5-11-2023

Exploring the ontological status of data: A process-oriented approach

Dragana Paparova
University of Agder, dragana.paparova@uia.no

EXPLORING THE ONTOLOGICAL STATUS OF DATA: A PROCESS-ORIENTED APPROACH

Research Paper

Dragana Paparova, University of Agder, Norway, dragana.paparova@uia.no

Abstract

Information systems scholars have been inferring data as ontologically unstable and epistemilogically uncertain and mobile. Data have been conceptualized as distinctive from digital technologies and possessing properties to relate with other data, digital technologies, actors and socio-political environments. Across such relations, data stabilize into larger objects, but also change as part of actors' value-creation processes. However, data have been predominantly understood as open-ended, and the ability of data to simultaneously acquire structures and change has not been sufficiently explored – this requires an ontological investigation. The research question this paper seeks to address is "how can data, understood as both process and structure, be ontologically accounted for?". The paper offers two contributions. First, it unpacks the process ontology of assemblage theory to account for data as dualities of structure and change. Second, it provides an understanding of data as irreversible historical productions which simultaneously engage in enduring and changing processes.

Keywords: data, process ontology, assemblage theory.

1 Introduction

Data took on increasing significance in the information systems (IS) field, initially within discussions around "big data" stemming from pervasive digitalization and datafication (Lycett, 2013), and more recently concerning advanced technological developments such as machine learning and artificial intelligence (AI) (Faraj et al., 2018). The early works characterized data's volume – processing vast amounts of data; velocity – speed of processing data; variety – the heterogeneity of data sources and forms; veracity – credibility and reliability of the data sources; and value-creation, as data were used to fulfill various actor goals (Constantiou & Kallinikos, 2015; Lycett, 2013). This understanding of data came nonetheless from organizations collecting, sharing and using social media data, (Constantiou & Kallinikos, 2015) where data were collected without a pre-determined purpose and data's value was explored a-posteriori in actors' meaning-making processes. However, as Günther et al. (2017) raised, the collection, sharing and usage of other data types, such as personal health data, must be justified by a pre-defined purpose; therefore data's value potential is not always open-endedly explored by actors. These works imply that there are different degrees to which data can relate to other data, digital technologies, actors and socio-political environments.

IS research has also shown how data's value potential can be constrained by technical, organizational, or legal structures. The ability of data to acquire structure, have been discussed around data's capacity to aggregate and form objects (Aaltonen et al., 2021; Alaimo & Kallinikos, 2022) but also as part of larger phenomena, such as data governance. The literature stream on data governance explored the rules, rights, roles and responsibilities for governing data followed by conceptual (Abraham et al., 2019; Benfeldt, 2017) and empirical works (Parmiggiani & Grisot, 2020; Van den Broek & Van Veenstra, 2015). These works imply how data acquire structure as they are collected, shared and used within and across organizations, by following specific frameworks, rules and regulations.

Data are also a central area of interest to practitioners. For instance, as of 2019, the European Commission adopted the General Data Protection Regulation (GDPR) aiming to regulate the storage, processing and usage of personal data (European Commission, 2016). More recently, the European Commission (2020) also encouraged the development of "common European data spaces" as shared infrastructures which can accelerate access, sharing, processing and usage of data for innovation across various industries and sectors, such as healthcare, finance, energy. Such practical developments highlight the centrality of data in today's organizational work, (cross-)sector collaborations, national and international political and regulatory contexts, where data are not solely a resource for value creation, but also resources that require dedicated governance approaches.

Data, as a distinctive IS phenomenon, have been ontologically understood as unstable (Alaimo, Kallinikos, and Aaltonen, 2020); however the simultaneous processes across which data produce change and acquire structure have not been ontologically unpacked. The aim of this paper is to clarify the ontological status of data, by arguing for a process-oriented ontology, where fluidity and stability are dualities, instead of dualisms (Farjoun, 2010). Ontology, as "the science of what is, of the kinds and structures of objects, properties, events, processes, and relations in every area of reality" (Smith, 2003; p. 155), is concerned with the entities committed to theorizing and the relations with which such entities form larger wholes. The focus of this paper is to set the stage for re-examining *what is data*, by exploring the following research question: "*How can data, understood as both process and structure, be ontologically accounted for?"*. To answer this research question, I build on the realist, process-oriented ontology of assemblage theory (AT) (DeLanda, 2006, 2016; Deleuze & Guattari, 1987), where space and time, structure and change, stability and fluidity, entities and relations are mutually-enabling, instead of exclusive. By engaging with data's ontological status, this paper contributes to calls on producing novel theoretical and philosophical contributions in IS (Grover & Lyytinen, 2015), and IS debates on data (Aaltonen et al., 2021; Aaltonen & Tempini, 2014; Alaimo et al., 2020; Jarvenpaa & Essén, 2023; Tuomi, 1999) by providing an understanding of data as irreversible historical productions which simultaneously engage in enduring and changing processes.

The paper is organized as follows. Next, I provide an overview of IS research on data and the underlying ontological assumptions. In section three, I unpack data's ontology as relational and its implications for how we study data's structure and change. Section four introduces AT as a realist, process-oriented ontology, and presents the concepts of assemblages, virtuality and multiplicities as a useful vocabulary for understanding data's ontology. Section five argues how AT's process-oriented ontology can bring an understanding of data as irreversible historical productions simultaneously engaging in enduring and changing processes. Section six discusses the main contributions of the paper. Finally, section seven highlights the implications for IS research and practice.

2 IS research on data

2.1 Data as value

The early understandings of data in the IS field date back to the late 1990s, covering debates on the distinction between data, information and knowledge. As noted by Tuomi (1999), *data* were assumed to exist first as raw isolated facts or symbols, which are then interpreted or assigned meaning, relevance and purpose to become *information*. Knowledge is then extracted from data as a higher form of information. Tuomi (1999) challenged this hierarchical view of data very early on, by showing that data emerge last, only after structure and semantics are fixed to represent information. Therefore, data do not acquire structure, but are made by a structure, that is used to model, represent, and process them. Data took on larger significance in the IS discourse with the advent of big data stemming from more pervasive digitalization and datafication (Lycett, 2013), and the value creation potential of data has been central in many of those studies (Abbasi et al., 2016, 2016; Constantiou & Kallinikos, 2015; Günther et al., 2017; Kallinikos & Constantiou, 2015; Woerner & Wixom, 2015). The conceptual understandings of value creation from big data came not the least from organizations starting to use social media data. Constantinou and Kallinikos (2015, p. 54) describe the “heterogeneous, unstructured, agnostic, trans-semiotic nature of big data” – as differing from the well-structured data traditionally collected and used within a centrally controlled scheme – as social media data were captured so that they can be used a-posteriori.

As an empirical phenomenon, data and value were early on investigated in Aaltonen and Tempini's (2014) study on a mobile network operator. Their work shows how individual data tokens get re-grouped into larger audience-making events, which are incrementally formed and shaped to acquire meaning. The authors show how the actual meanings of data change over time, as “the employees perceived and acted on the assumption that there is more information in the data than that which is being actualized by the current metrics and reporting information” (ibid., p. 104). The meanings of data were also stabilized through the following mechanisms: 1) *semantic closure*, in which data were interpreted for a specific purpose; 2) *pattern-finding* or setting the parameters on how data were filtered in and out; as well as 3) *framing*, or reporting and presentation of the insights from such data. As the authors conclude, valuable information is only potential in the actual data, as the data pool is not deterministically useful or meaningful unless mechanisms on actualizing such value are set in place.

As empirical works followed, the focus started shifting from organizational perspectives on exploiting the potential of big data, to exploring the potential of data in environments where multiple stakeholders engage in simultaneous value creation processes, such as online platforms. For example, Barrett et al. (2016) explored how different forms of value were created in an online community for sharing healthcare data, as coevolving with stakeholders’ value creation processes. Over time, value was created in different forms, such as financial, service, ethical, epistemic, reputational, or platform value; resulting in four different value propositions of the online community, such as rating, connecting, tracking, and profiling. These value propositions were not given, or pre-existing, yet emerging as new stakeholder relations were established by analyzing and repurposing data in the community.

The non-linear value creation processes of data and their simultaneous evolution with digital technologies, was furthermore explored by Tempini (2017) in the empirical setting of a social media patient community. His study shows how value gets entangled as multiple stakeholders relate across a data-intensive infrastructure; thus bringing in non-linear and multi-dimensional value creation processes. As he notes, “the repeated updating and expansion of the web-based data-intensive infrastructure, exercised with a view to gradually improve and refine data practices across the network, repeatedly ignited cycles of value creation disruption. When an innovation disrupts shared practice, actors need to resituate data use processes in a way that is valuable according to any of the value dimensions they have stake in. Some dimensions of value creation could be enabled while others are hampered or shifted.” (p. 206). These works (Barrett et al., 2016; Tempini, 2017) clearly show how creating value from data across multiple actors is non-linear and dominated by fluid and unstable processes.

The epistemic role of data was furthermore explored around data’s mobility and ability to change on their “data journeys”, where some data are lost, others get merged and acquire different forms and meanings (Leonelli & Tempini, 2020). However, empirical accounts also show that beyond these fluid characteristics, data can acquire stable states, as they relate to other data in different sociotechnical environments (ibid.). Data get accommodated in larger infrastructures, standardized across digital technologies and governed to manage their dissemination and interpretation across multiple actors (Aanestad et al., 2014; Parmiggiani & Grisot, 2020; Peukert et al., 2022; Winter & Davidson, 2020). I now turn to unpack works exploring data's potential to stabilize into larger objects and acquire (relative) structure.

2.2 Data as objects

Beyond exploring data's epistemological status, IS scholars have also accounted for data's ontological status. Studies have defined data's properties, such as being *editable* (continuously revised, renewed and expanded), *portable* (shared across various digital technologies) and *re-contextualizable* (distanced from their origin and re-assigned meaning) (Alaimo, Kallinikos, and Aaltonen 2020). Data have also been characterized as being *comprehensive* (collections of tokens, behaviors), *granular* (can be aggregated, aligned and juxtaposed) and *unbounded* (have open-ended potential for acquiring meaning) (Aaltonen & Tempini, 2014). These properties provide data with a use-agnostic nature, as although they can be gathered for a purpose, their meaning is constantly explored, instead of reading pre-defined metrics. Data are, thus, not ready-made for usage, rather are often ambiguous and indeterminate, and need to be worked on, produced (Østerlie & Monteiro, 2020; Parmiggiani et al., 2021), aggregated and transformed (Aaltonen et al., 2021). As noted by Alaimo, Kallinikos, and Aaltonen (2020), these qualities bring "ontological instability" around what data are and "epistemological uncertainty" (p. 166) around how they are produced and what do they convey.

Moreover, IS scholars have also conceptualized data as being able to assemble into larger objects and commodities (Aaltonen et al., 2021; Alaimo, 2021; Alaimo & Kallinikos, 2022). For example, Aaltonen et al. (2021) show how data tokens relate according to certain criteria to be formed into objects and then assigned meaning as commodities, as they gain and lose properties. As they elaborate, data are actively produced in their journey of becoming commodities across which they constantly transform. In a study on social media data objects, Alaimo (2021) also shows how data and data objects are mutually coevolving. Data objects, as she elaborates, define what classifies as data and define interaction patterns, but such patterns are not fixed, yet constantly assembled. Another study by Alaimo and Kallinikos (2022) also shows how data objects come into being by aggregating data under a certain structure to acquire larger knowledge entities. However, these data objects are not the final output, but the intermediate step onto developing more complex organizational processes.

These studies show that data can acquire structure; however, such structure is not fixed, but relatively stable and characterized by an on-going process of change. Additionally, these works show how data enter relations not simply with other data, but also digital technologies, and organizational environments, in which they interact across complex patterns of structures, relations, processes, entities, agencies. This requires an ontological understanding of data's spatial dimensions, i.e., the structures generated by, and generating data, and the temporal dimensions across which data transform as they form larger wholes. I now turn to ontologically unpacking data's relationality across space and time.

3 The ontological status of data: Data as relations

Scientific ontology deals with foundational beliefs of what the world we are researching is comprised of. To be more concrete, what kinds of *entities, relations, processes, structures,* exist in such a world. There are two opposed ontological stances on what does the structure of the world consist of – substantialism and relationism (Cooper, 2005; Dainton, 2014; Kempton, 2022). Substantialists view the world as a *container* in which everything else exists and occurs. The ocean contains water, fish, algae, microorganisms, and other sea life. Data pools contain data, their relations, meanings, which exist as independent entities, fixed and finished forms, and can be clearly separated from their environment. Categorizing data as independent entities, would mean that data have clear boundaries, and although they can relate with the wider environment, such relations would be conceptually treated as secondary. This paper adheres with the other view – relationism – where it is the relations between entities, rather than the entities themselves, that are central.

Relations can be defined as connections, interactions, sequences, causes and effects, or as spatio-temporal (Kempton, 2022; Pentland et al., 2020). IS research on data, particularly in qualitatively-oriented studies, has been predominantly underpinned by a relational ontological stance. However, ontological views differ on whether entities can be separated from the relations they are in, i.e., whether data change through their relations with other entities, or if such changes can (also) be caused by an internal structure (Kempton, 2022). As raised by Kempton (2022) "[a]ssuming separation and stability can be problematic when studying contemporary digital phenomena", such as machine learning and data analytics, as such technologies learn and change over time (ibid., p. 02). As he continues, "it can be difficult to establish clear boundaries between the agencies of people and the agencies of machines, as the lines between them are blurred" (ibid., p.02). For instance, IS research has shown how humans and AI do not solely substitute, or complement each other, yet can function as integrated dynamic systems which get reconfigurated over time – assemblages (Grønsund & Aanestad, 2020). Therefore, data, algorithms, humans are not fixed, independent entities separated from one-another, but as a sociotechnical phenomenon, they are always in relation. Understanding data as ontologically always in relation to other entities, processes, actors, has implications for how we study data's structure and the changes they undergo, which I elaborate below.

First, *data's relations over space and time*. Space and time, as top-level ontological concepts (Dainton, 2014), have implications for how we study data's relationality. The dimension of space provides the conditions for the existence of some relations over others, i.e. some data relations get enabled while others get constrained. For example, GDPR enables the processing of personal data within the scope defined by law, or (informed) consent, but constrains the processing of personal data outside the purposes defined by law or consent. However, bringing in space without the dimension of time, would indicate that all possibilities in space are given at once; i.e., data are always in immediate relations with all possible data, digital technologies, actors, socio-political environments. Instead, by understanding time and space not as separate entities, but always in relation with each-other, we can show how as some relations actualize, new opportunities emerge; the relations data enter are always in the process of change. For example, different ways of processing data could be adopted, creating new opportunities for data usage, and triggering changes in the existing regulatory frameworks. Time, therefore, has an ontological significance for the movements of data in space, as it brings space in a constant process of production.

Second, *data as dualities of structure and change*. A relational approach to studying data also has implications for how we study data's structure and change. By putting relations, and not data as independent entities at the centre, we can focus on parts and wholes, data and the forms they take, not as separate, but mutually-enabling processes, i.e. dualities (Farjoun, 2010). Therefore, data's stable forms are not opposed to change, but both, a medium and an outcome of change. For instance, legal frameworks, such as GDPR, simultaneously limit and enable how actors collect, share and use personal data by promoting coordination, setting up a common set of rules, but also constraining that data are used for purposes other than the ones they were collected for. Understanding the structures data acquire and the changes they undergo as mutually-enabling, instead of exclusive, can help us show how data are constantly produced by, and simultaneously generating structures, but such structures are not static, yet change over time. As Langley and Tsoukas (2022) point out, stability and change are conceptually different, but ontologically inseparable, as they mutually interpenetrate to a point where one includes elements of the other. I now turn to elaborating on the specifics of assemblage theory as a process-oriented ontology.

4 Assemblage theory as a realist, process-oriented ontology

Realist ontologies in the social sciences have been traditionally structure-oriented and focused on causation stories, where entities with essences possess causal capacities to produce a certain outcome

(Rutzou & Elder-Vass, 2019). Realism, therefore, was demanding form, order, and clarity, as opposed to chaos (Rutzou, 2017). As such, it would not deny the existence of relations between entities, but these relations would be causal and relating structure to structure. This understanding of realism has been the dominating assumption in the social sciences. However, recently, DeLanda (2006, 2016) brought in an alternative approach to the structure-process divide, as per the work of Deleuze and Guattari (1987), by introducing the ontology of assemblage theory (AT). As noted by Rutzou (2017) "this is a realism which affirms the world without necessarily affirming our representations of that world" (p. 405). This take on realism is different and important for studying data as a sociotechnical phenomenon for the following reasons.

First, the realism in assemblage theory forefronts *difference*, heterogeneity, and change (DeLanda, 2006; Rutzou, 2017; Rutzou & Elder-Vass, 2019). The traditional takes on realism have been advocating for clarity and simplicity. Rutzou (2017) argues that the authenticity of this traditional realism can be questioned, as it can be perceived as too abstract of a substitute for the messiness of reality. As he notes, "there are different degrees, intensities, and balances between order and chaos in phenomena just as, we might argue, there are different degrees of openness and closure in systems" (p. 404). He further on elaborates how the realism in AT is heterogeneous, but not completely chaotic, as it is also characterized by interdependence, conditions, processes, forces, and structures. The realism in AT presents an ontology in which the complex, dynamic and open world is not settled enough to be reducible to independent entities, such as things and categories. However, it still recognizes that heterogeneous parts can form relatively stable wholes. As Rutzou and Elder-Vass (2019) clarify, assemblage theory as an "ontology is a complex interplay between heterogeneity and homogeneity, dynamism and recurrence, but heterogeneity and dynamism always seem to have the upper hand" (p. 406). This was exemplified by DeLanda (2000) by referring to genes. Genes are not a blueprint for the generation of organic structure and function. Rather, genes act as constraints on a variety of processes that spontaneously generate order in organisms, in a way teasing out a form from them. Therefore, genes do not predict the structure that processes form, but provide patterns of behavior across which multiple structures can unfold. Similarly, in a data world, algorithms do not predict the data outcomes, but provide a set of variables across which various data outputs can unfold.

Second, AT focuses on becoming and formation stories, in contrast to composition and causation stories (Rutzou & Elder-Vass, 2019). In structure-oriented ontologies, such composition and causation is assigned to the role of mechanisms. Mechanisms are isolated parts which are homogeneous and generated by structures that possess causal power, but mechanisms are also producing structures. Therefore, mechanisms relate structures to structures. The central element in AT, on the other hand, are not mechanisms but *multiplicities*. These multiplicities are not independent entities, yet "everything is always an active production, and processes of production, that inextricably flow and bleed together in vast interconnected networks and assemblages" (Rutzou 2017; p. 407). As elaborated in the original text by Deleuze and Guattari (1987), multiplicities resemble rhizomes. They do not have points of departure like roots, nor do they have an end, yet look more like a map which is open and connectable on all of its dimensions. Multiplicities can be broken, shattered, can intersect and merge into larger wholes, but will always connect and start time and time again. The multiplicities are, therefore, not only connecting structures to structures, but continuously entangle a variety of entities, structures, processes, and forces, which are not produced by causes, but become contingent through historical evolution. DeLanda (2000) exemplifies this by referring to the unfolding of events during the industrial revolution. Technology should not be viewed as evolving in a straight line, as if the advent of steam power and factory production were the inevitable outcome of the evolution of machines. Instead, mass production techniques were only one alternative among several and the fact that they dominated the development is itself in need of explanation. Similarly, surveillance capitalism, as conceptualized by Zuboff (2019), was not the inevitable outcome of datafication and social media networks, nor was it an unfortunate accident. Instead, it unfolded through selecting choices among possible alternatives, where the deterioration of personal privacy was not a determinate outcome.

Third, AT promotes a *flat ontology* in which all there is, are assemblages. Structure-oriented ontologies commonly differentiate between hierarchical levels for the existence of entities. In critical realism, such hierarchy is assigned to the empirical (observable events), the actual (all events, whether experienced or not), and the real (the actual and the causal mechanisms which have not been instantiated in the actual) (Rutzou & Elder-Vass, 2019). In AT, entities are not caused by deep structures and do not simply form at two levels – the micro and the macro. Instead, heterogeneous sets of entities can interconnect, and form an emergent whole, but such wholes are not a new ontological entity, i.e. a totality, instead are a unique entity operating at a different scale. DeLanda (2000) uses the term “scales” to indicate that in AT, entities such as institutions, are not totalities which act causally on lower-level entities, e.g. organizations aiming to process data. Instead, institutions have the same ontological status as data, emerging from the relations between smaller scale entities – such as data, organizations, digital technologies – just operating at different spatio-temporal scales. Similarly, knowledge is not hierarchically produced by creating information out of data (Tuomi, 1999), yet, data, information and knowledge can relate at different spatio-temporal scales. I unpack these concepts of assemblage theory in more details in the sections that follow.

4.1 Assemblages

The main concept of assemblages translates from the French word “agencement”, implying that the assemblage is not an outcome, but a process of assembling. However, recent interpretations of AT, with DeLanda (2006, 2013, 2016) as a central contributor, have adopted the term “assemblages”. The process of assembling is one of a double articulation. It starts, as DeLanda (2016) elaborates, from a set of heterogeneous parts which relate to form larger wholes, which wholes are then stabilized by repeating, enduring processes. Acknowledging that wholes stabilize through enduring and recurring processes does not indicate that assemblages are unities in which the identity of parts is dependent on the whole. Instead, the parts are autonomous and can be detached from one assemblage and attached to another where different sets of relations can be established. DeLanda (2006, 2016) calls these *relations of exteriority*, to refer to the relations established between autonomous parts, which can change, without the identity of the whole changing.

Taking heterogeneity as a starting point is what distinguishes AT from other realist ontologies. This heterogeneity does not arise by a mere arrangement of distinctive parts and their *properties*. Instead, it arises due to the heterogeneity of relations such parts can establish, or their *capacities*. In AT, the properties of the whole are not given, but emerge as the parts interact; the whole is not formed by arranging properties, but by parts exercising their capacities. Therefore, if the parts stop interacting, the whole will also not be formed. As DeLanda (2006) explains, “relations of exteriority also imply that the properties of the components parts can never explain the relations which constitute a whole, that is relations do not have as their causes the properties of the components parts between which they are established” (ibid., p. 11). Therefore, it is the relations between the parts, and not their properties, what brings identity to the whole.

Let us translate this to data. Data are heterogeneous components. They range from social media data generated about people’s behavior online, data about physical components such as pumps on oil and gas platforms, health data created to record people’s healthcare, or wellness. Data also posess certain properties, such as being editable, portable, recontextualizable (Alaimo, Kallinikos, and Aaltonen 2020). These properties enable data to relate to other components, such as digital technologies, including social media platforms, sensors, electronic patient record (EPR) systems; contexts, such as finance, healthcare, energy, oil and gas. Due to their properties, data *can* be edited and ported across digital technologies, organizations, and contexts; however, *whether* such relations will take place, i.e., whether data will be shared, also depends on the interoperability of digital technologies, organizational practices, or the regulations enforced to protect personal or non-personal data, among other forces. Therefore, the relations data exercise are not caused by their properties, but also depend on the properties and capacities of other entities they interact with.

4.2 Multiplicities

As DeLanda (2000) elaborates, a realist ontology cannot only incorporate the processes which bring entities into being, but also the processes which keep their identity over time. The term multiplicity, originating from matemathics, refers to the measurements of geometrical space. In AT, a multiplicity denotes how an emergent space (consisting of part-to-whole relations) is to be measured. As such, multiplicities are a core concept in the ontology of AT, as they replace what is essentialism in other realist ontologies. Essentialist ontologies are composed of fully formed entities, called unities, which possess a core set of properties that define what they are, as well as causal capacities to constrain their parts. Multiplicities, on the other hand, structure the possibility space of the assemblage by defining "spaces of possibilities" as the possible ways in which an assemblage can change.

Therefore, multiplicities define the degree to which assemblages can be formed and change, i.e. the *structure* of the assemblage. The structure of assemblages is defined by the distribution of two multiplicities: 1) invariant (more stable, recurrent, shared by many parts), and 2) variant (more unstable, prone to change). DeLanda (2013) refers to these multiplicities as universal and individual singularities, respectively. The distribution of multiplicities which are both, invariant and variant, stable and changing, brings certain regularities to the structure of the assemblage, but such structure is not strata, i.e., in absolute stability, nor is it in complete fluidity, yet always oscilates between stability and fluidity. This allows the structure of assemblages to accommodate entities, relations, processes, actors, which keep on changing as they interact over certain patterns of regularities over time.

The structure gives direction, but does not predict the relations assemblages will establish, as such relations are not a copy of the structure, yet correspond to it only to a certain degree. DeLanda (2006, 2016) explains how such relations unfold across simultaneous processes which stabilize or destabilize the assemblage, where the degree of stability is defined by two parameters, territorialization and coding. Territorialization refers to defining and sharpening the spatial boundaries of actual assemblages. Coding refers to increasing the degree of internal homogeneity of components and relations of an assemblage. The more territorialized and coded the assemblage is, the higher the degree of stability. At the same time, the assemblage is engaging in processes of destabilization; "[a]ny process which either destabilizes spatial boundaries or increases internal heterogeneity is considered deterritorializing" (DeLanda, 2006; p. 14). Therefore, assemblages simultaneously engage in processes of stabilization and destabilization, territorialization and deterritorialization, but never reach equilibrium, as their structure, and the relations they establish keep on changing. What is relevant is not what led to the relatively stable states, but the actual processes across which assemblages form and change.

Let us exemplify this through institutions, such as legal bodies regulating data-sharing. Institutions stabilize their identity by enacting and enforcing laws, and are more stable when they have well-defined spatial boundaries within which their jurisdiction applies. Any process which brings this jurisdiction in question blurrs the spatial boundaries and destabilizes institutions, making them more prone to unlawful behavior. For instance, overlapping laws about the collection, sharing and usage of data, as in the case of cloud technologies, blur the territories within which data jurisdictions apply (Daskal, 2015). In this case, data travel across countries' borders and there is a physical disconnect between where data are stored, where they are accessed from, and who owns them. Another example for destabilization of institutions are the discrepancies between law formulation, and actual implementation. For instance, laws formulated to protect the individual-level privacy of data subjects, cannot reflect population-level data processing perfomed by advanced digital technologies, such as AI, where data are related to each-other (Viljoen, 2021). Therefore, institutions can be stabilized by the continuous enforcement of laws within their jurisdiction, reducing the gap between law formulation and actual implementation, and their ability to sanction unlawful behavior within defined territorial boundaries.

4.3 Virtuality

Assemblages form along a structure of possible forms, but also keep on changing; this brings the need to unpack how assemblages unfold. AT has a flat ontology, defined around *the real, the actual, and the virtual.* The actual includes all relations that *are* actualized; e.g., all the ways in which data are shared across actors. DeLanda (2013) points out that "Deleuze speaks not of *realization*, but of *actualization* and introduces a novel ontological category to refer to the status of multiplicities themselves: *virtuality*" (p. 24). The virtual consists of all relations that *can* be actualized, out of which some will and others will not; e.g., all the ways in which data can be shared across actors. As per the words of Deleuze (2014) "the virtual must be defined as strictly a part of the real object – as though the object had one part of itself in the virtual into which it is plunged as though into an objective dimension" (p. 272). Therefore, the real consists of both, the actual and virtual; e.g., all the ways in which data *are* shared, *can* be shared, and *could* be shared across actors.

By introducing the concept of the virtual, AT shows how the actual can be realized in a variety of ways, which include mechanisms, but also reasons, and motives – producing non-linear causality. The non-linear causality comes from the status of multiplicities, which are distributed and meshed, and not sharply distinguishable from one-another, as essences are. In AT, the multiplicities are not given all at once – as is the case of essences – but *unfold progressively*, not by producing finalized forms, but by giving form to processes. Therefore, some relations actualize over others, not by being causally produced, or logically necessary, but by becoming historically contigent (DeLanda, 2006), as they could have unfolded otherwise. Due to the concept of virtuality, the focus in AT is not solely on how the actual unfolds, but also how it could unfold in conditions that may or may not occur.

Relating back to the example of the industrial revolution. The sequence of events that have taken place, such as using steam power, inventing machines, or the organization of workers into factories can be followed in the actual, but in the virtual, the possibilities for these events actualizing or not, and moving from manual to machine manufacturing, coexisted. Similarly, the sequence of events across which surveillance capitalism emerged can be followed in the actual, but in the virtual, they could have taken another route. Laws could have been adapted to reflect the technological advancements and constrain big tech companies to utilize personal data for commercial gains. Surveillance capitalism, therefore, was not caused by big tech companies exploiting data in a lawless space, but emerged through complex interactions among global adoption and usage of social media, large network effects, and destabilized legal frameworks where the laws were lacking behind the rapid pace of technological advancements and innovation – among other reasons.

5 Assemblage theory and its implications for data as a sociotechnical phenomenon

The implications of the concepts from AT – assemblages, multiplicities, virtuality – discussed in the sections above, can be summarized in the following ways. First, assemblages are always heterogeneous in the actual, as they are historically produced and unique. Second, assemblages get formed around a defined set of regularities which correspond to the virtual, but are not a copy of it. Adding in the virtual is significant in showing that as historically produced, the assemblages unfold across relations which are subsequent only in the actual, but in the virtual multiple possibilities coexist on how such relations can actualize (DeLanda, 2000). As DeLanda elaborates, "the ontological status of assemblages is two-sided: as actual entities all the differently scaled social assemblages are individual singularities, but the possibilities open to them at any given time are constrained by the distribution of universal singularities, the diagram of the assemblage, which is not actual, but virtual" (DeLanda 2013; p. 41). Building on these concepts, the realist, process-oriented ontology of AT can provide an understanding of data as being: *1) irreversible historical productions;* which *2) simultaneously engage in enduring and changing processes*.

5.1 Data as irreversible historical productions

The vocabulary of AT helps us understand data as *irreversible historical productions* which unfold progressively. Unfolding progressively means that not all the possible ways in which data can relate are given at all times. Instead, as some relations actualize, new possibilities open up, while others coexist. As data's forms and transformations progress in the actual, they become historically contingent, but not logically necessary, as in the virtual, they could have unfolded otherwise. Therefore, the changes data undergo cannot be logically decomposed solely by looking at the data forms in the actual; instead, multiple possible data transformations coexist, making the forms data acquire and the changes they undergo irreversible over time.

Let us consider AI-human-data assemblages. Algorithms pattern data into meaningful relationships using a set of variables. There are a variety of ways in which data can relate across these variables; for instance, the same set of data can be manipulated by the same algorithm and produce different outcomes; the same algorithmic outputs can be read by different humans and have distinct meanings. Therefore, the data outputs are neither random and chaotic, nor can they be calculated with absolute certainty. Instead, the actualized outcomes are one possibility among others, and the probability of each data output happening depends on what comes before and what can come after. This dependency can include the data used to train the algorithm, the previous set of operations performed on data, and the outputs of such operations, among other factors. Therefore, the data-human-AI assemblages resemble *possible* data outputs, but algorithms cannot *predict* the outcomes that will take place across the assemblage with complete certainty. As the outcomes are only probabilistic, data-human-AI assemblages cannot be logically followed back by solely looking at the data outputs, or the algorithmic patterns. Instead, these outputs involve historical contingencies, including the forms and meanings data have acquired, the human interpretations and alterations of algorithms, and the changing probabilities for each output happening over others. Therefore, the formations and transformations of data-human-AI assemblages become irreversible as they are not certain, but probabilistic and change among multiple outcomes over time.

5.2 Data as simultaneous enduring and changing processes

Understanding data as irreversible historical productions does not imply that data simply change over time; they also endure. By using the vocabulary of AT, we can understand data as simultaneously engaging in enduring and changing processes, where structure and change are a matter of degree. There will always be enduring processes across which data acquire forms and keep their identity over time; and changing processes across which data transform from one form to another, not by being chaotic, but by following certain patterns of regularities.

Let us, again, take data-human-AI assemblages as an example. Data, humans and AI relate across a set of parameters, programmed as algorithms. Algorithms pattern data using a set of variables, not by working as independent deep structures which cause data outputs, but by providing a set of instructions across which data outputs can unfold. However, as algorithms produce data outputs, they learn from the data and the operations they perform get altered. Some processes data engage in become contingent and enduring – certain steps in the algorithm get repeated unless instructed otherwise, or standards are developed for possible interpretations of the algorithmic outputs; other processes keep on changing – new data can come in, algorithmic rules are altered, humans re-interpret the algorithmic outputs. Therefore, over time, data acquire stable forms, as they stabilize across standards or data outputs; but also keep on transforming, as they can always be aggregated otherwise, and assigned another meaning; i.e. data engage in on-going processes of endurance and change.

Another example are data-intensive infrastructures (Jarvenpaa & Essén, 2023; Tempini, 2017). In healthcare, such data-intensive infrastructures hold legacy data, i.e. administrative or treatment-related health data about patients stored across various EPR systems. These legacy data endure over time by being standardized and interoperable across systems (Fossum et al., 2019), routinely produced and used as part of clinical work practices (Grisot et al., 2019), shared across systems, actors, organizations and regulations (Paparova et al., 2023). However, health data are not simply openly shared, as the conditions under which they are collected, processed or used is determined by healthcare personnel's official need in using such data for treatment purposes, the organizational rules and practices, legal and regulatory frameworks; i.e. processes which endure over time. At the same time, new data can become part of such infrastructures (e.g. sensor-based data from wearables), used in different ways across organizational work practices, requiring alterations of the legal frameworks, creating new possibilities for using health data for prevention and prediction of diseases; i.e., processes which change the data-intensive infrastructure over time.

Therefore, data can enter repetitive, regular processes, acquire larger forms and objects (Aaltonen et al., 2021), get standardized across digital technologies (Tempini, 2021), or patterned across algorithms (Grønsund & Aanestad, 2020); however these forms are not fixed and finished, as although data endure, the changing processes they enter always have the upper hand over time.

6 Discussion: Data, structures and change

This paper answers the research question: "*how can data, understood as both process and structure, be ontologically accounted for?*" by introducing the realist, process-oriented ontology of assemblage theory. With this, the paper contributes to calls on producing novel theoretical and philosophical contributions in IS (Grover & Lyytinen, 2015), by bringing in an ontology which can provide an understanding of data as irreversible historical productions which simultaneously engage in enduring and changing processes. The philosophical stances in IS, particularly in qualitatively-oriented studies, have been commonly centered around interpretivism and critical realism as the two poles, and their idealist and realist ontologies respectively. The former has been predominantly focused on processes and flow, and the latter on outcomes and stable entities. This paper presents assemblage theory (DeLanda, 2006, 2016; Deleuze & Guattari, 1987) as an alternative realist, process-oriented ontology, able to account for both, processes and outcomes, structures and change, as dualities (Farjoun, 2010). Assemblage theory has so far received only limited attention in IS (Hanseth & Rodon Modol, 2021; Tarafdar & Kajal Ray, 2021), and such focus was predominantly on applying its conceptual vocabulary. The ontological potential of AT has been recognized by other fields (Hodges, 2008; Rutzou, 2017; Rutzou & Elder-Vass, 2019), but has not been discussed in the IS discipline so far. Bringing in the ontological assumptions of AT in IS can drive the field further by providing an ontology which accounts for complex sociotechnical phenomena, which deal simultaneously with structures, processes, relations, entities, actors, oscillating across multiple dimensions (Rutzou & Elder-Vass, 2019), instead of solely from structures to structures.

The predominant assumptions in IS have so far been focused on the instability (Parmiggiani et al., 2021) and open-ended potential of data in actors' value-creation processes (Alaimo et al., 2020; Barrett et al., 2016; Tempini, 2017). Indications of data's ability to acquire structure have also been present, such as in studies on data objects (Alaimo, 2021; Alaimo & Kallinikos, 2022), or data governance (Abraham et al., 2019; Van den Broek & Van Veenstra, 2015; Winter & Davidson, 2020). More importantly, Leonelli and Tempini (2020) conceptualized data as historical productions, or data lineages as "not static objects whose significance and evidential value are fixed, but objects that need to be transformed in order to travel and be re-used for new goals" (p. 07). The authors foreground data's mutability and transformations as they travel, although they acknowledge how across those journeys data can have various degrees of stability.

This paper builds on this work (Leonelli & Tempini, 2020), and argues for understanding data as irreversible historical productions which simultaneously engage in enduring and changing processes. With this, the paper contributes to IS debates on data (Aaltonen et al., 2021; Alaimo et al., 2020; Parmiggiani et al., 2021), and the call by Jarvenpaa and Essén (2023) who encouraged novel theoretical approaches on "data sustainability" as "data's capacity to endure across technological and human generations" (p.10). This paper show how AT as a realist, process-oriented ontology can accommodate time *and* space, process *and* structure, fluidity *and* stability, endurance *and* change, to understand the transformations data undergo, and the forms they acquire over time.

Moreover, the process-oriented ontology of AT could help IS researchers to account for data as a sociotechnical phenomenon which has varying degrees of structure and change across different organizational, technological and legal contexts. For instance, the processes data enter can be more enduring or prone to change, depending on whether data are personal or non-personal; sensitive health data about patients, or maintenance data about physical components in oil and gas platforms; regulated by law or organizational contracts. By understanding data as having varying degrees of structure and change, we could study data as being more open-ended if shared without restrictions – such as open-government data; or following specific rules if they are sensitive and regulated by law – such as personal health data. This could bring a more comprehensive understanding that goes beyond data's properties of being editable, recontextualizable (Alaimo et al., 2020), or mobile (Leonelli & Tempini, 2020), to also encompass the structures across which data enter processes with other data, actors, digital technologies, socio-political environments.

This paper does not argue that AT as an ontology should be the preferable choice of IS scholars aiming to study data as a sociotechnical phenomenon. However, the concepts presented could be useful for scholars aiming to study data along the interplay of organizational, technological, legal contexts where data fluctuate across stability and instability over time.

7 Implications for IS research and practice

As noted by Little (2016), it is not possible to research a domain well if we do not know what things or processes it consists of. Metaphysical debates, such as "what is data?", and "are data separate, independent, fixed entities, or are they always in relation?", could provide rich foundations for developing the IS field's own metaphysics (Hassan et al., 2018). This paper explores the ontological status of data by arguing for a realist, process-oriented ontology which can accommodate structure and change as dualities. The ideas raised in this paper could contribute to IS researchers and practice in the following ways.

Data have become a central debate in IS, commonly conceptualized across larger IS phenomena such as data platforms (Alaimo & Kallinikos, 2017), data infrastructures (Tempini, 2017), data governance (Abraham et al., 2019; Parmiggiani & Grisot, 2020), data network effects (Gregory et al., 2021), artificial intelligence (Faraj et al., 2018; Grønsund & Aanestad, 2020). Understanding data's ontology could bring clarity around the role of data as part of these larger phenomena, but also to data as a phenomenon in itself. This could stimulate additional works focused on the distinctiveness of data as an IS phenomenon (Aaltonen et al., 2021; Alaimo et al., 2020), instead of solely treating data as a by-product of IT governance, digital platforms, or digital ecosystems. Moreover, understanding data's ontology as relational and produced through dualities of structure and change, could bring distinct research streams in IS closer, such as data-driven value creation (Alaimo et al., 2020; Tempini, 2017) and data governance (Abraham et al., 2019; Benfeldt, 2017), and encourage them to learn from each-other. Furthermore, the understanding of data's relations as historically contingent and irreversible, but not logically necessary, could ontologically ground phenomena such as algorithmic unfairness and bias (Schulze et al., 2022), as IS scholars and practitioners question the degree to which decisions can be automated and delegated to algorithms with or without human supervision.

Moreover, this paper provides contributions that bring the IS field closer to practical debates around data. One example is the work done by the European Commission on building trusted data spaces (2016, 2020) aiming to encourage actors to explore data's value potential across sectors and industries, while preserving the European laws, rules and regulations. The ideas presented in this paper could help IS researchers and practitioners understand data as simultaneously being governed by rules and regulations, and creating value, instead of treating governance as opposite to value. Furthermore, by focusing on data's relations, instead of data as independent entities, this paper could help practitioners understand data not as finished, fixed products, but as resources whose value needs to be worked on and continuously produced, as data, digital technologies, organizational practices and socio-political contexts are assembled.

References

Aaltonen, A., Alaimo, C., & Kallinikos, J. (2021). The Making of Data Commodities: Data Analytics as an Embedded Process. *Journal of Management Information Systems*, *38*(2), 401–429.

Aaltonen, A., & Tempini, N. (2014). Everything Counts in Large Amounts: A Critical Realist Case Study on Data-Based Production. *Journal of Information Technology*, *29*(1), 97–110.

Aanestad, M., Eike, M. C., Grünfeld, T. B., & Undlien, D. (2014). Building an Infrastructure for Personalized Medicine: Connecting the local knowing in practice to global "circuits of knowledge.". International Conference on Organizational Learning, Knowledge and Capabilities, Oslo, Norway.

Abbasi, A., Sarker, S., Aalto University, & Chiang, R. (2016). Big Data Research in Information Systems: Toward an Inclusive Research Agenda. *Journal of the Association for Information Systems*, *17*(2), 01-32.

Abraham, R., Schneider, J., & vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management*, *49*, 424–438.

Alaimo, C. (2022). From people to objects: the digital transformation of fields. *Organization Studies*, *43*(7), 1091-1114.

Alaimo, C., & Kallinikos, J. (2017). Computing the everyday: Social media as data platforms. *The Information Society*, *33*(4), 175–191.

Alaimo, C., & Kallinikos, J. (2022). Organizations Decentered: Data Objects, Technology and Knowledge. *Organization Science*, *33*(1), 19–37.

Alaimo, C., Kallinikos, J., & Aaltonen, A. (2020). Data and Value. In *Handbook of digital innovation*. 162–178.

Barrett, M., Oborn, E., & Orlikowski, W. (2016). Creating Value in Online Communities: The Sociomaterial Configuring of Strategy, Platform, and Stakeholder Engagement. *Information Systems Research*, *27*(4), 704–723.

Benfeldt, O. (2017). A Comprehensive Review of Data Governance Literature. *Selected Papers of the IRIS*, 120–133.

Constantiou, I. D., & Kallinikos, J. (2015). New Games, New Rules: Big Data and the Changing Context of Strategy. *Journal of Information Technology*, *30*(1), 44–57.

Cooper, R. (2005). Peripheral Vision: Relationality. *Organization Studies*, *26*(11), 1689–1710.

Dainton, B. (2016). *Time and space*. Routledge.

Daskal, J. (2015). The Un-Territoriality of Data. *The Yale Law Journal*, 326–559.

DeLanda, M. (2000). *A Thousand Years of Nonlinear History*. Zone Books.

DeLanda, M. (2006). *A New Philosophy of Society: Assemblage Theory And Social Complexity* Continuum.

DeLanda, M. (2013). *Intensive science and virtual philosophy*. Bloomsbury Publishing.

DeLanda, M. (2016). *Assemblage Theory*. Edinburgh University Press.

Deleuze, G. (2014). *Difference and Repetition*. Bloomsbury Publishing.

Deleuze, G., & Guattari, F. (1987). *A Thousand Plateaus: Capitalism and Schizophrenia*. University of Minnesota Press.

European Commission. (2016). Regulation EU 2016/679 of the European Parliament and of the Council of 27 April 2016, General Data Protection Regulation. *Official Journal of the European Union*.

European Commission. (2020). *European data strategy*. European Commission.

Faraj, S., Pachidi, S., & Sayegh, K. (2018). Working and organizing in the age of the learning algorithm. *Information and Organization*, *28*(1), 62–70.

Farjoun, M. (2010). Beyond dualism: Stability and change as a duality. *Academy of Management Journal*, *35*(2), 202–225.

Fossum, K., Fossum, S. M., Hanseth, O., & Sanner, T. A. (2019). Adaptive Networked Governance of E-Health Standards: The Case of a Regional Health Information Infrastructure in Norway. *ECIS 2019 Proceedings*, 16.

Gregory, R. W., Henfridsson, O., Kaganer, E., & Kyriakou, H. (2021). The role of artificial intelligence and data network effects for creating user value. *Academy of management review*, *46*(3), 534-551.

Grisot, M., Moltubakk Kempton, A., Hagen, L., & Aanestad, M. (2019). Data-work for personalized care: Examining nurses' practices in remote monitoring of chronic patients. *Health Informatics Journal*, *25*(3), 608–616.

Grønsund, T., & Aanestad, M. (2020). Augmenting the algorithm: Emerging human-in-the-loop work configurations. *The Journal of Strategic Information Systems*, *29*(2), 101614.

Grover, V., & Lyytinen, K. (2015). New State of Play in Information Systems Research: The Push to the Edges. *MIS Quarterly*, *39*(2), 271–296.

Günther, W. A., Rezazade Mehrizi, M. H., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. *The Journal of Strategic Information Systems*, *26*(3), 191–209.

Hanseth, O., & Rodon Modol, J. (2021). The Dynamics of Architecture-Governance Configurations: An Assemblage Theory Approach. *Journal of the Association for Information Systems*, *22*(1), 130–155.

Hassan, N. R., Mingers, J., & Stahl, B. (2018). Philosophy and information systems: Where are we and where should we go? *European Journal of Information Systems*, *27*(3), 263–277.

Hodges, M. (2008). Rethinking time's arrow: Bergson, Deleuze and the anthropology of time. *Anthropological Theory*, *8*(4), 399–429.

Jarvenpaa, S. L., & Essén, A. (2023). Data sustainability: Data governance in data infrastructures across technological and human generations. *Information and Organization*, *33*(1), 100449.

Kallinikos, J., & Constantiou, I. D. (2015). Big Data Revisited: A Rejoinder. *Journal of Information Technology*, *30*(1), 70–74.

Kempton, A. M. (2022). The digital is different: Emergence and relationality in critical realist research. *Information and Organization*, *32*(2), 100408.

Langley, A., & Tsoukas, H. (2022). *The SAGE Handbook of Process Organization Studies*.

Leonelli, S., & Tempini, N. (2020). *Data Journeys in the Sciences*. Springer Nature.

Little, D. (2016). *New directions in the philosophy of social science*. Rowman & Littlefield.

Lycett, M. (2013). ‘Datafication’: Making sense of (big) data in a complex world. *European Journal of Information Systems*, *22*(4), 381–386.

Østerlie, T., & Monteiro, E. (2020). Digital sand: The becoming of digital representations. *Information and Organization*, *30*(1), 100275.

Paparova, D., Aanestad, M., Vassilakopoulou, P., & Bahus, M. K. (2023). Data governance spaces: The case of a national digital service for personal health data. *Information and Organization*, *33*(1), 100451.

Parmiggiani, E., & Grisot, M. (2020). Data Curation as Governance Practice. *Scandinavian Journal of Information Systems*, 38.

Parmiggiani, E., Østerlie, T., & Almklov, P. G. (2021). In the Backrooms of Data Science. *Journal of the Association for Information Systems*, 64.

Pentland, B. T., Mahringer, C. A., Dittrich, K., Feldman, M. S., & Wolf, J. R. (2020). Process Multiplicity and Process Dynamics: Weaving the Space of Possible Paths. *Organization Theory*, *1*(3), 263178772096313.

Peukert, C., Bechtold, S., Batikas, M., & Kretschmer, T. (2022). Regulatory Spillovers and Data Governance: Evidence from the GDPR. *Marketing Science*, mksc.2021.1339.

Rutzou, T. (2017). Finding Bhaskar in all the wrong places? Causation, process, and structure in Bhaskar and Deleuze. *Journal for the Theory of Social Behaviour*, *47*(4), 402–417.

Rutzou, T., & Elder-Vass, D. (2019). On Assemblages and Things: Fluidity, Stability, Causation Stories, and Formation Stories. *Sociological Theory*, *37*(4), 401–424.

Schulze, L., Trenz, M., Cai, Z., & Tan, C.-W. (2022). *Algorithmic Unfairness on Digital Labor Platforms: How Algorithmic Management Practices Disadvantage Workers*.

Smith, B. (2003). Philosophical Ontology. In *Blackwell Guide to the Philosophy of Computing and Information*.

Tarafdar, M., & Kajal Ray, D. (2021). Role of Social Media in Social Protest Cycles: A Sociomaterial Examination. *Information Systems Research*, *32*(3), 1066–1090.

Tempini, N. (2017). Till data do us part: Understanding data-based value creation in data-intensive infrastructures. *Information and Organization*, *27*(4), 191–210.

Tempini, N. (2021). Data curation-research: Practices of data standardization and exploration in a precision medicine database. *New Genetics and Society*, *40*(1), 73–94.

Tuomi, I. (1999). Data Is More than Knowledge: Implications of the Reversed Knowledge Hierarchy for Knowledge Management and Organizational Memory. *Journal of Management Information Systems*, *Vol. 16*(3), 103–117.

Van den Broek, T., & Van Veenstra, A. F. (2015). Modes of Governance in Inter-Organizational Data Collaborations. *ECIS 2015 Proceedings*, 13.

Viljoen, S. (2021). A Relational Theory of Data Governance. *The Yale Law Journal*, 82.

Winter, J. S., & Davidson, E. J. (2020). *Harmonizing Regulatory Spheres to Overcome Challenges for Governance of Patient-Generated Health Data in the Age of Artificial Intelligence and Big Data*. TPRC48: The 48th Research Conference on Communication, Information and Internet Policy.

Woerner, S. L., & Wixom, B. H. (2015). Big Data: Extending the Business Strategy Toolbox. *Journal of Information Technology*, *30*(1), 60–62.

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (1st edition). PublicAffairs.