

Critical Data Studies Meet Sociology

LIFE TODAY IS data driven – and the basis of data collection. Industries and institutions gather more digital data for their potential to generate value and action than ever before. These range from health data (Stevens *et al.* 2018; Vezuridis & Timmons 2021) to environmental data (Deitz 2023; Vardy 2020) to “alternative” data which are cleaned, repackaged and sold for use by industries in need of predictive models, like finance (Hansen & Borch 2022; see also Plantin 2019).

This special issue provides a window onto the meeting of critical data studies and sociology in Scandinavia to explore the situated, localised modes of “being and becoming” with data (Lupton 2018:9) in a particular moment in time, a time when in the Northern European context, public administrators, institutions, software developers and scientists are busy imagining and trialling arrays of data-driven solutions to automate public services, intellectual work and infrastructure operations. Data, in these contexts, often figure as an element of already ongoing, much older practices of classifying, categorising, quantifying and producing actionability on the social. However, data bring to these the novel element of the imagined possibility for feeding these practices into machine learning, automation algorithms and computational models to fulfil ideas of efficiency, increased productivity or predictive potentials. Data become, in this context, part of a landscape of instrumenting actions and socio-technical structures that order social life at a distance and according to formal rules (Savolainen & Ruckenstein 2024).

This is not unique in the Nordic context, but there is a particular “early adopter” spirit here that drives development in the welfare state and other state/private entanglements. Likewise, there are significant funds being allocated to this development from public and private actors. Therefore, it is interesting to see how this impetus to adopt and adapt data-driven technological solutions is playing out in our area. And while not every article in this double issue is specifically focused on Nordic cases, many of them are, and engage a sensitivity for the situated structural entanglements of data, here.

One particular concern shared by many of the pieces in this special issue is care for the tensions that emerge as different actors try to locate themselves as users and developers in relation to these technologies and the values of efficiency they promise (Zakharova *et al.* 2024). The technologies are, unsurprisingly, employed with different means and ends in contexts as diverse as public schools, streets and scientific labs. Situated empirical analyses allow one to raise important questions: What does it mean when authorities use experimental, non-scientific methods to detect the age

of unaccompanied minors who have arrived as refugees to Sweden, and render them into datasets of biometric data that inform institutional decisions on citizenship and migration? And how does this interact with changing refugee policies? What does it mean when software developers gain the authority to shape the truthfulness of datasets that inform medical professionals' decisions on whether someone has cancer? And what does that do in a context of state-funded healthcare for all? How do intersectional and geographic differences matter in the context of making education and security in Sweden data and algorithmically driven? And what role does the particular ways our cities have become segregated in Sweden play in this development? We see the granularity of the studies in this special issue as valuable in analysing these questions and using them to foreground the ongoing negotiations and transformations of agencies and power, knowledge and public values with data-driven technologies.

Context is important for each of these studies, yet we also think the observations made about the social-data-political entanglements can trigger research in other cultural contexts, too. For example, when Baki Cakici and Alena Thiel explore what it means to become an identifier – a number – useful for data collection and addressability, they are exploring the relationship to Nordic personal numbers that appears at the moment of assignment, at birth. But their conceptual nuance of addressability and communicative positions allows an affective analysis of this reductive practice that is also useful for surveillance literature at large. Amanda Lagerkvist, Jenny Eriksson-Lundström and Maria Rogg's use of *crip theory* to discuss biometric indexing at the nexus of commercial use of the hand palm will be relevant in many different contexts engaging biometrics, commercial or otherwise. And Francis Lee's discussion of *agencing* as a verb will, we are sure, prove useful for analysis of datafied scientific practice in many domains. We posit that these papers may act as a foil to observations in contexts dominated by different infrastructures and systems, and by different imaginaries. It is our hope that this special issue will serve up some inspirational tools and terms for many colleagues engaging data as an object of study.

Meets ...

The articles presented here speak to the broader domain of scholarly work loosely labeled as “critical data studies”, a relatively young, multidisciplinary field concerned with the inquiry into the politics of producing, organising, moving and materialising data in the world (Dalton *et al.* 2016). In a recent book, Rob Kitchin defines critical data studies as revolving around questions about “data forms, processes, and purposes, and consider what is at stake when data are produced and deployed to create knowledge, manage society, derive business value, and perform numerous other tasks” (Kitchin 2024:2). Much of the work in the field has discussed “data” and its corollary, “datafication” in the domains of economy, mythology and power that enforces a universalising epistemology which amplifies social discrimination, inequalities and toxicities (e.g. Benjamin 2019; D'Ignazio & Klein 2020; Douglas-Jones *et al.* 2021; Elish & boyd 2018; Milan and Treré 2019; Thylstrup 2019). Moreover, as data-driven solutions

and tech companies expand into different professional spheres, scholars explore the pressures that these practices put on welfare state values and on expertise for instance in domains such as public administration and healthcare (Dencik & Kaun 2020; Stevens *et al.* 2024).

These questions are, of course, also central to sociology and its diverse strands, flavours and modulations, perhaps especially science & technology studies (STS). Here one finds early insights into how scientific knowledge was produced as objective (Fleck 1929 [1986]) and mobilised (c.f. Latour & Woolgar 1979; Latour 1987; Law & Hassard 1999) and how categorisation practices (Bowker & Star 1999) and ways of knowing (Barad 2007) create and form the basis of “science”, even when it is anything but objective (Shapin 2010). Sociology and critical data studies share many tools, concerns and insights.

For us, it also seemed natural to put together a special issue with critical data studies and sociology, because of the way that critical data studies have, from the very start, been sensitive to space and context (Dalton *et al.* 2016). That sensitivity is also found in sociology and neighbouring fields, where the research toolbox is filled with methods and theories to address it. We feel there is a productive synergy that arises in these disciplinary overlaps, one which allows us to engage nuanced and situated approaches to making structural power dynamics visible, and conduct work that is engaged with partial perspectives and the processes of othering that data involve.

The notion of human-data assemblages is another shared theoretical concept that feeds into this special issue. It directs attention to processes of transformation of social relations, the material contexts that produce them, and how data is not a “thing” but “a mode of being and becoming” (Lupton 2018:9). This analytical stance, which often analyses data and humans as entangled relations rather than ontologically discrete objects, is common in the borderlands between STS, critical data studies and sociology (c.f. Armstrong 2019; Kim *et al.* 2024; Leonelli *et al.* 2017; Ribes 2019; Stevens *et al.* 2018; Thompson 2020; Vardy 2020). It helps us see how the value of data emerges from the intra-action between matters and materials, discourses and knowledge, human and more-than human bodies and agencies. These, too, are contingent and situated, to borrow a word common in feminist technoscience (c.f. Suchman 2023). Studying them means paying attention to the details of how they are situated, with what implications and for whom. This care for context and contingencies is what can make a special issue in a regionally obscure and geographically peripheral journal interesting.

Actually, at this point in writing the Introduction, we sighed and wanted to stop going on about disciplines and borders. We, the guest editors, are both located in an interdisciplinary department (Department of Thematic Studies (TEMA), at Linköping University) and work in very undisciplined environments. For us, the whole mess of border-crossing, non-disciplined work we see daily, and also highlight in this special issue, provides us with really interesting analytical terms, tools and methods to try to make sense of data and what it is doing. As Metcalf notes, critical data studies “foregrounds questions about how data are produced, how they are maintained and mobilized over time, and in what ways they are interpreted. In doing so, it illuminates

how data themselves—rather than just the actors who rely on them—can drive shifts in larger social formations” (Metcalf 2024:3). Such a stance is almost more methodological than disciplinary. It is a stance that runs through many of the works we find inspiring and many of the studies this special issue gives a forum to – but these are studies which call many different types of departments home. We are grateful that *Sociologisk Forskning* makes space for this undisciplined stance at their hearth.

Making society legible to machines

During a recent seminar, a colleague of ours with a background in computer science exclaimed: “Data is not data until it is made machine-processable.” Their statement highlights the crucial, if self-evident fact, that machines require formatted and ordered data. Put simply, messy social realities and processes have to be made legible to algorithms. Legibility demands reductions and cuts, formatting or excluding those facets of the social that do not fit data-processing logics (c.f. Plantin 2019; Nafus 2024). Several contributors in this special issue explore the forms that reduction can take and their implications.

Katarina Winter and Klara Hermansson discuss the reduction of the much broader issue of security and vulnerability produced via a child-tracking app that works in more resource-rich geographical areas of Stockholm but not in other areas. Their analysis helps us see how apps operate as agents in structuring the ‘individualization and privatization’ of responsibility for crime preventions (p. 61), while also normalising and enforcing insecurity and risk. Winter and Hermansson’s example shows how data-driven apps are redefining the very meaning of security and safety, changing it “from a matter of passive vulnerability to one of active engagement aided by tracking children’s location data”. (p. 67).

Amanda Lagerkvist, Jenny Eriksson-Lundström and Maria Rogg explore another facet of reduction and legibility by following the implications of a palm-reading device developed and sold by Amazon. They discuss how the device requires making the human palm machine readable through biometric data, an act that discards the experiential, affective dimensions of being a body in the world through hand-touch while “indexing and fixing the self as a purchasing being” (p. 26). In this case, reduction favours values of “convenience” and ability, a generative act that produces new subjects in the commercial world of data while it simultaneously disavows being a “human in diversity and disability” (p. 19).

These examples expose the extent to which making bodies, lives and societal problems legible to algorithms demands reductive adjustments between algorithms and humans. Large societal issues are reduced to what algorithms can cope with or what their developers ask for (e.g. a paying subject). Crucially, the methods through which these reductions collapse and register messy social realities create the very realities that they try to capture (c.f. Bowker & Star 1999; Law 2004).

Why is this important? A lot of the critical work coming out of the Anglo-Saxon context shows that a machine-readable society enhances those qualities and procedures

that are easier to fit into machine logics, while simultaneously evacuating and discarding those that simply do not fit machine learning models and algorithmic logics (c.f. Benjamin 2019; Amaro 2023). Lagerkvist, Eriksson-Lundström and Rogg (this issue) posit that this leads to a crip society, with core qualities and values like care, humanness and shared responsibility removed from our understanding of the social, and replaced by convenience and trackability. To be sure, these are not deterministic or irreversible processes. We know that people are not passive protagonists, but regularly redefine and even abandon technologies that fail to speak to their concerns, as Winter & Hermansson show, too, in this issue. Examples from our Northern European context are a helpful reminder that data-driven futures are mouldable and open-ended, falling apart and re-emerging, by and with *us*. They remind us that while discursively these futures and the sociotechnical relations that produce them might seem durable, even inevitable, they are always already partially breaking down, abandoned and discarded (Velkova & Plantin 2023). And yet, what contributors show is also how what sticks in the friction between emergence and falling apartness, to paraphrase Anna Tsing (2005), is a society reduced to its machine-readable qualities.

Reordering machine-human agencies through data work

This brings us to the issue of shifting human-machine agencies. Much of the ongoing enthusiasm and capital vested in data-driven technologies stems from the belief that these can perform better than human insight. This belief in technologies' abilities is often coupled with an imaginary of obedient assistants and servants who do all that undesirable, undignifying or difficult work that (some often very privileged) people do not want to do (see Suchman 2007). These promises are not new (Armstrong 2019; Kaijser *et al.* 2024). But today, machine learning and "artificial intelligence" are the technologies that people in positions of power task with expectations to fulfil these promises. Of course, artificial intelligence is neither artificial nor intelligent (Crawford 2021). At best, the term is an empty signifier which diverts attention from the power relations and ordering logics through which it is produced (Suchman 2023). Machine learning algorithms gain their power as humans grant or revoke their agency.

Our view on agency has its base in feminist technoscience and the sociology of science and technology, which often approaches agency as a result of relations made between humans and nonhumans, while also acknowledging that there always exist power differentials between them (c.f. Barad 2007; Suchman 2007). Lee (this issue) proposes calling the process of authorising data-driven technologies to have agency in professional settings *agencing*, and reminds us that technologies do not land, fully formed and unchanging, into society but are actively socialised, negotiated and eventually discarded in already existing practices. Drawing on examples of data work from the biosciences, he shows how agencing can deny machines agency in certain contexts that are important for scientists, and enrol them in others as assistants or as co-explorers of knowledge, in the hope that they could deliver greater insight. Lee's example shows how these changes are not in human or machine "intelligence" as such, but in the

material and epistemic cultures of scientific work. Specimens get replaced with their computational representations and scientific work shifts towards machine-generated predictions based on such representations.

These shifts are tightly connected to data work. Data work can take many forms but generally refers to work that is crucial for maintaining, producing or simply demonstrating the value of data-driven technologies. This value is relational and context specific. In online environments, data work is gendered, racialised, and categorised in highly political ways, often by gig or temporary workers, including prisoners and refugees (D'Ignazio & Klein 2020; Kaun et al. 2023; Lehtiniemi and Ruckenstein 2022). This kind of work includes tasks like word annotation of images for machine learning algorithms, cleaning errors from scanning and digitising documents, or differentiating displayable from non-displayable content to different publics online. Data work of this type contributes to the public legitimacy of algorithmic systems and to the myth of human-free machinic capabilities. In the scientific laboratory, data work might be gendered and low-status, too, but it also has to be “truthful” and trustworthy, in order to be credible. The cultural practice of crafting and negotiating the “pristineness” or “rawness” of datasets becomes an art and a virtue of data work in the sciences (Harrison 2024; Plantin 2019).

The contexts and power dynamics surrounding data work matter for understanding how human-machine relations and agencies shift with data-driven technologies. In online platforms, powerful companies are in the position of agencing algorithms and redistributing agencies by positioning human workers in the role to behave like machines ... and machines in a position to behave like humans, deprecating positive human abilities such as empathy and emotion (Ruckenstein and Turunen 2019).

Scientific practices engage different values and professional cultures, where human-machine agencies emerge from data work that might boil down to the search for credible data referents for scientific “truths” and the negotiation of knowledge between computer scientists and domain experts. Trials of machine-learning or automated decision-making algorithms draw data scientists and computer engineers out of their spheres of competence, granting them the position of experts in an expanding array of domains that they might know little about, as when computer engineers find themselves in a position to judge cancer through a range of pixel values on a digitised image, as Charlotte Högberg writes in this issue. She shows how health experts make medical problems doable for computer scientists, who in turn are expected to perform as health experts. Are health experts then in turn forced to become a kind of maintenance workers for computer engineers?¹ The articles in this issue suggest it may be time to pause and ask how shifting expertise and machine-human agencies impact professional domains.

1 The question draws on M. Ruckenstein (2024). *Why Focus on Repair?* Presented at Swedish STS Days 2024 3-4 October, Norrköping, Sweden.

Making subjects/subjectivities

Outside scientific laboratories and computer engineering offices, data work can also enact ordinary people as governable subjects. This, too, is nothing new (Armstrong 2019), but highly relevant today. Data work is an indelible part of statecraft and biopolitical governance, and of categorising people, for instance as “Europeans” (M’Charek *et al.* 2014). Statisticians and census workers, Ruppert & Scheel (2021) show, make people through “inscription devices and practices” (c.f Law 2004) such as procedures, registers, categories, standards, materials, infrastructures and more recently, big data and algorithmic technologies to sort people into categories and eventually make them into subjects and members of a “people”, or deny them this status. Always political, this kind of work means that inclusions and exclusions have profound repercussions for people (Bowker & Star 1999; D’Ignazio & Klein 2020). Excluded subjects, for instance, have to negotiate with machine and human gateways to secure a place in core data categories in order to be recognised as addressable subjects with rights to the most mundane state services, such as a marriage certificate, as recent studies of large-scale roll-outs of biometric identification systems in India show (Singh 2023). In this special issue, Baki Cakici and Alena Tiel theorise further the notion of “addressability” for unpacking the communicative dimension of the work of subject construction and identification through data. Rather than issues of address and entitlement, they discuss addressability as “a reductive transformation that breaks chains of meaning, and ... constructs the communicative position” (Cakici and Tiel, this issue: p. 116).

The work of making subjects addressable to statecraft through communication involves many actors. An obvious but often neglected example is media. Media co-produce addressability and have the power to legitimise for the public the state abuse of vulnerable people, as Matilda Tudor shows in this issue. She analyses media discourse on the ethically controversial medical age assessments that were conducted on unaccompanied minors who sought asylum in Sweden after the 2015 migration “crisis”. Medical age assessments were used as a test of whether migrants had the right to protection by the state. Tudor notes the powerful truth-making effect that turning a body into data has, and how this process was used by the media to reframe the discourse on migration in Sweden, essentially weaponising and othering child refugees, denying them the right of protection by claiming they were adults. Her argument is a context sensitive example that shows how datafied bodies become objectified, racialised and othered, made into technical units and denied existence even *before* they become recognised as subjects by statecraft. Thinking with Ramon Amaro (2023:13–14) we see Tudor’s work as an example of how data-driven decision-making systems operate as technologies of objectivity but are already produced through social relations that condition and racialise people, negating their value.

In another close study of how technology, the state, commercial actors and values are entangled in the process of making (future) citizens, Ingrid Forsler’s nuanced research on Estonian and Swedish school policies for digital tools, in this issue, shows how

different collective imaginaries about the role of tech in a nation-state's future and the role of schools in social policy impact approaches to education. Through a series of "future workshops" which engaged actors from these different domains in Estonia and Sweden, distinct differences emerged in how technology was imagined to be useful in a school setting between the two countries, differences which also reflected how young students were imagined and addressed as future adult citizens. But Forsler's work demonstrates that using tech and data to address future citizens or today's students also means making them understandable to machines, returning our conversation to legibility and reductionism. These demands can impact how and where the classroom and learning is imagined.

What to do with all of this? With entanglements around data that engage state actors, laws & policies, commercial interests, scientific practices, our bodies, our lived experiences of the world, norms & values, ground truths and representations ... an interdisciplinary approach does seem near to hand. We decided to close the special issue with Magdalena Kania-Lundholm's piece, which calls for sociological analyses of the systematic glitchiness of data-driven systems. Glitches, she writes, thinking of work on maintenance, repair and feminist data studies, are mirrors of systemic exclusion and discrimination that provide windows for intervention. Following the glitch is, then, a methodology for analysing transforming structures, agencies and power relations with data-driven technologies, and for "reimagining technologies with and for people" (Kania-Lundholm, this issue, p. 182). Her work intentionally bridges sociology and critical data studies, showing the essential political saliency of engaging data as an object of critical study, as the other theoretical approaches that comprise this special issue also do.

What's next?

We see potential for new lines of inquiry that emerge from the work presented in this special issue. One topic on the horizon is the relationship between ground truths and the generated (synthetic) data used by many AI systems. In this issue, Högberg starts to examine this with her work on medical AI and we think this exploration is only the beginning of a new line of discussion. Synthetic data produced by generative AI are intentionally different from original data. They are produced to assure privacy, create portable data, avoid regulations and – most interestingly – contribute new and additional "information" for machine learning algorithms training on the data. But the intersectional hallucinations (Johnson & Hajisharif 2024) these ontological differences inevitably introduce also trigger interesting questions about representation and reference, as well as silences on context, production and maintenance. Critical data studies, sociology and STS have many well-sharpened tools for discussing these and we look forward to future work that probes the contexts and contingencies of real and synthetic data, and hybrid versions of both. We suspect critical studies coming from all these areas (and the areas between, where most of us in this special issue are working) will have valuable tools for probing the "new" relationship to ground the truth and

representation it may trigger in some fields, the way boundaries are drawn around and between real and synthetic, for examining how the politics of categorisation are even more strongly obscured by the shimmer (and tarnish) of synthetic data and AI knowledge, and on the emerging practices that claim to assure provenance and pedigree.

These future analyses could also engage the constituent materialities and transforming computation and communication infrastructures of (synthetic) data, asking how they sustain, reinforce or transform inequalities and power structures in society. After all, synthetic data, and any digital data processing depends on minerals and metals, water, land, energy and human care. These have to be assembled in computers and communication systems, stacked in data centres, connected to fibre optic cables and power grids, financed, and maintained to materialise the digital computing infrastructure that makes digital data practices possible. A domain of fervent speculation and sociomaterial transformation, decisions about these matters result in environmental and energy “crises”, as they, too, generate new regimes of worth and inequalities through imaginative relations with uncanny elements, such as the heat generated from data processing that today warms many Nordic citizens (Velkova 2016). We therefore believe that (new) materialist sensitivities could fruitfully complement sociological analyses in tracking the different decisions, imaginings and (de-)materialisations of computing infrastructure for data in places, contexts and over time. These analyses could unearth how new elements, such as atoms and gases, wind and the sun are reimagined to sustain computing infrastructure, how they become invisibilised in everyday data-processing contexts, and how these relations co-produce and normalise societal inequalities. Such analyses, we suspect, might push scholars to craft new conceptual tools and intellectual pathlines that explore how the promise of making every issue in the world computable by machines remakes differently and differentially the very categories, economies and infrastructures on which society stands.

Acknowledgements

We thank the anonymous peer reviewers for supporting this special issue with their time and critical engagement. We thank all contributors for their articles, and for their dedication and effort to keep up with a tight publication timeline. Julia Velkova’s work is supported by the Profutura Scientia program of Riksbankens Jubileumsfond and SCAS, as well as by the *Reimagine ADM* project funded by FORTE in Sweden (GD-2022/0019) and the European Union’s Horizon 2020 Research and Innovation Programme, Grant Agreement no 101004509. Ericka Johnson’s work is funded by Linköping University, the Wallenberg AI, Autonomous Systems and Software Program – Humanity and Society (WASP-HS), Vinnova, and VR 2024-01837.

References

- Amaro, Ramon (2023) *The black technical object*. Cambridge: MIT Press.
- Armstrong, David. (2019) "The social life of data points: Antecedents of digital technologies." *Social Studies of Science* 49(1) 102–117.
- Barad, K. (2007) *Meeting the universe half way*. Durham: Duke University Press.
- Benjamin, R. (2019) *Race after technology*. London: Polity.
- Bowker, G.C. and S.L Star (1999) *Sorting things out: Classification and its consequences*. Cambridge: MIT Press.
- Crawford, K. (2021) *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. New Haven: Yale University Press.
- Dalton, C. M., L. Taylor & J. Thatcher (2016) "Critical data studies: A dialog on data and space." *Big Data & Society* 3 (1):1-9.
- Deitz, S. (2023) "Outlier bias: AI classification of curb ramps, outliers, and context." *Big Data & Society*, 10 (2):1-14.
- Dencik, L. and A. Kaun (2020) "Datafication and the welfare state." *Global Perspectives* 1 (1): 12912.
- D'Ignazio, C. & L. Klein (2020) *Data feminism*. Cambridge: MIT Press.
- Douglas-Jones, R., A. Walford & N. Seaver, N. (2021) "Introduction: Towards an anthropology of data." *Journal of the Royal Anthropological Institute* 27 (S1):9–25.
- Elish, M.C. & D. Boyd (2018) "Situating methods in the magic of Big Data and AI." *Communication Monographs* 85 (1):57–80.
- Fleck, L. (1929 [1986]). "On the crisis of 'Reality'", 47–58 in R. Cohen & T. Schnelle (Eds.) *Cognition and fact – Materials on Ludwik Fleck*. Dordrecht: Reidel.
- Hansen, K.B. & C. Borch (2022) "Alternative data and sentiment analysis: Prospecting non-standard data in machine learning-driven finance", *Big Data & Society* 9 (1):1-14.
- Harrison, K. (2024) *Behind the science: The invisible work of data management in big science*. Bristol: Bristol University Press.
- Johnson, E. & S. Hajisharif (2024) "The intersectional hallucinations of synthetic data", *AI & Society*.
- Kaijser, Arne, Gustav Sjöblom, Johan Gribbe, & Per Lundin (2024) *Maktens maskiner*. Lund: Arkiv.
- Kaun, A., A. Logsdon, P. Seufferling, F. Stiernstedt (2023) "Serving machines and heterotopias: Data entry work in prisons and refugee camps in the US and Uganda", 144–161 in L. Parks, J. Velkova & S. De Ridder (Eds.) *Media backends: Digital infrastructures and sociotechnical relations*. Urbana: University of Illinois Press.
- Kim, Y., M. Finn, A. Acker, B. Chaudhuri, S. Wedlake, R. Ellis & J. Srinivasan (2024). "Epistemologies of missing data: COVID dashboard builders and the production and maintenance of marginalized COVID data", *Big Data & Society* 11 (2): 1-14.
- Kitchin, R. (2024) *Critical data studies: An A to Z guide to concepts and methods*. Hoboken: Polity.
- Latour, B. (1987) *Science in Action*. Cambridge: Harvard University Press.

- Latour, B. & S. Woolgar (1979) *Laboratory life*. Princeton: Princeton University Press.
- Law, J. & J. Hassard (1999) *Actor network theory & after*. New Jersey: Wiley.
- Law J (2004) *After Method: Mess in Social Science Research*. London: Routledge.
- Lehtiniemi, T. and M. Ruckenstein (2022) "Prisoners training AI", pp. 184–196 in S. Pink, M. Berg, D. Lupton, M. Ruckenstein (Eds) *Everyday automation: Experiencing and anticipating automated decision-making*. Abingdon: Routledge.
- Leonelli, Sabina, Brian Rappert & Gail Davies (2017) "Data shadows: Knowledge, openness, and absence", *Science, Technology and Human Values* 42 (2):191–202.
- Lupton, D. (2018) "How do data come to matter? Living and becoming with personal data", *Big Data & Society* 5 (2): 1–11.
- M'charek, A., K. Schramm & D. Skinner (2014). "Topologies of race: Doing territory, population and identity in Europe", *Science, Technology, & Human Values* 39 (4):468–487.
- Metcalf, K. (2024) "Categorical misalignment: Making autism(s) in big data biobanking", *Social Studies of Science*. Online First.
- Milan, S. and E. Treré (2019) "Big data from the South(s): Beyond data universalism", *Television & New Media* 20 (4):319–335.
- Nafus, D. (2024). "Unclearing the air: Data's unexpected limitations for environmental advocacy", *Social Studies of Science* 54 (2):163–183.
- Plantin, Jean-Christophe (2019) "Data cleaners for pristine datasets: Visibility and invisibility of data processors in social science", *ST&HV* 44 (1):51–73.
- Ribes, David (2019) "STS, meet data science, once again", *ST&HV* 44 (3):514–539.
- Ruckenstein, M. and L.L.M. Turunen (2019) "Re-humanizing the platform: Content moderators and the logic of care", *New Media & Society* 22 (6): 1026–1042.
- Ruppert E. and S. Scheel (eds) (2021) *Data Practices: Making up a European People*. London: Goldsmiths Press.
- Savolainen, L. and M. Ruckenstein (2024) "Dimensions of autonomy in human–algorithm relations", *New Media & Society* 26 (6):3472–3490.
- Shapin, S. (2010) *Never pure. Historical studies of science as if it was produced by people with bodies, situated in time, space, culture and society and struggling for credibility and authority*. Cambridge: Harvard University Press.
- Singh, R. (2023) "The backend work of data subjects: Ordinary challenges of living with data in India and the US", 229–244 in L. Parks, J. Velkova & S. De Ridder (Eds.) *Media backends: Digital infrastructures and sociotechnical relations*. Urbana: University of Illinois Press.
- Stevens, Marthe, Rik Wehrens & Antoinette de Bont (2018) "Conceptualizations of big data and their epistemological claims in healthcare: A discourse analysis", *Big Data & Society* 5 (2):1–21.
- Stevens, Marthe, S.R. Kraajievelde and T. Sharon (2024) "Sphere transgressions: Reflecting on the risks of Big Tech expansionism", *Information, Communication & Society* 27 (15):2587–2599.
- Suchman, L. (2007) *Human-machine reconfigurations*. Cambridge: Cambridge University Press.

- Suchman, L. (2023). "The uncontroversial 'thingness' of AI", *Big Data & Society* 10 (2): 1-5.
- Thompson, T. L. (2020). "Data-bodies and data activism: Presencing women in digital heritage research", *Big Data & Society* 7 (2): 1-7.
- Thylstrup, N. B. (2019) "Data out of place: Toxic traces and the politics of recycling", *Big Data & Society* 6 (2): 205395171987547.
- Tsing, A.L. (2005) *Friction: An ethnography of global connection*. Princeton: Princeton University Press.
- Vardy, Mark. 2020. "Relational agility: Visualizing near-real-time Arctic sea ice data as a proxy for climate change", *Social Studies of Science* 50 (5):802–820.
- Velkova, J. (2016) "Data that warms: Waste heat, infrastructural convergence and the computation traffic commodity", *Big Data & Society* 3 (2):1–10.
- Velkova, J. and J.-C. Plantin (2023) "Data centers and the infrastructural temporalities of digital media: An introduction", *New Media & Society* 18 (3): 273-286.
- Vezyridis, Paraskevas & Stephen Timmons (2021) "E-Infrastructures and the divergent assetization of public health data: Expectations, uncertainties, and asymmetries", *Social Studies of Science* 51 (4):606–627.
- Zakharova, I., J. Jarke & A. Kaun (2024). "Tensions in digital welfare states: Three perspectives on care and control", *Journal of Sociology* 60 (3):540–559.

Author presentations

Ericka Johnson is a professor in Gender & Society at the Department of Thematic Studies, Linköping University.

Contact information: Ericka.johnson@liu.se

Julia Velkova is an associate professor of media and communication studies, and senior lecturer in technology and social change at the Department of Thematic Studies, Linköping University. She is also a Profutura Scientia fellow (2023-2028), in residence at the Swedish Collegium for Advanced Studies during the fall 2025 and spring 2026.

Contact information: julia.velkova@liu.se