Smarter Work? Promises and Perils of Algorithmic Management in the Workplace Using Digital Traces

Inauguraldissertation
zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaften
durch die Wirtschaftswissenschaftliche Fakultät
der Westfälischen Wilhelms-Universität Münster

Vorgelegt von Joschka Andreas Hüllmann

Dezember 2020

Dekan:	Prof. Dr. Gottfried Vossen
Erstgutachter:	Prof. Dr. Stefan Klein
Zweitgutachterin:	Prof. Dr. Mary Beth Watson-Manheim
Datum der Disputation:	08. Februar 2021

Foreword

The present dissertation was crafted during my stay at the University of Münster and reports on my research progress at that time. This published version was revised and contains several updates compared to the first edition. I added two paragraphs in sections 3.1 and 3.5. I included this foreword. A curriculum vitae with a list of publications was appended at the end, and I updated the people analytics paper in the appendix.

My dissertation was made possible by the continuous encouragement and unwavering support of many friends and colleagues. In particular, my extraordinary gratitude goes to Stefan Klein, who provided me with the opportunity to work and follow my research ideals. His trust granted me the autonomy and freedom to embark on a highly individualized PhD journey. Stefan Klein sees his candidates' ideas from each their own perspective. He shares my wisdom that a fruitful discussion is one where the participants learn something. Stefan provided me (and still does) with timely, relevant, and valuable feedback, which in some cases, I should have listened to closer. However, as outlined in a seminal article by Martin A. Schwartz (2008) titled "importance of stupidity in scientific research", I choose to make my own errors and learn accordingly. Stefan Klein taught me how to do meaningful research and develop academic contributions beyond only publishing papers.

Furthermore, I would like to thank Mary Beth Watson-Manheim, who shared her vast experience with me and, together with Stephan Nüesch, agreed to sit on my advisory committee. I also want to mention Kai Riemer, who inspired me to pursue an academic path, and Mathias Fischer, who gave me the opportunity and wisdom to kickstart my journey.

This journey would have never come to a successful end without my dear colleagues and friends from the IOS research group with their day-to-day insights and ideas for my and our research. A special thank-you goes to my office roomie Jana Mattern, who was the most wonderful support I could have imagined during my time. On our joint therapist's couch, she made skilful use of her psychological background to motivate and cheer me up. If that proved a dead-end, both she and Tobias Berthold Hoge provided me with a couple of pints in my most dire times.

Last but not least, I want to express my appreciation for my family. I can always depend on my sister's support, Linda Hüllmann, who got married during my PhD journey. My parents, Andreas and Maria, who provide eager advice and always have my back. My mother might even have suggested, "why not pursue a PhD?" I followed their advice and here I am.

Joschka Hüllmann

2021-08-05

(note: this redacted version does not contain the full papers in the appendix, please find them online or on my website)

Widmung

Content

Tables

1 Algorithmic Management

Headlines about excessive data collection and algorithms that make decisions about people's lives are common. Insurance companies want access to health data[1], elections are influenced by targeted advertisements (Persily 2017), and algorithms decide who receives medical attention in COVID-19 triage[2]. Data, tracking, and algorithms are everywhere. Algorithmic management concerns business, politics, and society alike—every single citizen is affected. We are now living in a world where everyday life is increasingly digitised, and citizens' behaviours are tracked, stored, and analysed for profit. The ongoing controversies surrounding data tracking and analysis emphasise the need to explore the societal implications of algorithmic management. The emergence of new technologies and technological approaches, in particular, advances in computing power and data analysis, storage, and sharing, enable the rise of the "algorithmification."

This thesis addresses the algorithmification of the workplace, focusing on opportunities and challenges of algorithmic management at work, and digital traces research on work. When are the two activities appropriate? How can algorithmic management and digital trace research be implemented effectively, in terms of validity and meaningfulness? I contribute insights on these two questions, arguing that they espouse a duality and affect each other, resulting in a need to balance validity, appropriateness, and meaningfulness.

The structure of the manuscript is as follows. First, I outline recent research on algorithmic management and digital traces, pointing out where my research contributes. Then, I depict my research questions and the methodology of how I address the questions. A critical reflection, and discussion follow, concluding with the implications of my research and open questions for the future.

1.1 Nascent Research in Algorithmic Management

For the information systems discipline, algorithmic management is of immense interest. At the intersection of organisation, management, decisions, technology, and people, it understandably garners enthusiasts and sceptics from a wide range of academics in and beyond this discipline.

The enthusiasts emphasise the benefits of algorithmic management and its applications in all aspects of society, for example, advances in medical diagnosis, education, employment, entertainment, and the workplace (e.g., Berente et al. 2019; Chen et al.

---

[1] https://www.tk.de/techniker/unternehmensseiten/unternehmen/die-tk-app/tk-app-tk-fit-2023654 (accessed 2020-11-14).
[2] https://www.ft.com/content/d738b2c6-000a-421b-9dbd-f85e6b333684 (accessed 2020-11-14).

2012). Conversely, the sceptics point to the issues that surround implementing algorithmic management in practice, such as ethics, biases, discrimination, and privacy violations (e.g., Gal et al. 2020). In this thesis, I focus on the validity and implications of algorithmic management in the workplace.

Rapid advances in technology and the increasing deployment of algorithmic management call for more research on the promises and perils of algorithmic management to inquire how we can reap the benefits while addressing the problems that it poses. Multiple bodies within our discipline espoused such a call for research. The *Management Information Systems Quarterly* published a special issue call for papers in 2019 (Berente et al. 2019), while *Information Systems Frontiers* published a special issue call for papers in 2020 (Abedin et al. 2020). The premier International Conference on Information Systems (ICIS) in 2020 offered various workshops dedicated to the topic, such as "AI Beyond the Hype," and "The Future of Digital Work." Besides the large institutions, distinguished individuals within our discipline endorse this call for research, such as Shoshana Zuboff (2015, 2019) and Erik Brynjolfsson (Brynjolfsson and Mitchell 2017). In Germany, prominent funding bodies such as the State of North-Rhine Westphalia Foundation[3], or the BMAS[4], also support research in this topic.

1.2 Defining Algorithmic Management

With many researchers working on this topic, it is not surprising that definitional differences of "algorithmic management" occur. For this manuscript, I use it as an umbrella term that includes information systems that automate, augment, support, and inform decisions. In this context, these decisions are always directly related to people data or outcomes, which renders the topic sensitive and subject to ethical issues. In my use of the term "algorithmic management," I exclude decisions that are not related or only indirectly related to people. Despite the term "management," the decisions are not exclusive to an occupational context. Although this manuscript does emphasise the workplace, the learnings can be translated to other contexts, and later sections address a consumer perspective in more detail.

Initially, an "algorithm" depicts a well-defined sequence of computational instructions to solve a deterministic computational problem (Fernández-Macías 2018, citing the Merriam-Webster dictionary). In the context of algorithmic management, however, the term algorithmic is understood in a broader sense. While I do not exclude rigid decision

---

[3] https://www.ki.nrw/zertifizierung/ (accessed 2020-12-06).
[4] German Federal Ministry for Labour and Social Affairs (BMAS), https://testing-ai.gi.de/ (accessed 2020-12-06).

rules from my definition, my focus lies on probabilistic models, which are estimated from big data and then used to predict outcomes of interest (Crowston and Bolici 2020; Duggan et al. 2020). The models are based on multivariate statistical analyses, from basic regressions to complex machine learning algorithms such as random forests or convolutional neural networks. The data consists of a high number of observations, depicting logs of behavioural actions called digital traces. The combination of big data and complex statistics led to the euphemism "black box" because non-experts cannot comprehend how an algorithm arrives at its decision. In extreme cases, even experts and the algorithm developers themselves cannot explain how it arrives at its conclusion (e.g., Simonite 2019).

Technological advances and cultural changes are both drivers of algorithmic management. The technological advances are threefold. First, the processing power of computers is increasing exponentially, according to Moore's law, enabling complex statistical computations in a feasible time. Second, data storage is getting larger and cheaper, so organisations can track and save data on people's behaviours for a long time. Third, the diffusion of technology leads to increased use of smartphones, wearables, internet of things, as well as digital collaboration tools, for example, chats and video-conferencing, both of which generate digital traces. Combined, the three factors lead to the ubiquity of data (Brynjolfsson and McAfee 2014), with humans being called "walking data generators" (McAfee and Brynjolfsson 2012, p. 5).

Cultural changes parallel the technological advances; that is the willingness of organisations to share, collect, and analyse data, because they see value in big data and algorithmic decision-making (McAfee and Brynjolfsson 2012). Reinforcing the diffusion of collaboration tools is the shift towards virtual and multi-team forms of organising (Hüllmann 2019; Hüllmann and Kroll 2018). A contemporary example is the significant extent of working from home due to the COVID-19 crisis (Mattern et al. 2021). More generally, the widespread digitisation of social phenomena leads to the availability of digital traces of many situations. Table 1 lists selected examples, where algorithmic management has been applied. In the next section, this manuscript provides more in-depth insights into algorithmic management in the workplace.

Type of application	Authors
Criminal recidivism	(Faraj et al. 2018; Lee 2018)
Political micro targeting	(Faraj et al. 2018)
Personalised pricing	(Badmaeva and Hüllmann 2019; Hupperich et al. 2018)
Credit scoring	(Abedin et al. 2020; Goad and Gal 2018)
News recommendations	(Trielli and Diakopoulos 2019)
Performance evaluation	(Gal et al. 2020; Lee 2018)
Hiring	(Dastin 2018; Lee 2018)
Ride-hailing	(Lee 2018; Möhlmann and Zalmanson 2017)
Predictive policing	(Faraj et al. 2018; Waardenburg et al. 2018)
Medical diagnosis	(Abedin et al. 2020; Lee 2018; Walorska 2020)

Table 1. Applications of Algorithmic Management.
Shaded rows depict algorithmic management in the workplace, non-shaded outside of the workplace. However, the boundaries if workplace or not are blurry, e.g., criminal recidivism is work for the legal justice, but not for the affected criminal.

1.3 Algorithmic Management in the Workplace

The digitisation of work processes with the resulting availability of digital traces facilitates the application of algorithmic management in the workplace. According to Duggan et al. (2020, p. 119), such technology “changes how organisations manage work.” As a result, the deployment of algorithmic management at work raises questions: How do the roles of managers, workers, and decisions change (Crowston and Bolici 2020)? How do we manage the algorithms (Berente et al. 2019)? As algorithms perform some jobs, while workers continue to perform others, established premises about management and organising may be challenged. What jobs do the algorithms take over?

Inform, Support, Augment, Automate

As I mentioned earlier, the term algorithmic management subsumes decision-support and automation of decisions. Gronsund and Aanestadt (2020, p. 2) put it well when they assert that “automation is not an all-or-nothing” phenomenon. Instead, they argue that algorithmic management occurs alongside a continuum from “manual performance to automatic performance”.

Crowston and Bolici (2020, p. 2) distinguish the levels of this continuum into four areas:

– inform with no automation (manual information analysis),
– decision support (with automated information analysis),
– blended or augmented decision making (a decision is suggested and accepted by user), and
– fully automated (algorithm decides and implements decision action).

The level of automation that can be achieved with algorithmic management depends on job and task characteristics. Jobs consist of multiple tasks and are characterised into *routine* versus *non-routine* and *cognitive* versus *non-cognitive*. Non-routine jobs have a high task variety, and cognitive jobs have low analysability. High task variety and low analysability make automation difficult. Typically, jobs are partially automated, in which selected tasks are automated while others are not. Tasks with well-defined input and output are likely candidates for automation if enough data is available (Brynjolfsson and McAfee 2014; Crowston and Bolici 2020). Automating only single tasks of a job requires attention to the interdependencies of tasks, and the required coordination mechanisms to ensure proper alignment of the automated and non-automated tasks (Crowston and Bolici 2020).

If tasks are not suited to automation, algorithmic systems might produce invalid results or decisions with low validity. For example, prediction of temporal rhythms or social outcomes is notoriously difficult, although vendors claim otherwise (Hüllmann and Krebber 2020; Hüllmann and Kroll 2018). Since humans are affected, ensuring proper automation and valid decision-support is crucial (Hüllmann and Mattern 2020).

The benefits of automating tasks through algorithmic management include positive economic effects (Brynjolfsson and McAfee 2014; Faraj et al. 2018) in lowering prices and "boosting productivity" (Brynjolfsson and Mitchell 2017, p. 22). A contested proposition is the reduction of mundane jobs. On the one hand, authors argue that reduction of mundane jobs is beneficial because workers can perform meaningful tasks instead (Grønsund and Aanestad 2020). On the other hand, getting rid of mundane and straightforward tasks leaves only the complex tasks, which can negatively affect work quality if there is a skills-task mismatch (Riemer and Peter 2020). Other drawbacks of automation include the dehumanisation of interactions and lower perceived job security (Riemer and Peter 2020).

Since many jobs are only partially automated, algorithms and humans must work together. They influence each other. Humans must work with the results of algorithms and feed input into algorithms as well as build the algorithms. Lee (2018) suggests a "transition period" in which more and more algorithms are implemented to address managerial decisions, leading to a complex work environment which exhibits algorithmic decisions and human judgement (Grønsund and Aanestad 2020). The "interactions between algorithms and humans are under researched," according to Abedin et al. (2020, p. 1; via Harper 2019). How is this co-work effective? (Abedin et al. 2020; via Seeber et al. 2020). If tasks are automated, how does it alter tasks that are not automated? If tasks are augmented, how does it change the task and the role of worker?

Such constellations lead to a changing workplace, as the jobs and tasks that humans perform are altered. Beyond existing roles changing, new jobs and roles may be introduced. An example of such a new role is the "algorithmist" (Gal et al. 2020; Waardenburg et al. 2018), who mediates between algorithms and subject matter experts (e.g., employees and managers as users of the algorithm), having expertise in both domain knowledge and data analysis. The role acts as a translator. It assesses the plausibility of the algorithmic output, acknowledges the uncertainty of the decision, and communicates these things to users. As algorithms assume managerial duties, and responsibilities are shifted to new roles, the organisational power structures change (Lee 2018). What are the agency and the accountability of the algorithm and the new roles, especially for erroneous outcomes (Faraj et al. 2018; Østerlund et al. 2020)?

Despite new roles such as the "algorithmist," Lynne M. Markus (2017, p. 234) is sceptical that "keeping humans in the loop is the antidote for algorithmic intelligence risks." She argues that complacency ("non-vigilance of what an algorithm does") and bias ("users trusting results of algorithms more than they should") hinder an acknowledgement of the algorithm's shortcomings. This is exacerbated by a lack of expertise and errors in using, designing, and implementing algorithms (Markus 2017).

Expertise and transparency are relevant for assessing the validity of algorithms internally by organisations, but also externally through algorithmic audits. Suppose managers were looking to introduce algorithmic management into their organisation by buying software from the market. They discover that the inner mechanisms and logics of proprietary software solutions are opaque (Hüllmann et al. 2021a). Hence, they cannot evaluate the validity of the product. A remedy may be legislative regulations that enforce third-party auditing of algorithmic validity. However, it may be difficult for experts to properly assess the algorithmic outcomes—in particular as they are susceptible to the specific context (Howison et al. 2011). More theoretical and empirical work is required to address these areas.

Accept or Reject

Technological literacy is not only relevant for experts formally assessing algorithms. Users require technical literacy to make an informed decision about whether they accept or reject an algorithm's output. Burton et al. (2020) systematically review the literature, inquiring under which conditions people accept or reject the decision-aid of algorithms.

Technological literacy shapes users' acceptance because low technological literacy prevents an understanding of the algorithm (Burton et al. 2020; Crowston and Bolici 2020). Poor technological literacy results in users who are unable to recognise implausible results, and thus are unaware when human intervention may be required (Crowston and Bolici 2020; Faraj et al. 2018).

A perceived lack of control and autonomy lead to the rejection of algorithmic output (Burton et al. 2020). Actual lack of control leads to the same rejection, as the mechanisms of the algorithm may be opaque (Crowston and Bolici 2020; Hüllmann and Mattern 2020). Conversely, transparency can enhance the legitimacy and acceptance of algorithms, moderated by perceptions of meaningfulness of algorithmic decisions (Goad and Gal 2018). If workers do not trust valid algorithmic output, they will not follow recommendations, while trying to determine how the algorithm works (Lee et al. 2015).

The Opacity of the Black Box

Determining how an algorithm works and ensuring its validity seems to be a recurring problem.

With predictive algorithms, people are judged by "their propensity to act rather than on their actions" (Faraj et al. 2018, p. 3), while at the same time, the "logics of quantification" are reductionist of people's intentions, desires, and behaviours (Mayer-Schönberger and Cukier 2012). The algorithms are based on digital traces, historical behaviours of individuals, which do not take into account the individual's complex nature. The historical data is, by definition, incomplete. The labelling of the historical data to make the algorithm work is imbued with theoretical assumptions about the world, potentially leading to confounded outcomes, bias, and discrimination (Berente et al. 2019). The algorithms reinforce the biases through a feedback loop, exacerbating the relevance of questions in ethics and fairness (van den Broek et al. 2019).

An illustrative example is the movie "Minority Report," in which people are arrested because an algorithm predicts their criminal behaviour—which is not always correct. Another example from real life is predictive policing and criminal recidivism, which may promote and reinforce racial biases (Table 1).

For accurate predictions and meaningful decision support, algorithms need sufficient data. For finding patterns in the data that can be predicted, the patterns must be sufficiently stable (Crowston and Bolici 2020; Hüllmann and Krebber 2020). If not enough information is available, the algorithm should defer the decision-making to the human. Otherwise, implausible or erroneous results can occur (Crowston and Bolici 2020; Hüllmann and Krebber 2020). Recognising such erroneous results requires transparency of the algorithmic black box (Crowston and Bolici 2020), which is often not available for proprietary products (Hüllmann et al. 2021a; Hüllmann and Krebber 2021a;

Hüllmann and Mattern 2020). Goad and Gal (2018) suggest various ideas to address these issues and improve the accuracy and validity of algorithmic management—subsequently mitigating adverse effects on organisations. Algorithmic management should be advertised as fallible and reductionist, and the opacity of the internal mechanisms should be reduced. Instead, the people responsible for the algorithm should seek to explain its internal mechanisms.

Despite their ideas, further research is needed. Some algorithms are too complex to understand even for experts and technologically literate people. For example, top Google engineers still cannot fix the Gorilla image recognition software (Simonite 2019). The validity of algorithmic management using digital traces, in particular, requires further research (Howison et al. 2011; Hüllmann 2019).

Excursion into People Analytics

"People analytics" is the buzzword for applying algorithmic management to the human resources function, to optimise hiring, retention, engagement, among other goals. Levenson (2018) defines people analytics more inclusively, incorporating qualitative approaches in his definition. However, the dominant understanding is that people analytics concerns quantitative approaches based on multivariate statistical analyses (Hüllmann et al. 2021a; Hüllmann and Krebber 2021a; Hüllmann and Mattern 2020). There is little academic literature on people analytics because it is a practitioner-driven phenomenon. Nevertheless, the scientific discourse on algorithmic management expresses many issues relevant to the people analytics hype. Hence, I argue, the topic would benefit from considering the insights of the algorithmic management discourse (Hüllmann and Krebber 2021a).

1.4 Summary

The previous section outlines selected areas of research that deal with algorithmic management in the workplace. Early fieldwork shows that the introduction of algorithms into the workplace changes how people work, with broad implications for the individual, group, and organisation. The implications, and the sensitivity of algorithmic management, give rise to novel research questions that I address in this manuscript:

- In which scenarios is it appropriate and meaningful to implement algorithmic management?
- What procedures must be established for the correct, fair, and ethically sound implementation of algorithmic management?

The questions are closely intertwined with the validity of the algorithm. For example, is the algorithm powerful enough to decide, or does it merely provide a partial and uncertain indicator to inform a decision? Acceptance or rejection depends on the perceived validity of the algorithm. Discrimination and bias are the symptoms of an invalid algorithm.

Assessing the validity of algorithms is contingent on understanding the mechanisms of an algorithm, which is based on multivariate models that are estimated from digital traces. With surging interest in collecting and analysing digital traces, the information systems discipline initiates discussion on proper procedures for working with digital traces, which I will address in the following section.

2 Digital Traces

2.1 The Tracing of Everything Digital

Already in 2006, Eagle and Pentland (2006, p. 255) claimed that "mobile phones are wearable sensors." Today, in 2020, the significance of their claim becomes overwhelmingly clear, as the contact tracing for COVID-19 using mobile phones is of extraordinary importance in our fight against the pandemic (e.g., Ahmed et al. 2020; Martin et al. 2020).

Despite the widespread use of digital traces, the predictability of human behaviour is a highly contested question. While Princeton professor Arvind Narayanan puts algorithmic management using digital traces down as "… essentially an elaborate random number generator" (Narayanan 2019), others suggest that the social behaviour of humans is predictable to a high degree given the necessary data (Lazer et al. 2009; Pentland 2015; Song et al. 2010; Stewart 2019).

Research using digital traces has spawned a whole new community called computational social science, expressing interests in, inter alia, big data analytics, social media analytics, and network analysis; at the intersection of information systems, computer science, social sciences, and statistics. Recent publications and commentaries in the information systems discipline show the relevance and controversy of the topic.

Proponents of digital traces research hold high expectations (e.g., Maass et al. 2018), claiming that with the ubiquity of algorithms and the "algorithmisation" of the workplace, digital traces provide novel avenues and opportunities for research (Flyverbom and Murray 2018; Rahwan et al. 2019). The "algorithmisation" not only mediates change in the workplace but also makes it visible (Hüllmann and Kroll 2018). Hence, previously unobservable phenomena may be observed (Jackson et al. 2020). It is suggested that the computational analysis of digital traces offers new ways of theorising (Hedman et al. 2013) and facilitates the generation of novel theory (Berente et al. 2018). For example, as "most work practices involve digital technology" (Orlikowski and Scott 2016, p. 2), digital traces can enable processual research into the micro-practices at work (Hüllmann 2019; Østerlund et al. 2020). What actions do people perform? How do they work? Notably, this has already been suggested in 2001 by Barley and Kunda (via Hüllmann 2019).

In 2016, the *Management Information Systems Quarterly* announced a special issue on big data and analytics, followed by the *Journal of Management Information Systems* in 2018. Despite the expectations, cautious voices express concerns regarding the establishment of proper procedures for analysing digital traces that ensure the validity of inferences (e.g., Grover et al. 2020). My following questions complement my previously mentioned questions on algorithmic management, reiterating the correspondence between the two topics:

- In which scenarios is it appropriate and meaningful to analyse digital traces?
- What procedures must be established to ensure valid inferences from digital traces analysis?

The information systems discipline is perfectly suited for discussing and applying the analysis of digital traces (Agarwal and Dhar 2014; Howison et al. 2011). In 2018, digital traces research occurred in 16% of the papers in the basket of eight journals (Grover et al. 2020). Despite increasing numbers, Xu et al. (2020, p. 1258) posit that "… so far, the number of studies using these digital footprints in management research remains relatively small." Although our discipline is perfectly suited for digital traces research, Andersen et al. (2016) argue that we have yet to fully adopt digital traces, and alter our methods and theories accordingly. Conversely, various authors caution that elevating any problem that can be addressed with digital traces to an information systems problem will dilute the core values of the discipline (Grover et al. 2020). Instead, we should acknowledge that other disciplines, e.g., psychology (Landers et al. 2016) or media studies (Kneidinger-Müller 2018), utilise digital traces. In this regard, multiple people (e.g., Landers et al. 2016; Pachidi 2020[5]) have referenced Abraham H. Maslow's quote, "If the only tool you have is a hammer, you tend to see every problem as a nail" (Maslow 1966). They caution that we should not narrowly focus on digital traces. Instead, digital

[5] Personal conversation at the "AI and the changing nature of work" Cambridge Network webinar (2020-12-07).

traces should complement existing methods (Hüllmann 2019; Hüllmann and Krebber 2020).

2.2 Defining Digital Traces

Digital traces are historical, longitudinal logs of human behaviours and actions that are generated through routine technology use (Hüllmann 2019). While Berente et al. (2018) follow this definition and depict digital traces as longitudinal records of information technology use with timestamps, Andersen et al. (2016) suggest a narrower definition. They define digital traces as dyadic logs of interactions; that is, exclusively social media data. However, digital traces must not necessarily be relational nor interactional. Chaffin et al. (2017, p. 3) and Hedman et al. (2013) consider that digital traces are useful to capture arbitrary "behavioural constructs [at] the individual and team level." Digital traces exist in public and enterprise contexts, e.g., Facebook versus Facebook Workplace (Chen et al. 2012).

Newell and Marabelli (2015) dichotomise the generation of digital traces into passive and active. Passive generation of digital traces understands them as unobtrusive measures (Berente et al. 2018) that are a by-product of using information technology and that are not deliberately generated for research (Grover et al. 2020). In this case, Olteanu et al. (2019, p. 2) call them "digital exhaust," for example, logs from communication and collaboration systems such as email, chat, or social network platforms. The logs are stored centrally in the hosted locations of the software, and analysts extract them without end-user interaction, e.g., on premise or in the cloud (Hüllmann 2019). The traces can entail the complete history of using a particular technology (Howison et al. 2011; Hüllmann 2019), and people are not necessarily aware that these traces are generated and used for research.

Example sensors	Example content	Example methods
(Hardware) Camera Microphone GPS location tracker Bluetooth location tracker Infrared **(Software)** Desktop capture Click monitoring Time tracker	Locations Peers Time Tool use	Network analysis Parametric statistical analysis Linguistic analysis Web/mobile analysis Machine learning

Table 2. Examples of Sensors, Content, and Methods.
Adapted from Chen et al. (2012), Eagle and Pentland (2006), Berente et al. (2018), Hüllmann (2019).

On the other hand, active digital traces are generated through the deliberate deployment of sensors (Berente et al. 2018). Sensors can be hardware sensors, e.g., sociometric badges, fitness bracelets (Breiter and Hepp 2018), or software sensors, e.g., dedicated applications running in the background that collect activity data (Hüllmann 2019). Active digital traces are typically generated and collected directly from the individual's device.

Hedman et al. (2013) subsume digital traces under "big data" with high velocity, volume, and variety; petabytes of data on a high number of individuals with a high number of observations (Breiter and Hepp 2018; Chen et al. 2012). For example, "… 233,444 interactions from 12,017 users" (Hüllmann and Kroll 2018, p. 5). However, digital traces also occur as "little data" (Galliers et al. 2017, p. 185) that produce relevant insights and focus on a few individuals with a high number of observations. For example, "… 13,062 sent emails over 2.5 years" from five individuals (Hüllmann and Krebber 2020), or data from a single individual with 44,971 observations (Perer et al. 2006). Beyond the size in terms of number of observations and participants, I distinguish the data granularity by breadth ("number of measureable properties") and depth ("level of aggregation within each property") (Hüllmann 2019).

Structured metadata seldom leads to petabytes of data. For example, the publicly available archive of GitHub in 2019 has a size of 1.68 terabytes with 588,187,302 observations[6]. Although the primary form of digital traces is structured activity logs (metadata), unstructured data (rich media) can be attached to the activity data, inflating the size. With social media, for example, the structured metadata provides information about who is the sender and target of an interaction and at what time the interaction occurred. Attached rich media could then be the text content of the message, an audio message, or a video message. In this way, digital traces can lead to massive databases of video, text, and audio (Christin 2020; Grover et al. 2020) that can be used for qualitative research (Thapa and Vidolov 2020). In practice, however, rich media is often stripped from the datasets, as organisations seldom grant access due to the sensitivity of the data and the potential to leak personal information (Hüllmann 2019).

Different trace data requires different analysis methods. For example, social data may be explored using network analysis (Hüllmann and Hentschel 2021; Hüllmann and Kroll 2018) or natural language processing (Cetto et al. 2018), while non-parametric machine learning approaches process clickstream data (Rothmeier et al. 2020). Digital traces may also be counted in terms of frequency measures, which are then analysed by parametric methods such as linear regression (Hüllmann et al. 2021b; Hüllmann and Kroll 2018).

[6] https://www.gharchive.org/ (accessed 2020-11-18). You can check the size using Google BigQuery, where you can select a year and view the tab "details."

Table 2 depicts some more examples for illustrative purposes. Table 3 depicts a taxonomy of digital traces.

Characteristic	Instance		
Nature of Data	Historical, longitudinal logs of routine technology use		
Relationality	Monadic	Dyadic	
Generation	Passive	Active	
Sensors		Hardware	Software
Size	Big Data	Little Data	
Structure	Structured	Structured with media attached	

Table 3. Taxonomy of Digital Traces.
It is not exhaustive but includes the relevant types for this manuscript.

Despite being historical event log data, digital traces are not objective, but rather, reflect the values and motivations of the individual. They must be seen in the context of the social situation. The information technology artefact generating the traces with its software features affects the data (Andersen et al. 2016). The occurrence and absence of traces must be interpreted with theory and limitations of operationalisation in mind (Hüllmann 2019; Østerlund et al. 2020).

2.3 Datafication is the Driver of Digital Traces Research

Similar to algorithmic management, the increasing digitisation of social phenomena drives the production and analysis of digital traces. With the diffusion of mobile devices, wearables, and internet of things (Eagle and Pentland 2006; Newell and Marabelli 2015; Tonidandel et al. 2018), along with computational advances (Grover et al. 2020), digital traces become ubiquitous and cheap (Chen et al. 2012; Lazer et al. 2009; Shapiro and Varian 1999). In particular, work processes are digitised increasingly and observable in digital traces. Combined with novel computational tools, digital traces present opportunities for new research and insights (Berente et al. 2018). Chaffin et al. (2017) remark that these advances in mobile technology will enable data collection at an unprecedented scale and granularity. Breiter and Hepp (2018, p. 388; citing Damkjær 2015) provide an extreme example—a person's first digital trace is generated before their date of birth, and the last one is "beyond their death." The opportunities for analysing digital traces are ample.

Scharkow (2016) asserts that self-reported data is less accurate compared to digital traces. Perceptual responses, such as surveys or interviews, are subjective and sparse[7]. For example, surveys lack continuity even if they are panelled (Eagle and Pentland 2006).

[7] Scharkow does not criticise rich qualitative methods per se. Researchers often seek subjective and perceptual data deliberately.

The changing nature of work towards distributed and virtual work, crossing spatial and temporal boundaries, hinders interviews and ethnographic accounts (Barley and Kunda 2001; Hüllmann 2019). It is challenging to observe distributed individuals over extended periods (Hüllmann and Krebber 2020). Digital traces promise a less skewed and more complete account of the historical behaviours of humans than surveys[8] (Hüllmann 2019).

Recent research proposes how digital traces allow the combination of variance and process theoretical perspectives[9] (Pentland et al. 2020)[10]. Lindberg (2020) alludes to Shirley Gregor's (2006) theory types of explaining and predicting. He proposes that with digital traces, information systems research can combine both types of theory, blending a process and variance theoretical perspective, and produce novel and richer theories. For other forms of combinations, see Ortiz de Guinea and Webster (2017).

Digital traces and computational methods are researched to include contextual information and facilitate theory development (Lindberg 2020; Pentland et al. 2020). Both theory-driven and data-driven analyses approaches are feasible (Maass et al. 2018). Inductive and deductive approaches are possible (Galliers et al. 2017). Lindberg (2020, p. 108) claims that the gap between qualitative and quantitative inquiry is closing. He cites an argument from DeLanda (2005) about "extensive and intensive properties of phenomena." Rich theoretical constructs can increasingly be measured, or quantified, through detailed digital traces. At the same time, digital traces with attached media allow for qualitative assessments, and can be used together with other means of qualitative inquiry (Lindberg 2020). Despite often being depicted as strictly positivist (e.g., Pentland 2015), mixed methods with digital traces can assume different epistemological paradigms. For example, digital trace research can be conducted under interpretivist or pragmatist paradigms (Lindberg 2020), and grounded theory approaches are possible (Berente et al. 2018). Oesterlundt et al. (2020) suggest a sociomaterial approach to digital trace research together with interviews.

Despite the opportunities, digital traces should not be considered a panacea. Instead, they are a complementary tool in the academic's toolbox, which should be used together with other approaches. The analysis of digital traces poses various challenges and issues that should be heeded.

---

[8] A thought experiment: if you had unlimited resources and endurance, would an ethnographic account be more complete than the entirety of digital traces?

[9] See also Lindberg (2020), as well as Cloutier and Langley (2020). Original work from Langley (1999, p. 693), and Mohr (1982).

[10] Please note that Brian Pentland and Alex Pentland, both cited in this thesis, are two different people.

2.4 Issues and Challenges of Digital Traces

Two central issues of digital traces research are associated with theory and measurement. To make a meaningful theoretical contribution, researchers should avoid the trap of empiricism. Instead, the digital traces must be linked to theoretical constructs while addressing construct validity. Addressing construct validity is complex and requires caution because the traces are generated in context and cannot be interpreted objectively.

Theory in Research

Landers et al. (2016, p. 480) put it pejoratively and insist on avoiding "brute force empiricism". Such studies may suffer from the "streetlight effect" in favouring problems for which data is readily available over those that need substantive research (Rai 2017). These studies try to impress with big data sets instead of addressing meaningful problems that add to cumulative knowledge (Grover et al. 2020). The theory is of secondary importance and chosen ad-hoc or post-hoc, to fit the patterns in the data. Howison et al. (2011) call this "theoretical fitting." In worse cases, empiricist work may randomly test all possible relationships and generate hypotheses *a posteriori* (known as "data dredging"), because by chance you may find something significant, but the findings are only spurious correlations (Xu et al. 2020). Studies without robust theory tend to be "incremental and narrowly empirical" (Grover et al. 2020, p. 277), and do not generalise well from the specific context in which the data was collected. They are merely addressing "local" problems.

Digital traces research should not be conducted agnostic to theory. Instead, theory should be established explicitly (Landers et al. 2016), because only with theory do mere numbers become exciting information (Greenstein 1983, p. 4). The empirical setting of the study should be linked to a "broader knowledge goal," and preferably an information systems research archetype (Rai 2017). It should be considered, however, that theory from traditional studies in the offline context does not necessarily generalise to a digital context, as long as it is not evidenced through empirical data (Howison et al. 2011; Hüllmann et al. 2021b). For longitudinal studies, researchers should consider whether the phenomenon under study is stable for the observation period (Howison et al. 2011; Hüllmann and Krebber 2020).

Context, Interpretation, and Situations

Based on their literature view, Grover et al. (2020) claim that digital traces are often analysed under an implied positivist paradigm. Despite a high number of observations, digital traces remain reductionist, and should be understood as signals or indicators, not truth (Freelon 2014; Howison et al. 2011). They provide only a reflection on behaviours and are not objective (Østerlund et al. 2020). Instead, they can be interpreted in multiple—potentially opposing—ways (Freelon 2014).

As a result, the context and conditions under which digital traces are generated must be considered for analysis (Flyverbom and Murray 2018). The routine use of complex hardware and software systems generates digital traces. When, how, and how much these systems are used depends on the organisation in which they are deployed—the structures, context, and specific situation shape the human behaviour and the subsequent trace generation (Howison et al. 2011).

One example showing that digital traces are not objective is impression management, where individuals try to game the systems by deliberately producing performance data and metrics (Pachidi et al. 2016). This can have a detrimental effect on their performance. By doing this, they spend resources on impression management instead of on producing profitable outcomes (da Cunha 2013).

Further, absence of digital traces does not mean an absence of activity (Hüllmann 2019). The system's capabilities define what actions and behaviours generate a trace. A sudden absence of digital traces may occur due to shutting down a system, cleansing log files, faulty storage, or a system outage (Xu et al. 2020). Incomplete traces may bias the results and render them inconsistent, useless, or at worst, misleading. Software and hardware systems are dynamic, and which behaviours are logged may change (Howison et al. 2011). Thus, the validation of digital traces as a measurement instrument is an ongoing effort (Chaffin et al. 2017). Proprietary hardware and software systems exacerbate the problem of the measurement instrument's quality because the internal mechanisms are opaque, and researchers cannot check how the system works and generates digital traces.

Beyond the systems that researchers can investigate, backchannel systems or "shadow IT" may generate data that is not analysed (Hüllmann et al. 2021b). To address this, researchers suggest looking for anomalies or sudden changes in the data (Howison et al. 2011), checking the quality and reliability of the hardware and software systems generating the data (Chaffin et al. 2017), and consulting experts with "intimate knowledge" of the hardware and software systems (Howison et al. 2011, p. 20).

Ensuring the quality and reliability of measurements depends on the specific device being deployed (Tonidandel et al. 2018). Different vendors, or different devices from the same vendor, yield different measurements. For example, fitness bracelets that measure your pulse frequency can vary, and the generated timestamps for different chat software can also vary. Hardware sensors suffer from noise, sensor sensitivity and configuration, or misapplications of the sensor (Chaffin et al. 2017; Hüllmann 2019). Ultimately, attributing variance in digital traces that is caused by the measurement instrument to individuals and their behaviour may render a study's result invalid (Chaffin et al. 2017).

Construct Validity—Linking Digital Traces to Theoretical Constructs

Digital traces are logs of historical human actions and behaviours that are not necessarily generated for research purposes. To derive theoretical insights from digital traces, the measurement constructs (digital traces) must be linked to higher-level theoretical constructs (theory) (Chaffin et al. 2017; Howison et al. 2011). Establishing this link is called "operationalisation," while assessing the quality of this link is known as "construct validity." Do the digital traces measure what you meant to measure substantively (Braun and Kuljanin 2015, p. 521)?

Establishing the link between digital traces and theoretical constructs is not "straightforward" (Chaffin et al. 2017, p. 6). The operationalisation of the theoretical constructs is based on human design choices and must be argued and tested (Chaffin et al. 2017; Lindberg 2020). However, an argument for the operationalisation is often missing, and the construct validity in digital trace research is seldom addressed (Braun and Kuljanin 2015; Howison et al. 2011).

For other instruments, such as surveys, construct validity is meticulously established through tests, for example, in the psychology discipline. An instrument is deemed valid for measuring a theoretical construct when it correlates strongly with other instruments for the same theoretical construct. Braun and Kuljanin (2015, p. 523) formulate it in statistical terms: "construct validity is established when measures of the same theoretical construct, measured with different methods (e.g., self-report and behavioural trace) correlate more strongly with one another (i.e., convergent validity) compared with measures of different constructs measured with the same method (e.g., all self-report) and different (e.g., self-report and behavioural trace) methods (i.e., discriminant validity)."

However, establishing construct validity solely through quantitative instruments and statistical means is difficult, because it is unclear what the baseline or "ground truth" is (Braun and Kuljanin 2015). Instead, researchers suggest consulting subject matter experts, or triangulating with qualitative measures (Braun and Kuljanin 2015; Howison et al. 2011; Hüllmann and Krebber 2020).

Preprocessing and Analysis Decisions

Working with data sets, and in particular big data and digital traces, always requires preprocessing to avoid the "garbage in, garbage out" fallacy[11]. Preprocessing includes decisions on cleansing, normalising, or transforming the data; for example, removing all outliers beyond a certain threshold. But how do you determine this threshold? According to Xu et al. (2020), such thresholds, or preprocessing decisions, are often arbitrary and not theoretically guided. Researchers should either argue for their preprocessing decisions, grounded in theoretical assumptions, or report how other decisions would have affected the results (e.g., Hüllmann et al. 2021b).

Analysing large data sets of digital traces comes with peculiarities that require caution. Typically the studies are overpowered due to a high number of observations, resulting in conflated significance values (e.g., p-values). Effectively, given a sufficiently high number of observations, you will always find significant effects. Instead, more attention should be paid to the estimated effect sizes and the data generating process (Mertens and Recker 2020). For dyadic trace data, the assumptions of parametrical statistical analysis are often ignored and violated (Howison et al. 2011). For example, social interactions on an enterprise social network or chat platform are by definition not identically and independently distributed, because if one sends a message, another one will receive a message. Thus, the two measures—send and receive for two individuals—are not independent.

Privacy, Discrimination, and Ethics

Privacy concerns depend on the mode of digital trace generation. The active and deliberate generation of digital traces for research purposes can include informed consent by the study participants *a priori*. Participants can voluntarily share their active traces from only a limited observation period, for example, they wear a sociometric badge for two weeks. Passive digital traces, on the contrary, are historical logs generated as a by-product of using a software or device. They potentially span long observation periods, and may not be deleted or cleared for years. Researchers can only ask for consent *a posteriori*, after the data has been generated. When researchers use passive traces, individuals may not know that their data is being used for research, which poses ethical concerns (Markus and Marabelli 2017; Tonidandel et al. 2018). For example, studies conducted on social media platforms such as Twitter or Yammer typically do not inform each user account that research is going on. The reason is often that researchers only get

[11] „Garbage in, garbage out" is a phrase from computer science that says if you provide invalid or „bad" input to an algorithm that you will receive invalid or „bad" output.

deidentified (anonymised) data and thus are unable to contact the users. Despite analysing deidentified data sets, the privacy threat of re-identification exists. Re-identification works, because the parts of the data set that are not deidentified may be unique to such a degree that they can be matched with publicly available user information (Cavoukian and Castro 2014; Narayanan et al. 2011; Rocher et al. 2019). For example, we may conduct a survey and remove the names of the individuals in the data set, but keep the gender. If there is only one male in the data set and we know who the male person is. Then, we can look up what this person has answered in the survey.

Long term and detailed accounts of human behaviour are sensitive, and workplace data is proprietary. As a result, the data cannot be shared for open access, preventing independent researchers from replicating the results. Another issue is that the probabilistic models based on historical traces include historical stereotypes, leading to biased estimators. Such biased estimators invalidate the results of a study and are a recurring problem for algorithmic management (Tonidandel et al. 2018).

2.5 Summary

Digital traces are observations of human behaviour and come in different forms, for example, log files from social media platforms, or measurements from hardware sensors. They are longitudinal and often collected without the knowledge of the subjected individuals. With the increasing digitisation of social phenomena at work, digital traces provide an appealing data source for research and analysis.

However, the pitfalls should be considered when conducting research with digital traces. Merely looking for patterns in the data does not necessarily produce generalisable insights that contribute to broader knowledge. Instead, research with digital traces should keep theory in mind. Linking digital traces to theoretical constructs requires caution, argumentation, and evidence, acknowledging that they are not generated in a vacuum, but in social situations with context. Ensuring the validity of digital traces in one context does not necessarily generalise to other contexts.

Digital traces provide a partial view of human behaviour based on the specific types of traces used. Here, the instrument used, for example, sensor data, logs from social media, and its peculiarities, such as measurement noise or technical configuration, should be recognised. The validity problems are exacerbated by the fact that the computational methods and algorithms are becoming more complex, leading to an opacity of internal mechanisms and a lack of explainability (Rahwan et al. 2019). Olteanu et al. (2019) provide an overview of what can go wrong with analysing social traces.

Generally, researchers recommend triangulating and performing mixed-methods research with digital traces, complementing it with qualitative data which provides the necessary contextual insights and reduces validity concerns. Calls for triangulation are plentiful (Breiter and Hepp 2018; Freelon 2014; Grover et al. 2020; Howison et al. 2011; Hüllmann and Krebber 2020; Østerlund et al. 2020; Tonidandel et al. 2018). All cited authors in the previous paragraphs are proponents of triangulation and mixed methods.

Besides theoretical and validity issues, the extent and granularity of data collection raise concerns of privacy and spawn questions about ethical implications. Individuals cannot avoid generating traces and are subject to research without their knowledge. Some claim that with big data, established methods become obsolete, due to the "big data revolution" (cited from Breiter and Hepp 2018, p. 391). Nevertheless, theorising is as vital as always, since digital traces are reductionist and only indicators of social phenomena.

Requiring many branches of expertise, e.g., statistics, computer science, organisation, and sociology, our discipline of information systems is well suited to inform and guide digital traces research. I have previously argued that the validity of digital traces research does not necessarily translate from one context to another. When and how digital traces are a proper measurement instrument requires further research. If and how insights may be translated to other contexts is unclear, and it remains to be seen how well we can rely on the validity of prior work, requiring intuition and future research. I suggest a dedicated effort to validate digital traces in independent studies as a new type of study (e.g., Hüllmann and Krebber 2020). This thesis provides the first step in that direction.

3 My Research Questions and Outline

Algorithmic management and digital traces are linked in that digital traces analysis is used as a means for implementing algorithmic management. Recent work on digital traces understands the topic as a methodological matter, presenting opportunities and challenges for academia in the application of digital traces analysis. Conversely, algorithmic management, although not exclusively based on digital traces, makes great use of digital traces in practice. Here, research addresses the implications of domain-specific applications of digital traces analysis in society, presenting questions of ethics, validity, and fairness. Both topics are linked, as proper procedures for digital traces are contingent on the application context. In contrast, the ethics, validity, and fairness of algorithmic management depend on the concrete implementation of the digital traces analysis.

As a result, I devise two research questions. The first seeks to identify situations and domains that are suitable for algorithmic management and digital traces research. The second deals with the mechanisms and methods, how proper implementation of algorithmic management and digital traces research can be achieved. These questions guide my research and this thesis:

- In which scenarios is it appropriate and meaningful to implement algorithmic management and analyse digital traces to derive insights?
- What procedures must be established to ensure valid inferences and appropriate mechanisms from digital traces analysis, and the implementation of algorithmic management?

When referring to "meaningful," I mean that the implementation or analysis yields significant, relevant, or valuable insights. For example, the application of algorithmic management in practice can yield improved decision-making, resulting in more efficiency, productiveness, competitive advantage, or wellbeing for employees; digital traces analysis in research can lead to original, or novel contributions that add to the lasting and cumulative knowledge of our discipline. When referring to "appropriate," I mean that the individuals involved, including stakeholders such as analysts, developers, data subjects, managers, and researchers, are accepting of the moral, ethical, fairness, privacy, and accountability implications of algorithmic management and digital traces analysis. Under "appropriate" I subsume measuring the extent of data collection, tracking, and algorithmic management, as it is a prerequisite to discussing appropriateness. The dimensions "appropriate" and "meaningful" are not binary, but continuous and can assume different levels of values. Reaching high levels is a tradeoff because meaningful insights can hurt appropriateness. The first question, thus, balances appropriateness and meaningfulness.

While the first question addresses in what scenarios and when to analyse digital traces, the second question emphasises how to do it. What guidelines should be followed when analysing digital traces to achieve a high validity of the results? Here, the key is finding the right balance of validity and appropriateness. For example, excessive collection of digital traces leads to higher prediction accuracy, but less privacy.

There are not only tradeoffs within the two questions, but also between them. How to effectively implement an analysis depends on the scenario in which it is going to be deployed. As a result, the questions will neither be answered exhaustively nor definitely. Instead, the answers will be partial, as they depend on the context, application, and implementation of algorithmic management using digital traces. Since we can only address and evaluate these questions per context, this thesis presents selected studies. Generalising from these studies requires human intuition and judgement, and is thus subjective. However, this thesis provides my views on the generalisability of the insights from the studies in the discussion section.

This thesis does not present a philosophical argument, or an empirical argument on the morality or ethics of algorithmic management and digital traces. Rather, the focus lies on the methods and validity of the conclusions drawn, with references to issues of appropriateness. For the people analytics part, this thesis assesses whether the practitioners and academics address the topics of appropriateness and validity; for the consumer tracking part, it quantifies the extent of tracking.

3.1 My Research Quadrant: We and Others—Workplace and Beyond

In this thesis, I classify the selected studies by domain and analyst. The domain refers to the application context in which the digital traces are collected and analysed; that is, the workplace, or beyond the workplace (consumers). The analyst refers to the party that collected and analysed the digital traces (who uses the traces?). An overview of the studies is depicted in Table 4.

The selected studies focus on the three dimensions of meaningfulness, appropriateness, and validity of digital traces' analysis in the workplace. People analytics is such a workplace topic that is gaining attention, and explores how practitioners use digital traces and apply them for algorithmic management. A prerequisite to addressing the research questions is uncovering how practitioners analyse digital traces, including what data they collect and what purpose their analysis serves. Better understanding of people analytics' underlying mechanisms allows assessing the meaningfulness, appropriateness, and validity of digital traces analysis in practice. However, my studies show that practitioners typically do not disclose sufficient information about their analysis. As a result, their mechanisms' cannot be assessed against the three dimensions. This infeasibility of sound assessment is problematic. For example, lacking validity may lead – intentionally or unwittingly – to erroneous and potentially harmful conclusions drawn from the analysis about employees or anyone subject to people analytics (e.g., gig workers, job applicants) (Gal et al. 2017). Furthermore, people analytics may violate workers' rights but the inscrutability of the underlying mechanisms makes it hard to hold companies accountable for their practices (Hüllmann and Mattern 2020).

Responding to the lack of information, I have independently assessed the meaningfulness, appropriateness, and validity of digital traces analysis through explorative studies. Since these three dimensions of digital traces analysis depend on the application domain (Howison et al. 2011), I have created variance and diversity in the domains to determine, which analysis yields meaningful, appropriate, and valid results. Based on these insights, I have derived recommendations for digital traces analysis across different application domains. The application domains are chosen from the context of *changing nature of* *work*, as the new ways of working are enabled by the increasing deployment of digital tools and communication and collaboration software (Hüllmann 2019). Such software and tools generate the digital traces, which are being analysed (Hüllmann 2019). My research has implications not only for the methodology of digital traces analysis and algorithmic management, but also for the theory in the given domain, where it has been applied.

Retrieving data from organisations to explore the dimensions of meaningfulness, appropriateness, and validity of digital traces analysis is restricted by privacy regulations and organisational constraints on data sharing. As a result, this manuscript looks beyond the workplace at consumer data to test the boundaries of digital traces analysis' for the three dimensions, particularly the validity. The topics beyond the workplace are diverse and include tracking by game publishers, scientific publishers, and online retailers, as well as data driven business models.

		Domain	
		Workplace	**Beyond Workplace**
Who uses the traces?	**We**	**Changing Nature of Work** Hüllmann & Kroll (2018) Hüllmann (2019) Hüllmann & Krebber (2020) Hüllmann & Hentschel (2021) Hüllmann, Krebber, Troglauer (2021)	**Machine Learning** Rothmeier et al. (2020)
	Others	**People Analytics** Hüllmann & Mattern (2020) Hüllmann, Krebber, Troglauer (2021) Hüllmann & Krebber (2021)	**Consumer Tracking** Klein & Hüllmann (2018) Badmaeva & Hüllmann (2019) Hüllmann & Krebber (2021)

Table 4. My Research Quadrant.

3.2 Digital Traces and the Changing Nature of Work

The *changing nature of work* community within the information systems discipline investigates new forms of work. Contemporary trends in organisational structuring include dynamic, collaboration-intensive work with a growing number of peers. It is characterised by multi-team or multi-project settings in globalised, distributed, virtual work that crosses spatial, temporal, and formal boundaries. In this type of work, individuals assume high autonomy and responsibility. The new forms of work are enabled by the diffusion of digital technology and the digitisation of work (Hüllmann 2019; Schellhammer and Watson-Manheim 2019).

Dating back to a call for research by Barley & Kunda (2001), the application of digital trace research for inquiring about the changing nature of work is catching on. Barley & Kunda (2001) suggest looking at work practices by analysing longitudinal data traces, complemented with qualitative data. In Hüllmann (2019), I explain that “most work practices involve digital technology” (via Orlikowski and Scott 2016, p. 2), and argue that digital traces provide fertile ground for data analysis to yield insights into digital work activities. Previous research on digital work used a variety of methods, such as participant observations, content analyses, semi-structured interviews, surveys, or video recordings (e.g., Espinosa et al. 2003; Poels et al. 2017). However, these methods fall short in the distributed and dynamic settings of digital work, where it is difficult to observe how people interact and collaborate. How do you observe people working together in ten different locations at different times? Given limited resources, you cannot possibly do fieldwork in multiple locations at the same time[12]. As a result, I claim that “reliable observations of people are difficult, who perform computer-based work or are part of dispersed teams” (Hüllmann 2019, p. 1).

The study thus suggests digital traces as a method. It provides an overview of different types of digital traces (similar to Table 3), and proposes several guidelines for the application of digital traces research. The study addresses the question of what proper procedures must be established for the valid application of digital traces analysis. Subsequently, this thesis includes four studies that explore different scenarios for the application of digital traces.

In Hüllmann and Kroll (2018), we explore the aspect of dynamic team assignment and collaboration. Looking at communication in enterprise social networks, we predict the success of social onboarding, the socialisation process of new hires adopting the norms and culture of the assigned team or project. The study triangulates two different data sources of digital traces and shows implications for construct validity and semi-automating the process of team assignment.

In Hüllmann and Krebber (2020), we identify temporal rhythms from email trace data. The changing nature of work leads to more complex schedules with less management control. Individuals need awareness of their colleagues’ temporal rhythms and must align their temporal structures accordingly to enable effective collaboration. Following the “little data” paradigm, we looked at over 13,800 emails sent between 2017 and 2019 for five individuals and triangulated the results with semi-structured interviews. We were

[12] Technically, with surveillance cameras it would be possible to visually observe the people at the same time. See also my previous thought experiment.

interested in knowing whether digital traces present an appropriate and valid approach for inquiring about temporal rhythms.

Our study Hüllmann et al. (2021b) puts forward that communication tools are essential to effective collaboration in distributed settings. It appropriates a reductionist interpretation of the media repertoire (Watson-Manheim and Bélanger 2007), which we call media collection. We identify the prevalent set of media collection in use at a wholly distributed organisation and quantify which factors lead to the choice of a media collection. This study can be understood as a replication study, that seeks to corroborate existing theory on media collections and media choice by using digital traces in a distributed setting.

In Hüllmann and Hentschel (2021), we look at digital traces from an enterprise social network. We quantify whether the organisation’s communication tends to follow the formal organisation structure, or the informal structure. The formal structure is based on the organisational chart and intended communication paths, whereas the informal structures is the communication that does not follow the chart, or was not deliberately intended. We quantify which antecedents, e.g., homophily, proximity, hierarchy position, attitude, have a significant effect on the distribution of informal and formal communication in the organisation.

3.3 Digital Traces and People Analytics

The second topic in the domain of the workplace is people analytics. This topic is gaining traction as companies recognise the potential of using the digital traces of their employees for analysing and ultimately managing them more effectively.

My three studies on people analytics focus on practitioners, using digital traces and other forms of data, to inform organisational decision-making for people-related outcomes. Hence, we investigate others’ use of digital traces in the workplace. We consider the appropriateness and validity of their approaches. As a nascent field, we want to understand how academics and practitioners understand people analytics.

In Hüllmann and Mattern (2020), we introduce the topic of people analytics from an information systems perspective. We highlight three challenges of the field with regards to people analytics. First, the term “people analytics” is loaded and ambiguous, because different people mean different things when using this term. Second, the value propositions of people analytics lack empirical backing; the internal algorithmic mechanisms are often opaque. Third, we expose that the extant literature does not sufficiently address the appropriateness of people analytics, such as privacy, ethical, moral, and other implications. The study calls for further research on people analytics and provides a categorisation schema, grounded in management information systems and decision support systems literature, that we use in the two subsequent studies (Hüllmann et al. 2021a; Hüllmann and Krebber 2021a).

In Hüllmann and Krebber (2021a), we present the results of an exhaustive literature review across practitioners and academic literature to capture all the different mental conceptions people may have when referring to "people analytics." To achieve mutual understanding when using this term, and enable future research, we provide a morphological box on people analytics. The morphological box is a multi-dimensional categorisation for non-quantifiable characteristics that allows the selection of multiple characteristics per dimension. By using the morphological box, people can ensure that they have the same understanding of "people analytics" when discussing the topic. A mutual understanding is crucial because judging the appropriateness and meaningfulness of people analytics endeavours is contingent on the selected understanding.

While the study Hüllmann and Krebber (2021a) examines the literature of practitioners and academics, thus capturing what people express about people analytics, Hüllmann et al. (2021a) presents the results of a market and vendor analysis. It looks at the IT artefact of people analytics. As opposed to summarising what consultants say about people analytics, the study depicts what solutions are discussed and sold under the term "people analytics."

Together, both studies produce a holistic view on the state of the art of people analytics. They question the opacity of proprietary solutions and services which render it challenging or impossible to judge the validity and privacy, among other issues, of algorithmic management of the human resources function. Interestingly, the consultancies do address the question of whether people accept or reject algorithmic decisions. They recommend different ways to convince the employees by sophisticated change management approaches.

3.4 Consumer Tracking and Churn Prediction

Despite the focus on algorithmic management in the workplace, the remaining four studies address algorithmic management and digital traces analysis in consumer domains such as online retail, online scholarly databases, and online multiplayer games.

The accelerating digitisation of society and social phenomena drives the topics in this thesis. Humans cannot avoid generating digital traces. Putting this thesis into a societal and economic context, the short essay (Hüllmann and Krebber 2021b) outlines the evolution of the data economy. "Data economy" is a neutral term that we use to depict an economic perspective that understands data as an economic good that can be used for analysis (e.g., algorithmic management), or that can be traded. Others have coined different terms to emphasise selected issues surrounding the data economy, e.g., surveillance capitalism (Zuboff 2015), or data capitalism (West 2017). The essay traces the history of the data economy and reports the current market value and extent of the data economy in Europe. It raises several questions regarding the implications of digitisation, including the increase of tracking and personalisation. In the following, I use the phrase "tracking and personalisation of services" synonymously to algorithmic management in consumer domains[13].

The acceptance and implications of tracking consumers online for the implementation of personalised services such as news, content, or prices, is part of an ongoing debate (Table 1). Due to the opacity of online service providers regarding their data tracking and analysis practices, research is concerned with identifying and quantifying the extent of personalisation of online services. The two studies Klein and Hüllmann (2018), and Badmaeva and Hüllmann (2019) belong to this type of research. They are similar to the people analytics studies in that they seek to expose and understand what others are doing.

Klein and Hüllmann (2018) expose and criticise the changing business models of scientific publishers. Using Elsevier as an example, the study illustrates how scientific publishers integrate their vertical value chain of academic value generation. They gather excessive amounts of data on scholars and act as analytics service providers, not publishers. The study calls for a debate on Elsevier's business practices regarding data privacy and the abuse of its oligopolistic market position.

Likewise, Badmaeva and Hüllmann (2019) seek to expose price discrimination practices in German online retail. The study examines the extent of personalised price discrimination using a triangulated approach of automatic crawler, survey, and documentation review. It does not find any evidence for personalised prices.

Due to the algorithmic opacity, neither study assesses the validity of the algorithmic systems. Instead, the goal was to get more information about data tracking and personalisation in practice, which is the prerequisite for discussing the questions of appropriateness and meaningfulness. Without this information, the questions can only be addressed hypothetically from a normative standpoint. Such an explicit ethical argument is not the goal of this thesis, as mentioned earlier.

[13] "Tracking and personalisation" and "algorithmic management" are both broad phenomena. I focus on the intersection of the two.

The last study is Rothmeier et al. (2020). Unlike the two former studies, it does not try to identify the extent of tracking on the internet. Instead, it is comparable to the "changing nature of work" studies, because it seeks to establish proper and valid means for churn prediction. To this end, Rothmeier et al. (2020) evaluate machine learning models with varying configuration regarding their prediction accuracy. Actions related to user retention can be automated or semi-automated depending on the prediction accuracy.

3.5 Summary

People analytics is an emerging topic. It depicts the algorithmisation of decision-making through quantitative analysis of employee behaviours in people-related organisational processes such as hiring, retention, or staffing. Driven by commercial software vendors and service providers, the market segment of people analytics' tools and services can be characterised as opaque with respect to the collected data and employed algorithms. To shed light on this market segment, I have conducted two studies looking at people analytics tools and consultancy services. These studies provide a novel classification of people analytics, contributing to a new understanding of how vendors and consultancies vary in goals, collected data, and underlying mechanisms. The studies find that employee behaviours are analysed using digital traces, enabled by increasing data availability, behavioural visibility, and advances in computational analyses (Leonardi and Treem 2020). However, transparency with respect to the inner workings of the mechanisms remains scarce. Vendors and service providers offer little empirical backing that their digital traces analyses are meaningful, appropriate, and valid. They hide the exact mechanisms of how the employed methods work and what data is collected. The overall logic is data-driven and inductive, and there is lack of theorization. Only few vendors refer to established theories to substantiate or justify their analysis results. It remains unclear whether the data is appropriate to the companies' goals and whether the sought analysis results are valid. A reflection on boundary conditions, that is, the scope of generalisability and applicability of findings to other data sets, is missing.

Since the meaningfulness, appropriateness, and validity depend on the context and prediction task, I conduct multiple studies to generate insights on digital traces analysis. The selection of studies is partly opportunistic, partly exploratory, addressing different application domains to create diversity and variance. First, I depict different types of digital traces that enable different kinds of analyses. Akin to people analytics, I focus on social and behavioural traces. I use social trace metadata without any content to predict the onboarding success of prospective hires. The prediction quality is mediocre. As a result, I evaluate whether interviews or increasing the amount of data improves the quality of the results. Triangulation with interviews, indeed, provides a rich context and leads to an improved understanding. Increasing the extent of data collection is achieved by looking at churn prediction in an online multiplayer game where every single user activity is logged. Again, more data has significantly improved the prediction accuracy.

The juxtaposition of the different contexts, domains, and perspectives ("we versus others using traces") allows me to evaluate the generalisability of my results regarding meaningfulness, appropriateness, and validity of algorithmic management using digital traces. I discuss generalisations from the individual studies beyond each specific study context, drawing higher-level conclusions about the nature and impact of algorithmic management using digital traces.

I frame the answer to these questions as "it depends". Algorithmic management is neither the panacea, that is, digital traces are not the magical instrument to provide novel insights of unprecedented magnitude; nor is people analytics an approach to turn the workplace into an Orwellian dystopia (even though some may try to do that). Rather, digital traces will provide a useful and complementary data source for academic and managerial inquiry, in so far its limitations, the extent of its validity, and the potential for unintended side effects are properly taken into account.

4 Methods

Addressing my two research questions requires a mixed-methods approach. We explore selected problems in the changing nature of work, and develop proper procedures for digital traces research by analysing digital traces and triangulating the results with interviews and secondary data sets of traces. The purpose of triangulation is to corroborate and confirm the quantitative findings and to strengthen the robustness of the inferences (Venkatesh et al. 2013). Following this approach, the data is collected from distributed organisations, with a global IT services provider supplying most of the data. I analysed the data using parametric statistics (e.g., regression analyses), nonparametric statistics (e.g., decision trees, neural networks), and social network analysis in the open-source statistical programming language R. Interviews were coded manually.

We review websites, policy documents, as well as practitioners' and academics' literature, to explore the extent of tracking and the mechanisms of people analytics. The literature was coded inductively by at least two researchers using pen and paper and Excel. In one study, we programmed software and conducted a survey. Both instruments measure the extent of tracking and personalisation. The details of the methods are not provided here, as they are found in the respective papers. I give a critical reflection of the methods and discuss opportunities for future research in the discussion section.

There is the question of commensurability of qualitative theory generated by ethnographic and other means with analysis results based on digital traces. This question is rooted in the differences between the interpretivist and positivist epistemologies. Whereas the theory generation followed interpretivist or relativist paradigms, the origins of digital traces research (and computational social science) are rooted in Auguste Comte and Emile Durkheim, with proponents today advocating buzzwords such as social physics (Pentland 2015). Pentland argues that with just enough data, we can observe and explain the entire social world objectively. Østerlund et al. (2020) posit a differing stance and suggest that digital traces research can fit the interpretivist paradigm and merely provides a lens into complex social phenomena. While I do not resolve this debate or make a philosophical argument, it is helpful to understand that the differences between these two paradigms are a contributing factor to the issue of construct validity. Rich and contextual theoretical propositions cannot always be operationalised using basic digital traces and require triangulation with other data sources, hence my use of mixed-methods research.

5 Discussion

For the discussion section, I advise my students to think about future work. What should future work include in the methods or theory section, based on your results? What are the methodological and theoretical implications? In the case of this thesis, merely following my advice may be counterproductive, because I have already included my broader insights in the previous sections. Instead, I might opt for highlighting the insights of every single study included in this dissertation. However, this would merely be repeating the respective studies and would be overly specific.

Geletkanycz and Tepper (2012) note that the discussion section is somewhat paradoxical. It presents an end to one study, but offers a beginning for others, spawning new ideas, questions, and problems. They suggest looking for a middle ground; that is, the collective results of the study. Based on the collective results, authors should argue for the broader meaning and significance of their study, without “meandering” or “overreaching” (p. 259). Why should the study be published or shared? What are the implications for scholars, practitioners, citizens?

In the following, I attempt to find such a middle ground. In doing so, I deviate from Geletkynycz and Tepper (2012) and present a critical summary of my results, because this thesis is a collection of papers, and revisiting the results to discuss them strikes me as beneficial for the reader. After reflecting on the results, I string them together and address the two research questions, emphasising the tradeoffs among appropriateness, meaningfulness, and validity.

5.1 Critical Reflection

Although each manuscript dedicates a section to its limitations, the manuscripts do not elaborate on the limitations in detail due to space restrictions. Here, I expand on the limitations, and discuss the key takeaways of the studies' results. I address the manuscripts in the same order as I introduced them earlier. For each manuscript, I summarise its results, depict its methods and theoretical implications, and show its takeaways for the overarching two research questions of this thesis (four paragraphs per manuscript). For brevity, I will refer to the research questions and dimensions using the keywords in Table 5.

Keywords	What I refer to
Research Questions	
Scenarios, Applications, or Domains (When?)	In which scenarios is it appropriate and meaningful to implement algorithmic management and analyse digital traces to derive insights?
Methods, Procedures, or Guidelines (How?)	What procedures must be established to ensure valid inferences and appropriate mechanisms from digital traces analysis, and the implementation of algorithmic management?
Dimensions	
Appropriate	For example, ethics, morality, privacy, fairness, accountability, transparency.
Meaningful	For example, profitable, efficient, contributing to wellbeing, satisfactory, original, novel.
Valid	For example, accurate numbers, correct results.

Table 5. My Research Questions and Dimensions.

Hüllmann (2019) – Construction of Meaning

The manuscript provides a brief conceptual overview of digital traces. It details what digital traces are, how they can be generated and collected, and provides guidelines for their analysis. In other words, it proposes proper procedures for the valid application of digital traces analysis. The manuscript (Hüllmann 2019) informs the taxonomy of digital traces that is found in the previous section of this thesis.

The manuscript is preliminary: the literature review is not exhaustive. The length was limited to five pages by the workshop[14], and other works about the application of digital traces research have only appeared in 2020; for example, the works mentioned in this thesis (e.g., Grover et al. 2020; Pentland et al. 2020). As a result, the manuscript substantiates its arguments about proper procedures for digital traces research with

[14] Pre-ICIS 2019, International Workshop on The Changing Nature of Work.

examples from two pilot studies, empirically. Since the guidelines are based on empirical cases, future research should clarify how they generalise to other contexts. Further empirical evidence is required to corroborate the propositions and discuss the benefits and drawbacks of the suggested guidelines, towards a methodological consensus in the information systems community.

Beyond the methodological considerations, the manuscript provides an earlier version of the digital traces taxonomy compared to this thesis. It focuses on the data sources and extraction methods of digital traces as well as on the purpose of analysis. However, other aspects of digital traces, such as devices for collection, types of traces, software, or the data content of traces are not addressed (cf. section 2.2). As a result, the picture drawn of digital traces in Hüllmann (2019) is not exhaustive, but presents a starting point which I have picked up in the theory part of this thesis.

The manuscript shows that digital traces provide prospects for asking questions in the domain of the changing nature of work. Future research can extend the taxonomy of digital traces to illuminate when and how digital traces are appropriate, meaningful, and valid to use, contingent on the type of digital traces chosen. Different types of traces are suited to address different types of research questions. Using the wrong type for addressing a research question renders the results invalid or at least misleading. Hence, future research with digital traces should consider the taxonomy and guidelines to avoid such mistakes. The taxonomy can be extended by conducting a systematic literature review. Besides implications for academic methodology, practitioners should heed the lessons learned. Digital traces do not provide an *orbuculum* or *clairvoyance*. While they complement existing data sources to gain differentiated insights about organising work, the epistemological limitations should be considered when implementing organisational change based on digital trace analysis or algorithmic management. In particular, vendors selling software solutions and consulting services based on digital traces should disclose their approaches and potential shortcomings (Hüllmann et al. 2021a).

Hüllmann and Kroll (2018) – Social Onboarding

The study on enterprise social networks and socialisation estimates the effects of seven user roles on the onboarding success of existing hires onto new project teams. Despite mixed results, it shows that two of the selected roles can be used as indicators to inform the onboarding process. It highlights the importance of construct validity for digital traces research.

The operationalisation of the theoretical constructs from Bauer & Erdogan (2011) is based on user roles which are derived from the literature on enterprise social networks (e.g., Angeletou et al. 2011; Smith et al. 2009). This literature uses explorative analyses, such as cluster analysis or exploratory factor analysis, to identify user roles from the digital traces of enterprise social network platforms. For example, a high level of activity, that is, posting or liking content, represents a cluster of “power users” (Hacker and Riemer 2020, p. 14). In our study, we mapped these user roles to the behaviours described in the socialisation model of Bauer & Erdogan (2011) based on an argument that the descriptions of the roles and the model constructs describe the same behaviours. However, the mapping of these user roles to the socialisation model of Bauer & Erdogan (2011) has not been empirically validated. The Bauer & Erdogan (2011) model has been operationalised through surveys, and it is unclear whether digital traces, from which the roles are derived, measure the theoretical construct in the same way. We ensured construct validity through a rational argument, but not through empirical evidence. Future studies would benefit from empirical instrument validation and a longitudinal study design. The analysis in Hüllmann & Kroll (2018) is cross-sectional and lacks data on the temporal changes in communication behaviours.

The study finds that focussed communication with peers has positive effects on socialisation success. Conversely, being widely connected and popular within the organisation as a whole does not help an employee to socialise with a newly assigned team. Mere connectedness does not lead to cohesion and trust. The results indicate that it is harder to onboard onto bigger teams than smaller ones.

The manuscript shows that using only digital traces from enterprise social networks provides a limited lens for analysing complex social processes, and that social factors are not necessarily linear. Operationalising the theoretical constructs into meaningful measures remains challenging, and suggests that a mixed-methods approach, including triangulation of data sources, may be needed. Triangulation promises better reliability of findings and remains a valuable approach that—moving forward—the information systems discipline has the expertise to implement. Future studies on onboarding should validate the user roles and enterprise social network traces against established survey metrics. Results from an analysis of enterprise social network traces should be seen as indicators, and interpreted cautiously. Predicting onboarding success with digital traces requires further research. Future studies can look into other types or more fine-grained digital traces in the future; for example, rich media (text content) of the enterprise social media (e.g., Cetto et al. 2018). Generalising from social onboarding to other social processes, various tools in practice promise valuable insights from such analyses (Hüllmann et al. 2021a). However, context, impression management, and other obstacles to validity are not dealt with transparently. It is unclear to prospective users whether the metrics have been empirically validated in an appropriate manner.

Hüllmann and Krebber (2020) – Temporal Rhythms

The study on temporal rhythms reviews existing literature on temporal rhythms and explores how the triangulation of digital traces and interviews can generate novel insights about temporal rhythms.

The study explores 13,800 sent emails of five participants between 2017 and 2019. The results are then triangulated with semi-structured interviews. Only emails were used as digital traces, as other traces were unavailable. The number of collected emails enables estimating stable and aggregate temporal patterns. The estimation of dynamic patterns requires more data in breadth and depth (Hüllmann 2019). The validation of the email traces via interviews depends on the subjective memories and statements of the participants. Beyond interviews, working time diaries and timekeeping software provide data sources for validating the digital traces that are not skewed by the memories of the participants.

The nine-to-five workday is still relevant for the selected sample. However, the insights should be interpreted within the restriction of the sample, as it is comprised of five people from a specific domain of knowledge work and, therefore, generalising beyond the respective organisations requires caution.

The study illustrates a novel algorithm for detecting breaks at work, and estimates higher-level patterns that are stable across the timeframe, such as start, end of working day, and breaks. Future work could estimate what level of data granularity is needed to achieve high predictive power, to predict dynamic patterns. For example, a simulation model could estimate the predictive power as a function of data granularity that depicts the breadth, depth, and number of digital traces (Hüllmann 2019). Meaningful and valid digital traces research, in most cases, requires a stable phenomenon, as predicting dynamic patterns requires excessive amounts of data (Howison et al. 2011). Activities and actions not logged as digital traces are agnostic to the analysis, but the absence of traces does not mean an absence of activity. Rough data yields rough precision ("garbage in, garbage out"). The context is missing with only traces. Triangulation provides context and helps us understand patterns; for example, participants explained that the absence of traces were due to holidays or vacations. Workplace analytics tools relying on similar analyses should be met with caution. Nevertheless, digital traces provide a meaningful data source for future investigations of phenomena like temporal rhythms in the workplace.

Hüllmann et al. (2021) – Media Collections

The study corroborates established theory on media choice in a purely distributed work setting by using a sample of digital traces.

It uses monthly aggregated statistics of digital traces because daily detailed logs were unavailable in the data set. As a result, we performed a high-level correlation analysis that opts for a parsimonious measurement model, considering high-level social effects. On the contrary, the study does not open the “black box,” does not take a process view, and does not examine the mechanisms of the analysed social effects (Zuckerman 2017). Limitations of the study, therefore, include that the measurement constructs do not capture how and why the effects occur, but they capture the extent of the effects instead. The analysis does not consider the context of media choice, which has previously been investigated in qualitative studies such as Watson-Manheim & Belanger (2007). Since only data on Microsoft 365 tools is available, the analysis misses any form of tool use beyond Microsoft 365. Dyadic and group-level effects are not addressed, and the analysis is only cross-sectional, as it does not contain longitudinal measurements.

The large and unique data set of digital traces from a distributed global services provider to corroborate existing theory is meaningful. The digital traces substantiate that the supervisor correlates with an employee’s media choice, also in distributed teams. Expectedly, the physical location does not matter in purely distributed settings. Being in the same organisational subunit does not show a significant association with media choice, presumably because informal peers, as well as the specific tasks, are more important (Fulk et al. 1990). The study shows that despite new tools, such as chat and enterprise social networks, email remains the core communication tool in 2019. Smaller media collections, that is, fewer tools used routinely, prevail over bigger ones.

For future studies with digital traces on tool use, more traces in breadth and depth are encouraged (Hüllmann 2019). Instead of monthly aggregated data, discrete event-log data of single user actions with timestamps provide longitudinal insights based on computationally intensive algorithms and elucidate the tool usage behaviour of users in great detail, for example, how tools are used (cf. Rothmeier et al. 2020). Dyadic data and triangulation with other data sources enable the inclusion of context sharply into the analysis. Shadow IT and backchannel systems must be considered when interpreting the results. The extraction of digital traces from production-level systems is costly. As opposed to temporal rhythms and social onboarding, construct validity is not an issue for measuring the extent of media use. Deriving meaningful and valid insights is straightforward, as the digital traces describe routine tool use on a basic level. Tool use is not a higher-level theoretical construct that requires sophisticated operationalisation.

Researchers can complement the study with other data sources to identify the antecedents and effects of tool use.

Hüllmann and Hentschel (2021) – Informal Drivers

The study examines the antecedents of formal and informal communication. It infers the informal social structures of the organisation by analysing the digital traces from the enterprise social network of a global distributed services provider. These informal structures are compared to the formal organisational chart to compute the distribution of formal and informal communication. Then, the study estimates the influence of the antecedents, e.g., homophily, proximity, hierarchy position, and attitude, on the distribution of formal and informal communication.

Methodologically, this study combines regression analyses with permutation-based equivalence tests of enterprise social network data. To the best of my knowledge, this has not been computed before, as I could not find a study in the reputable information systems outlets, except for Goh and Bockstedt (2013) who perform an equivalence test but not a permutation-based one. Our study's approach gracefully handles the violated assumptions mentioned by Howison et al. (2011), cited previously. It is informed by multiple disciplines. Permutation testing of network data comes from the sociology of animal networks (Farine 2017; Farine and Whitehead 2015), also suggested in human sociology (Borgatti 2004), with equivalence tests coming from psychology (Lakens 2017, 2020; Lakens et al. 2018).

Preliminary results indicate that the hierarchy position, attitude, and size of the organisational unit have a significant effect on the distribution of formal and informal communication. Being high up in the organisational hierarchy correlates with formal communication. This is expected given that senior managers are at the top of the organisational line chart. An open attitude and small organisational units correlate with informal communication. If you have only a few people on your team, you are likely to talk to other teams to perform your job.

Enterprise social network traces have been used previously to examine informal structures in organisations. However, the combination with active directory data, such as the organisational chart, may open avenues to inquire about the interplay of formal and informal structures, which is under researched. Project management and enterprise resource planning software contains data about work-, task-, and team-assignment and may yield meaningful insights for future studies about formal and informal structures.

Hüllmann and Mattern (2020) – Three Issues

The manuscript provides a brief overview of people analytics and criticises the status quo. It argues that people analytics papers, both academic and consultancy papers, lack in theory, and are not corroborated by empirical evidence. According to the manuscript, the field suffers from conceptual and terminological ambiguity.

The manuscript is a research commentary based on a literature review of academic and practitioners' papers, and its arguments are substantiated by examples from the literature.

Addressing the identified issues, the manuscript proposes a research agenda and contributes a coding scheme, grounded in decision support systems and management information systems literature.

The issues are rooted in the opacity of vendors and consultancy practices regarding the internal mechanisms of their solutions and services. Therefore, it is impossible for consumers and customers to judge the validity or appropriateness of the mechanisms.

Hüllmann et al. (2021) – IT Artifact

The manuscript elucidates the role of information technology for people analytics by surveying the existing tools on the market, listing the available tools and their focus at the time of data collection. It identifies five archetypes of people analytics tools and relates the methods and role of information technology to the relevant discourse in the information systems literature.

We collected the list of tools by monitoring social media related to people analytics for five months (from August to December 2019). The list is inclusive as it includes all tools that were labelled as people analytics on social media. Two researchers conducted a two-cycle coding process. The first cycle included explorative coding, inductively. Diverging codes were resolved jointly through bilateral conversations during the second cycle. From the codes, the archetypes were derived intuitively.

The five archetypes are *technical monitoring*, *technical platforms*, *employee surveillance*, *social network analytics*, and *human resources analytics*. Technical monitoring subsumes tools for monitoring adoption, usage, and performance levels of collaboration software. Technical platforms are general-purpose analytics tools that can be used for people analytics but also for other analytical domains. Employee surveillance includes invasive employee monitoring and tracking, such as desktop or video screen capture. Social network analytics uses digital traces from communication and collaboration software to improve collaboration processes. Human resources analytics supports the human resources function through data analytics.

Depending on the archetype, different issues emerge. For example, employee surveillance raises severe privacy and ethical concerns, whereas technical monitoring and platforms have little to do with people analytics. The manuscript corroborates the claim from Hüllmann and Mattern (2020) that the vendors do not discuss privacy and validity concerns sufficiently while marketing their tools. Buyers should perform a critical assessment of the tools before using them productively. However, the opacity renders an assessment of appropriateness and validity difficult. Applications of such tools in the workplace and the resulting effects on employee behaviour are unclear and require further research. Questions of accountability are unresolved. Digital traces are relevant for the people analytics tools according to the vendor descriptions. However, no transparent information on what exactly is collected and analysed, or on the internal mechanisms of the algorithmic systems are available.

Hüllmann and Krebber (2021) – Strategic and Operational People Analytics

The manuscript presents an exhaustive and descriptive literature review on the conceptions of people analytics in the academic and practitioners' literature, and is the follow up of the two studies Hüllmann and Mattern (2020) and Hüllmann et al. (2021a). From the variety of conceptions in the literature, we derive a morphological box of people analytics that helps practitioners and academics to establish mutual understanding for subsequent research and applications of people analytics in practice.

We performed a systematic keywords search across three scholarly search engines, filtering by title, abstract, and keywords, to identify academic literature. Twenty consultancy websites were searched for input. The coding scheme developed in Hüllmann and Mattern (2020) was used to code the literature and identify the conceptions. We derived the taxonomy through comparing, synthesising, and condensing the findings from the coding scheme.

The nucleus of people analytics is the human resources function, which increasingly digitises its processes and uses data analytics to inform decisions in hiring, retention, staffing, and employee management. The dominant conception of people analytics includes, thus, the application of data analytics in the operations of the human resources function. However, another prevalent understanding is that, with people analytics, the human resources function expands its problem area and addresses broader issues related to work and collaboration processes. This conception emphasises the shift from human resources as an operational function towards a strategic function in the organisation. Both conceptions come with different foci and issues. Future research can discuss these issues, contingent on the selected understanding of people analytics via the morphological box.

Discussing when algorithmic management is appropriate and meaningful requires transparency about the internal mechanisms of the algorithms. What data is collected? How is it analysed? For what purpose is data collected? The answer to these questions depends on the underlying conception and implicit assumptions about people analytics. As a result, this manuscript makes an important contribution with the morphological box to enable future inquiries into these questions.

Hüllmann and Krebber (2021) – Data Economy

The manuscript offers a brief introduction to the data economy as an economic perspective. It outlines the history of the data economy and associated business models, and raises several questions about the societal implications of the data economy.

In this manuscript, we review recent articles about the data economy and explain how the ideas of the information economy precede the data economy. The extent and relevance of the data economy are demonstrated by referring to a recent large-scale data market study of the European Union.

Theoretically, the data economy understands data as an economic good that can be integrated into the value chain or traded on the market. This presents opportunities; for example, organisations can optimise processes and decisions, offer personalised products and services, or develop entirely data-driven business models. Challenges include the risk of data monopolies (e.g., Google, Facebook, and Amazon), discrimination, polarisation, and the invasion of informational self-determination rights.

The manuscript puts the other works of this thesis into an economic and societal context. It asks broad questions about the appropriateness, meaningfulness, and validity of algorithmic management in public that require debate from politicians and society. Academia and future research can inform such a debate.

Badmaeva and Hüllmann (2019) – Price Discrimination

The study investigates selected German online retailers for indicators of personalised price discrimination, triangulating three different methods.

The first method is a literature review and email inquiry to the selected shops, while the second is a survey in which participants use their personal devices to extract prices manually. The third is an automated crawler, which is a software program that simulates user behaviour and automatically accesses the online retailers' websites to extract the prices. The second and third method acquire the price information from an *ex ante* defined list of products. Before checking the prices, the crawler software simulates regular user activity to build profiles that the online retailers can identify and use to personalise its prices. We validated the profiles manually by checking the profiles that Google and Facebook inferred about us[15], but two online retailers recognised our software as a bot. It is unclear whether the other shops also recognised the crawling. As a result, we cannot distinguish between whether the shops just showed default prices because we were identified, or if they did not recognise our crawler and do not personalise prices. We sampled the biggest shops by revenue in Germany, so smaller and specialised shops were not considered for the study. They may yield different results.

The study makes several contributions to the information systems discipline. It provides survey items for examining personalised price discrimination via human participants. The survey items were collected from an extensive review of tracking vendors and the data and information they offer their clients. The study provides software for simulating user profiles and automatically extracting the prices from the selected shops. The study also contributes evidence that the selected German online retailers did not discriminate their prices based on personalised user information in 2018. Although we found no price discrimination, the study shows that it is technically feasible and suggests continuous monitoring in the future, in case retailers start to discriminate prices based on personal information.

There is an unresolved information asymmetry between online retailers and consumers, which generalises to people analytics, other business models of the data economy, and algorithmic management in general. Consumers do not know which data is collected from them, and they do not know if their data is being analysed. They can only trust the public data privacy policies and official statements from the vendors—or shops in this case. As a result, identifying the extent of such practices makes a significant contribution as a prerequisite of judging appropriateness and validity of algorithmic systems in public.

Klein and Hüllmann (2018) – Datenkapitalismus (Data Capitalism)

The manuscript comments on the transformation of scientific publishers towards data analytics as service companies. It depicts Elsevier as an example and illustrates its

---

[15] https://adssettings.google.com/authenticated (accessed 2020-12-04).
https://www.facebook.com/adpreferences/ad_settings/ (accessed 2020-12-04).

platform business model, elucidating potential issues about business practices, privacy, and ethics.

We performed a web search, collecting and analysing the official press documents and policies of Elsevier, along with reviewing literature and opinion pieces on the topic.

The manuscript points out Elsevier's opaque business practices and monopolistic position. It calls for a necessary debate on the "datafication" of scientific publishing and potential abuse of personal data for commercial interests. A consumer can only believe the statements of Elsevier due to the information asymmetry between vendor and consumer. Elsevier also computes scholarly performance metrics (PlumX), but it is unclear how they are calculated and how valid they are.

Judging appropriateness and meaningfulness given the opacity of Elsevier's business practices and proprietary products and services is difficult. Further research can use means beyond the official documents to inquire about the extent of Elsevier's tracking and data collection. Other means include, for example, crowdsourcing the results from personal information request, enabled by the general data privacy regulation, or monitoring software similar to the crawler in price discrimination.

Rothmeier et al. (2020) – Churn Prediction

The study on churn prediction looks at consumers, specifically gamers. It leverages discrete event-based tracking, meaning that every user action is logged and analysed. Compared to the workplace studies, it shows a difference in results quality given the magnitude of available data. The study compares more than 1,000 features compared to the small number of features in the other studies (e.g., sent emails or aggregated tool use). Most features are directly related to game actions such as clicking something, while some are non-game related features (e.g., login, activity, language). In total, 203,999 events were collected and analysed.

The study explores various labelling approaches, which is important for classification. In the context of churn prediction, it refers to what measured user behaviour constitutes churn. In terms of workplace analytics, this can be translated into what constitutes desired behaviour. The study experimented with different thresholds and hyper parameters, which govern how the algorithms are applied to the data. This can be understood as the configuration of the analysis, where different configurations yield different outcomes. The study reports on the configuration that yields the best results. The best configuration enables churn prediction accuracies of 97%, which at the time of publication had not been reached before. It validates the results with an independent data set, which was only available after the first part of the study had concluded. Future work may apply computationally intensive algorithms (e.g., deep neural networks) and reduce the prediction interval from 14 days towards real-time prediction.

Decision trees, random forest, and other non-linear estimation models reach the highest accuracy for predicting churn in the case of the tested game. Besides reaching a high prediction accuracy, the study identifies strong predictors for churning and retention. The success of the player, regular play with consecutive days of playing, and the player's game experience are indicators that a player is likely to continue playing in the future.

In the gaming domain, similar churn prediction accuracies have been reached with various approaches for different types of games before. Automating the implementation of nudges to retain players is feasible with such high accuracies, for example, they can be offered bonuses or discounts. The large data set enabled high prediction accuracy. Furthermore, games have rigid rules, limiting the possible behavioural actions that players can take. Playing the game can be considered a routine, cognitive task. The game logs all actions with minuscule detail. It is purely a virtual multiplayer game, and outside actions are not feasible—except external chat programs. Prediction of churn and retention of people is not only meaningful for games but also for employees. Employee retention is a common objective of people analytics. The study's results could be transferred to routine jobs. Conversely, knowledge work is cognitive and non-routine; thus, prediction and automation are inherently more challenging because there are no fixed rules that limit behaviours. Future research can look into the feasibility of predicting churn for different types of jobs.

5.2 Balancing Appropriateness, Meaningfulness, and Validity

My first research question refers to situations where the application of algorithmic management and digital traces analysis yield appropriate and meaningful insights. The second question asks about proper procedures to ensure that valid insights are reached. Reaching appropriate, meaningful, and valid insights requires a tradeoff, balancing the three goals. The research questions can only be addressed together as they influence one another.

5.3 Appropriateness

Algorithmic management is disfavoured in public; for example, Ed Pilkington labels it a "digital dystopia" in The Guardian (2019)[16]. He reports that algorithms are trained on WEIRD[17] demographics, discriminating against minorities or poor people. Such algorithms would automate inequality (Eubanks 2017), for example, in healthcare. Gianfrancesco et al. (2018) assert that "existing health care disparities should not be amplified by thoughtless or excessive reliance on machines."

Where do these discriminative tendencies come from? Algorithms are trained on historical data and reproduce existing stereotypes. The New York Times headlined this cause multiple times, for example in "Biased Algorithms are Easier to Fix than Biased People" (Mullainathan 2019)[18], or in "We Teach AI Systems Everything, Including Our Biases" (Metz 2019)[19]. Fixing this issue, however, is a grand challenge. Even after multiple years, top Google engineers are unable to fix their face recognition algorithm that discriminates against black people (Simonite 2019). Microsoft experimented with a Twitter bot called Tay, which quickly turned racist (Victor 2016)[20]. Amazon scrapped its people analytics informing hiring because it discriminated against women (Dastin 2018).

Academia addresses these issues of fairness, bias, and discrimination (e.g., Selbst et al. 2019; Zarsky 2016). However, algorithms do not only reduce fairness but also help to improve and reduce bias and discrimination, for example, in the legal system (Kleinberg et al. 2018). As Kleinberg et al. (2018) outline, judging the appropriateness of algorithms requires transparency. Currently, we do not find this transparency in people analytics, scholarly publishers, price discrimination, or online tracking.

The internal mechanisms of proprietary people analytics software are opaque, hindering independent audits and validity assessments. The vendors themselves do not demonstrate that their solutions work. They often do not sufficiently address issues of privacy, fairness, and discrimination, or the dehumanisation of work (Hüllmann et al. 2021a).

The extent of tracking and data collection in consumer domains is opaque as well. Due to information asymmetries, people can only trust the official policies of the companies, and cannot judge for themselves. Independent audits and assessment have yet to occur. Even

---

[16] https://www.theguardian.com/technology/2019/oct/14/automating-poverty-algorithms-punish-poor (accessed 2020-11-05).
[17] "Western, Educated, Industrialized, Rich, and Democratic" (Henrich et al. 2010, p. 61).
[18] https://www.nytimes.com/2019/12/06/business/algorithm-bias-fix.html (accessed 2020-12-08).
[19] https://www.nytimes.com/2019/11/11/technology/artificial-intelligence-bias.html (accessed 2020-12-08).
[20] https://www.nytimes.com/2016/03/25/technology/microsoft-created-a-twitter-bot-to-learn-from-users-it-quickly-became-a-racist-jerk.html (accessed 2020-12-06).

if companies admitted to data collection, it would remain unclear what they did with it. Further, if companies explained what data they collected and analysed, it would remain unclear what internal mechanisms they implemented to analyse it. Uncovering the extent of data tracking is an ongoing effort because the practices are dynamic and can change anytime. At the same time, uncovering such practices is a crucial prerequisite towards a discussion on the appropriateness of algorithmic management. So far, judging appropriateness of proprietary tools is tough, and resembles a cat-and-mouse-game.

5.4 Validity and Meaningfulness

The discussion of appropriateness refers to validity and meaningfulness. If the algorithms do not work properly, because they discriminate or bias against minorities, the appropriateness is questioned. There is a general discussion on how well algorithms can predict social phenomena. Princeton professor Arvind Narayanan (2019) posits that predicting social outcomes is fundamentally dubious and flawed because the phenomena are non-stable and non-deterministic. On the contrary, proponents assert that given enough data, inferences about social phenomena are feasible (e.g., Lazer et al. 2009; Pentland 2015; Stewart 2019). Answering this question, whether it works or not, is non-trivial, and depends on the scenario.

Measuring and analysing the extent of media use based on digital traces is straightforward, because digital traces depict historical logs of routine media use, such as using tools, applications, and software. Other social phenomena, however, are more complex, such as onboarding or temporal rhythms. The analysis of these complex phenomena using digital traces benefits from triangulation through interviews, in which study participants point out implausible results, give context, and explain the patterns in the data. “Data in isolation are meaningless, a collection of numbers. Only in context … do they assume significance …” (Greenstein 1983, p. 4).

Studies on temporal rhythms and media collections show that questioning participants helps to clarify the plausibility of simple descriptive measures. However, it is unclear how well participants help to assess complex patterns in the data if they lack technological literacy. If participants are affected by algorithmic systems, their willingness to give truthful assessments is unclear (Pachidi et al. 2016).

Beyond providing context through triangulation, my research results continuously show that operationalisation through digital traces does not capture the theoretical propositions of social theory precisely. “The discrepancy between our mental models and the real world may be a major problem of our times; especially in view of the difficulty of collecting, analysing, and making sense of the unbelievable amount of data to which we have access today" (Bardi 2011, p. 104). The link between measurement constructs and higher-level theoretical constructs must be established, meticulously, through argumentation, or explicit validation through empirical evidence. To alleviate this issue, traces of finer granularity and attached rich media data, for example, text, audio, and video content, can capture the theoretical propositions more sharply. Sharper capture of theoretical propositions reduces the gap between measurement and theoretical construct that must be bridged. More granularity in data enables the step from variance models towards processual research models based on digital traces. Beyond estimating stable patterns, finer data powers longitudinal analyses to identify dynamic patterns and micro-practices, and elucidate the internal mechanisms of social processes.

Although finer and more extensive data collection improves the validity of the inferences, it coincides with serious concerns for privacy invasion. How much privacy invasion and tracking of our every action do we allow for valid algorithmic systems? It depends on the perceived meaningfulness of the algorithmic system. Future research can look into what level of data granularity is needed for valid functioning of algorithmic systems, and for which scenarios the invasion of privacy is appropriate (e.g., healthcare).

The churn prediction study complements the changing nature of work studies. It illustrates that social phenomena can be predicted given enough data and tracking. However, the study predicts churn in a game setting. The game follows rigid rules and is comparable to cognitive, routine tasks, whereas knowledge work typically does not follow such rigid rules.

Discussing the validity and appropriateness of algorithmic management and people analytics, I use an analogy from sports and games. People analytics is inspired by sports analytics; for example, baseball and the sabermetrics (Hüllmann and Mattern 2020). Recent examples include elite soccer analytics (Memmert and Rein 2018) or e-sports analytics (Hodge et al. 2019). Sports analytics shows advances in recent years in the variety and volume of data used, and in improvements in quality of prediction outcomes. Compared to people analytics, there are two differences. The first is about the rigid rules of sports, and the feasibility of generating meaningful and valid analytics. The second is about the considerations of appropriateness regarding promises and concerns of analytics.

First, sports follow rigid rules, yet the games play out dynamically, with different factors influencing the outcome (e.g., individual, team, opponent, and environment). As a result, devising tactics, including the relevant factors, is crucial for elite teams seeking to win. This setting can be compared to jobs with a high share of routine tasks but also elements of non-routine creative tasks that significantly alter the outcomes. People analytics, by definition, aims at analytics for people-related organisational outcomes, typically concerning office work (Hüllmann and Mattern 2020) and frequently addresses non-routine, cognitive work settings, as is suggested by the digital traces studies mentioned before. In these cases, further research should examine what types of tasks and jobs can be supported by people analytics in a meaningful way. Although elite soccer has rigid rules, team play and inter-individual and dyadic social cohesion are still factors, which is also exhibited by the previously mentioned studies on the changing nature of work. For example, is mere enterprise social network metadata (Hüllmann and Kroll 2018) sufficient to improve social onboarding? What social processes are candidates for valid automation, augmentation, or informing?

The second difference is that elite athletes are continuously being tracked in minuscule detail. At the highest level of play, squeezing out the last bit of performance outweighs potential concerns for privacy, as match data, training, and nutrition data of the individual players are tracked and analysed. While digital traces are valuable to investigate topics in organisational settings, the studies' results show that more data would improve the outcomes (Hüllmann and Krebber 2020). Deploying software sensors such as "monitoring software" (Hüllmann et al. 2021a) or hardware sensors like the Humanyze sociometric badges (Waber 2013) could provide such data. Nevertheless, the question remains, what level of data collection and privacy invasion is acceptable and appropriate? It depends on the scenario. So far, companies seem to follow the maxim "the more, the merrier" (Hüllmann et al. 2021a). Similar intuitions can be seen in the companies tracking consumers in the data economy. Consumer analytics, seem to gather as much data as possible, while sidelining potential privacy concerns (Badmaeva and Hüllmann 2019; Klein and Hüllmann 2018).

The study on churn prediction illustrates that selected human behaviours, including social behaviours, can be predicted given enough data, supporting the propositions that have been put forward by computational social scientists (e.g., Lazer et al. 2009; Pentland 2015; Stewart 2019).

5.5 Replication and Novel Insights

Grover et al. (2020, p. 275) argue that "mere passive application[s] of received theory (instantiation)" with digital traces do not deliver meaningful insights towards theory. This statement seems short-sighted given the replication crisis in other disciplines. Replicating established findings with novel data sources and in other contexts is crucial to corroborate existing knowledge going forward. For this reason, this thesis makes a dedicated effort to replicate established theory with digital traces (Hüllmann et al. 2021b; Hüllmann and

Hentschel 2021). Karl Popper (1963)[21] asserts that “a good theory makes “bold” predictions that withstand repeated attempts of falsification.” Balancing the need to explore original theory and to corroborate existing theory is non-trivial. “Science is the belief in the ignorance of experts … It is necessary to teach both to accept and to reject the past with a kind of balance that takes considerable skill.” (Richard Feynman 1969)[22]

Beyond lacking replication, another cause for the crisis is said to be questionable research practices (Fraser et al. 2018; John et al. 2012)[23]. This thesis contributes to establishing proper procedures for the analysis of digital traces in different contexts. Contingent on the scenario or social phenomenon under study, different types of traces are required. This thesis provides a taxonomy of traces to help researchers choose the right type of traces for a research question, and explores the validity of inquiring selected social phenomena. Subsequent work should explore extending the taxonomy and formalising the mapping between the type of question and type of digital traces required. To this end, future research should also establish construct validity of digital traces for different scenarios. Moving forward, the curricula of the information systems discipline could teach methods and pitfalls of digital traces research; for example, checklists and tutorials come to mind.

5.6 Meaningfulness versus Feasibility

I have previously mentioned that it is hard to collect digital traces from organisations in the field, because the data is sensitive, and it takes considerable resources and personnel to collect it (Hüllmann 2019). For passive digital traces, you need a senior manager sponsoring your project, and a technical manager to oversee the extraction, because the data is sourced from live production systems. Active traces require hardware and volunteers, but are typically easier to retrieve in terms of approval (Hüllmann 2019).

As a result, the implementation of digital traces projects is challenging. Devising a good research question for analysis with digital traces considering solely validity, meaningfulness, and appropriateness is insufficient. The researcher must consider the feasibility and practicality of data collection.

Although feasibility is important, researchers must not fall for the streetlight effect: “almost any problem is interesting if it is studied in sufficient depth ... the problem must

[21] Cited from Agarwal and Dhar (2014, p. 446). Originally from (ed.) H. Keuth, 2nd edition (2009).

[22] http://www.fotuva.org/feynman/what_is_science.html (accessed 2019-09-17). “[Feynman] presented at the fifteenth annual meeting of the National Science Teachers Association, 1966 in New York City. Reprinted from The Physics Teacher Vol. 7, issue 6, 1969, pp. 313-320.”

[23] Both cited from Uygun-Tunc (2020): https://medium.com/science-and-philosophy/trust-and-criticism-in-science-bbb62bd6890 (accessed 2020-12-04).

be such that it matters what the answer is—whether to science generally or to mankind" (Rai 2017)[24]. For digital traces projects, the researcher should address research problems that are relevant and can be answered.

To convince organisations to make an effort and extract digital traces, researchers must balance practical relevance and theoretical relevance; that is, striving for a contribution to a broader knowledge goal (Rai 2017). The "little data" approaches with digital traces should not be disregarded. They are easier to implement and triangulate with volunteers, while providing deep contextual insights. Although I argue for data of finer granularity, studies with a parsimonious measurement model provide evidence and should not be scrapped. However, again, a balance is needed. On the one hand, the absence of data can cause errors in understanding. On the other hand, "no analysis can overcome the unreliability of basic data" (Allen 1951).

6 Conclusion

6.1 Responsibility

Digital traces provide powerful means to inquire about the workplace and enable algorithmic management, but the application of digital traces analysis requires careful diligence. Advancing algorithmic systems requires thinking about the implications.

"Your scientists were so preoccupied with whether or not they could, they didn't stop to think if they should." (Jurassic Park 1993)

In 1939, Albert Einstein signed the Einstein-Szilard letter about nuclear bombs: "Now it appears almost certain that this could be achieved in the immediate future."[25] It would be presumptuous to compare Albert Einstein and nuclear energy research to digital traces research. However, digital traces analysis and algorithmic management are tools that can be meaningful when used for good but harmful when used with malicious intentions. Advancing digital traces and algorithmic management requires responsible research and innovation, similar to other technologies such as nuclear energy, or CRISPR[26].

Algorithmic management in the workplace using digital traces promises flexibility, improved wellbeing and productivity. At the same time, it leads to "more work—with

[24] Original quote is attributed to P. B. Medawar, Nobel Laureate in Medicine and Physiology, 1979; cited in Van de Ven 2007, p. 71; which I cited from Rai (2017).

[25] More information from The Franklin D. Roosevelt Presidential Library & Museum: www.fdrlibrary.marist.edu/archives/pdfs/docsworldwar.pdf (accessed 2020-12-08).

[26] Clustered Regularly Interspaced Short Palindromic Repeats," genome editing approach.

fewer boundaries"[27], potentially suffering from impression management with people who "work very hard to produce evidence that [they] constantly [do] work instead of, well, actually doing work"[28].

However, dismissing algorithms altogether is unreasonable. The world is overflowing with data and information, and finding your way around without algorithmic systems is impossible (Shapiro and Varian 1999; Varian 1995).

6.2 Quo Vadis?

Highlighting two key results of this thesis:

- Algorithmic management in practice is too opaque and requires more transparency and disclosure of mechanisms and implications with regard to meaningfulness, appropriateness, and validity.
- Digital traces research is valuable but non-trivial and requires the establishment of proper guidelines and the identification of research problems (scenarios) that are suitable to be addressed using digital traces analysis.

Where do we go from here?

More Datafication

The rise of the internet of things exemplifies the growing digitisation with increasing sensors, and ultimately, the tracking, collection, and analysis of personal data. Technical advances include tracking via speech recognition in smartphones (Kröger and Raschke 2019), or surveillance via motion and location sensors. For example, Amazon warehouse workers are being tracked to see if they are compliant with COVID-19 distancing policies[29]. Similar to the results in this thesis, it remains unclear if such motion tracking, with subsequent privacy invasion, leads to the desired outcomes; that is, fewer transmissions. Vendors are quick to provide novel people analytics solutions[30].

Beyond the workplace, algorithmic management makes inroads in many domains. Fivethirtyeight predicts political elections[31], sports analytics uses data analytics to optimise team tactics (Memmert and Rein 2018), delivery services and ride-hailing companies use nudging (Duggan et al. 2020), social media platforms automatise detection

[27] https://www.wired.com/story/how-work-became-an-inescapable-hellhole/ (accessed 2020-12-08).
[28] Ibid.
[29] https://www.wired.co.uk/article/coronavirus-work-office-surveillance (accessed 2020-12-08).
[30] https://www.humanyze.com/humanyze-announces-new-remote-workplace-analytics-solution (accessed 2020-12-06).
[31] https://projects.fivethirtyeight.com/2020-election-forecast/ (accessed 2020-12-06).

of hate speech (Jorgensen et al. 2020; Niemann et al. 2020), governments adopt algorithmic management[32], and an algorithm decides who gets medical attention in times of COVID-19 triage[33].

Concomitantly, defences are being set up. Internet browsers introduce technical barriers and anti-tracking features into their software, for example, deactivation of third-party cookies[34]. Parliaments pass new legislations that restrict data collection and analysis, for example, the general data protection regulation. Governments and independent bodies introduce guidelines for the deployment of algorithmic systems, for example, New Zealand[35], Germany[36], or the AI Now Institute[37] at New York University. To this end, pioneers of artificial intelligence work on "Responsible AI" that ensures fairness, reliability, safety, privacy, security, transparency, and accountability of algorithms; for example, Microsoft[38] and Google[39]. The municipalities of Amsterdam and Helsinki launched algorithm registers that list and explain all algorithms that are used in their respective administrations[40] [41]. Non-governmental organisations observe and monitor the deployment of algorithmic management in the wild, and criticise the opacity, dubbing it the "black box society." They call for more transparency, better technological literacy, and the auditing of algorithms (Chiusi et al. 2020).

Auditing

Currently, auditing algorithms is a prominent idea that is being discussed and researched to enable responsible algorithms. German politics instated a commission for algorithmic governance to work on regulating algorithmic management, standardise deployment, and develop norms and certifications for regulation (Beining 2020). Since algorithms are complex and proprietary, the public is unable to evaluate them, but professional third-party audits have been suggested (Beining 2020; Faraj et al. 2018; Latzer and Festic

---

[32] http://algorithmtips.org/2020/10/05/government-agencies-big-and-small-are-increasingly-adopting-controversial-algorithms-for-hiring/ (accessed 2020-12-06).
[33] https://www.ft.com/content/d738b2c6-000a-421b-9dbd-f85e6b333684 (accessed 2020-12-06).
[34] https://www.heise.de/news/Online-Ad-Summit-Was-tun-ohne-Cookies-4909399.html (accessed 2020-09-23).
[35] https://data.govt.nz/use-data/data-ethics/government-algorithm-transparency-and-accountability/algorithm-charter (accessed 2020-12-06).
[36] https://algorithmwatch.org/en/germanys-data-ethics-commission-releases-75-recommendations-with-eu-wide-application-in-mind/ (accessed 2020-12-06).
[37] https://ainowinstitute.org/ (accessed 2020-12-06).
[38] https://www.microsoft.com/en-us/ai/responsible-ai (accessed 2020-12-06).
[39] https://cloud.google.com/responsible-ai (accessed 2020-12-06).
[40] https://algoritmeregister.amsterdam.nl/en/ai-register/ (accessed 2020-12-06).
[41] https://venturebeat.com/2020/09/28/amsterdam-and-helsinki-launch-algorithm-registries-to-bring-transparency-to-public-deployments-of-ai (accessed 2020-12-06).

2019). Implementation of third-party audits could then be enforced *de jure*, providing a certified seal for qualifying algorithm systems that the public can see.

Compared to the traditional auditing of products and machines, the auditing of algorithms shows three differences, according to Beining (2020):

- Algorithms are subject to rapid change and updates;
- Algorithms are probabilistic systems, not deterministic;
- Algorithms are sociotechnical systems with different applications that must be seen in context.

Because algorithms are rapidly changing and rely on historical data as input, they require not only outcome-based auditing but process-based auditing as well (similar to ISO 9001[42]). The input data varies per context, so predetermined data samples for evaluating algorithms could be established—similar to how the academic computer science discipline judges the merits of novel algorithms. Future research is required to develop such data samples, among other quality criteria for auditing algorithms. Auditing algorithms requires not only specialised expertise in technology and statistics (Faraj et al. 2018), but also in domain knowledge, and social and legal knowledge (Beining 2020). In some cases, even experts are unable to properly understand and assess algorithms due to their complexity (Dourish 2016; Faraj et al. 2018). As a result, auditing algorithms and ensuring compliance requires a high investment of resources, and it is unclear how small- to medium-sized enterprises can achieve this investment. The auditing certified seal could start as a voluntary effort (Beining 2020).

Ongoing research efforts for algorithmic auditing include the German ExamAI project, which inquires how algorithmic management can be audited and certified amidst concerns for discrimination and bias, focusing on appropriateness[43]. It is funded by the German Federal Ministry of Labour and Social Affairs and supported by the German Informatics Society. Another project is “Trustworthy AI,” which develops procedures for third-party auditing of algorithmic management with a focus on validity and trust[44]. The project works on a combination of outcome-based and process-based auditing. It is funded by the German state North Rhine-Westphalia and supported by the Fraunhofer Institute for Intelligent Analysis and Information Systems together with the German Federal Office for Information Security.

[42] ISO 9001 is a norm to ensure quality standards in arbitrary processes through third-party auditing.
[43] https://testing-ai.gi.de (accessed 2020-12-06).
[44] https://www.ki.nrw/zertifizierung/ (accessed 2020-12-06).

Appropriate, meaningful, and valid algorithmic management, and digital traces analysis are recognised as top-level issues that will shape the future prowess of the Federal Republic of Germany in a digitised and algorithmified world.

6.3 Final Words

I have introduced this thesis with headlines of algorithmic management and data tracking. By the time this thesis is published, the newspapers will be littered with new headlines about algorithmic management—positive articles promising advances in artificial intelligence and digitalisation, and negative articles exposing data leaks, bias, discrimination, and other scandals. What all these headlines share, and what they demonstrate is that myriads of people are concerned with algorithmic management: scholars developing new algorithms, and researchers questioning the implications of algorithms in the workplace and beyond; commercial vendors bringing algorithms into practice, building value chains around them, improving their services and processes. Governments are funding research and instating regulations and policies, to balance the promises and perils of algorithmic management for their citizens. The citizens. Every single citizen, many unknowingly, is affected by algorithms in one way or another, and increasingly so.

The algorithmification is inevitable.

It is critical that we derive meaning from the algorithmification, whether in the form of original research findings or effective algorithmic systems in practice, while ensuring appropriate and valid mechanisms that balance privacy, fairness, and accuracy. The evaluation of these mechanisms should not happen *ex post*. Instead, it is the duty of information systems scholars, at the intersection of organisation, technology, and people, to proactively shape the discourse of algorithmic management and digital traces. They must share their expertise with the public and advise organisations, politicians, and citizens about appropriate, meaningful, and validated algorithms through audits, regulations, policies, recommendations, and education. I hope that I have made a small but meaningful contribution to this end.

The answer is not no algorithms, but better and smarter algorithms.

(Joschka A. Hüllmann 2020)

References

Abedin, B., Junglas, I., Meske, C., Motahari-Nezhad, H., and Rabhi, F. 2020. "Call for Papers: Special Issue on Designing and Managing Human-AI Interactions," *Information Systems Frontiers*, pp. 1–5. (http://repositorio.unan.edu.ni/2986/1/5624.pdf).

Agarwal, R., and Dhar, V. 2014. "Editorial—Big Data, Data Science, and Analytics: The Opportunity and Challenge for IS Research," *Information Systems Research* (25:3), pp. 443–448.

Ahmed, N., Michelin, R. A., Xue, W., Ruj, S., Malaney, R., Kanhere, S. S., Seneviratne, A., Hu, W., Janicke, H., and Jha, S. K. 2020. "A Survey of COVID-19 Contact Tracing Apps," *IEEE Access* (8), pp. 134577–134601.

Allen, R. G. D. 1951. *Statistics for Economists*, (1st ed.), London, UK: Hutchinson University Library.

Andersen, J. V., Lindberg, A., Lindgren, R., and Selander, L. 2016. "Algorithmic Agency in Information Systems: Research Opportunities for Data Analytics of Digital Traces," in *49th Hawaii International Conference on System Sciences (HICSS)*, Koloa, Hawaii, USA, pp. 4597–4605.

Angeletou, S., Rowe, M., and Alani, H. 2011. "Modelling and Analysis of User Behaviour in Online Communities," in *International Semantic Web Conference (ISWC 2011)*, Berlin, Heidelberg, pp. 35–50.

Badmaeva, T., and Hüllmann, J. A. 2019. "Investigating Personalized Price Discrimination of Textile- , Electronics- and General Stores in German Online Retail," in *Proceedings of the 14th International Conference on Wirtschaftsinformatik (WI)*, Siegen, Germany.

Bardi, U. 2011. *The Limits to Growth Revisited*, (1st ed.), (C. A. S. Hall, ed.), New York, New York, USA: Springer.

Barley, S. R., and Kunda, G. 2001. "Bringing Work Back In," *Organization Science* (12:1), pp. 76–95.

Bauer, T. N., and Erdogan, B. 2011. "Organizational Socialization: The Effective Onboarding of New Employees.," in *APA Handbook of Industrial and Organizational Psychology, Vol 3: Maintaining, Expanding, and Contracting the Organization.*, S. Zedeck (ed.), Washington: American Psychological Association, pp. 51–64.

Beining, L. 2020. "Vertrauenswürdige KI Durch Standards?," *Stiftung Neue Verantwortung e. V.*, Berlin, Germany, pp. 1–18.

Berente, N., Gu, B., Recker, J., and Santhanam, R. 2019. "Call for Papers MISQ Special Issue on Managing AI," *MIS Quarterly*, pp. 1–5.

Berente, N., Seidel, S., and Safadi, H. 2018. "Research Commentary—Data-Driven Computationally Intensive Theory Development," *Information Systems Research* (30:1), pp. 50–64.

Borgatti, S. P. 2004. "Testing Network Hypotheses. Presentation," *University of Kentucky Lectures*, Colchester, UK, pp. 1–23.

Braun, M. T., and Kuljanin, G. 2015. "Big Data and the Challenge of Construct Validity," *Industrial and Organizational Psychology* (8:4), pp. 521–527.

Breiter, A., and Hepp, A. 2018. "The Complexity of Datafication: Putting Digital Traces in Context," in *Communicative Figurations* (1st ed.), A. Hepp, A. and Breiter, and

U. and Hasebrink (eds.), Cham, Switzerland: Springer International Publishing, pp. 387–405.

van den Broek, E., Sergeeva, A., and Huysman, M. 2019. "Hiring Algorithms: An Ethnography of Fairness in Practice," in *Proceedings of the 40th International Conference on Information Systems (ICIS)*, pp. 1–8.

Brynjolfsson, E., and McAfee, A. 2014. *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*, (1st ed.), New York, New York, USA: W. W. Norton & Company.

Brynjolfsson, E., and Mitchell, T. 2017. "What Can Machine Learning Do? Workforce Implications," *Science* (358:6370), pp. 1530–1534.

Burton, J. W., Stein, M., and Jensen, T. B. 2020. "A Systematic Review of Algorithm Aversion in Augmented Decision Making," *Journal of Behavioral Decision Making* (33:2), pp. 220–239.

Cavoukian, A., and Castro, D. 2014. "Big Data and Innovation, Setting the Record Straight: De-Identification Does Work," *Office of the Information and Privacy Commissioner*, Ontario, Canada.

Cetto, A., Klier, M., Richter, A., and Zolitschka, J. F. 2018. "'Thanks for Sharing'—Identifying Users' Roles Based on Knowledge Contribution in Enterprise Social Networks," *Computer Networks* (135), pp. 275–288.

Chaffin, D., Heidl, R., Hollenbeck, J. R., Howe, M., Yu, A., Voorhees, C., and Calantone, R. 2017. "The Promise and Perils of Wearable Sensors in Organizational Research," *Organizational Research Methods* (20:1), pp. 3–31.

Chen, H., Chiang, R. H. L., and Storey, V. C. 2012. "Business Intelligence and Analytics: From Big Data to Big Impact," *MIS Quarterly* (36:4), pp. 1165–1188.

Chiusi, F., Fischer, S., Kayser-Bril, N., and Spielkamp, M. 2020. "Automating Society Report," *AlgorithmWatch*, Berlin, Germany.

Christin, A. 2020. "Algorithmic Ethnography, during and after COVID-19," *Communication and the Public* (5:3–4), pp. 108–111.

Cloutier, C., and Langley, A. 2020. "What Makes a Process Theoretical Contribution?," *Organization Theory* (1:1), pp. 1–32.

Crowston, K., and Bolici, F. 2020. "Impacts of the Use of Machine Learning on Work Design," in *Proceedings of the 8th International Conference on Human-Agent Interaction*, New York, New York, USA: ACM, November 10, pp. 163–170.

da Cunha, J. V. 2013. "A Dramaturgical Model of the Production of Performance Data," *MIS Quarterly* (37:3), pp. 723–748.

Damkjær, M. 2015. "Becoming a Parent in a Digitized Age: Facebook as an Agent of Change?: Performative, Dialogical, and Preservative Facebook Strategies in the Transition to Parenthood," in *NordMedia 2015: Media Presences - Mobile Modernities*.

Dastin, J. 2018. "Amazon Scraps Secret AI Recruiting Tool That Showed Bias against Women," *Reuters*. (https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G).

DeLanda, M. 2005. *Intensive Science and Virtual Philosophy*, (3rd ed.), London, UK: Bloomsbury Academic.

Dourish, P. 2016. "Algorithms and Their Others: Algorithmic Culture in Context," *Big Data & Society* (3:2), pp. 1–11.

Duggan, J., Sherman, U., Carbery, R., and McDonnell, A. 2020. “Algorithmic Management and App-work in the Gig Economy: A Research Agenda for Employment Relations and HRM,” *Human Resource Management Journal* (30:1), pp. 114–132.

Eagle, N., and Pentland, A. (Sandy). 2006. “Reality Mining: Sensing Complex Social Systems,” *Personal and Ubiquitous Computing* (10:4), pp. 255–268.

Espinosa, J. A., Cummings, J. N., Wilson, J. M., and Pearce, B. M. 2003. “Team Boundary Issues Across Multiple Global Firms,” *Journal of Management Information Systems* (19:4), pp. 157–190.

Eubanks, V. 2017. *Automating Inequality. How High-Tech Tools Profile, Police, and Punish the Poor*, (1st ed.), New York, New York, USA: St. Martin’s Press.

Faraj, S., Pachidi, S., and Sayegh, K. 2018. “Working and Organizing in the Age of the Learning Algorithm,” *Information and Organization* (28:1), pp. 62–70.

Farine, D. R. 2017. “A Guide to Null Models for Animal Social Network Analysis,” *Methods in Ecology and Evolution* (8:10), (R. Freckleton and S. Rands, eds.), pp. 1309–1320.

Farine, D. R., and Whitehead, H. 2015. “Constructing, Conducting and Interpreting Animal Social Network Analysis,” *Journal of Animal Ecology* (84:5), pp. 1144–1163.

Fernández-Macías, E. 2018. “Automation, Digitisation and Platforms: Implications for Work and Employment,” *Publications Office of the European Union*, Luxembourg. (http://eurofound.link/ef18002%0Ahttps://digitalcommons.ilr.cornell.edu/intl/636%0Ahttp://eurofound.link/ef18002).

Flyverbom, M., and Murray, J. 2018. “Datastructuring—Organizing and Curating Digital Traces into Action,” *Big Data & Society* (5:2), pp. 1–12.

Fraser, H., Parker, T., Nakagawa, S., Barnett, A., and Fidler, F. 2018. “Questionable Research Practices in Ecology and Evolution,” *PLoS ONE* (13:7), p. e0200303.

Freelon, D. 2014. “On the Interpretation of Digital Trace Data in Communication and Social Computing Research,” *Journal of Broadcasting & Electronic Media* (58:1), pp. 59–75.

Fulk, J., Schmitz, J., and Steinfield, C. 1990. “A Social Influence Model of Technology Use,” in *Organizations and Communication Technology*, J. Fulk and C. Steinfield (eds.), Newbury Park, California, USA: SAGE Publications Ltd, pp. 117–142.

Gal, U., Jensen, T. B., and Stein, M.-K. 2020. “Breaking the Vicious Cycle of Algorithmic Management: A Virtue Ethics Approach to People Analytics,” *Information and Organization* (30:2).

Gal, U., Jensen, T. B., and Stein, M. K. 2017. “People Analytics in the Age of Big Data: An Agenda for IS Research,” in *Proceedings of the International Conference on Information Systems (ICIS)*, Seoul, South Korea, pp. 1–11.

Galliers, R. D., Newell, S., Shanks, G., and Topi, H. 2017. “Datification and Its Human, Organizational and Societal Effects: The Strategic Opportunities and Challenges of Algorithmic Decision-Making,” *Journal of Strategic Information Systems* (26:3), pp. 185–190.

Geletkanycz, M., and Tepper, B. J. 2012. “Publishing in AMJ–Part 6: Discussing the Implications,” *Academy of Management Journal* (55:2), pp. 256–260.

Gianfrancesco, M. A., Tamang, S., Yazdany, J., and Schmajuk, G. 2018. “Potential Biases in ML Algorithms Using EHR Data,” *JAMA Internal Medicine* (178:11), pp. 1544–1547.

Goad, D., and Gal, U. 2018. "Understanding the Impact of Transparency on Algorithmic Decision Making Legitimacy," in *Working Conference on Information Systems and Organizations* (Vol. 543), IFIP Advances in Information and Communication Technology, U. Schultze, M. Aanestad, M. Mähring, C. Østerlund, and K. Riemer (eds.), Cham, Switzerland: Springer International Publishing, pp. 64–79.

Goh, K. H., and Bockstedt, J. C. 2013. "The Framing Effects of Multipart Pricing on Consumer Purchasing Behavior of Customized Information Good Bundles," *Information Systems Research* (24:2), pp. 334–351.

Greenstein, G. 1983. *Frozen Star*, (1st ed.), New York, New York, USA: Freundlich Books.

Gregor, S. 2006. *The Nature of Theory in Information Systems*, (30:3), pp. 611–642.

Grønsund, T., and Aanestad, M. 2020. "Augmenting the Algorithm: Emerging Human-in-the-Loop Work Configurations," *The Journal of Strategic Information Systems* (29:2), pp. 1–16.

Grover, V., Lindberg, A., Benbasat, I., and Lyytinen, K. 2020. "The Perils and Promises of Big Data Research in Information Systems," *Journal of the Association for Information Systems* (21:2), pp. 268–293.

Hacker, J., and Riemer, K. 2020. "Identification of User Roles in Enterprise Social Networks: Method Development and Application," *Business & Information Systems Engineering*, Springer Fachmedien Wiesbaden.

Harper, R. H. R. 2019. "The Role of HCI in the Age of AI," *International Journal of Human–Computer Interaction* (35:15), pp. 1331–1344.

Hedman, J., Srinivasan, N., and Lindgren, R. 2013. "Digital Traces of Information Systems: Sociomateriality Made Researchable," in *Proceedings of the Thirty Fourth International Conference on Information Systems (ICIS)*, pp. 1–17.

Henrich, J., Heine, S. J., and Norenzayan, A. 2010. "The Weirdest People in the World?," *Behavioral and Brain Sciences* (33), pp. 61–135.

Hodge, V., Devlin, S., Sephton, N., Block, F., Cowling, P., and Drachen, A. 2019. "Win Prediction in Multi-Player Esports: Live Professional Match Prediction," *IEEE Transactions on Games* (E.A.), IEEE, pp. 1–1.

Howison, J., Wiggins, A., and Crowston, K. 2011. "Validity Issues in the Use of Social Network Analysis for the Study of Online Communities," *Journal of the Association for Information Systems* (12:12), pp. 767–797.

Hüllmann, J. A. 2019. "The Construction of Meaning through Digital Traces," in *Proceedings of the Pre-ICIS 2019, International Workshop on The Changing Nature of Work*, München.

Hüllmann, J. A., and Hentschel, J. 2021. "Beyond the Formal: Drivers of Informal Communication in Enterprise Social Networks," *In Preparation*.

Hüllmann, J. A., and Krebber, S. 2020. "Identifying Temporal Rhythms Using Email Traces," in *Proceedings of the America's Conference of Information Systems (AMCIS)*, Salt Lake City, Utah, USA.

Hüllmann, J. A., and Krebber, S. 2021a. "Strategic and Operational People Analytics: Reviewing the Dominant Conceptions in Academia and Practice," *In Preparation*, pp. 1–38.

Hüllmann, J. A., and Krebber, S. 2021b. "The Data Economy: An Introduction," *In Preparation*.

Hüllmann, J. A., Krebber, S., and Troglauer, P. 2021a. "The IT Artifact in People Analytics: Reviewing Tools to Understand a Nascent," in *Proceedings of the 16th*

International Conference on Wirtschaftsinformatik (WI), Duisburg-Essen, Germany.

Hüllmann, J. A., Krebber, S., and Troglauer, P. 2021b. "Exploring Media Collections of Distributed Workers Using Digital Traces," in *In Preparation*.

Hüllmann, J. A., and Kroll, T. 2018. "The Impact of User Behaviours on the Socialisation Process in Enterprise Social Networks," in *Proceedings of the Australasian Conference on Information Systems (ACIS)*, Sydney, Australia.

Hüllmann, J. A., and Mattern, J. 2020. "Three Issues with the State of People and Workplace Analytics," in *Proceedings of the 33rd Bled EConference*, Bled, Slovenia.

Hupperich, T., Tatang, D., Wilkop, N., and Holz, T. 2018. "An Empirical Study on Online Price Differentiation," in *Proceedings of the Eighth ACM Conference on Data and Application Security and Privacy*, pp. 76–83.

Jackson, C., Anderson, L. C., and Kieliszewski, C. A. 2020. "Methodological Reinforcements: Investigating Work Through Trace Data and Text," in *International Conference on Applied Human Factors and Ergonomics*, Cham, Switzerland: Springer, pp. 70–76.

John, L. K., Loewenstein, G., and Prelec, D. 2012. "Measuring the Prevalence of Questionable Research Practices with Incentives for Truth Telling," *Psychological Science* (23:5), pp. 524–532.

Jorgensen, M., Choi, M., Niemann, M., Brunk, J., and Becker, J. 2020. "Multi-Class Detection of Abusive Language Using Automated Machine Learning," in *Proceedings of the 15. Internationale Tagung Wirtschaftsinformatik (WI 2020)*, pp. 1763–1775.

Klein, S., and Hüllmann, J. 2018. "Datenkapitalismus Akademischer Wissenschaftsverlage," *Wirtschaftsdienst* (98:7), pp. 477–480.

Kleinberg, J., Ludwig, J., Mullainathan, S., and Sunstein, C. R. 2018. "Discrimination in the Age of Algorithms," *Journal of Legal Analysis* (10), pp. 113–174.

Kneidinger-Müller, B. 2018. "Self-Tracking Data as Digital Traces of Identity: A Theoretical Analysis of Contextual Factors of Self-Observation Practices," *International Journal of Communication* (12), pp. 629–646.

Kröger, J. L., and Raschke, P. 2019. "Is My Phone Listening in? On the Feasibility and Detectability of Mobile Eavesdropping," in *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 11559 LNCS), pp. 102–120.

Lakens, D. 2017. "Equivalence Tests," *Social Psychological and Personality Science* (8:4), pp. 355–362.

Lakens, D. 2020. "The Practical Alternative to the P-Value Is the Correctly Used p-Value," *Perspectives on Psychological Science* (in press), pp. 1–18.

Lakens, D., Scheel, A. M., and Isager, P. M. 2018. "Equivalence Testing for Psychological Research: A Tutorial," *Advances in Methods and Practices in Psychological Science* (1:2), pp. 259–269.

Landers, R. N., Brusso, R. C., Cavanaugh, K. J., and Collmus, A. B. 2016. "A Primer on Theory-Driven Web Scraping: Automatic Extraction of Big Data from the Internet for Use in Psychological Research.," *Psychological Methods* (21:4), pp. 475–492.

Langley, A. 1999. "Strategies for Theorizing from Process Data," *The Academy of Management Review* (24:4), pp. 691–710. (http://www.jstor.org/stable/259349).

Latzer, M., and Festic, N. 2019. "A Guideline for Understanding and Measuring Algorithmic Governance in Everyday Life," *Internet Policy Review* (8:2), pp. 1–19.

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabasi, A.-L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., and Van Alstyne, M. 2009. "SOCIAL SCIENCE: Computational Social Science," *Science* (323:5915), pp. 721–723.

Lee, M. K. 2018. "Understanding Perception of Algorithmic Decisions: Fairness, Trust, and Emotion in Response to Algorithmic Management," *Big Data & Society* (5:1), pp. 1–16.

Lee, M. K., Kusbit, D., Metsky, E., and Dabbish, L. 2015. "Working with Machines: The Impact of Algorithmic and Data-Driven Management on Human Workers," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*, New York, New York, USA: ACM Press, pp. 1603–1612.

Leonardi, P. M., and Treem, J. W. 2020. "Behavioral Visibility: A New Paradigm for Organization Studies in the Age of Digitization, Digitalization, and Datafication," *Organization Studies* (41:12), pp. 1601–1625.

Levenson, A. 2018. "Using Workforce Analytics to Improve Strategy Execution," *Human Resource Management* (57:3), pp. 685–700.

Lindberg, A. 2020. "Developing Theory Through Integrating Human and Machine Pattern Recognition," *Journal of the Association for Information Systems* (21:1), pp. 90–116.

Maass, W., Parsons, J., Purao, S., Storey, V. C., and Woo, C. 2018. "Data-Driven Meets Theory-Driven Research in the Era of Big Data: Opportunities and Challenges for Information Systems Research," *Journal of the Association for Information Systems* (19:12), pp. 1253–1273.

Markus, M. L. 2017. "Datification, Organizational Strategy, and IS Research: What's the Score?," *The Journal of Strategic Information Systems* (26:3), pp. 233–241.

Markus, M. L., and Marabelli, M. 2017. "Researching Big Data Research: Ethical Implications for IS Scholars," in *The 23rd Americas Conference on Information Systems (AMCIS)*, pp. 1–5.

Martin, T., Karopoulos, G., Hernández-Ramos, J. L., Kambourakis, G., and Nai Fovino, I. 2020. "Demystifying COVID-19 Digital Contact Tracing: A Survey on Frameworks and Mobile Apps," *Wireless Communications and Mobile Computing* (2020), (K.-K. R. Choo, ed.), pp. 1–29.

Maslow, A. H. 1966. *The Psychology of Science: A Reconnaissance*, (1st ed.), New York, New York, USA: Harper & Row.

Mattern, J., Lansmann, S., and Hüllmann, J. A. 2021. "It's Not That Bad! Perceived Stress of Knowledge Workers During Enforced Working From Home Due to COVID-19," in *Proceedings of the 16th International Conference on Wirtschaftsinformatik (WI)*.

Mayer-Schönberger, V., and Cukier, K. 2012. *Big Data: A Revolution That Transforms How We Work, Live, and Think*, (1st ed.), New York, New York, USA: Houghton Mifflin Harcourt.

McAfee, A. P., and Brynjolfsson, E. 2012. "Big Data: The Management Revolution," *Harvard Business Review* (90:10), pp. 1–9.

Memmert, D., and Rein, R. 2018. "Match Analysis, Big Data and Tactics: Current Trends in Elite Soccer," *Deutsche Zeitschrift Für Sportmedizin* (69), pp. 65–72.

Mertens, W., and Recker, J. 2020. "New Guidelines for Null Hypothesis Significance Testing in Hypothetico-Deductive IS Research," *Journal of the Association for Information Systems* (21:4), pp. 1072–1102.

Möhlmann, M., and Zalmanson, L. 2017. "Hands on the Wheel: Navigating Algorithmic

Management and Uber Drivers Autonomy," in *Proceedings of the International Conference on Information Systems (ICIS)*.

Mohr, L. B. 1982. *Explaining Organizational Behavior*, (1st ed.), San Francisco, California, USA: Jossey-Bass.

Narayanan, A. 2019. "How to Recognize AI Snake Oil. Presentation," *Princeton Lectures*, pp. 1–21.

Narayanan, A., Shi, E., and Rubinstein, B. I. P. 2011. "Link Prediction by De-Anonymization: How We Won the Kaggle Social Network Challenge," in *The 2011 International Joint Conference on Neural Networks*, IEEE, July, pp. 1825–1834.

Newell, S., and Marabelli, M. 2015. "Strategic Opportunities (and Challenges) of Algorithmic Decision-Making: A Call for Action on the Long-Term Societal Effects of 'Datification,'" *The Journal of Strategic Information Systems* (24:1), pp. 3–14.

Niemann, M., Welsing, J., Riehle, D. M., Brunk, J., Assenmacher, D., and Becker, J. 2020. "Abusive Comments in Online Media and How to Fight Them," in *Proceedings of the Multidisciplinary International Symposium on Disinformation in Open Online Media*, pp. 122–137.

Olteanu, A., Castillo, C., Diaz, F., and Kıcıman, E. 2019. "Social Data: Biases, Methodological Pitfalls, and Ethical Boundaries," *Frontiers in Big Data* (2:13).

Orlikowski, W. J., and Scott, S. V. 2016. "Digital Work: A Research Agenda," in *A Research Agenda for Management and Organization Studies*, Northhampton, Massachusetts, USA: Edward Elgar Publishing, pp. 88–96.

Ortiz de Guinea, A., and Webster, J. 2017. "Combining Variance and Process in Information Systems Research: Hybrid Approaches," *Information and Organization* (27:3), pp. 144–162.

Østerlund, C., Crowston, K., and Jackson, C. 2020. "Building an Apparatus: Refractive, Reflective & Diffractive Readings of Trace Data," *Journal of the Association for Information Systems* ((in press)), pp. 1–43.

Pachidi, S., Huysman, M., and Berends, H. 2016. "Playing the Numbers Game: Dealing with Transparency," in *Proceedings of the International Conference on Information Systems (ICIS)*.

Pentland, A. 2015. *Social Physics*, (1st ed.), New York, New York, USA: Penguin Books.

Pentland, B. T., Recker, J., Wolf, J. R., and Wyner, G. 2020. "Bringing Context inside Process Research with Digital Trace Data," *Journal of the Association for Information Systems* (21:5), pp. 1214–1236.

Perer, A., Shneiderman, B., and Oard, D. W. 2006. "Using Rhythms of Relationships to Understand E-Mail Archives," *Journal of the American Society for Information Science and Technology* (57:14), pp. 1936–1948.

Persily, N. 2017. "The 2016 U.S. Election: Can Democracy Survive the Internet," *Journal of Democracy* (28:2), pp. 63–76.

Poels, T., Tucker, D. A., and Kielema, J. 2017. "The Development of a Theoretical Framework of Organisational Rhythm," *Journal of Organizational Change Management* (30:6), pp. 888–902.

Popper, K. 2009. *Vermutungen Und Widerlegungen. Das Wachstum Der Wissenschaftlichen Erkenntnis*, (2nd ed.), (H. Keuth, ed.), Tübingen, Germany: Mohr Siebeck.

Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J. F., Breazeal, C., Crandall, J. W., Christakis, N. A., Couzin, I. D., Jackson, M. O., Jennings, N. R., Kamar, E., Kloumann, I. M., Larochelle, H., Lazer, D., McElreath, R., Mislove, A., Parkes, D. C., Pentland, A. 'Sandy,' Roberts, M. E., Shariff, A., Tenenbaum, J. B., and Wellman, M. 2019. "Machine Behaviour," *Nature* (568:7753), Springer US, pp. 477–486.

Rai, A. 2017. "Avoiding Type III Errors: Formulating IS Research Problems That Matter," *MIS Quarterly* (41:2), iii–vii.

Riemer, K., and Peter, S. 2020. "The Robo-Apocalypse Plays out in the Quality, Not in the Quantity of Work," *Journal of Information Technology* (35:4), pp. 310–315.

Rocher, L., Hendrickx, J. M., and de Montjoye, Y.-A. 2019. "Estimating the Success of Re-Identifications in Incomplete Datasets Using Generative Models," *Nature Communications* (10:3069), Springer US, pp. 1–9.

Rothmeier, K., Pflanzl, N., Hüllmann, J. A., and Preuss, M. 2020. "Prediction of Player Churn and Disengagement Based on User Activity Data of a Freemium Online Strategy Game," *IEEE Transactions on Games* (E.A.), pp. 1–11.

Scharkow, M. 2016. "The Accuracy of Self-Reported Internet Use—A Validation Study Using Client Log Data," *Communication Methods and Measures* (10:1), Routledge, pp. 13–27.

Schellhammer, S., and Watson-Manheim, M. B. 2019. "The Isolated Individual in the Networked Organization: The Need to Regain Control," *GRF/NSF Joint Grant Proposal*, Münster, Germany, pp. 1–9.

Seeber, I., Bittner, E., Briggs, R. O., de Vreede, T., de Vreede, G. J., Elkins, A., Maier, R., Merz, A. B., Oeste-Reiß, S., Randrup, N., Schwabe, G., and Söllner, M. 2020. "Machines as Teammates: A Research Agenda on AI in Team Collaboration," *Information and Management* (57:2), pp. 1–22.

Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., and Vertesi, J. 2019. "Fairness and Abstraction in Sociotechnical Systems," in *Proceedings of the Conference on Fairness, Accountability, and Transparency - FAT* '19*, New York, New York, USA: ACM Press, pp. 59–68.

Shapiro, C., and Varian, H. R. 1999. *Information Rules. A Strategic Guide to the Network Economy*, Boston, Massachusetts, USA: Harvard Business School Press.

Simonite, T. 2019. "The Best Algorithms Struggle to Recognize Black Faces Equally," *WIRED*. (https://www.wired.com/story/best-algorithms-struggle-recognize-black-faces-equally/?verso=true).

Smith, M., Hansen, D. L., and Gleave, E. 2009. "Analyzing Enterprise Social Media Networks," in *2009 International Conference on Computational Science and Engineering* (Vol. 4), IEEE, pp. 705–710.

Song, C., Qu, Z., Blumm, N., and Barabasi, A.-L. 2010. "Limits of Predictability in Human Mobility," *Science* (327:5968), pp. 1018–1021.

Stewart, I. 2019. "Our Behaviour in Bulk Is More Predictable than We like to Imagine," *Aeon Essays*, pp. 1–6.

Thapa, R., and Vidolov, S. 2020. "Evaluating Distributed Leadership in Open Source Software Communities," in *Proceedings of the European Conference on Information Systems*, pp. 1–17.

Tonidandel, S., King, E. B., and Cortina, J. M. 2018. "Big Data Methods: Leveraging Modern Data Analytic Techniques to Build Organizational Science," *Organizational Research Methods* (21:3), pp. 525–547.

Trielli, D., and Diakopoulos, N. 2019. "Search as News Curator," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, New York, New York, USA: ACM, May 2, pp. 1–15.

Varian, H. R. 1995. "Information Economy - How Much Will Two Bits Be Worth in the Digital," *Scientific American* (2), pp. 44–46.

Venkatesh, V., Brown, S. A., and Bala, H. 2013. "Bridging the Qualitative-Quantitative Divide: Guidelines for Conducting Mixed Methods Research in Information Systems," *MIS Quarterly* (37:1), pp. 21–54.

Waardenburg, L., Anastasia, S., and Huysman, M. 2018. "Predictive Policing: How Algorithms Inscribe the Understanding of Crime in Police Work," in *Academy of Management Global Proceedings*.

Waber, B. 2013. *People Analytics: How Social Sensing Technology Will Transform Business and What It Tells Us about the Future of Work*, (1st ed.), Upper Saddle, New Jersy, USA: Financial Times Press.

Walorska, A. M. 2020. "The Algorithmic Society," in *Redesigning Organizations*, D. Feldner (ed.), Cham, Switzerland: Springer International Publishing, pp. 149–160.

Watson-Manheim, M. B., and Bélanger, F. 2007. "Communication Media Repertoires: Dealing with the Multiplicity of Media Choices," *MIS Quarterly* (31:2), pp. 267–293.

West, S. M. 2017. "Data Capitalism: Redefining the Logics of Surveillance and Privacy," *Business & Society* (58:1), pp. 20–41.

Xu, H., Zhang, N., and Zhou, L. 2020. "Validity Concerns in Research Using Organic Data," *Journal of Management* (46:7), pp. 1257–1274.

Zarsky, T. 2016. "The Trouble with Algorithmic Decisions: An Analytic Road Map to Examine Efficiency and Fairness in Automated and Opaque Decision Making," *Science Technology and Human Values* (41:1), pp. 118–132.

Zuboff, S. 2015. "Big Other: Surveillance Capitalism and the Prospects of an Information Civilization," *Journal of Information Technology* (30:1), pp. 75–89.

Zuboff, S. 2019. *The Age of Surveillance Capitalism*, (1st ed.), New York, New York, USA: Public Affairs, Hachette Book Group.

Zuckerman, E. W. 2017. "On Genre: A Few More Tips to Article-Writers," *MIT Sloan School of Management*, Cambridge, Massachusetts, USA, pp. 1–16.

Web References

Caren Chesler (2019). Coronavirus will turn your office into a surveillance state. WIRED UK (last access 2020-12-08). https://www.wired.co.uk/article/coronavirus-work-office-surveillance

Jeffrey Dastin (2018). Amazon scraps secret AI recruiting tool that showed bias against women. Reuters (last access 2020-12-08). https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G

Torsten Kleinz (2020). Online Ad Summit: Was tun ohne Cookies? Heise Online (last access 2020-09-23). https://www.heise.de/news/Online-Ad-Summit-Was-tun-ohne-Cookies-4909399.html

Cade Metz (2019). We Teach A.I. Systems Everything, Including Our Biases. The New York Times (last access 2020-12-08). https://www.nytimes.com/2019/11/11/technology/artificial-intelligence-bias.html

Sendhil Mullainathan (2019). Biased Algorithms Are Easier to Fix Than Biased People. The New York Times (last access 2020-12-08). https://www.nytimes.com/2019/12/06/business/algorithm-bias-fix.html

Anne Helen Petersen (2020). How Work Became an Inescapable Hellhole. WIRED UK (last access 2020-12-08). https://www.wired.com/story/how-work-became-an-inescapable-hellhole/

Ed Pilkington (2019). Digital dystopia: how algorithms punish the poor. The Guardian (last access 2020-12-08). https://www.theguardian.com/technology/2019/oct/14/automating-poverty-algorithms-punish-poor

Tom Simonite (2018). When It Comes to Gorillas, Google Photos Remains Blind. WIRED UK (last access 2020-12-08). https://www.wired.com/story/when-it-comes-to-gorillas-google-photos-remains-blind/

Ian Stewart (2019). Social physics. Aeon (last access 2020-12-08). https://aeon.co/essays/our-behaviour-in-bulk-is-more-predictable-than-we-like-to-imagine

Duygu Uygun-Tunc (2020). Trust and criticism in science, Part I: Critical rationalism instead of organized skepticism. Science and Philosophy (last access 2020-12-08). https://medium.com/science-and-philosophy/trust-and-criticism-in-science-bbb62bd6890

Daniel Victor (2016). Microsoft Created a Twitter Bot to Learn From Users. It Quickly Became a Racist Jerk. The New York Times (last access 2020-12-08). https://www.nytimes.com/2016/03/25/technology/microsoft-created-a-twitter-bot-to-learn-from-users-it-quickly-became-a-racist-jerk.html

Appendix

A Papers Included in This Dissertation

Citation	Hüllmann (2019)
Title	The Construction of Meaning through Digital Traces
Authors	Joschka Hüllmann (100%)
Outlet	Proceedings of the Pre-ICIS 2019, Workshop on The Changing Nature of Work, München
Citation	Hüllmann and Kroll (2018)
Title	The Impact of User Behaviours on the Socialisation Process in Enterprise Social Networks
Authors	Joschka Hüllmann (80%), Tobias Kroll (20%)
Outlet	Proceedings of the 29th Australasian Conference on Information Systems, Sydney, Australia
Citation	Hüllmann and Krebber (2020)
Title	Identifying Temporal Rhythms using Email Traces
Authors	Joschka Hüllmann (85%), Simone Krebber (15%)
Outlet	Proceedings of the America's Conference of Information Systems, Salt Lake City, Utah, USA
Citation	Hüllmann et al. (2021)
Title	Exploring Media Collections of Distributed Workers Using Digital Traces
Authors	Joschka Hüllmann (TBD), Simone Krebber (TBD), Patrick Troglauer (TBD)
Outlet	In preparation
Citation	Hüllmann and Hentschel (2021)
Title	Beyond the Formal: Drivers of Informal Communication in Enterprise Social Networks
Authors	Joschka Hüllmann (TBD), Julian Hentschel (TBD)
Outlet	In preparation
Citation	Hüllmann and Mattern (2020)
Title	Three Issues with the State of People and Workplace Analytics
Authors	Joschka Hüllmann (100%), Jana Mattern (0%)
Outlet	Proceedings of the 33rd Bled eConference, Bled, Slovenia
Citation	Hüllmann et al. (2021)
Title	The IT artifact in People Analytics: Reviewing the Tools to Understand a Nascent Field
Authors	Joschka Hüllmann (65%), Simone Krebber (20%), Patrick Troglauer (15%)
Outlet	Proceedings of the 16th International Conference on Wirtschaftsinformatik (WI), Duisburg-Essen, Germany
Citation	Hüllmann and Krebber (2021)
Title	Strategic and Operational People Analytics: Reviewing the Dominant Conceptions in Academia and Practice
Authors	Joschka Hüllmann (TBD), Simone Krebber (TBD)
Outlet	In preparation
Citation	Hüllmann and Krebber (2021)
Title	The Data Economy: An Introduction
Authors	Joschka Hüllmann (TBD), Simone Krebber (TBD)
Outlet	In preparation
Citation	Badmaeva and Hüllmann (2019)
Title	Investigating Personalized Price Discrimination of Textile- , Electronics- and General Stores in German Online Retail
Authors	Tsagana Badmaeva (40%), Joschka Hüllmann (60%)
Outlet	Proceedings of the 14th International Conference on Wirtschaftsinformatik (WI), Siegen, Germany
Citation	Klein and Hüllmann (2018)
Title	Datenkapitalismus Akademischer Wissenschaftsverlage
Authors	Stefan Klein (70%), Joschka Hüllmann (30%)
Outlet	Wirtschaftsdienst, Volume 98, Issue 7
Citation	Rothmeier et al. (2020)
Title	Prediction of Player Churn and Disengagement Based on User Activity Data of a Freemium Online Strategy Game
Authors	Karsten Rothmeier (50%), Nicolas Pflanzl (5%), Mike Preuss (5%), Joschka Hüllmann (40%)
Outlet	IEEE Transactions on Games, Volume 3, Issue 1

B Curriculum Vitae – Joschka Hüllmann

Personal Data

Name	Joschka Andreas Hüllmann
Date of Birth	
Nationality	German

Education

April 2017 – February 2021	Dr. rer. pol., Information Systems **Westfälische Wilhelms-Universität Münster** Advisors:
April 2015 – March 2017	Master of Science, Information Systems **Westfälische Wilhelms-Universität Münster** Advisors:
August 2011 – January 2015	Bachelor of Science, Information Systems **IBS IT & Business School Oldenburg** Advisors:
August 2011 – January 2015	Fachinformatiker Anwendungsentwicklung **IHK Oldenburg**
Until August 2011	Abitur

Professional Experience

August 2017 – September 2021	Research Associate **Westfälische Wilhelms-Universität Münster**
April 2017 – July 2017	Research Associate **Universität Hamburg**
February 2015 – March 2017	Software Engineer, Part time **AMCON GmbH, Cloppenburg**
August 2011 – January 2015	Junior Software Engineer (Duales Studium) **BTC Business Technology Consulting AG, Oldenburg**

C Academic Publications

10. **Hüllmann, J. A.**, Sivakumar, A., & Krebber, S. (2021). Data Management Platforms: An Empirical Taxonomy. In Proceedings of the 34th Bled eConference 2021, Bled, Slovenia.

9. **Hüllmann, J. A.**, Krebber, S., & Troglauer, P. (2021). The IT Artifact in People Analytics: Reviewing the Tools to Understand a Nascent Field. In Proceedings of the 16th International Conference on Wirtschaftsinformatik, Duisburg-Essen, Germany.

8. Mattern, J., Lansmann, S., & **Hüllmann, J. A.** (2021). Home Office due to COVID-19: It's not that bad! Enforced Working from Home and Perceived Stress. In Proceedings of the 16th International Conference on Wirtschaftsinformatik, Duisburg-Essen, Germany.

7. **Hüllmann, J. A.**, & Mattern, J. (2020). Three Issues with the State of People and Workplace Analytics. In Proceedings of the 33rd Bled eConference, Bled, Slovenia.

6. **Hüllmann, J. A.**, & Krebber, S. (2020). Identifying Temporal Rhythms using Email Traces. In Proceedings of the America's Conference of Information Systems (AMCIS), Salt Lake City, Utah, USA.

5. Rothmeier, K., Pflanzl, N., **Hüllmann, J. A.**, & Preuss, M. (2020). Prediction of Player Churn and Disengagement Based on User Activity Data of a Freemium Online Strategy Game. IEEE Transactions on Games, 13(1), 78-88.

4. **Hüllmann, J. A.** (2019). The Construction of Meaning through Digital Traces. In Proceedings of the Pre-ICIS 2019, International Workshop on The Changing Nature of Work, München, Germany.

3. Badmaeva, T., & **Hüllmann, J. A.** (2019). Investigating Personalized Price Discrimination of Textile-, Electronics- and General Stores in German Online Retail. In Proceedings of the 14th International Conference on Wirtschaftsinformatik (WI), Siegen, Germany.

2. **Hüllmann, J. A.**, & Kroll, T. (2018). The Impact of User Behaviours on the Socialisation Process in Enterprise Social Networks. In Proceedings of the 29th Australasian Conference on Information Systems (ACIS), Sydney, Australia.

1. Klein, S., & **Hüllmann, J. A.** (2018). Datenkapitalismus akademischer Wissenschaftsverlage. Wirtschaftsdienst, 98(7), 477–480.

Declaration of Authorship

I hereby declare that, to the best of my knowledge and belief, this Doctoral Thesis titled "Smarter Work? Promises and Perils of Algorithmic Management in the Workplace Using Digital Traces" is my own work. I confirm that each significant contribution to and quotation in this thesis that originates from the work or works of others is indicated by proper use of citation and references. This thesis has not yet been part of another examination.

Münster, 06 August 2021

Joschka Hüllmann