester and his definition:

“The mental image of the world around you which you carry in your head is a model. One does not have a city or a government or a country in his head. He has only selected concepts and relationships which he uses to represent the real system. A mental image is a model. All of our decisions are taken on the basis of models. All of our laws are passed on the basis of models. All executive actions are taken on the basis of models. The question is not to use or ignore models. The question is only a choice among alternative models” (1971: 112).

Forrester explains that a mental model is not precise, but fuzzy, and is not complete but fragmentary. This is especially the case in social systems which is why knowledge about social systems may be insufficient. Forrester claims that “we do know enough to make useful models of social systems” (1971: 111). Consequently, the function of a mental model is to parallelize the thinking of individuals in order to achieve learning and knowledge that reach across individuals and do not solely exist within the individual’s mind (Senge 1990).

The interaction between people and big data will particularly be in dire need of mental models, in order to generate alternatives, be able to gain a better understanding and improve knowledge about this complex interaction. I will develop the mental model of big data within organizations, as shown in Figure 6.

Figure 6: Mental Model



It is essential first to describe certain core assumptions regarding big data within organizations. Big data will have an impact on organizations and this impact will not be static but rather highly dynamic. These core assumptions are unique in all organizations, but influenced by temporal, factual, and social dimensionalities (following Stein 2000 on the basis of Kluckhohn & Strodtbeck 1961). Their arrangement, too, is unique for every organization and will be dynamic. These core assumptions will, therefore, merge into one distinct cross-sectional dimensionality. This unique dimensionality will act as the situational parameter on which an organization will

depend, but will not be able to change in real-time; the organization will need to deal with it.

On the basis of the cross-sectional dimensionality, the general environment, and the influence of big data within organizations, I propose homeodynamic organization as a novel organizational type. It is derived from the homeodynamic concept introduced by Yates (1994) and, therefore, rooted in complex systems theory, but is expanded towards the need of dealing with big data. Consequently, any organization facing big data will transform into a homeodynamic organization and needs to react on this change. The driving force in dealing with big data will be the HR department. It is essential to highlight that big data in organizations will focus on the effect on the actors of said organizations, which means that employees are at the heart of my research.

The changes enforced by homeodynamic organization and, therefore, by big data, trigger a reorientation by the HR department. This reorientation will lead to new roles for the HR department. These new roles are oriented on the categorization of Ulrich et al. (2013), however, adapted to the unique settings within a homeodynamic organization. Therefore, I present six unique roles (HR konstruktor, canon keeper, theorycrafter, built-in Schumpeter, data maker, and data geek) and one cross-sectional role being the big data watchdog. All these roles are tackling certain aspects of the homeodynamic organization as well as the cross-sectional dimensionality introduced by big data.

However, this is just a response of the HR department to these fundamental changes, but as introduced in chapter 2.3, big data will increase the complexity within the organization. Reacting and changing the roles will not be sufficient, consequently, the HR department will create new structures within the organization. These new structures will mostly work in the background or the 'backend'. This construct that deals with big data and is created as well as implemented by the HR department will be called HR daemon.

The HR daemon comprises of a data farm that generates, cultivates, and harvests big data for organizations. The concept of a fog of big data is concerned with the problem that big data are not always precise and the challenges this impreciseness entails. It consists of a big data baloney detection which discovers faulty big data and big data tinkering which creates the possibility of exploring and searching for big data. Following this, the big data risk governance will be able to evaluate the risk of big data and combine it with the general risk, thus enabling the HR department to obtain a better sense of the potential risks and empowering senior management to make better decisions. The next component of the HR daemon is big data immersion dealing with certain aspects that are essential for handling big data and, consequently, required for any homeodynamic organization. It consists of big data authorship, big data curation, and big data literacy. Big data authorship tackles the question of data copyright and data privacy, thereby creating a solution that may work for organizations. Big data curation needs to keep big data in order and organized in a certain way, so that organizations do not drown in data. Finally,

big data literacy refers to the HR department training employees in the adequate use of big data.

Big data are bound to change the organization as well as the role of the HR department extensively. But big data and homeodynamic organization depend on the proactive usage by all people within the organization. As depicted in chapter 2.4, big data are nothing that will be helpful if they stay in the backend; consequently, big data augmentation is a proactive goal of the HR department to increase the usefulness of big data within the organization. The homeodynamic organization requires a 'frontend' implementation that deals with interaction interface between people and system, in this case the HR centaur.

The HR centaur will enable employees to utilize big data for their purposes and increase the effect of big data on the organization as a whole. It is a way to make big data available and usable for everybody within the organization and, by that, to transform big data into a resource of pro-activity for homeodynamic organization, rather than the organization just reacting on big data. A big data membrane constrains the border of organizations. Big data are everywhere and there are no boundaries to them. The goal is to achieve a way of protecting certain parts of the big data and keeping them secure. Other parts of big data can be shared freely and openly.

Finally, homeodynamic Goldilocks will emphasize that a data-augmented homeodynamic organization will only perform well within a certain range and will be stable only if certain criteria are upheld. Big data will help achieve this goal but will likewise impede the process depending on the core assumptions made concerning big data in the beginning. For that reason, Goldilocks will be different for every organization.

3.2 Methodology

From a theoretical perspective, big data is still a relatively novel phenomenon. Big data have, however, a great impact on today's society, organizations, and individuals. Big data are currently lacking a concise theoretical foundation. Many researchers limit their view on big data to the perspective of a certain academic field. The prime goal of theory in general is to describe and explain (Whetten 1989), but big data challenge researchers due to their vastness. Researchers are unable to fully grasp big data; there will always be certain blind spots in any theoretical conceptualization.

The foundation of understanding big data is generated in the use of data. But data are not theories and will not automatically lead to theories (Sutton & Staw 1995). It is also evident that big data will not ever be understood entirely and that big data are too big for one grand theory alone. Any theory will always be an approximation (Weick 1995), so does one about big data. Big data and the concept of the homeodynamic organization both have complex and dynamic definitions, and any theory will be a lengthy interim struggle (Runkel & Runkel 1984). As Weick (1995) explains, there are few fully-fledged theories, and, therefore, big data cannot

be made tangible by any comprehensive theory. It may be more fitting to ‘theorize’ big data and by that understanding big data as a more dynamic phenomenon. Weick describes theorizing as follows:

“The process of theorizing consists of activities like abstracting, generalizing, relating, selecting, explaining, synthesizing, and idealizing. Those emergent products summarize progress, give direction, and serve as placemarkers. They have vestiges of theory but are not themselves theories. [...] The key lies in the context – what came before, what comes next?” (Weick 1995: 389).

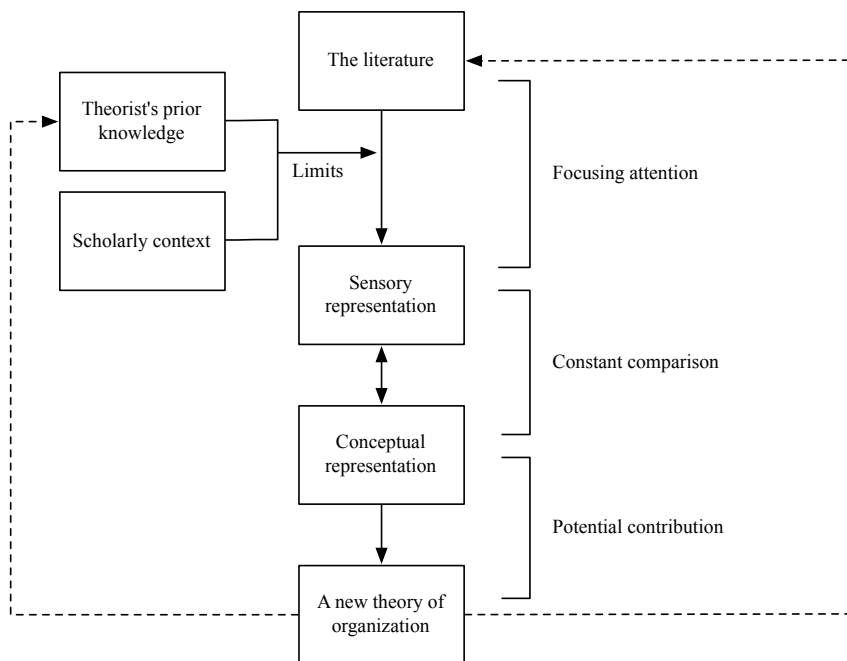
If big data are all about data, it may seem obvious to consider the grounded theory (Glaser & Strauss 1967) and analyze data in order to create theories rooted in a positivistic view (Martin & Turner 1986). This may be especially fitting, as there is no theoretical framework available, since grounded theory does not depend on a theoretical framework (Allan 2003). It remains debatable, however, whether grounded theory leads to a theory or even contributes to theorizing (Suddaby 2006, Thomas & James 2006).

Another way of theorizing big data lies in a thought experiment or an experiment-in-imagination (Hempel 1965) that would anticipate the impact of big data on the basis of certain general rules and derive the outcome by means of deductive inference. In the context of big data, however, deduction may not be sufficient. Although the premise is to derive conclusions from the general to the specific (Samuels 2000), the experiment calls for the question: what is ‘the general’ in big data? Obviously that would be $n = \text{all}$, but that is not achievable (Junqué de Fortuny 2013, Ekbia et al. 2014, Forsyth & Boucher 2015). The basis of the literature is also highly dynamic (Thompson 1956) which is especially true for big data. Consequently, deduction in the case of big data would take place from the bigger specific to the smaller specific. Induction may, therefore, be more suitable, as it moves from special observations to general ones (Samuels 2000). That, however, sounds relatively similar to the social-constructivism or the proposed data-constructivism. A third form is abduction. The term was coined by Peirce (1958) and Hanson (1958). Gregory and Muntermann describe abduction as the method of “creating a theory [...] based both on real-world observations that are inductively observed as well as theoretical viewpoints, premises, and conceptual patterns that are deductively inferred” (2011: 8). The term gained popularity in the field of artificial intelligence (Bylander et al. 1991). Abduction is, in a sense, a way of combining induction and deduction. That, however, would be an oversimplification (Mayer & Pirri 1996). Induction and abduction highlight the data and deduction and abduction focus on knowledge creation (Shepherd & Sutcliffe 2011). Induction and deduction, however, are not sufficient to theorize big data. Abduction reveals the potential of bridging both elements.

On the premise of bridging induction and deduction, Shepherd and Sutcliffe (2011) developed the inductive top-down theorizing approach in order to establish a method of deriving new organizational theories. The goal is to connect induction, deduction, and abduction in a coherent approach. However, it may be debatable

whether the authors are highlighting induction rather than abduction, especially as they state that the approach is “consistent with abduction” (Shepherd & Sutcliffe 2011: 361). Consequently, the name inductive top-down theorizing reveals a link to induction and deduction (through top-down) but no reference to abduction, although theorizing is concretized as “abductive theorizing” (Shepherd & Sutcliffe 2011: 371). The authors’ intention was to incorporate earlier literature as well as data and new literature to build a new theory. Contrary to a solely deductive approach, the data and new literature “speak to the theorist (through the formation of gists) to focus attention so as to detect tensions, conflicts, or contradictions” (2011: 362). They also follow the general understanding that theorizing is an iterative process (Thompson 1956) and theories merely milestones. Consequently, the theorizing process becomes more critical, and as Weick argues: “We cannot improve the theorizing process until we describe it more explicitly, operate it more self-consciously, and decouple it from validation more deliberately” (1989: 516). The model of such a theorizing process is shown *explicitly* in Figure 7.

Figure 7: Inductive Top-Down Theorizing (Shepherd & Sutcliffe 2011: 366)



For Shepherd and Sutcliffe (2011), academic literature is the basis of research. It underlies constant change, however, and can consist of papers, books, presentations, working papers, and so on. Such a body of literature is massive and, therefore, research focuses a researcher's attention so that it be influenced by both a theorist's prior knowledge and the scholarly context. From a self-reflecting perspective, I tried to keep the literature I used as extensive as possible, especially as I have a background in organizational behavior, HRM, and information systems. I also made the acquisition of literature ongoing and took literature notes (Eisenhardt, 1989). As Shepherd and Sutcliffe (2011) describe, the scholarly context represents another influencing factor to a researcher. It is for that reason that working at a German university in the field of HRM and organizational behavior also had some influence on my theorizing process.

The focus of attention gravitates around the influence of big data within organizations and the influence on the actors within organizations. Technological elements of big data are reduced to social influences, and are not described in detail. Sensory representation is focused on humans and big data within organizations. In order to derive a new theory from this sensory representation, a step towards conceptual representation is required:

“This conceptual representation refers to general abstract statements of relationships between constructs – incorporating explanations of ‘how’ and ‘why,’ boundary conditions of values, and assumptions of time and space – that allow for a more coherent resolution of the theorist’s sensory representation” (Shepherd & Sutcliffe 2011: 366–367).

Both representations are compared to each other constantly in order to achieve a coherent picture. The authors claim that through the use of thought experiments and metaphorical reasoning, a convergence between both representations is possible. Thought experiments are similar to experiments-in-imagination (Hempel 1965). I will apply several thought experiments to existing examples of big data use and compare them to the general concept I have derived from the literature. Due to the vast disciplinary variety of sources of literature, metaphorical reasoning (Tourangeau & Sternberg 1982) will likewise become necessary to converge sensory and conceptual representation. Especially in the description of certain behaviors of big data, several metaphors as well as exemplary cases are utilized to describe big data more precisely. Big data are hard to grasp, and probably even more so is the way in which big data create reality. Metaphors are needed to describe the phenomena in more detail.

As a result, the thesis derives a new theory of organization through the inclusion of big data which makes it a potential contribution to the theoretical discussion on the effect of big data on society. Shepherd and Sutcliffe (2011) described four attributes that make a theory strong. The first one is its broadness and that it goes beyond one disciplinary field (Kilduff 2006). The second one is its simplicity and that it depends on few assumptions only. The third is the theory's concern with interconnections and interrelatedness. The fourth is that a theory has only few different explanations. A thesis may reach a certain outcome, but more importantly becomes a “stimulus for new theorizing” (Shepherd & Sutcliffe 2011: 374).

