Introduction to the special issue on *Data Justice*

Author info:

Lina Dencik, Arne Hintz, Joanna Redden and Emiliano Treré run the Data Justice Lab at Cardiff University's School of Journalism, Media and Culture, UK.

In May 2018 over 200 people from around the world met at Cardiff University in the United Kingdom to participate in a conference on 'Data Justice' hosted by the Data Justice Lab and the DATAJUSTICE project[i]. They included scholars from a range of disciplines stretching across media studies, geography, computer science, law, philosophy, sociology and politics as well as civil society groups and professionals working at the intersection of technology and society. The conference marked a clear recognition that the way data is generated, collected and used in society and everyday life has become an increasingly prominent and contentious issue. Developments in 'smart' technologies, machine learning and Artificial Intelligence are now an integral part of how societies are organized and decisions made, both in rhetoric and in practice. How we come to understand the world, what services we are able to access, where we are able to go, what we are able to do, and the way we are governed now all potentially feature data practices that shape the terms and conditions for our participation in society.

Yet the conference also marked a notable shift in the framing and understanding of what is at stake with such developments. It explored how the transformations associated with the 'datafication' of society (Mayer-Schönberger and Cukier 2013) entail power dynamics that require investigation and critique. It recognised that this is not primarily a technical question but relates to long-standing social, political, economic and cultural issues. The processing of data from across our lives can fundamentally shape social relations, the kinds of information valued and what is 'knowable' and therefore acted upon. At the same time, data, and the way it is generated, collected, analysed and used, is a product of an amalgamation of different actors, interests and social forces that shape how and on what terms society is increasingly being datafied. This shifts the focus of the data-society nexus away from simple binaries that frame the debate in terms of trade-offs or 'good' vs. 'bad' data in which data is an abstract technical artefact. Instead, data is seen as something that is situated and necessarily understood in relation to other social practices. By focusing on the concept of 'data justice', the conference addressed questions of how our understanding of social justice is changing in the context of datafication, what concepts and practices are needed, and how social justice can be advanced in a datafied society. Participants discussed how, for example, questions of political change, labour

relations, colonial and North-South relations, policing and border control, civil society and democracy are all implicated by the rapid increase in data processes, and considered what this means for core tenets of social justice. Together, they explored the breadth as well as the boundaries of 'data justice' as a concept, practice and approach. This special issue brings together a selection of papers presented at the conference, speaking to key themes from a range of contexts and perspectives.

The notion of data justice connects different approaches, disciplines and concerns, as both the conference and this special issue demonstrate. As part of a nascent debate, it has led to different interpretations of the interplay between data and social justice, and to different strategies and responses. Often, data justice is a response to prominent and rather limited perspectives on the societal implications of data-driven technologies that have tended to focus on issues of efficiency and security on the one hand and concerns with privacy and data protection on the other (Dencik et al. 2016). The framework of data justice broadens the terms of the debate in a way that accounts for a host of issues that are compounded in the datafied society, as evidenced in recent scholarship relating to democratic procedures, the entrenchment and introduction of inequalities, discrimination and exclusion of certain groups, deteriorating working conditions, or the dehumanization of decision-making and interaction around sensitive issues. These discussions suggest a need to position data in a way that engages more explicitly with questions of power, politics, inclusion and interests, as well as established notions of ethics, autonomy, trust, accountability, governance and citizenship.

The Data Justice Lab emerged out of a recognition for the need to advance research and practice that can take account of these concerns, building on the work of others and shifting the lens through which we understand both developments and implications of the growing digital economy. The Snowden leaks of 2013 provided a pivotal moment for exploring the significance of data-centric technologies in our everyday lives, and moved our attention towards the nature of transformations in governance and the limitations in our ability to contend with them. A reframing towards data justice initially offered a way to situate questions of data in relation to ongoing social justice concerns that could engage a wider political mobilisation (Dencik et al. 2016) and could inform debates on citizenship more broadly (Hintz et al. 2018). Furthermore, it provided an umbrella under which to explicitly identify and illustrate the nature and diversity of harms that are caused by uses of new data systems (Redden and Brand 2017) as a way to foreground the politics of data and the significance of the contexts in which data

systems are implemented. This continues to be a key concern as we focus efforts more to understand actual developments on the ground and how these play out in practice, not least at the local level where people's participation in society is often oriented (Dencik et al. 2018). Importantly, a key objective of the Data Justice Lab has been to advance knowledge and debate on digital technologies outside of what has been a very US-focused debate that has shaped understandings of both challenges and responses. Whilst we can learn from this debate and need to be informed about it (not least as the digital economy is dominated by US-based companies), we also need a better understanding of developments happening elsewhere. Europe, for example, provides a significantly different context for understanding both what is at stake and what appropriate responses might be, not to mention the rapid and complex growth of data collection in countries in Asia and Africa, and the intricate and rich histories of engaging with data in Latin America. More broadly, it is about fostering the recognition and the exploration of ways of thinking and using data from the margins, and the many Souths inhabiting our world, as a way to promote a reparation to the cognitive injustice that fails to recognize non-mainstream ways of knowing the world through data (Milan and Treré 2019; Treré 2019).

The Data Justice Lab therefore sits in connection with a range of research and practice that engages with the relationship between data and social justice in different ways. In scholarship, we have seen the concept of data justice used to denote an analysis of data that pays particular attention to structural inequality, highlighting the unevenness of implications and experiences of data across different groups and communities in society. This has, in some interpretations, led to new articulations of principles to underpin data governance that can better account for such inequalities (Heeks 2017; Taylor 2017), or practices in the handling of data that make asymmetries in the representation and power of data explicit (Johnson 2018). In other interpretations, conceptions of justice have been foregrounded in ideas about the design process and the conditions within which data infrastructures emerge, leading to calls for more participatory design practices that emphasize the involvement of communities and that seek to build alternative bottom-up infrastructures to empower rather than oppress marginalized groups (Costanza-Chock 2018). Related to this, debates on data justice have emerged at the intersection of activism and technology in which data is seen as an avenue to revert or challenge dominant understandings of the world, (re)creating conditions of possibility for counter-imaginaries and social justice claims to emerge (Milan and van der Velden 2016; Gray 2018).

Grassroots groups and social justice campaigns have also started to apply a more comprehensive and critical approach to datafication and, in some cases, have done so within a 'data justice' framework. The Center for Media Justice in the United States recently created their own Data Justice Lab as part of the annual Data for Black Lives conference, dedicated to thinking through ways to bridge research, data, and movement work relating to issues like surveillance, carceral tools, internet rights, and censorship. The Detroit Digital Justice Coalition has worked with local residents in identifying potential social harms that may emerge through the collection of citizen data by public institutions, situating these within the on-going criminalization and surveillance of low-income communities, people of colour and other targeted groups. As a result, they have developed a set of guidelines for equitable practices in collecting, disseminating and using data. The US/ Canadian Environmental Data & Governance Initiative EDGI has preserved vulnerable scientific data in the aftermath of the US election of Trump in 2016, and in the process has developed an 'environmental data justice' framework that considers the politics, generation, ownership and uses of environmental data. Themes pertaining to data justice are also prevalent in the growing 'platform cooperativism' movement that sets out to challenge the nature of business ownership and governance emerging under platform capitalism, building on the values of cooperativism to create a fairer future of work in a digital economy. Such themes also inform a growing mobilisation towards more citizen-centered data infrastructures in public governance structures, such as the visions expressed in the 'Roadmap Towards Technological Sovereignty' outlined by the local administration in Barcelona.

These varied approaches and interpretations of how data and social justice relate illustrate the potential richness, but also the complexity with bringing these notions into dialogue. Indeed, in light of the centrality of data in contemporary forms of capitalism and power asymmetries, a notion of data justice could seem instinctively like an oxymoron. It might therefore make little sense to position it as an end-goal, something that we should be striving to achieve. At best, this would merely serve to legitimize and strengthen fundamentally unjust social structures that need to be overturned before we can begin to conceptualise forms of 'just data'. Instead, we should use data justice as a form of critique, a framework for shifting the entry-point and debate on data-related developments in a way that foregrounds social justice concerns and ongoing historical struggles against inequality, oppression and domination. The question remains whether data infrastructures can ever be extracted and redirected from the current conditions of injustice. For others, however, holding on to a substantial notion of data justice

as an ideal serves as a fruitful avenue to seek reforms that can better uphold justice claims pertaining to particular contexts. On this reading, data justice can advance principles and practices that emerge from existing conditions of possibility in order to facilitate processes that may lead to emancipatory outcomes. The aim of data justice in this sense, therefore, is to pinpoint where and how changes in data developments need to come about.

In bringing together scholars, practitioners and activists under the thematic umbrella of data justice, we inevitably come to see that as a concept, field of inquiry, set of practices or approach, data justice lends itself to different strands of analysis. Moreover, as a theme it incorporates a range of topics that in different ways grapple with the societal transformations that are associated with datafication and the implications these have on people's lives. It is no surprise, therefore, that the Data Justice conference featured presentations on a range of issues from social welfare to policing to migration to labour to activism (to name a few). As an integrated part of contemporary power relations, it is also no surprise that frameworks for analyzing such transformations and implications of datafication included ethics, law, sovereignty, feminism, and decoloniality amongst others. The scope of these topics illustrates the extent to which data as a social justice issue, in whatever form, stretches across social categories and types of knowledge and action. In essence, it exposes the need for a much more comprehensive dialogue on what is actually at stake with datafication, and a much wider range of stakeholders involved in asserting the nature of both challenges and possibilities than we have had so far.

With this special issue we have sought to capture some of the diversity and nuances that were prevalent at the conference and, together, form the debate on data justice. The eight articles that are collected for the issue were all presented at the conference and speak to some of the key themes that were discussed. They also provide a range of approaches and understandings of both the possibilities and challenges of the concept and how it may be applied in different contexts. Encompassing both conceptual and empirical work from across the globe, the articles help us advance the meaning of data justice, the kinds of debates and sites of study such a notion invites, and steer us towards future directions of research.

In the article from Seeta Peña Gangadharan and Jedrzej Niklas we are invited to reflect on the positioning of data in discussions on data justice and ways in which this is reflected amongst European civil society groups. Based on interviews with different human rights organisations,

the article argues for a 'decentering' of technology, drawing on insights from Nancy Fraser's notion of 'abnormal justice' as a way to understand technology's role in the production of social inequalities and the interconnections between maldistribution, misrecognition and misrepresentation. Human rights organisations frequently 'see through' technology, Gangadharan and Niklas argue, understanding technological developments as connected to traditional forms of injustice. Importantly, the article therefore draws our attention to the way long-standing experiences of discrimination and inequality need to be foregrounded in our engagement with technology, highlighting how technology is situated in relation to larger systems of institutionalised oppression. In making this point, we are asked to reflect on the limitations of focusing on the technology as the central component of contemporary injustices and the fallacies of therefore proposing technological or data-centric solutions to such injustices. In light of a growing emphasis on 'bias' and 'fairness' as computational concerns, and a prevalent adoption of 'responsible' data governance as a way to contend with disparities in impact, this contribution also serves as a pertinent reminder of the dangers of ahistorical and disassociated engagements with data issues, even within a framework of data justice.

Indeed, with Anna Lauren Hoffmann's contribution we are provided with further insights into the lessons we can learn from long-standing struggles around issues such as fairness and discrimination, pre-dating prominent recent focus on the way data-driven technologies might further discrimination and unfair treatment of certain groups. Outlining the limitations of a liberal rights-based discourse, as evidenced in US anti-discrimination law, Hoffmann's article calls on us to avoid repeating fundamental problems with such frameworks when applying them to new contexts involving data-driven technologies. Thinking about bias and discrimination as isolated instances, as one-dimensional issues that can only incorporate a single form of identification, or as something that can be addressed through distributive means, all ignore the structural conditions that underpin the problems that have been brought to light in the context of data systems. Moreover, Hoffmann asserts the importance of the role of data and algorithms in the production of particular kinds of meaning, reinforcing certain discursive frames over others. In this, she points to the need for frameworks engaging with data justice to also account for the normalisation and production of systematic advantage, resisting the temptation to think there is an easy 'fix' to these problems. Such thinking, she asserts, will at best leave us with "little more than a set of reactionary technical solutions that ultimately fail to displace the underlying logic that produce unjust hierarchies of better and worse off subjects in the first place."

Taking us from an engagement with discourses around data and discrimination, the contribution from Silvia Masiero and Soumyo Das moves us towards an engagement with the ways in which data is integrated into systems of governance, and with what social justice implications. Their article sheds light on the controversial dynamics of datafication within social protection schemes designed specifically for poor people. Drawing on the incorporation of Aadhaar - India's biometric population database - in the national agenda for social protection, Masiero and Das identify a techno-rational perspective that frames Aadhaar as a means to enhance the effectiveness of anti-poverty schemes. This vision is contrasted with the narratives of the beneficiaries of India's Public Distribution System collected through several interviews carried out in various Indian urban and semi-urban settings. Their analysis illustrates three problematic forms of data injustice - legal, design-related and informational - that were not in place before datafication and that are underpinned by Aadhaar's functionality with a shift of the social protection agenda from in-kind subsidies to cash transfers. Based on their research, Masiero and Das argue for a politically embedded view of data, showing how data is shaped by specific choices that can have multiple, potentially adverse implications for anti-poverty programme recipients. What emerges is a multifaceted and contested picture of datafication from the Global South, with data being used as a force that contributes to profoundly reform existing anti-poverty schemes, urging researchers to critically examine the political design behind datafied forms of social protection.

Continuing the discussion on how data is transforming social protection and welfare, Sora Park and Justine Humphry provide case studies from Australia that illustrate how exclusion can be embedded in the design and implementation of social welfare technologies and services. Drawing on the cases of Australia's Centrelink automated Online Compliance Intervention System and the government's National Disability Insurance Agency's (Nadia) efforts to develop an 'intelligent' avatar interface for users, their study demonstrates how punitive policies and a disciplinary logic were embedded in the design of Centrelink's system. They show how even when effort is taken to enhance inclusion through user participation in design development, as with Nadia, such efforts can be undermined at the stage of implementation. With Park and Humphry's article, therefore, we can see a continuation of previous debates on the digital divide, advancing the understanding of this by detailing how important it is that we recognize the ways that automated systems are making exclusion and inequality worse. This is a recognition that becomes increasingly important the more these systems are used to make

judgements and decisions about users and also mediate access to services and benefits. With the examples from these very different national contexts, India and Australia, we can see how data justice as a theme engages with certain transcending logics and practices that are prevalent across different parts of the world in very different manifestations.

With Dorothy Kidd's contribution we move to thinking about data's role not just in governance, but also in resistance, playing a central role in on-going social justice struggles. Her article puts data justice into historical context, reminding us that data collection and control is key to the European imperialist project of resource extraction and colonization. Resistance to such data control has a long history. As she notes, contests over maps are one of the 'longest-running examples of data activism'. Kidd's article details Indigenous resistance through the use of counter-mapping. She locates counter-mapping as an effective strategy, not on its own, but when tied to long-term political organizing. Her article demonstrates how essential organization is to data justice by detailing how counter-mapping has been used to mobilize. Her study details the ways Indigenous uses of counter-mapping in the 1970s contributed to new collective imaginaries and identities. She discusses how counter-mapping is being used more recently as part of a 'trans-media approach' to contest extractivism, provide a de-colonial education, circulate information, foster networks and work toward Indigenous sovereignty. Her study provides a real world illustration of how those who own or control data can use it to exert power and control but also how activists are using data to challenge this power and work toward social transformation. With this contribution, we are asked to consider how the cases of Indigenous counter-mapping provide important lessons about the need for a more comprehensive understanding of data justice, one that sees it as part of ongoing efforts for redistributive, transformative and restorative justice.

Furthering the theme of how data may advance an engagement with different forms of injustice, Jonathan Gray provides a contemporary example of this with his article, in which he scrutinizes how Amnesty International's practices of documenting and responding to abuses have been extended, modified and redistributed through data and digital technologies. Gray examines Amnesty's Decoders initiative that merges two tenets of the organisation's mission: documentation and volunteer-driven mobilisation. In order to characterize how data is involved in attending to situations of injustice within this kind of project, Gray introduces the notion of 'data witnessing' as a collective, distributed accomplishment. In contrast to accounts which place emphasis on individual forms of witnessing, with Gray's contribution, we are presented

with an illustration of how the work of Amnesty Decoders involves a choreography of human and non-human actors to attend to the systemic scale of injustices at a distance across time and space. Data witnessing encodes and makes possible diverse social, cultural and political approaches to injustice through the creation of media objects such as structured databases, maps, visualisations, and machine learning algorithms. These projects, Gray argues, can be regarded as experimental apparatuses for witnessing situations of injustice with data, revealing new dynamics of data politics and data activism, and suggesting innovative ways in which care, concern and solidarity may be built and structured by digital data, even if the projects do not go quite far enough in shifting the focus from individual to systemic injustices. This suggests that the possibilities of data witnessing remain very much open to experimentation, discussion and research.

The final two articles explore the practical implementation of data justice principles, while continuing a conversation with other contributions to this special issue on themes such as extraction, activism, and data governance. Richard Heeks and Satyarupa Shekhar investigate the role of data within international development through a focus on marginalised urban communities. Analysing four mapping initiatives in cities of the global South, they trace how data is captured about residents living in slums and other informal settlements. As a way of making these residents and their problems visible to policymakers, Heeks and Shekhar note that such data collection can deliver gains for target communities. Yet, and in accordance with other contributions in this special issue, they find that the context of datafication is crucial for understanding opportunities and challenges. Their research shows how external actors and wealthier communities generate more substantial gains which, in turn, increases rather than diminishes relative inequality; how 'datafied' communities may lose control of their own representation; and how data that might challenge political elites or hold them accountable remains largely invisible. They conclude that legibility is thus ambivalent. In exploring these case studies, Heeks and Shekhar apply an analytical framework that connects procedural, instrumental, rights-based, distributive and structural dimensions. They thus situate their work within a concern for the relationship between data and structural inequality and propose a means of systematically analysing data justice in development contexts.

Moving from international development to environmental science, the contribution by Vera, Walker, Murphy, Mansfield, Siad and Ogden explores the role of data justice in relation to environmental data collection and the concept of environmental justice. Grounded in a critical

examination of the extractive logics of datafication, their contribution develops and discusses the concept of 'environmental data justice'. The authors engage this approach by examining how it informs the work of the Environmental Data & Governance Initiative (EDGI), a distributed organization that formed in response to the 2016 U.S. presidential election and the resulting threats to national environmental programmes. Through grassroots archiving of data sets, monitoring federal environmental and energy agency websites, and writing rapid-response reports about how federal agencies are being undermined, EDGI mobilises environmental data justice to explore constructive data uses and create new data infrastructures that are participatory and embody equitable, transparent data care. In their dual role as both scholars and participants, the authors demonstrate the merits but also the tensions of environmental data justice in reflecting on a case of EDGI's public advocacy against changes proposed by the U.S. Environmental Protection Agency. They bring an engaged and, at the same time, self-reflexive perspective into the exchanges of this special issue that presents data justice as a practice with both promises and shortcomings and as a path towards a more just form of negotiating datafication.

Together, the contributions to this special issue therefore constitute a rich polyphony of engagements, understandings and interpretations around the notion of data justice. They make a strong case for using this concept as a powerful lens to illuminate the contradictions, the challenges, along with the social, political, economic and cultural implications associated with the process of datafication. They do so providing empirically grounded analyses that display the diversity of contexts in which data are experienced and processed, and explore the wide range of settings where varied entanglements between human and non-human actors contribute to shaping our datafied futures. Finally, they lead the way in stimulating dialogue and developing research around the pressing injustices of our time, beginning to trace the conceptual and empirical horizons of how social justice can be advanced in a datafied society.

References

Costanza-Chock, S. (2018). Design Justice, A.I., and Escape from the Matrix of Domination. *Journal of Design and Science.* n.p.

Dencik, L., Hintz, A., & Cable, J. (2016). Towards data justice? The ambivalence of anti-surveillance resistance in political activism. *Big Data & Society, 3*(2), 1–12. doi:10.1177/2053951716679678

Dencik, L. Hintz, A., Redden, J., and Warne, H. (2018). *Data Scores as Governance: Investigating uses of citizen scoring in public services*. Project Report. Open Society Foundations and Cardiff University. Available at: https://datajustice.files.wordpress.com/2018/12/data-scores-as-governance-project-report2.pdf

Gray, J. (2018). Three Aspects of Data Worlds. *Krisis: Journal for Contemporary Philosophy*. Issue 1.

Heeks, R. (2017). A Structural Model and Manifesto for Data Justice for International Development. *Development Informatics Working Paper Series*, No. 69.

Hintz, A., Dencik, L. and Wahl-Jorgensen, K. (2018). *Digtial Citizenship in a Datafied Society*. Cambridge: Polity Press.

Johnson, J. A. (2018). *Toward Information Justice*. Springer International Publishing.

Mayer-Schönberger, V. and Cukier, K. (2013). *Big Data: A Revolution That Will Transform How We Live, Work and Think*. New York: John Murray.

Milan, S. and Treré, E. (forthcoming 2019). Big Data from the South(s): Beyond Data Universalism. *Television and New Media*.

Milan, S. and van der Velden, L. (2016). The Alternative Epistemologies of Data Activism. *Digital Culture & Society,* 2(2), pp. 57-74.

Redden, J. and Brand, J. (2017). *Data Harm Record*. Available at: https://datajustice.files.wordpress.com/2017/12/data-harm-record-djl2.pdf

Taylor, L. (2017). What is data justice? The case for connecting digital rights and freedoms. *Big Data & Society*, 4(2).

Treré, E. (2019). *Hybrid Media Activism: Ecologies, imaginaries, algorithms*. London and New York: Routledge.