Book of Abstracts

Conference

Big Data Discourses

Communicating, Deliberating, and Imagining Datafication

1 & 2 February 2024

Zeppelinhaus Leipzig
Nikolaistraße 27-29
04109 Leipzig

https://www.sozphil.uni-leipzig.de/en/institut-fuer-kommunikations-und-medienwissenschaft/professuren/chair-of-media-and-communication/big-data-discourses-communicating-deliberating-and-imagining-datafication

Approaching datafication through discourse means to understand and to engage with the eminent reality-making power of communication, deliberation, and imagination. It foregrounds the work that goes into rendering datafication a socially relevant phenomenon and problem.

The conference sets out from the idea that the public understanding of datafication is driven by discourses in the media and among policymakers and the imaginaries they evoke. The event invites us to look at what datafication is or should be for a variety of publics and speakers and how they discuss, criticize, or envision the collection and use of data at different places, speaking from different situations, and at different times. That way, the conference does not merely interrogate the status quo of big data analytics. Rather, discourses also involve prospective ambitions and normative stances about potential, desirable, or unwanted innovations. The conference turns its attention to discourses whose programs of thought actively shape the social constitution of Big Data and translate into practices, organizational forms, policies, and institutions. Discourses are in fact integral to how we come to engage with datafication. Inquiring into the semantics, interpretations, and cultural values that prelude, accompany, and surround investments and innovations into Big Data requires by definition interdisciplinary work. This includes, among others, critical data studies, STS, sociology, communication, linguistics, political science, cultural studies, geography and education, as well as security studies and gender studies.

By taking the understanding of datafication as a matter of contingent articulation, the conference helps to dismantle claims about the given and irrevocable facticity of data formats and data analytics so as to explore ways of reimagining their status and implications. In doing so, it seeks to gain leverage in critically examining how datafication's social imaginations are shaped and to enable alternative readings.

AGENDA

Thursday, February 1, 2024

08:30–09:00 am	Welcome, Registration and Coffee
09:00–10.15 am	Keynote **Minna Ruckenstein (University of Helsinki, Finland)** *Algorithmic Folklore at the Edge of the Future*
10:15–10:25 am	Coffee Break
10:30 am –12:00 pm	**I: Discourses of Datafication and AI (Chair: Charlotte Knorr)** • Laura Liebig (University of Bremen, Germany) *Chasing the hype: AI Discourse in German Policy and Media* • Michael Dokyum Kim (State University of New York at New Paltz, USA) *Big Data Discourse in International Development* • Holger Pötzsch, Zoheb Mashiur, Melania Borit (UiT The Arctic University of Norway, Norway) *Between 'Stochastic Parrots' and 'Skynet': Data Imaginaries in Popular Cultural Expressions on AI and Autonomous Technologies*
12:15–12:55 pm	Lunch
01:00–02:30 pm	**II: Contexts of Datafication (Chair: Christian Pentzold)** • Alison Powell (London School of Economics and Political Science, UK) *Deceptive Stories about scale: Data-based health startups and the consequence of technology-driven innovation discourse* • Sebastian Pranz (Darmstadt University of Applied Sciences, Germany) *"It still tries to kill me, but less often." How Tesla Drivers Imagine Automation* • Nadja Schaetz, Laura Laugwitz, Juliane A. Lischka (University of Hamburg, Germany) *Domestication of Data in Journalism* • Göran Bolin (Södertörn University, Sweden), Rita Figueiras (Universidade Catolica Portuguesa, Portugal), Veronika Kalmus (University of Tartu, Estonia) *Towards a Datafied Mindset: Conceptualising Digital Dynamics and Analogue Resilience*
02:30–02:55 pm	Coffee Break

3:00–04:30 pm	**III: Criticizing Datafication (Chair: Emma Hughes)**
	• Lisa Reutter (Copenhagen University, Denmark)
	  No, Data is (not) the new oil? Fueling post-extractivist anxieties
	• Dayei Oh (University of Helsinki, Finland), John Dowey (Loughborough University, UK)
	  Does content regulation AI promote democratic discourse? Feminist critique of toxic language AI
	• Ewa Dąbrowska (Freie Universität Berlin, Germany)
	  Datafication and data exploitation as a means of economic development. Critical data discourses in India
	• Mengqi Li Sam (Birmingham City University, UK), Yuxuan Zhang Frida (BNU–HKBU United International College, China)
	  ChatGPT's China: A Cultural Critique of ChatGPT's Representations of China

04:30–04:40 pm	Short Break

04:45–06:00 pm	**IV: Participation in Datafication (Chair: Christian Pentzold)**
	• Judith Fassbender (Alexander von Humboldt Institut für Internet und Gesellschaft, Germany)
	  Reviewing Participatory Dimensions in Data-Centric Projects
	• Tuukka Lehtiniemi (University of Helsinki, Finland)
	  What if it doesn´t go wrong when prisoners train AI
	• Felix V. Münch, Gregor Wiedemann, Jan Philipp Rau, Phillip Kessling, Jan-Hinrik Schmidt (Leibniz-Institute for Media Research - Hans-Bredow-Institut, Germany)
	  Preconditions, Challenges, and Solutions for Enabling Large Scale Online Discourse Observation

07:00 pm	Dinner at Auerbachs Keller Mädler Passage Grimmaische Strasse 2-4 D-04109 Leipzig, Germany Google-Maps: https://maps.app.goo.gl/qT9KNiyyFmPN3GkZA"

Friday, February 2, 2024

08:00 am	Welcome and Coffee

08:30–10:00 am	**V: Datafication and Public Services (Chair: Christian Pentzold)**

- Maris Männiste (Södertörn University, Sweden)
  Transforming interaction between the state and the citizen - the case of ""Siri of the public sector services"
- Amela Muratspahić (Södertörn University, Sweden)
  From Decisions to Data – Datafication and Automation in the Swedish Public Services
- Jannie Møller-Hartley (Roskilde University, Denmark), Anne Kaun (Södertörn University, Sweden)
  Automation as Mediation: Making sense of messy entanglements in AI-driven ADM
- Astrid Mager, Doris Allhutter (Australian Academy of Services, Australia)
  Infrastructures of welfare. Narratives and counter-narratives of data infrastructures in the context of public health insurance and open commons

10:00–10:25 am	Coffee Break

10:30–12:00 am	**VI: Datafication and Control (Chair: Charlotte Knorr)**

- Veronika Nagy (Utrecht University, The Netherlands)
  Outsourcing security intelligence: the risks of digital litter in migration control practices
- Jasper van der Kist, Silvan Pollozek (European University Viadrina Frankfurt (Oder), Germany)
  Datafication Discourses in Migration and Asylum Governance in Germany
- Irina Zakharova (Leibniz Universität Hannover, Germany)
  Datafied Welfare: Discourses about Regulation and Care
- Mateusz Trochymiak, Robert E. Zajonc (Warsaw University, Poland), Sebastian Sosnowski (Polish Academy of Science (IFIS PAN), Poland)
  False promises of a paradigm shift in welfare administration. Deconstructing digital transformation of care service policy in Warsaw, Poland

12:00–01:00 pm	Lunch

Book of Abstracts

Thursday, February 1, 2024

09:00–10.15 am	Keynote **Minna Ruckenstein (University of Helsinki, Finland)** *Algorithmic Folklore at the Edge of the Future*

Keynote: Algorithmic Folklore at the Edge of the Future

Minna Ruckenstein (University of Helsinki, Finland)

This talk mobilizes the notion of algorithmic folklore to delve into vernacular culture, encompassing beliefs and stories surrounding contemporary media technologies. While algorithmic folklore provides a lens to study the cultural life of algorithms, addressing historically rooted structures of thought and feeling, imaginaries, practices, and anticipations, it aids in resisting the urge to reduce people's lives to mere landing zones for technologies. By introducing instability and complexity into discussions that often lean towards utopian and dystopian ends, algorithmic folklore opens up possibilities to rethink technological and political-economic aspects of algorithmic systems and to trace alternative collective futures and imaginaries. Attending to these alternatives allows for the consideration of concerns that are currently neglected, and facilitates the bridging of elements that may seem unbridgeable.

10:30–12:00 am	**I: Discourses of Datafication and AI (Chair: Charlotte Knorr)** • Laura Liebig (University of Bremen, Germany) *Chasing the hype: AI Discourse in German Policy and Media* • Michael Dokyum Kim (State University of New York at New Paltz, USA) *Big Data Discourse in International Development* • Holger Pötzsch, Zoheb Mashiur, Melania Borit (UiT The Arctic University of Norway, Norway) *Between 'Stochastic Parrots' and 'Skynet': Data Imaginaries in Popular Cultural Expressions on AI and Autonomous Technologies*

Chasing the hype: AI Discourse in German Policy and Media

Laura Liebig (University of Bremen, Germany)

The recent hype surrounding Artificial Intelligence (AI) systems conceals the long history of technological advancements categorized under this umbrella term, including the low points, often referred to as "AI winters" (Cardon, 2018). The emergence of a new era for AI systems applied across various sectors like mobility, healthcare, and public administration triggered the need for political attention. National governments began formulating strategies and investing in research and development to enhance productivity and ensure prosperity, prompting what has been termed a "new space race" (Ulnicane, 2021). In German media, AI technologies are portrayed in an exceedingly positive light, promising economic growth and prosperity (Fischer & Puschmann, 2021). Simultaneously, political and media entities are actively disseminating ethical and legal considerations, including human-centeredness and discrimination (Cools et al., 2022; Ossewaarde & Gülenç, 2020). Despite the extensive engagement with national AI strategies and their media coverage, the interplay between these two spheres concerning AI technologies remains largely unexplored. This encompasses not only legacy media coverage and policy discourse but also communication about this topic on social media platforms. In light of these

considerations, this project investigates the discursive structures of the AI debate using an innovative approach that combines agenda-setting (McCombs & Shaw, 1972) and the concept of discursive power (Jungherr et al., 2019). To accomplish this, the study leverages a combination of data sources. It involves the analysis of journalistic articles from leading German legacy media outlets and policy documents from the German government. Additionally, tweets from German journalists and politicians are collected to uncover the discursive dynamics between these two spheres. Given the increasing attention and engagement with AI in recent years, data is gathered from 2012 to 2021. Methodologically, the media, policy, and social media texts employs are analysed through deductive content analysis and dynamic keyword-assisted topic modelling (keyATM; Eshima et al., 2021) to map agenda-setting dynamics among sub-issues, channels, and actors contributing to the discourse on AI. Preliminary findings reveal a distinct hype cycle between 2018 and 2020. In the policy sphere, the development of the German AI strategy led to a surge in policy documents from German ministries responsible for executing the strategy. Consequently, German politicians and political parties began tweeting about AI, primarily from 2017 onwards. Notably, the focus of legacy media coverage regarding AI underwent a transformation. Before 2015, media discourse predominantly revolved around research-related topics, reflecting the exploratory phase of AI in public discourse. However, around 2015 and 2016, a paradigm shift occurred, with media coverage increasingly emphasizing economic and political matters. Importantly, ethical and societal considerations remained a consistent theme throughout this timeline, underscoring their enduring importance in AI development and deployment in legacy media articles. In summary, the strategic political discourse surrounding the issue led to a shift in the media's emphasis toward economic aspects. However, it is important to note that the media occupied a pioneering role in addressing research findings and ethical considerations associated with these technologies from an early stage.

Big Data Discourse in International Development

Michael Dokyum Kim (State University of New York at New Paltz, USA)

References to Big Data are ubiquitous, no less so in the context of "Big Data for development" (BD4D). BD4D is invested in finding ways to transform the abundant digital data into tailored information and actionable knowledge for strategic interventions in development (UN Global Pulse, 2012, p. 6). As the narratives of Big Data both ominously and promisingly suggest Big Data as "the future of humankind" (Couldry & Yu, 2018, p. 4473), the fervour of BD4D invites diverse actors to join the common mission of transforming socioeconomic conditions of humanity. Critical scholarship on Information Communication Technology for Development (ICT4D) and critical data studies have underlined significant insights for digital development practices, such as the critique of technological determinism (Klein & Unwin, 2009; Servaes, 2014), the modernization bias (Mansell, 2014), data colonialism (Couldry & Mejias, 2019), and decoupling of human agency from our 'data doubles' (Gurumurthy & Chami, 2019; Kitchin, 2014), and called for a more holistic approach with "a critical awareness and sensitivity to the globalized assumptions, discourses, power structure, possible corporate interests and ideological influences behind ICT4D projects" (Zheng et al., 2018, p. 4). However, the question remains as to how and to what extent these critical lessons are reflected in the field of international development. How do the significant actors of global development articulate about integrating Big Data into development? What emerges as the prominent assumptions of Big Data discussed in the context of addressing global problems and development? To what extent are the challenges and insights of ICT4D reflected in the BD4D?

Placing these questions at the centre, drawing on literature from ICT4D and critical data studies, this study conducts a discourse analysis to investigate a corpus of documents related to BD4D (policy papers, strategy papers, annual reports, books, action plans, and informational webpages) published by major agenda-setters of international development within the six-year range, from 2015 to 20201. It deconstructs the assumptions of Big Data articulated through the discussions of its opportunities and challenges for global development. Despite years of reflections on the enthusiasm for the data revolution, the analysis reveals the persistence of the materialistic and modernistic discourses of development within which Big Data becomes uncritically analogous to the new form of

'natural resource', 'knowledge' and an 'engine' necessary for addressing social problems and sustainable development. Amidst the growing enthusiasm for Big Data's potentialities in development, I argue that treating Big Data as such prevents BD4D from critical questioning while it endorses exogenous epistemologies and a technocentric development model. The prominent BD4D discourse, thus, depoliticizes the emerging ideology of dataism (Van Dijck, 2014) as well as the social relations of data colonialism (Couldry & Mejias, 2019), promoting continuous, or even irresistible, exploitation of human agency under the neoliberal logic of technology-and-development complex. Constructive deconstruction of BD4D discourse advances us to think beyond the generally asked question of what Big Data can 'do good for' development, toward questioning how development should also 'do better with' Big Data.

Between 'Stochastic Parrots' and 'Skynet': Data Imaginaries in Popular Cultural Expressions on AI and Autonomous Technologies

Holger Pötzsch, Zoheb Mashiur, Melania Borit (UiT The Arctic University of Norway, Norway)

Algorithmic recommendation engines, artificial intelligence, large language models, self-driving vehicles, and other emerging digital technologies all rely on huge amounts of data. The methods of acquisition, appropriation, and commodification of this data necessary to drive these new technologies have been studied and critiqued for quite a while in both academic and public discourses (Eubanks 2018, Zuboff 2019, Broussard 2019, Sadowski 2020, Crawford 2021, Cave & Dihal 2023). Recently, the potentials and dangers of massive data-use in chatbots such as ChatGPT have attracted particular attention (Bender et al. 2022, Depounti et al. 2023, Williamson et al. 2023).

In this paper, we employ approaches in critical future studies (Goode & Godhe 2017, Joshi 2023) and a sociology of imagined futures (Suckert 2022) to address data imaginaries in popular film and television. Both fields are vested in the idea that technological development and operationalization always happen in context, i.e., how a certain technology is understood by people predisposes how it will operate and intervene in societal and political processes regardless of its intrinsic qualities. In this perspective, the functioning of technology emerges as contingent upon societal and other conditions including such located at the level of imaginaries. How we think of and imagine new technologies is an important factor in implementation and its possible effects. And, inevitably, there is always a multitude of possible ways forward (Joshi 2023). To capture dynamics between technology, culture, societies, and individuals we will draw upon insights from cultural studies (Hall 1977, 1997) to understand the socio-political function of cultural expressions and adapt Jasanoff and Kim's (2009) notion of socio-technical imaginaries to a critical analysis of a series of widely received films and series taking up themes of artificial agents, artificial intelligence, and human-machine relations and inquiring about processes of datafication underlying the depicted technological advances. In doing so, we loosen Jasanoff and Kim's concept from the field of tech-development and tech-utopian futures and reconnect it to its conceptual roots in the work of Cornelius Castoriadis (1995) and his idea of an imaginary institution of society with its peculiar focus on the constitutive potentials and effects of human creativity and imagination. With critical future studies as a conceptual framework (Goode & Godhe 2017, Suckert 2022), we outline possible implications of the identified imaginaries including such factors that are regularly excluded from the filmic narratives. We conduct a category-driven qualitative content analysis of a series of movies and series featuring artificial intelligence, androids, or cyborgs as prominent protagonists and inquire into how underlying processes of data capture, appropriation, and commodification are presented. We select commercially successful as well as culturally valued examples based on box-office numbers as well as reviews and ratings on widely acknowledged channels. Applied categories are type of datafication at play, type of threat/conflict regarding collection and treatment of data, type of solution, type of relation between data, humans and machines, and more. The full category system will be developed in correspondence with the evolving analyses of the selected titles. After assessing what we term a mainstream imaginary of datafication in AI-based technologies, we contrast this with short analyses of titles breaking with the emerging conventions. Through our

analysis we systematize dominant tropes and motifs before contrasting these to counter-imaginaries and connecting them back to contemporary debates about datafication and data-use in artificial intelligence and autonomous technologies. This way, we increase understanding of how social imaginaries conveyed in a cultural sphere can contribute to negotiations of new technologies in socio-political contexts. Films and series included range from titles such as 2001: A Space Odyssey (1968), Westworld (1973), and the Terminator-series (1984-) via children's movies such as Wall-E (2008), Next Gen (2018), and The Mitchells vs the Machines (2021) to works such as Her (2013), Ex Machina (2014), and Ghost in the Shell (1995/2017).

01:00–02:30 pm	**II: Contexts of Datafication (Chair: Christian Pentzold)**

- Alison Powell (London School of Economics and Political Science, UK)
  Deceptive Stories about scale: Data-based health startups and the consequence of technology-driven innovation discourse
- Sebastian Pranz (Darmstadt University of Applied Sciences, Germany)
  "It still tries to kill me, but less often." How Tesla Drivers Imagine Automation
- Nadja Schaetz, Laura Laugwitz, Juliane A. Lischka (University of Hamburg, Germany)
  Domestication of Data in Journalism
- Göran Bolin (Södertörn University, Sweden), Rita Figueiras (Universidade Catolica Portuguesa, Portugal), Veronika Kalmus (University of Tartu, Estonia)
  Towards a Datafied Mindset: Conceptualising Digital Dynamics and Analogue Resilience

---

Deceptive Stories about scale: Data-based health startups and the consequence of technology-driven innovation discourse

Alison Powell (London School of Economics and Political Science, UK)

On September 5, 2023, Babylon Health, the 'unicorn' data-driven primary health-care app, filed for bankruptcy. When it launched in the UK in 2013, the company was celebrated as a pioneer of digital health provision. One of the main promises behind the company was that it could make the provision of public, free-at-point-of-use health care more efficient (and therefore cheaper); to do this, it could draw on the UK's 'resource' of population-scale health data. The founder of Babylon Health framed his intentions as being parallel to those held by other data-driven technologies - "to do with health care what Google did with information" (Browne, 2023). Between 2018 and 2022, the company developed two different AI-powered triaging tools and established partnerships with several UK hospital trusts. These technical tools were intended to be part of a rapid 'blitz-scale' expansion of the business. Yet a year later the company was insolvent and the partnerships canceled. The fate of Babylon illustrates risks inherent to the perceived value of 'scaling up' technology-driven projects and companies, particularly those assumed to be of public benefit. The discourse of 'scaling' creates deceptive distortions. These distortions include the appropriation of ideological features of management associated with technical innovation into other settings, including management of people and institutions (Irani, 2015). The transposition of discourses of 'scale' or 'scaling' is problematic in two ways: first, because many technology-driven projects aimed at social amelioration do not, in fact, scale up (Heeks 2008; Foster and Heeks 2013) and second, because a focus on scaling occludes other normative concerns including equity, fairness and justice. These are pressing concerns in health care as well as in other applications of datafied public services. This paper analyzes features and consequences of the discourse of 'scaling' in relation to data systems, using the example of Babylon Health. It identifies how discourses of scaling draw on interpretations of software functionality, and in particular on the idea that software uniquely facilitates

scaling and increased efficiency of organizations and practices. This dismisses existing or emerging organizational complexity as well as long-term, emergent or ethically complex normative considerations. In the case of Babylon Health, these considerations include the ethical value of 'efficiency' in health care delivery, which might exacerbate or create injustices which are unable to be addressed through a consideration of scale.

The paper employs a methodological strategy employing 'inter-scalar vehicles' as defined by anthropologist Gabrielle Hecht. Hecht describes how inter-scalar vehicles can connect stories and scales that are usually set apart. This method can be adapted to examine discursive gaps in the understanding of 'scale' created organizational mismatches between promised technological interventions and the complex interscalar institutions of the NHS. This approach makes it possible to address the scalar claims of powerful actors, for whom scale is a mechanism for measurement, organization and control, and to investigate the potential for inter-scalar approaches for data policy.

"It still tries to kill me, but less often." How Tesla Drivers Imagine Automation

Sebastian Pranz (Darmstadt University of Applied Sciences, Germany)

The car manufacturer Tesla has paved the way in building electric cars for the mainstream market and creating the needed infrastructure of batteries, chargers, solar panels, and stationary battery systems. In addition, in 2013, its CEO Elon Musk announced that Tesla would introduce a system of autonomous driving cars allowing to take 90% of the control away from the driver within the next three years (Carroll, 2013). Although the fully autonomous Tesla is still a dream of the future, the company has invested heavily in self driving technology during the last decade, starting with the launch of the so-called Autopilot in October 2014. The intelligent assistant for lane control was eventually updated to what the company calls 'Full Self Driving' (FSD), which should extend autonomous driving beyond the motorway, requiring the driver to take over only in certain situations. Although the technology has considerable flaws and resulted in several fatal accidents (Samineh Gillmore & Tenhundfeld, 2020), it has attracted a vital community of drivers, acting as beta-testers for new software releases, sharing their expertise, and offering deep insights into the feelings and imaginaries of what it means to pioneer automated driving.

This contribution draws on a corpus of 17.530 comments retrieved from the subreddit r/teslamotors, where drivers share their everyday experiences and imaginaries navigating the black box of Tesla's driving assistance. A twofold research design is developed to explore the data. First, a topic analysis (Angelov, 2020; Blei et al., 2003) is conducted to identify the discursive structures and key notions within the corpus. On this basis, a qualitative content analysis following Kuckartz (2014) is performed. Against its empirical background, this contribution presents a relational notion of automation, informed by Actor-Network Theory (Dant, 2004; Latour, 2012; Laurier & Dant, 2012; Tennant et al., 2021) and critical algorithm studies (Forelle, 2022; Lee et al., 2019; Morley, 2019). It is argued that drivers perform automation by immersing themselves in a socio-technical figuration of car, code, and data, building a mental model of the car's behavior – an embodiment of its data scapes, helping them anticipate their car's possible decisions. However, this model is challenged by both frequent updates affecting the car's nature and unexpected occurrences ranging from little 'quirks' to dangerous situations and near misses. While most drivers identify with Tesla as an innovator in driving and follow Elon Musk's narrative that possible flaws will eventually lead to a safer car, some struggle with phenomena such as phantom braking or reported crashes on autopilot. To cope with a car that is perceived as 99% safe and 1% deadly, as one Redditor puts it, it is argued that drivers have to develop an imaginary integrating this unpredictability into their driving experience. The friction between technical mastery and autonomous technology (Winner, 1977) is described as 'wildness', referring to the joy of using autonomous technology that is not entirely tamed. Accordingly, drivers perceive themself as trainers and tamers reporting bugs and critical situations to the developers and taking their car to specific challenges after an update to test if its performance has improved: "It still tries to kill me, but less often."

Domestication of Data in Journalism

Nadja Schaetz, Laura Laugwitz, Juliane A. Lischka (University of Hamburg, Germany)

Journalism is a comparatively data-rich field (Sørensen and Kosta 2019). In this field, data has been subject to domestication, offering journalists the means to understand audiences, optimize editorial workflows, and craft personalized recommendations for users. Narratives of domestication illuminate a process of making data, once considered foreign or unfamiliar, more accessible, controlled, and integrated (Silverstone and Hirsch 1992). This reflects a broader trend where the assimilation of data is narrated as a journey of adaptation and integration into established practices—a transformation essential for harnessing data's full potential.

The relationship between journalism and data, however, is not without its complexities. A critical lens warns against the chilling effects of incessant data collection on users. Addressing data-intense news recommender systems, Helberger (2019) warns that continuous data collection may impede critical and diverse thinking among users, which underscores the balance journalism must strike in the pursuit of data-driven insights. In a critical view, data infrastructures are seen as exploiting users, misrepresenting their lifeworld, and causing merely a "data fetish" instead of providing meaningful insights (Schwarzkopf 2020; Zuboff 2019). Contrastingly, a business-intelligence perspective perceives data as a valuable resource supporting the creation of value and informed decision-making (Phillips-Wren, Daly, and Burstein 2021). This pragmatic approach acknowledges the transformative power of data in shaping the business of journalism. Embracing not only the cautionary tales of data exploitation but also the potential for empowerment, datafication has the capacity to foster collaboration and create a more interconnected society (Micheli et al. 2020; Couldry and Powell 2014). This view opens avenues for journalists to engage with their audiences more effectively, tailoring content to meet specific information needs, and thus, enhancing the overall quality of journalistic output.

We examine the nuanced nature of how these views are integrated into domestication narratives by analysing job advertisements and interviewing data professionals in news organizations. The analysis of n = 69 job advertisements for data analysts in resource-rich news organizations in the US and UK reflect a spectrum from a critical interrogation of data bias to a pragmatic embrace of data as a business intelligence tool. Data has epistemological character that is structured by architectures of materialities, reinforced through social interlinkages with a broader network of data agents, and institutionalized as paradigmatic setting that produce knowledge. Expert interviews with n = 66 data professionals from diverse news organizations across the US, UK, Germany, and Kenya reinforce this diversity of perspectives. The discourse encompasses critical reflections on data ethics, practical considerations of data as a business asset, and optimistic visions of data's role in empowering journalists and underrepresented audiences, fostering societal connections. In conclusion, the domestication of data in journalism is a narrative that integrates data-critical, business-intelligence, and partly audience-empowering narratives—each with its own set of promises. The future of journalism also lies its responsible navigation of its data-rich terrain and discussing implications of big data for journalism and society.

Towards a Datafied Mindset: Conceptualising Digital Dynamics and Analogue Resilience

Göran Bolin (Södertörn University, Sweden)
Rita Figueiras (Universidade Catolica Portuguesa, Portugal)
Veronika Kalmus (University of Tartu, Estonia)

People born around the same time, and influenced by the social, cultural, and technological context in which they were socialised, tend to share distinctive practices and mentalities. Supposedly, the rapid transformations of society that follows from digitalisation and big data management produces different responses among generational cohorts. In previous debates on the implications of datafication, it has been suggested that the mindsets of media users could progress to a quantified reasoning, leading to a "big data mindset" or a "metricated mindset". Such a

mindset could be described by evaluating social actions and relations according to the logics of datafied media. Indications of this can be found in small-scale experimental research, but these studies need to be further complemented and validated.

This paper explores the ways in which what we call the "analogue and the datafied mindsets" perceive the functioning of the datafied world. Based on a qualitative interview study of two generations of media users – one younger and one older in Estonia, Portugal, and Sweden –, we present and analyse underlying patterns in participants' analogue and digital media practices and related attitudes. We show that belonging to a media generation does not always produce a homogeneous mindset or a uniform attitude towards media technologies. The mindsets, being ideal-typical constructs, are not bound to individuals: the same person can display features of the analogue and the datafied mindset in relation to different parts of the digital world. One mindset does not replace the other, but rather adds another layer to the social action of the individuals. The mindsets are multi-dimensional and moulded by contrasting understandings and negotiations, indicating that the tenacious structures of the analogue world linger on in the datafied social space.

03:00–04:30 pm	**III: Criticizing Datafication (Chair: N.N.)**
	• Lisa Reutter (Copenhagen University, Denmark) *No, Data is (not) the new oil? Fueling post-extractivist anxieties*
	• Dayei Oh (University of Helsinki, Finland), John Dowey (Loughborough University, UK) *Does content regulation AI promote democratic discourse? Feminist critique of toxic language AI*
	• Ewa Dąbrowska (Freie Universität Berlin, Germany) *Datafication and data exploitation as a means of economic development. Critical data discourses in India*
	• Mengqi Li Sam (Birmingham City University, UK), Yuxuan Zhang Frida (BNU–HKBU United International College, China) *ChatGPT's China: A Cultural Critique of ChatGPT's Representations of China*

No, Data is (not) the new oil? Fuelling post-extractivist anxieties

Lisa Reutter (Copenhagen University, Denmark)

"Data is the new oil of the digital economy […] Like oil, for those who see Data's fundamental value and learn to extract and use it, there will be huge rewards." proclaimed Wire magazine in a 2014 article, reinforcing an idea central to recent developments in the public and private sector across the globe. Data is regarded as a driver of societal change and described as a raw material in need of refinement to fuel key transformative technologies such as artificial intelligence and data platforms. This metaphor has sparked wide discussions and been challenged both within and outside the technological sphere. By regarding data as oil, it is naturalized as a commodity in need of extraction that is hard to control due to its fluidity (Couldry and Yu 2018, Taffel 2023). Untangling the relation between power, ideology, data and technology, scholars associated within critical data and algorithm studies have continuously argued against an understanding of data as a raw material (Gitleman 2013, Puschmann and Burgess 2014). In addition to scholarly critique, the notion of oil itself draws increasingly negative connotations in the wake of the climate crisis. Following up on their 2014 proclamations, even Wire magazine argued in a 2019 article that "No, Data Is Not the New Oil" echoing several voices in the tech industry. Does that indicate that data has already lost its importance? Despite the recent pushback on ideas of data as a raw material or data as the new oil, the basic assumption these metaphors draw upon persist heavily in society today: Data is regarded as a key resource

for value creation in an increasingly datafied society. A resource that is in need of extraction and processing to advance and improve society according to policy makers, the tech industry and economist alike. Data mining, extraction and refinement is blooming, despite actors distancing themselves discursively from 'data as the new oil'. In this essay, I explore the pervasive power of extractivist metaphors and their connection to anxieties of a fossil free future. Why are the ideas of data gold mines or data as the new oil so powerful? Is it possible to imagine a world without extractivism? Reflecting upon these questions helps us uncover various tensions: Between materiality and immateriality, commercialization and de-commercialization, economic value and public value, between data markets and data governance.

Does content regulation AI promote democratic discourse? Feminist critique of toxic language AI

Dayei Oh (University of Helsinki, Finland)
John Dowey (Loughborough University, UK)

As social media becomes increasingly entwined with everyday politics, concerns about online harm and abuse are growing (Gagliardone, 2019). These issues threaten the internet's inclusivity and safety for all users. To address this, many social media platforms employ machine learning artificial intelligence (AI) to monitor and regulate what they deem as 'toxic language.' Google Counter Abuse team and Jigsaw developed the Perspective Application Programming Interface (API) to identify and control such language, defined as 'rude, disrespectful, or unreasonable comments likely to drive someone away from a discussion' (Dixon et al., 2018: 68). Critiques of AI censorship have highlighted its technical limitations that AI censorship is prone to biases and misclassifications (Thiago et al., 2020; Zhou, 2021). AI communities are researching ways to mitigate such biases by experimenting with new ways to deal with these problems imagined as 'technical glitches,' minor, incidental errors that can be patched up through machinery itself. Efforts to mitigate these issues include adjusting training and testing datasets, diversifying coder backgrounds, and using additional mathematical methods (Zhou, 2021). However, the fundamental problems with AI censorship lie deeper than the incidental technical 'glitches' that can be easily identified and solved with easy fixes. The problems are embedded in the very logic of content regulation that certain forms of language are 'toxic' and must be censored to promote democratic and productive discourse. The challenge arises from the inconsistent and loose definition of 'toxic language' within tech communities. Perspective API, for instance, lumps together various socially undesirable behaviours, from rudeness to hatefulness, without distinguishing their nuances. This broad definition results in misclassifications, failing to differentiate between incivility and intolerance and their impacts on democratic discourse.

This paper offers a feminist critique of AI content regulation, examining it through normative theories of inclusive and progressive public spheres and empirical evidence. The paper challenges the political liberalist (e.g., Rawls, 1996) perspective of civility embedded in AI's logic for regulating toxic language, advocating against regulating incivility. Instead, it distinguishes incivility from intolerance, acknowledging their differing implications for democratic discourse. The paper recognizes the potential expressive and symbolic value of incivility, particularly for marginalized groups (Edyvane, 2020; Elias, 2000; Zerilli, 2014), and argues for the regulation of intolerance and hate speech that perpetuates participatory inequalities (Gelber, 2021). It stresses the importance of identifying various forms of hate speech conveying intolerant and exclusionary ideas, even without explicit hateful or emotional language.

The paper then empirically assesses the effectiveness of toxic language AI in regulating incivility and intolerance on a large Twitter dataset. The findings reveal that current content regulation AI does not promote democratic discourse but hinders it by silencing uncivil yet pro-democratic voices of marginalized individuals and failing to detect intolerant messages with nuanced rhetoric. The paper calls for the development of new content regulation AI aligned with democratic theories informed by anti-racist, feminist, and other critical perspectives.

Datafication and data exploitation as a means of economic development. Critical data discourses in India

Ewa Dąbrowska (Freie Universität Berlin, Germany)

In India, the government has strategically used the discourse on data colonialism to justify data protection and localization bills. Data colonialism was conflated with colonialism tout court to show how Big Tech set out to exploit data of Indians. The new regulation sought to privilege local business actors vis-à-vis the Big Tech and put few restrictions on data surveillance by the state. In general, datafication and data exploitation are parts of a vision of economic development promoted by India domestically and at G20. Due to the powerfulness of this vision, parts of the Indian civil society reacted remarkably uncritically towards the data regulation strategy chosen by the Indian government. When data is the new oil, the strategy to rely on data for development in an immensely populous country such as India promises to deliver modernization of an unknown dimension. Other civil society organizations pointed to the dangers of state surveillance and did not buy into the "data democracy" and "data-based development" visions promoted by the Indian government. The proposed paper examines datafication and data exploitation discourses in Indian civil society and their nexus with development. It uses the argumentative discourse analysis from deliberative policy studies by Hajer (1993, 1995) and Hajer, Wagenaar (2003) to determine the matrix of discursive positions and to show how "discourse coalitions" play with and move between them. While data governance, regulation, and localization efforts of Indian authorities received attention in the literature (Thomas 2019; Seoane, Facundo 2021; Prasad 2022), the "datafication as development" discourse and its critiques were not yet explicitly examined.

ChatGPT's China: A Cultural Critique of ChatGPT's Representations of China

Mengqi Li Sam (Birmingham City University, UK)
Yuxuan Zhang Frida (BNU–HKBU United International College, China)

This research aims to explore the cultural impact of AI-generated content, using ChatGPT as an example. Drawing upon the concepts of Orientalism (Said, 1978) and Sinological Orientalism (Vukovich, 2012), this study argues that AI-generated content perpetuates and reinforces a Western-centric perspective towards China. Previous research extensively scrutinises biases in large language models (LLMs) in topics about politics (Hartmann et al., 2023; Blodgett et al., 2021; Weidinger et al., 2021) and gender (Lucy and Bamman, 2021) from the social and ethical perspectives. Scant attention has been given to its cultural implications. This study addresses this gap by employing Orientalism and Sinological Orientalism as a theoretical framework. It aims to deconstruct and reconstruct AI-generated content related to China, examining it within the power dynamics between China and the Western world. Orientalism and Sinological Orientalism contend that the Western perspective often portrays other cultures solely from a Western standpoint, reinforcing a Western-centric universalist ideology. In pursuit of this objective, the study employs Fairclough's (1989, 1995) three-step critical discourse analysis, intending to explain and situate the ChatGPT-generated content within a comprehensive socio-cultural framework.

At this preliminary stage, our analysis suggests that ChatGPT's content reflects a combined Orientalist and Sinological Orientalist perspective. For instance, when discussing factors influencing economic development in China and America, ChatGPT provides an initially objective and comprehensive analysis. However, when exploring whether China's approach or America's approach can be replicated by other nations, ChatGPT outlines more potential challenges in adapting China's approach compared to America's. Notably, ChatGPT's "inspiration" from China highlights aspects such as infrastructure development, technological innovation, and poverty reduction while neglecting the significance of China's political, social policies, and international stance, which ChatGPT identifies as contributing factors to China's economic growth. This selective emphasis on China's experience perpetuates the universalist ideology that the capitalist experience is more applicable than "Chinese

characteristics." Therefore, we argue that these generated contents have the potential to reinforce a Western perspective, possibly challenging the cultural identity and perception of other cultures. This echoes and further develops the famous question raised by Spivak (1985): Can the subaltern speak in the digital age? Considering Vukovich's research (2012), it is essential to recognise that the power relations between China and the Western world differ from those with other nations due to China's substantial power. Thus, while China is not subaltern in the same way as some other countries, it is still depicted and influenced by Western representations, as evident in ChatGPT's responses. From a different perspective, previous studies have contributed to amplifying the voices of marginalised groups, highlighting the issue of the "under-sampled majority" (Raji, 2020) and how systematic unequal power structures within culture are "captured by data, learned by LMs, and perpetuated by their predictions" (Hampton, 2021).

04:45–06:00 pm	**IV: Participation in Datafication (Chair: N.N.)**
	• Judith Fassbender (Alexander von Humboldt Institut für Internet und Gesellschaft, Germany) R*eviewing Participatory Dimensions in Data-Centric Projects*
	• Tuukka Lehtiniemi (University of Helsinki, Finland) W*hat if it doesn´t go wrong when prisoners train AI*
	• Felix V. Münch, Gregor Wiedemann, Jan Philipp Rau, Phillip Kessling, Jan-Hinrik Schmidt (Leibniz-Institute for Media Research - Hans-Bredow-Institut, Germany) *Preconditions, Challenges, and Solutions for Enabling Large Scale Online Discourse Observation*

Reviewing Participatory Dimensions in Data-Centric Projects

Judith Fassbender (Alexander von Humboldt Institut für Internet und Gesellschaft, Germany)

Participation regarding data governance is a topic which has been increasingly gaining attention (Ada Lovelace Institute, n.d.; Hintz et al., 2022; Mozilla Common Voice, n.d.). In conceptual perspectives, democratic ideals are often explicitly or implicitly connected to participation. At the same time, it is critiqued that participatory approaches in data-related practices are an insufficient approach to solve the problems they are meant to (cf. Sloane et al., 2020). Whilst participatory practices have been realised in data-centric projects in the past, participation was a less prominent feature and presumably implemented for a variety of reasons besides democratisation. To be able to differentiate those cases which include outside parties, I investigated: What approaches to data handling are discussed in the academic literature which include participatory elements?

Via a systematic literature review, following Kitchenham & Charters (2007) I sourced 31 cases from database requests from Scopus, ACM Digital Library and IEEE Xplore. Using the Participatory Science Cube by Schrögel & Kolleck (2018) – an adaptation of the Democracy Cube by Archon Fung (2006) – the cases were mapped along three dimensions: The reach dimension describes who participated, the normative dimension describes the level of participation in rulemaking for the project, and the epistemic dimension describes participation in practically handling the data. Additionally, the mode of data handling; the motivation to implement a participatory process and what changed due to the process were analysed. The preliminary results show: The most prominent mode of data handling was joining existing data repositories (45.2%) followed by collecting new data (29%) and the (re)organisation of data (16.1%). Motivations in the sample can be divided into four groups: safeguarding values and enabling self-determination; increasing functionality; increasing accessibility and data quality/completeness; drawing resources. In the application of the cube two clusters of cases emerge: The first are research databases with expert participation (reach). With two exceptions, researchers are providers and users of the data(bases) and

use the data independently (epistemic). At a minimum, a small degree of say in regards to rulemaking is given to the participants (normative). The second cluster consists of citizen science projects with a wide reach of participants, where people participate in labour aspects of data processing (epistemic) – usually according to strict rules. Participants are not given a say regarding the rulemaking process (normative). The analysis suggests: 1. Ownership of research data by expert participants resulted in participation in regards to rulemaking; 2. A large number of participants often occurs in combination with labour-intensive tasks regarding data handling but no participation in rulemaking. 3. When a large number of participants has a say in rulemaking it is likely a core objective of the project. This kind of project is present but underrepresented within the sample. At this point it remains to investigate why participation in rulemaking by non-expert participants is underrepresented in the sample, especially in regards to both collective interests and interests of data subjects.

What if it doesn´t go wrong when prisoners train AI

Tuukka Lehtiniemi (University of Helsinki, Finland)

When we talk about our research examining an unconventional data labor arrangement – Finnish prisoners training AI for a tech company – the audience tends to start nodding, as if they already knew what's coming. After all, when prisoners train AI, the prison-industrial complex meets the global data extraction machinery. What could be a better showcase of digital capitalism wreaking havoc? During discussions, we are expected to confirm expectations about ongoing exploitation of the marginalized. Our audiences, of course, have much to draw on: ten years of scholarship has examined datafication's social relevance in terms of its universal tendency to reshape lives, organizations, and societies. For Zuboff (2019), surveillance capitalism and its global apparatus of data production unilaterally claim human experience as raw material for the monetization of behavioral prediction and control. For Couldry and Mejias (2019), data colonialism similarly exploits humans by appropriating social and individual lives as data. As Markham (2021) has argued, imaginaries about datafication are constrained in that current tech trajectories tend to be naturalized to the extent that they seem inevitable. Datafication, then, gets rendered societally relevant especially in terms of harms and threats, and it is difficult to consider data-driven technologies without reproducing the hegemonic trends. I will use the prisoners training AI case to discuss how expectations about universality in datafication discourses subsume political, cultural, and socioeconomic particularities. Milan and Treré (2019) point out how critical data studies scholarship tends to assume that datafication and its politics result in similar outcomes everywhere, and call for approaches that "grasp the obscure developments, the cultural richness, and the vibrant creativity emerging at the margins of the 'empire'". Arora (2020) suggests that the current pessimistic undercurrent of this scholarship is not always reflected empirically, and we need "new framings to make sense of the complex matrix and flow of humans and technology". While the politics of datafication are undoubtedly hard at work also when prisons train AI, these politics should not be treated as isolated from local specificity: regulations, socially negotiated expectations, practices, creativity, and aspiration. The Finnish prison, as I will discuss, is a place that tech companies and their projects cannot simply transform to their liking. It is a strictly structured and regulated environment with its own rules, policies, and values. AI training and data labor do not simply turn the welfare state's prisons into sweatshops for data production. In fact, instead of just verifying the exploitative features of prison data labor, the case allows examining the conditions for more ambivalent and potentially optimistic narratives and imaginaries. In the prison, AI training settles in the frame of the aspirational (Author, 2022): data work is associated with a curious form of techno-hope, a hope that digital work can help with rehabilitation, that prisoners' lives might get a little better if they work on a computer. Data-driven technologies often come with problems, but so do universalizing discourses about datafication. At worst, they make us assume that nothing can be done, that data-driven technologies simply cannot be anything different, that there are no local specifics. Ultimately, as I will discuss, lumping everything together can mean passing up the opportunity to carefully examine the obscure – here, the welfare state's prison – to help form new ideas and imaginaries about what data-driven technologies could do for us.

Preconditions, Challenges, and Solutions for Enabling Large Scale Online Discourse Observation[1]

Felix V. Münch, Gregor Wiedemann, Jan Philipp Rau, Phillip Kessling, Jan-Hinrik Schmidt (Leibniz-Institute for Media Research - Hans-Bredow-Institute, Germany)

Regardless of social media being a (distorting) mirror, an amplifying sounding board, or a root cause of social developments, it takes place in, connects, and shapes the public sphere(s) of today (cf. Bruns and Highfield 2016). Especially the polarization of people's opinions and, thus, potential impacts on political and social processes has been one of the earliest observations researched on social media platforms (cf. Yardi and Boyd 2010). Acknowledging this very special role of social media communication, the project Social Media Observatory (SMO) started its work in the summer of 2020 as a virtual infrastructure provider for the newly funded German Research Institute Social Cohesion (RISC). RISC is a decentral, networked research organization, spanning eleven institutions in Germany from a diverse set of disciplinary and methodological perspectives to study and monitor phenomena related to social cohesion. Its strong roots in the social sciences pose a challenge to the goal of developing a focus on social media research with big datasets and longterm monitoring which, as an interdisciplinary endeavour by itself, requires both computational as well as social analytics skills.

Instead of building up the required knowledge and analytical skills separately in the widely dispersed parts of RISC, it appeared reasonable to combine large efforts for this task in one unit. This contribution puts up for discussion the basic concept, ethical and legal considerations, technical implementation as well as resulting tools and data collections of the SMO, derived from the challenges faced. Since 2020, the SMO is developed as an open science research infrastructure. It focuses on (the support of) long-term monitoring of public communication on selected platforms SMO and online news media to answer social science related research questions. Based on systematically compiled lists of public speaker categories, such as parliamentarians or media organizations, it collects statistics as well as content data to study the German social media discourse in comparison to mass media. Aggregated results are published via interactive dashboards. Raw data is published as ID lists for reproduction or shared with researchers upon request. The SMO further provides various tools, curated datasets, and documented workflows, for instance, to run thematic ad-hoc data collections. As a main feature, it maintains a curated knowledge base in wiki format to enable other researchers to perform systematic social media observations on their own. In summary, the SMO as a centralized infrastructure within the RISC provides a facilitated entry point into large-scale analysis for social media research. Our answer to the challenges of (big) social media data research is to support scholars using our infrastructure services in a do-it-yourself (DIY) fashion, enabling them to build their own solutions. Based on more than three years of experience following that mission, we will share how the SMO supports typical research design decisions, data collection, and analysis steps throughout the social media research process.

---

[1] Disclaimer: This abstract is based on and sometimes follows verbatim an already published article

Friday, February 2, 2024

08:30–10:00 am	**V: Datafication and Public Services (Chair: Christian Pentzold)**

- Maris Männiste (Södertörn University, Sweden)
  Transforming interaction between the state and the cititizen - the case of ""Siri of the public sector services"
- Amela Muratspahić (Södertörn University, Sweden)
  From Decisions to Data – Datafication and Automation in the Swedish Public Services
- Jannie Møller-Hartley (Roskilde University, Denmark), Anne Kaun (Södertörn University, Sweden)
  Automation as Mediation: Making sense of messy entanglements in AI-driven ADM
- Astrid Mager, Doris Allhutter (Australian Academy of Services, Australia)
  Infrastructures of welfare. Narratives and counter-narratives of data infrastructures in the context of public health insurance and open commons

Transforming interaction between the state and te cititizen - the case of "Siri of the public sector services"

Maris Männiste (Södertörn University, Sweden)

In recent years, chatbots have emerged as a ubiquitous presence in both private organizations and public sector agencies (see, e.g., Simonsen et al., 2020; Verne et al., 2022; Mehr et al., 2017), serving as initial points of contact for users seeking assistance with frequently asked questions. In public administration, chatbots are often employed to alleviate the street-level bureaucrats' burden of answering citizens' information needs. Strategic adoption of chatbots is driven by the objective of enabling public agencies to allocate more time to more personalized and value-added services for their citizens (Sun & Medaglia, 2019). Research findings (Simonsen et al., 2020) illuminate the role chatbots play as "digital mediators" in disseminating administrative information and services. While chatbots are imagined to liberate street-level bureaucrats from time-consuming tasks, they concurrently impose certain challenges on citizens who may lack comprehensive domain knowledge, such as specific benefit terminology, or the ability to craft effective prompts (Simonsen et al., 2020; Verne et al., 2022). Such deficiencies can result in misleading information or unsuccessful interactions with public agencies. Consequently, the quality of assistance provided by chatbots in their role as representatives of public agencies is contingent on citizens' administrative literacy, as well as factors like spelling accuracy and the precise utilization of keywords. This complexity is further compounded when chatbots are designed to emulate the functions of popular voice assistants, as exemplified by the Estonian chatbot currently in development, which aspires to be the "Siri of public sector services" (Grzegorczyk, 2021). Contemplating chatbots as the initial point of contact for citizens with public agencies, the accuracy and relevance of chatbot responses hold significant implications for citizens' life situations. This, in the end also impacts the trust towards particular organisations.

The present paper draws on the Estonian case study investigating the automatization of the Estonian welfare state, particularly the implementation of the named state-wide chatbot, which is supposed to reimagine the interaction between citizen and state. This study draws upon a qualitative case study where the combination of document analysis and interviews with key stakeholders engaged in the development and implementation (product owner, developers, local government workers) of the specific chatbot are used. Initial findings suggest that the introduction of the chatbot has imposed additional time-consuming responsibilities on street-level bureaucrats, who are now tasked with curating the data upon which the chatbot relies. In addition, analysis indicates the

mismatch between the imaginaries and public discourses state officials have expressed and how it is implemented in selected public sector organisations (ministry, local government).

From Decisions to Data – Datafication and Automation in the Swedish Public Services

Amela Muratspahić (Södertörn University, Sweden)

Ten years ago, there was a surge of interest in big data, both in the press and research following a number of tech scandals (Pentzold & Knorr, 2022). However, recently the interest in big data has increasingly given way to artificial intelligence (AI) (Elish & boyd, 2018), perhaps most notably with the introduction of the generative AI chatbot "ChatGPT". Yet, the development and implementation of AI is not limited to the private sector. Increasingly, public organisations are implementing artificial intelligence and automated decision-making (ADM) in their operations (Veale & Brass, 2019; Katzenbach & Ulbricht, 2019; Larasati, Yuda & Syafa'at, 2023; SCB, 2023). In Sweden, four governmental agencies were set to investigate how the country could foster AI in the public services, highlighting AI's potential for (cost)-efficiency (Digg, 2023). The conclusion focused mostly on coordination and information-sharing between public organisations on AI initiatives, including development of a trust model for public AI and a policy lab focusing on the new AI Act from the EU (Digg, 2023). However, previous research has found (Eubanks, 2018; Kaun and Dencik, 2020; Andreassen, Kaun & Nikunen, 2021) that certain use of AI and ADM in the public sector could increase inequalities between social groups and exacerbate marginalisation of already affected groups. Thus, this PhD project aims to explore the gaps between groups, conceptualised as the macro-, meso-, and micro levels and their involvement with increased automation and artificial intelligence in the Swedish public services. The macro level consists of decision-makers, allocating resources and implementing technological systems. The meso level includes both developers and caseworkers, executing decisions from the macro level. Lastly, the micro level consists of citizens, who are exposed to the systems in their use of public services. Drawing on the conceptual levels, the theoretical framework is rooted in Bourdieu's field theory since it describes society as structured into fields with their own logics, as well as highlighting inequalities in the concepts of capital and habitus (Bourdieu & Wacquant, 1992). Furthermore, as Sterne (2003), and more recently Airoldi (2022) highlight, Bourdieusian theory, and specifically the notion of habitus, is apt for researching technologies, including AI and machine learning (Airoldi, 2022). In this presentation, I will focus on how Bourdieu can be used theoretically and methodologically to analyse the gaps and inequalities between the macro-, meso-, and micro levels in regard to the embeddedness of AI in the Swedish public services. Further, my presentation will explore how gaps between actors can be conceptualised in line with Bourdieu's triad of field, capital and habitus, e.g. how a lack of field-specific knowledge of how AI actually works can have potential consequences for agents whose working- and daily lives are embedded with AI technologies. Finally, I will end my presentation with a discussion on the broader aim of this research, namely how the democratic implications of increased datafication and automation in the (Swedish) public services can be understood, focusing and critically examining ideals such as transparency and the notion of the informed citizen (Ananny & Crawford, 2018).

Automation as Mediation: Making sense of messy entanglements in AI-driven ADM

Jannie Møller-Hartley (Roskilde University, Denmark)
Anne Kaun (Södertörn University, Sweden)

In the aftermath of the hype around big and open data, public and private institutions are increasingly experimenting and implementing automated decision-making, allowing AI driven systems to automatically analyze and make decisions based on large amounts of data on and from citizens. Definitions of AI and Automated decision making often refer to the notion of delegation, i.e. the delegation of complex, cognitive tasks to algorithmic systems. However, rather than delegation, studies increasingly turn towards an emphasize of entanglement of machines and humans, arguing that rather than replacing humans, humans will labor and live with

machines in messy entanglements (Airoldi, 2023). In this paper we ask how we can make sense of these entanglements if we conceptualize automation as communicative mediation rather than the currently dominant understanding of automation as delegation to algorithmic and AI-powered systems. For illustrations we draw empirically on two case-studies.

The first case zooms in on an automation tool to profile long-term unemployed in Sweden. Based on some 30 variables concerning individual characteristics and contextual, the job seekers are profiled and sorted into three categories that distinguish between different probabilities of re-entry into the job market. The category is also determining which kind of programmes and coaching companies will be offered to the job seeker. The pre-assessment is motivated with higher efficiency and accuracy than manual processing and assessment by case workers. The second case investigates a project in a large Danish tabloid media organisation, who embarked on a project to develop ethically sustainable NLP models and recommender systems for the automated distribution of news. Using, sorting and modelling data on content and data on users, the organizations developed and implemented algorithmic recommendation of news, and this project was equally motivated by goals of efficiency and accuracy in the distribution of news, which is otherwise manually done by editors and journalists. In both cases, we consider the implementation of an AI system as a form of mediation. For example, in the first case, the translation of job market re-entry is mediated through specific variables that count differently for the probability of acquiring work. The specific variables are forms of mediating fitness for the contemporary job market while they also transform what is considered necessary training to be successful. In the case of news organisations, it can be observed in the processes of building the algorithmic recommendation systems (both using AI and more simple algorithms) that how the discursive imaginary of what the system can do remains in the project, even when the projects fail or are changed along the way. Using these case studies as illustrative examples, we argue that automation as mediated communication highlights the productive and transformative effects in three separate, but interconnected dimensions. These are 1) Anticipation 2) Perceived affordances and 3) Adaptations and second order effects.

Infrastructures of welfare. Narratives and counter-narratives of data infrastructures in the context of public health insurance and open commons

Astrid Mager, Doris Allhutter (Australian Academy of Services, Australia)

(Semi-)Automated decision-making (ADM) systems are on the rise in public sectors including employment, public health, and education. They are often driven by values of efficiency, effectivity, or the combat against fraudulent behavior and tend to underestimate the social implications they cause in the respective institutions, but also in society at large (Allhutter et al. 2020, Sztandar-Sztanderska and Zielenska 2022, Geiger 2023). At the same time, alternative imaginaries have started to take shape in the European context trying to re-imagine and rebuild digital technology and data infrastructures for the public good (Mansell 2012, Lehtiniemi and Ruckenstein, 2019, Kazansky and Milan 2021, Mager 2023, Macgilchrist et al. 2023). European values such as data protection, digital sovereignty, or transparency are mobilized to promote large-scale infrastructures in the areas of research (Mahfoud 2021, Mobach and Felt 2022), cloud computing (Baur 2023), and web search (Mager 2023). Against this background, the project Automating Welfare (FWF I 6075) examines the implications of datafication and automatization for the welfare state and the flourishing of citizens in eight European countries using a case-study approach and a mix of different methods (data journeys, interviews, short-term ethnographies, citizen workshops). In Austria, two case studies are carried out focusing

on data infrastructures in the context of public health insurance and open commons. The first case study explores the use (and reuse) of health insurance data for the detection of social fraud, but in tandem tries to envision potential future applications oriented towards the public good such as risk prevention or public health initiatives. The second case study focuses on the Open Commons Linz initiative, which is trying to open up data to citizens for communal welfare and educational purposes- fieldwork of both case studies has just been started, first insights will be shared

at the conference. How "infrastructures of welfare" are imagined, built, and co-produced with larger digital transformations of the welfare state and how counter-narratives could take shape focusing on social justice rather than social fraud detection will be discussed in the presentation.

10:30–12:00 am	**VI: Datafication and Control (Chair: Charlotte Knorr)**

- Veronika Nagy (Utrecht University, The Netherlands)
  Outsourcing security intelligence: the risks of digital litter in migration control practices
- Jasper van der Kist, Silvan Pollozek (European University Viadrina Frankfurt (Oder), Germany)
  Datafication Discourses in Migration and Asylum Governance in Germany
- Irina Zakharova (Leibniz Universität Hannover, Germany)
  Datafied Welfare: Discourses about Regulation and Care
- Mateusz Trochymiak, Robert E. Zajonc (Warsaw University, Poland), Sebastian Sosnowski (Polish Academy of Science (IFIS PAN), Poland)
  False promises of a paradigm shift in welfare administration. Deconstructing digital transformation of care service policy in Warsaw, Poland

Outsourcing security intelligence: the risks of digital litter in migration control practices

Veronika Nagy (Utrecht University, The Netherlands)

With the datafication of human mobilities and the growing commodification of surveillance practices, such as high-tech border control, securitization of migrants has become a core subject of policing studies. National law enforcement practitioners are under growing pressure of the public discourse and tend to increasingly rely on information of third parties, like global corporate security companies and transnational NGOs to collect and on up to data information. With the growing attention to migration and the increasing use of mobile technologies, new service strategies have been developed by different stakeholders like UNHCR in order to provide fast, transparent and efficient services to people across the borders. Many of these technologies, like digital payment methods are introduced in the framework of humanitarian support of migrants, however, as many biometric experiments in refugee camps have been illustrated, these digital data infrastructures are also regularly commodified as a so-called 'security intelligence'. Though basic registration was always part of the tasks of humanitarian organisations, their presence and first-hand contact with refugees or asylum seekers also contributed to their complex role in digitized administrative decision-making in which data is often unverified, manipulated, or outdated. The neoliberal pressure of governments towards NGOs increasingly reinforces the need for efficient and sustainable service provisions by increased data processing and identification practices, including biometric authentication systems. However, this data is also accessible for third party tech companies that may use such data for commercial purposes in the context of state corporate migration control measures. Based on a discourse analysis, this paper aims to explore the risks and limitations of outsourced security intelligence practices established by nongovernmental data infrastructures, in particular considering the role of 'digital waste'.

Datafication Discourses in Migration and Asylum Governance in Germany

Jasper van der Kist, Silvan Pollozek (European University Viadrina Frankfurt (Oder), Germany)

The public and political debates on the so-called "refugee crisis" in 2015 exposed the limitations of the Federal Office for Migration and Refugees (BAMF) and its inability to "proceed with business as usual". The appointment of Weise as the new president of BAMF signalled a commitment to not only reorganise the procedures, but also to transform the organisation and its culture. A powerful vision of a new and future-oriented BAMF was articulated, based on digitalisation, technologies, and innovation. In the summer of 2016, the digitalisation agenda 2020 was launched, which aimed to "lead the BAMF into the future" by turning it into a "digital breathing administration".

The premise of this article is that migration problems are increasingly being solved by resorting to new forms of big data and methods (Dijstelbloem and Broeders 2015). Rather than assuming that this technological innovation in asylum administration as a 'natural' (or 'ideological') process, we ask how datafication has become the common solution to a wide range of problems. Michel Foucault (1984) underscored the idea that any solution to a particular problem is inherently shaped by the initial perception of the problem itself, a "work of thought", a process he termed "problematisation". Building on Foucaultian governmentality studies and science and technology studies, the paper conducts a discourse analysis to trace the shift in the problematisation of migration in Germany between 2015 and 2020, which effectively reconfigured how the BAMF thinks about its government.

Our aim is to show how datafication has rapidly become a common answer to this variety of problems pertaining to the governance of migration in Germany. Based on analysis of official documents and popular media in Germany, we highlight this 'work of thought' around the digitisation agenda to get a sense of what "problems" those digital devices should provide solutions for: how a particular problem recognised and singled out as a target for government; and why digitalisation is seen as the 'natural' way forward in all this (Beer 2016; Kitchin 2017). Moreover, by focusing on these problematising moments – or 'historical conjuncture' (Rabinow 2003, 55) – the paper also speaks to the contentious politics of datafication and migration.

Our findings suggest that it was the linking up of migration with rhetoric of (i) security, (ii) efficiency, and (iii) objectivity that brought the notion of datafication to the fore. The first pertains to narratives of crime prevention, particularly identity fraud, alongside security and safety considerations. The second involves discussions about the efficiency and operationality of migration governance. The last concerns the uncertainty and discretion in asylum determination, alongside debates about biased and prejudiced human perception.

Datafied Welfare: Discourses about Regulation and Care

Irina Zakharova (Leibniz Universität Hannover, Germany)

With the ongoing datafication of welfare services provision, digital data are used for governance, decision-making, scoring, fraud prediction, calculation of welfare benefits and more. Both public and private actors are promoting their products and services, not at least to increase their public acceptance. In public discourses datafication of public welfare is, hence, presented as a 'solution' to certain problems, often addressing both individual citizens and communities. These discourses about datafication processes in the public welfare often promote that public actors care for others. For example, they aim to foster well-being, and various commercial technology providers such as Moodle – frequently used by public educational institutions - promise to even improve the whole world (https://moodle.com). Research across various domains of media and communication studies has shown, however, that such caring promises can have opposite, regulatory implications or write out complexities and contingencies of relations between the state and the citizens (Broomfield & Reutter, 2022; Redden et al., 2020; van Schie et al., 2020). Overall, critical data researchers and scholars of technology researching public education and, more broadly, digital welfare, are concerned with the ways public and commercial actors promote datafication processes e.g., relying on imaginaries, metaphors, and popular discourses, e.g., such about advantages of data or citizen

participation (Mager & Katzenbach, 202; Puschmann & Burgess, 2014). In such studies, however, the notion of 'care' is yet to be discussed more extensively. This paper aims to do so by analysing discourses about care in relation to datafication processes and the political implications of care-ful or care-less configurations.

This paper therefore asks, what is the role of care in current discourses about datafied welfare and what implications can these discourses have. Conceptually, the paper draws on feminist science and technology studies attending to the concept and ethics of care (Lindén & Lydahl, 2021; Tronto, 2016) and feminist data studies (D'Ignazio & Klein 2020). From that standpoint care can be considered both as a a practice of working with, maintaining, repairing, or resisting datafication processes, and a normative obligation some public actors have towards another. Such care, however, also has its 'dark side of care' (Martin et al. 2015) – paternalistic, toxic and controlling aspects of care that hinder rather than foster flourishing and self-determination. This paper attends to these contingencies of care and its dark sides in the public discourses about datafied welfare. Empirically, the paper is based on collaborative work from different domains of welfare such as education and ageing in Germany, while interviews with actors in public institutions as well as analysis of relevant documents produced by involved public actors and commercial technology providers serve as research material. This contribution offers an analysis of public welfare actors' discourses in regard to their values, the infrastructures their caring promises address, and the work required to fulfill the caring promises. By attending to these discourses and speculating about subsequent changing relations between the state and its citizens through the analytical lens of care, this contribution discusses the extent to which datafied transformation contributes to more just and equitable futures and what (other) relations are required to achieve these.

False promises of a paradigm shift in welfare administration. Deconstructing digital transformation of care service policy in Warsaw, Poland

Mateusz Trochymiak, Robert E. Zajonc (Warsaw University, Poland)

Sebastian Sosnowski (Polish Academy of Science (IFIS PAN), Poland)

When it comes to the digitalization of public sector and administration, many authors have claimed that the ICT brought a significant change in a way that policies are designed and delivered, often implying that we are witnessing a governance paradigm shift (Dunleavy, 2005, Alston 2019, Dencik, Kaun 2020, Van Zoonen 2020, Yeung 2023, Kempeneer 2023). The shift means that automatization, algorithmic governance, and datafication of the decision-making and service delivery process lead to a deep change in the instrumental logic, organizational settings, goals, and objectives behind policy (Sewerin et al., 2022). In other words, digitalization affects policy to the core, transforms values and institutional conditions to the end where a new system-level bureaucratic order (Bovens 2002) and new regimes of welfare distribution emerge (Eubanks 2018).

As we do not neglect the overall and long-term impact of ICT on the public sector, we claim that its power to change the policy process to its core is often overestimated. Policy implementation is an incremental process (Lindbloom 1959) and innovations, like ICT, are rather adjusted to the existing institutional framework (policy goals, regulations, work organization procedures etc.) and sociopolitical environment than the other way around (Bekkers 2013). System-level changes require significant effort from policymakers and can endanger the status quo of multiple actors involved in the process of decision-making and service delivery (Peters 2015). Therefore, policy leaders are expected to minimize the risk of loss and introduce small, often superficial changes, with limited impact on the policy process, in order to preserve the existing institutional framework and power relations between actors.

However, in the same time public sector digital transformation projects are presented as "game changers" as they promise significant improvement in quality of service (more accessible, user-friendly and tailored service) and more efficient process of decision-making (automated, faster, bureaucracy and administrative burden reduction). This overoptimistic perception of the "technology as a solution to everything" (Nielsen 2023) is useful for

policymakers, who build narration about "big change for better", while in fact they act in order to strengthen the value of the "old" policy regime (Bekkers 2013).

The goal of our research is to deconstruct "the rhetoric of change" that accompanied the Implementers of project aimed at digitalization of the care service delivery process in Warsaw and neighbouring counties. We aim to discover what is hidden under the digital transformation promises (goals, objectives, instrumental logic, organizational settings) and how it is supposed to change existing policy (institutional framework and power relations). Our methods include individual in-depth interviews with actors engaged in project design, implementation, and desk research analysis. Initial findings revealed that the official narration for the project is built on the promise of "better quality of service for welfare recipients". However, the project leaders and policy makers see the E-Care system mainly as a way to "fil the gaps" of the existing system of service monitor and improve the supervision over the care workers.

Contact:

Prof. Dr. Christian Pentzold
Universität Leipzig

Dr. Charlotte Knorr
Universität Leipzig
Nikolaistraße 27-29, Room 5.13
+49 341 97-35717