Researching the Human Rights Impact of New and Emerging Digital Technologies

Lorna McGregor

1. Introduction

In the last decade, academics, civil society organisations and international and regional organisations have published research on the human rights impact of new and emerging digital technologies. The types of technologies falling under this label are manifold, ranging from internet connectivity to 'artificial intelligence'. For example, the UN Human Rights Council Advisory Committee employed the term new and emerging digital technologies 'to refer to technological innovations that transform the boundaries between virtual, physical and biological spaces. They include new technologies and techniques for datafication (the process of transforming subjects, objects and practices into digital data), data distribution and automated decision-making, such as artificial intelligence, the Internet of things, blockchain technology and cloud computing, among others'.[1] Artificial intelligence itself constitutes a broad term with no settled definition. For example, in its Recommendation on the Ethics of Artificial Intelligence, the UN Educational, Scientific and Cultural Organization (UNESCO) explains that it 'does not have the ambition to provide one single definition of AI, since such a definition would need to change over time, in accordance with technological developments … [it] approaches AI systems as systems which have the capacity to process data and information in a way that resembles intelligent behaviour, and typically includes aspects of reasoning, learning, perception, prediction, planning or control'.[2]

Layers of opacity have presented a central obstacle to research on the human rights impact of new and emerging digital technologies. One dimension to opacity relates to the oft cited 'black box' problem of seeing and understanding the inner workings of technologies such as machine learning algorithms.[3] However, opacity extends beyond how a technology works. In contrast to other areas in which the human rights effects of a policy or practice are either immediately

[1] UN Human Rights Council, 'Possible impacts, opportunities and challenges of new and emerging digital technologies with regard to the promotion and protection of human rights: Report of the Human Rights Council Advisory Committee' A/HRC/47/52 (19 May 2021), at §3.
[2] UNESCO, *Recommendation on the Ethics of Artificial Intelligence* (23 November 2021) at 10.
[3] Jennifer Cobbe, 'Administrative Law and the Machines of Government: Judicial review of automated public-sector decision-making' 39 *Legal Studies* 636 (2019) at 639.

visible or made visible through human testimony, the power and control of businesses and states over new and emerging digital technologies have rendered many aspects of their use opaque. This includes the fact of their employment at all, the reasons for their uptake and the consequences of their integration within wider political and social systems, including in decision-making in critical areas of people's lives. Such opacity can render it difficult for individuals and groups affected by technological use to exercise their right to a remedy by meaningfully challenging the role of technology in wider decisions about them and for oversight, monitoring and regulatory bodies to hold actors to account for human rights harm in individual cases or systemically. It can also create obstacles to the development of effective governance frameworks to prevent future human rights harm as the design of such frameworks flows from the characterisation of the problem(s) to be governed.

Despite these obstacles, in both academia and practice, human rights researchers have developed a core body of research demonstrating the human rights effects of the design, development and deployment of new and emerging digital technologies. Some of this research recognises the potential for new and emerging digital technologies to advance human rights. This is most clearly illustrated by research on the need to overcome digital divides[4] and on the role of disaggregated data in making systemic patterns of discrimination and inequality visible and therefore actionable.[5] Claims are also made about the potential for AI-enabled technologies to facilitate the enjoyment of human rights. For example, researchers, including this author, posit that certain technologies could play a role in advancing the right to independent living and to autonomous decision-making for older people and people with disabilities.[6] However, this research also recognises the thin evidence-base to support such claims at present; the simultaneous risk to other human rights; and the current structural power disparities which mean that it is rarely rights-holders who design – or influence the design of - new and emerging

[4] Daniel Paré, 'The Digital Divide: Why the 'The' is Misleading' in Mathias Klang and Andrew Murray (eds), *Human Rights in the Digital Age* (2005); UN, 'Our Common Agenda Policy Brief 5: A Global Digital Compact – an Open, Free and Secure Digital Future for All' (May 2023) at 6.

[5] *See,* for example, UN General Assembly, 'Report of the Working Group of Experts on People of African Descent on its twenty-third and twenty-fourth sessions' A/HRC/42/59 (15 August 2019); UN Human Rights Council, 'Human rights of older persons: the data gap: Report of the Independent Expert on the enjoyment of all human rights by older persons', A/HRC/45/14 (9 July 2020).

[6] Neil Crowther and Lorna McGregor, *A Digital Cage is Still a Cage: How can new and emerging digital technologies advance, rather than put at risk, the human rights of older people who draw on social care?* Human Rights, Big Data and Technology Project Research Paper (June 2022); Donald Macaskill, 'TechRights: Human Rights, Technology and Social Care', *Scottish Care* (2018); UN Human Rights Council, 'Report of the Independent Expert on the enjoyment of all human rights by older persons' A/HRC/36/48 (21 July 2017); UN Human Rights Council, 'Report of the Special Rapporteur on the rights of persons with disabilities' A/HRC/49/52 (28 December 2021).

digital technologies or make the choice about whether and how technologies feature in their lives.[7] In this regard, a larger body of research focuses on the human rights risks presented by new and emerging digital technologies. While earlier research highlighted risks to the rights to privacy, non-discrimination and freedom of expression,[8] it has now widened to recognise the potential harm to all human rights depending on the type of technology, the purpose(s) of use, the context in which it is deployed and the governance framework in place, particularly with regard to red-lines, the application of the legality, necessity and proportionality test, and the adequacy and effectiveness of procedural safeguards, accountability measures and remedies.[9]

This shift in emphasis from a narrow focus on technologies to the purpose(s) and contexts in which they are used holds the potential to connect with research in related fields, such as critical data studies,[10] which reject the abstraction of technologies from their societal context in preference for the study of technologies as part of socio-technical systems or assemblages.[11] At the same time, human rights research has not fully embodied this conceptual framing and

[7] Crowther and McGregor ibid at 84-85.
[8] For example, Article 19 and Privacy International, *Privacy and Freedom of Expression in the Age of Artificial Intelligence* (2018); Anja Seibert-Fohr, 'Digital Surveillance, Meta Data and Foreign Intelligence Cooperation' in Joseph E. David, Yaël Ronen, Yuval Shany and J.H.H. Weiler (eds), *Strengthening Human Rights Protections in Geneva, Israel, the West Bank and Beyond* (2021); UNGA, 'Promotion and protection of the right to freedom of opinion and expression: Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression' A/73/348 (29 August 2018) (discussing the impact on the rights to freedom of opinion, freedom of expression, and privacy and the 'obligation of non-discrimination' in relation to artificial intelligence'); UK House of Commons and House of Lords, 'Joint Committee on Human Rights' *The Right to Privacy (Article 8) and the Digital Revolution* (30 October 2019) (focusing on privacy and non-discrimination); Sarah West, Meredith Whittaker, and Kate Crawford, (2019). *Discriminating Systems: Gender, Race and Power in AI*, AI Now Institute (2019); Mark Latonero, *Governing Artificial Intelligence & Upholding Human Rights & Dignity* Data and Society (2018) (emphasising nondiscrimination and equality, political participation, privacy, freedom of expression and disability rights approach and accessible design as a human rights frame).
[9] For some of the initial reports, *see,* Access Now, *Human Rights in the Age of Artificial Intelligence* (2018); Lorna McGregor, Vivian Ng, Ahmed Shaheed et al, *The Universal Declaration of Human Rights at 70: Putting Human Rights at the Heart of the Design, Development, and Deployment of Artificial Intelligence* (2018).
[10] Andreas Hepp, Juliane Jarke and Leif Kramp, 'New Perspectives in Critical Data Studies: The Ambivalences of Data Power – An Introduction' in Andreas Hepp, Juliane Jarke, Leif Kramp (eds), *New Perspectives in Critical Data Studies: The Ambivalences of Data Power* (2022) at 4 (discussing the origins of the term 'critical data studies' and its development by other academics and explaining it as 'concerned with the significance (and power) of digital data in contemporary society and how it relates to societal transformation').
[11] Rob Kitchin, 'Thinking Critically About and Researching Algorithms' 20 *Information, Communication & Society* 14 (2017) at 25; Andrew Iliadis and Federica Russo, 'Critical data studies: An introduction' *Big Data & Society* 1 (2016), at 3; Nick Couldry, 'Recovering critique in an age of datafication', 22 *New Media & Society* 1135 (2020), at 1141; *See also,* literature on regulation and technology emphasising to focus on actors rather than technology as the subject of regulation. For example, Lyria Bennett Moses, 'How to Think about Law, Regulation and Technology: Problems with 'Technology' as a Regulatory Target' 5 *Law, Innovation and Technology* 1 (2013). *See also,* similar discussions in the field of human-computer interaction, Michael Katell, Meg Young, Dharma Dailey, Bernease Herman, Vivian Guetler, Aaron Tam, Corinne Binz, Daniella Raz, and P. M. Krafft. 2020. Toward Situated Interventions for Algorithmic Equity: Lessons from the Field. In Conference on Fairness, Accountability, and Transparency (FAT* '20), January 27–30, 2020.

as a result has not engaged in extensive analysis of the power structures and political economy that drive technological uptake and facilitate the production of human rights harm. Particularly with the view to the prevention of future human rights harm,[12] in this chapter, I argue that the deepening of human rights research to capture not only the human rights effects of new and emerging digital technologies but also the structural causes of human rights harm is critical to thickening understanding of the nature of the problem to be governed. From a methodological perspective, this chapter therefore proposes greater engagement with other fields and disciplines, such as critical data studies, which have long interrogated these questions.[13]

Expansions in the framing of human rights research will nevertheless continue to experience tensions with varying forms of opacity. However, innovative research design, including through the employment of multi-method approaches that integrate new and traditional human rights research methods can facilitate the assemblage of knowledge from disparate sources and thus create friction with layers of opacity and power imbalances in the design, development and deployment of new and emerging digital technologies. Such research can contribute to efforts for accountability and remedies as well as provide a critical lens to assessments of the adequacy and effectiveness of existing and proposed governance frameworks which themselves are unlikely to be static or contained within one approach, but part of an evolving ecosystem.[14]

While this chapter focuses on the thickening of human rights research on new and emerging digital technologies that have already been deployed, the challenges associated with researching existing systems without clear governance frameworks in place also points to the importance of forecasting technologies likely to be developed and deployed in order to

[12] Much of this research has focused on the contribution of human rights principles to the governance of new and emerging digital technologies. *See,* for example, Lorna McGregor, Daragh Murray and Vivian Ng, 'International Human Rights Law as a Framework for Algorithmic Accountability' 68 *International and Comparative Law Quarterly* 309 (2019); Kate Jones, 'AI Governance and human rights: Resetting the relationship' *Chatham House* (January 2023); Karen Yeung, Andrew Howes, and Ganna Pogrebna, 'AI Governance by Human Rights Centred-Design, Deliberation and Oversight: An End to Ethics Washing' in Markus Dubber, Frank Pasquale and Sunit Das, *The Oxford Handbook of Ethics of AI* (2020); Dafna Dror and Yuval Shany, 'It's the End of the (Offline) World as We Know It: From Human Rights to Digital Human Rights – A Proposal Typology' 32 *European Journal of International Law* 1249 (2021); Alessandro Mantelero, 'AI and Big Data: A blueprint for a human rights, social and ethical impact assessment' 34 *Computer & Law Security Review* 754 (2018); Lottie Lane, 'Clarifying Human Rights Standards through Artificial Intelligence Initiatives' 71 *International and Comparative Law Quarterly* 915 (2022).

[13] *See,* Part 3(B) below.

[14] *See,* Gary E. Marchant, 'Governance of Emerging Technologies as a Wicked Problem' 73 *Vanderbilt Law Review* 1861 (2020) AT 1862 (arguing that '[n]o single optimum solution exists, but rather a collection of second-best strategies intersect, coexist, and – in some ways – compete').

facilitate the prevention of human rights harm. This point is not meant as the promotion of existentialist research that diverts political attention and resource away from dealing with current human rights challenges[15] but rather proposes that the forecasting of near-term technological design and deployment may avoid reactiveness behind the curve of technological roll-out which can be difficult to reverse.[16] Although space prohibits a detailed discussion of this proposition, I include discussion of forecasting as part of a future research design agenda focused on prevention of future human rights harm.

This chapter first identifies the barriers to human rights research produced by layers of opacity into the design, development and deployment of new and emerging digital technologies. Second, notwithstanding these barriers, the chapter maps the substantial body of research on the human rights effects of new and emerging digital technologies but suggests that human rights research into the drivers of technological uptake remains underdeveloped. It proposes a more central focus on structural causes of human rights harm, including through engagement with research in other fields, such as critical data studies. In the third part of this chapter, I discuss how a multi-methods approach to human rights research can help to assemble disparate pieces of knowledge into the human rights causes and consequences of the design, development and deployment of new and emerging digital technologies. The length of this chapter prohibits a detailed interrogation of all possible methods and their different configurations. In the space available, I first acknowledge and problematise the role of new and emerging digital technologies themselves as new methods for human rights research. In this regard, I highlight how such methodological engagement can overly fetishise[17] and exceptionalise research on new and emerging digital technologies, and thereby abstract such technologies from their social context, as well as undermine the role and relevance of traditional human rights research methods in this space. As such, I argue that while technological methods may have their place, research led by groups disproportionately impacted by the design, development and deployment of new and emerging digital technologies should be foregrounded. Moreover, I

[15] William Douglas Heaven, 'How existential risk became the biggest meme in AI' *MIT Technology Review* (19 June 2023); Tate Ryan-Mosley, 'It's time to talk about their real AI risks' *MIT Technology Review* (12 June 2023).
[16] *See,* Part 5 below.
[17] Morgan Currie, Jeremy Knox and Callum McGregor, 'Introduction: Data Justice and the Right to the City' in Morgan Currie, Jeremy Knox and Callum McGregor, *Data Justice and the Right to the City* (2022) at 1 (discussing the origins and meaning of 'data fetishism' as 'the habit of endowing objects and entities with an almost magical, or at least intrinsic, power to shape the world around us … a form of post-politics par excellence that casts aside theory and ideology and reduces various social problems to systems engineering conundrums solvable with sufficient data and processing capacity').

suggest that well-established human rights research methods, such as qualitative interviews and stakeholder meetings should remain central within multi-methods research in this field. Finally, reflecting on the initial reactiveness of human rights research both due to the newness of the field and the opacity constructed by powerful actors such as states and businesses, I highlight the importance of technological forecasting as part of human rights research design into new and emerging digital technologies.

2. Layers of Opacity as an Ongoing Barrier to Human Rights Research on New and Emerging Digital Technologies

As discussed in the introduction, layers of opacity have obfuscated the documentation of existing deployments of new and emerging digital technologies and have thus created obstacles to the identification of human rights harm and to accountability and redress. Opacity does not take one form.[18] In earlier literature, researchers employed opacity to refer to what is often framed as the 'black box' of machine-learning algorithms or other technological systems.[19] Within framings of opacity focused on technologies, Jennifer Cobbe identifies different layers of opacity as:

> 'The first is *intentional* opacity, where the system's workings are concealed to protect intellectual property. The second is *illiterate* opacity, where a system is only understandable to those who can read and write computer code. And the third is *intrinsic* opacity, where a system's complex decision-making process itself is difficult for any human to understand. More than one of these may combine – for example, a system can be intentionally opaque and it be the case that even if it wasn't then it would still be illiterately or intrinsically opaque. The result of algorithmic opacity is that an automated system's decision-making process may be difficult to understand or impossible to evaluate even for experienced systems designers and engineers, let alone non-technical reviewers. In many cases it will be virtually impossible to determine how or why a particular outcome was reached.'[20]

[18] Cansu Safak and Imogen Parker, 'Meaningful transparency and (in)visible algorithms' *Ada Lovelace Blog* (15 October 2020).
[19] Frank Pasquale, *The Black Box Society: The Secret Algorithms that Control Money and Information* (2015) (as one of the earlier framings of the 'black box' problem).
[20] Cobbe (3) at 639; *see also,* Kitchin (n11) at 20; Paul B. De Laat, 'Algorithmic Decision-Making Based on Machine Learning from Big Data: Can Transparency Restore Accountability' 31 *Philosophy & Technology* 525 (2018); Lilian Edwards and Michael Veale, 'Enslaving the Algorithm: From a "Right to Explanation" to a "Right to Better Decisions"?' *AI Ethics* 46 (2018).

While some researchers have argued that there is limited utility in trying to uncover how such algorithms function,[21] where they are used as part of a decision or process which produces human rights effects, the inability to interrogate how they did so could impair the essence of the human rights concerned as well as the right to an effective remedy. Accordingly, the opacity of technological systems remains an important line of inquiry for human rights researchers, including by raising questions of whether 'unexplainable' technologies should be used at all in decisions or processes which are likely to impact the enjoyment of human rights.

However, opacity is not limited to the inner workings of a specific technology. Other layers of opacity can create additional barriers for human rights researchers, as well as for individuals and groups affected and actors regulating their design, development and deployment. First, at the most basic level, states and businesses may fail to be transparent about the fact of their use of new and emerging digital technologies, including in some of the most serious decisions about people's lives. This can stem from a lack of an explicit legal basis and process for deliberation on the proposed introduction of a particular technology.[22] In this regard, states and businesses do not typically announce – particularly ahead of time – their plans to integrate technological systems, such as predictive analytics, into decision-making.[23] Thus, unless physically visible – such as surveillance systems employed at borders[24] or facial recognition cameras in public spaces[25] – individuals and groups may not know that technological systems are being used in ways which may impact their human rights. For example, in their study of the use of new and emerging digital technologies in the child welfare system in the UK, Joanna Redden et al. observe that data subjects are not directly notified, but rather public sector bodies rely on 'duty of care responsibilities' to justify data processing in situations such as 'protect[ing] someone in an emergency, and to fulfil their legal obligation to safeguard or

[21] Lilian Edwards and Michael Veale, 'Slave to the Algorithm? Why a 'Right to an Explanation' is Probably Not the Remedy you are Looking For' 16 *Duke Law & Technology Review* 18 (2018); Kitchin (n11) at 21; Mike Ananny and Kate Crawford, 'Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability' 20 *New Media & Society* 973 (2018), at 982; Joshua A. Kroll, Joanna Huey, Solon Barocas, Edward W. Felten, Joel R. Reidenberg, David G. Robinson and Harlan Yu, 'Accountable Algorithms' 165 *University of Pennsylvania Law Review* (2016).
[22] UN General Assembly, 'Extreme poverty and human rights: Report of the Special Rapporteur on extreme poverty and human rights' A/74/493 (11 October 2019) at §43.
[23] Lina Dencik, Arne Hintz, Joanna Redden and Harry Warne, *Data Scores as Governance: Investigating uses of citizen scoring in public services: Project Report'* Data Justice Lab (December 2018) at 14.
[24] Mijente, Just Futures Law & No Border Wall Coalition, *The Deadly Digital Border Wall* (2021) at 11.
[25] Pete Fussey, Bethan Davies and Martin Innes, ''Assisted' Facial Recognition and the Reinvention of Suspicion and Discretion in Digital Policing' 61(2) *The British Journal of Criminology'* 325 (2021).

support the wellbeing of children and young people' and serve notice through 'general privacy notice[s] published online'.[26] They note that one justification for not providing direct notification is that 'releasing details about the system may prejudice potential interventions and compromise the commercial interests of the company involved.'[27] However, it may also impact the ability of the subjects of such data processing to challenge its role in reaching decisions about their lives.

Second, in other contexts, some level of awareness may exist about the deployment of new and emerging digital technologies. However, insufficient detail on the reasons for their introduction, the precise type of technology or how they work,[28] the 'human systems around the technology',[29] and the arguments for its introduction (including over other technological and non-technological models) can impede doctrinal analysis of the compatibility of technological deployment with human rights principles, such as the pursuit of a legitimate aim and the necessity and proportionality of the technology. Moreover, despite the frequent claim that new and emerging digital technologies support evidence-based decision-making, a lack of transparency can make it difficult to identify the evidence supporting the deployment of such technologies.[30] For example, the UN Special Rapporteur on the Right to Education has noted that '[a]s debates around the digitalization of education tend to focus on the effectiveness of methods, tools and strategies – often in search of evidence-based solutions of "what works" – it is important to emphasize the lack of evidence and contextualised evaluation supporting the claimed added value of digital technologies in many respects'.[31]

Third, further layers of opacity can attend to the governance systems surrounding socio-technical assemblages. These layers of opacity may make decisions involving technologies difficult for an individual affected to understand and act upon.[32] They may also impede 'expert

[26] Joanna Redden, Lina Dencik and Harry Warne, 'Datafied child welfare services: unpacking politics, economics and power' 41 *Policy Studies* 507 (2020) at 516.
[27] Ibid at 516. See also Ryan Calo and Danielle Keats Citron, 'The Automated Administrative State: A Crisis of Legitimacy' 70 *Emory Law Journal,* 797 (2021).
[28] Lorna McGregor, 'Regulating Digital and AI Technologies: Lessons from the Digitisation of Contact Tracing during the COVID-19 Pandemic' 3 *Yearbook of International Disaster Law* 35 (2022) at 65-66.
[29] Margot Kaminski, 'Understanding Transparency in Algorithmic Accountability' in Woodrow Barfield (ed), *The Cambridge Handbook of the Law of Algorithms* 121 (2020) at 127.
[30] Karen Yeung, 'The New Public Analytics as an Emerging Paradigm in Public Sector Administration' 27(2) *Tilburg Law Review* 1 (2022) at 11 (discussing the turn to datafication in the public sector).
[31] UN Human Rights Council, 'Impact of the digitalization of education on the right to education' A/HRC/50/32, (19 April 2022), at §5.
[32] Margot Kaminski (n29), at 129.

oversight' or systemic analysis of 'error, bias, and discrimination in both machine and human systems … These information flows go to more than one person or actor'.[33]

As discussed in the next part of this chapter, a strong body of research has developed on the human rights effects of new and emerging digital technologies. However, the extent and specificity of such research, particularly empirical, is severely impacted by the power of state and private actors to obscure and control what information reaches the public domain and is thus researchable.[34] Where opacity constitutes a central blockage, at least with regard to public sector use, regulation requiring publication of technologies already under consideration and adopted may aid greater transparency. As such, some actors have called for audits and public registers of new and emerging digital technologies in addition to the publication of data protection and human rights impact assessments at different stages in the consideration of the adoption of a technology and once it is deployed.[35]

However, the content of audits, registers and impact assessments will dictate the level of insight and utility of their publication. For example, they will be of limited value where they do not specify the type of problem or issue that the actor, such as a state agency, seeks to address; include assessment of why a particular category of technology meets its objectives, is necessary, and has the least invasive impact on human rights when compared to other technological and non-technological approaches; set out the safeguards in place to mitigate against harm; and explain the evidence-base both for the assertion that a problem exists that requires resolution and that technology offers the optimal vehicle for doing so. Publication of impact assessments can also have limitations where they do not specify the actor providing the technology (whether through financial arrangement or as a donation) and the other actors that will be involved in its operation. Accordingly, regulatory approaches may assist in addressing opacity, however they will not necessarily solve the full extent of the challenges faced, meaning that opacity may continue to stand in tension with human rights research. As such, while in the following section I discuss the importance of expanding the framing of human rights research,

[33] Ibid at 129.
[34] Iliadis and Russo (n11) at 1 (arguing that '[d]ata are a form of power … a lack of data is another indication of power, the power not to look or to remain hidden. In their presence and absence, data are always-already active and never neutral').
[35] Swee Leng Harris, ''Data Protection Impact Assessments as rule of law governance mechanisms' 2 *Data & Policy* 1 (2020); Alessandro Mantelero, *Beyond Data: Human Rights, Ethical and Social Impact Assessment in AI* (2022); 'Our Response to the European Commission's Consultation on AI', Algorithm Watch, 12 June 2020, available at https://algorithmwatch.org/en/response-european-commission-ai-consultation/

in the final part of this article, I examine how this can be achieved through the adoption of a multi-methods approach to chip away at remaining layers of opacity as a means to continue to assemble new knowledge on the human rights impact of new and emerging digital technologies. As an iterative process, the encountering of barriers to research as a result of opacity should not be overlooked. Rather, they should be treated as a research finding and systematically documented in order to continue to challenge manifestations of opacity which create obstacles to effective human rights protection.

3. Thickening the Framing of Human Rights Research Aimed at the Prevention of Future Harm

Notwithstanding the barriers to research presented by the layers of opacity discussed in the previous section, a dominant trend in human rights research has been to analyse the human rights effects of the design, development and deployment of new and emerging digital technologies either by assessing the technology in the abstract or by analysing interferences with human rights where actors employ a technology in a particular context. This research has been central to the engagement of the international human rights system and for making the case for a human rights-based approach to the governance of new and emerging digital technologies. However, in this chapter, I argue that gaps remain in the structural analysis of the drivers for technological uptake that are critical to address, particularly with the view to preventing future harm. As such, after mapping current research on the human rights effects of new and emerging digital technologies, I propose greater engagement with related fields in which the power structures and political economy to new and emerging digital technologies have been the subject of longstanding critical analysis.

A. Assessing the human rights effects of new and emerging digital technologies

A substantial body of current research into the human rights effects of the design, development and deployment of new and emerging digital technologies adopts a case study approach from one of two angles. First, research draws on studies in other fields or directly demonstrates how inaccurate technologies, technologies trained on or reliant upon discriminatory input data (through over or underrepresentation of particular groups within data sets), and the weighting of data points can produce discriminatory outcomes. For example, ground-breaking research in academia and in practice has shown how different face recognition models have high levels

of inaccuracies on the faces of non-white males.[36] Similarly, research on emotion recognition technologies highlights the lack of scientific basis for the technologies as well as 'a weak association of emotions with facial expressions … facial expressions vary across cultures and contexts, making emotion recognition susceptible to bias and misinterpretations'.[37] In both cases, researchers underscore the risks that these technologies result in discriminatory stop and search, deprivation of liberty and impinge on the right to a fair trial.[38]

Further, research illustrates how the design and development of machine learning algorithms, such as algorithmic risk assessments, can result in discriminatory findings, which may be extended through a lack of challenge or acquiescence by the human decision-maker relying upon the algorithmic assessment.[39] Such research plays a critical role in revealing the fundamental incompatibility of specific technological products with human rights due to their inaccuracies and/or discriminatory design and/or operation.

A second body of research analyses new and emerging digital technologies through the lens of a specific right or group of rights such as privacy or freedom of expression or thought.[40] For example, the UN Special Rapporteur on Freedom of Religion or Belief has examined how a range of technologies from 'content curation',[41] to microtargeting,[42] to predictive technologies such as polygraphs and emotion-recognition,[43] can 'affect how we think, feel and behave'. Others have built on this research to examine the human rights implications of new and

[36] Joy Buolamwini and Timnit Gebru, 'Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification 'Proceedings of Machine Learning Research, Conference on Fairness, Accountability and Transparency (2018); Clare Garvie, 'Garbage In, Garbage Out: Facial Recognition on Flawed Data', Georgetown Center on Privacy and Technology (16 May 2019); Pete Fussey and Daragh Murray, *Independent Report on the London Metropolitan Police Service's Trial of Live Facial Recognition Technology* (2019).
[37] UN Human Rights Council, 'The right to privacy in the digital age: Report of the United Nations High Commissioner for Human Rights' A/HRC/48/31 (13 September 2021) at §28 (citing wider research on these technologies, such as Lisa Feldman, Ralph Adolphs, and Seth Pollak, 'Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements' 20 *Psychological Science in the Public Interest* iii (2019)); *Article 19 Emotional Entanglement: China's emotion recognition market and its implications for human rights* (2021).
[38] OHCHR ibid.
[39] For a discussion of this point, *see,* Reuben Binns, 'Human Judgment in algorithmic loops: Individual justice and automated decision-making' 16 *Regulation and Governance* 197 (2022).
[40] For example, *see,* UN Human Rights Council, 'The right to privacy in the digital age: Report of the United Nations High Commissioner for Human Rights', A/HRC/48/31 (13 September 2021); Human Rights Council, The right to privacy in the digital age :Report of the United Nations High Commissioner for Human Rights', A/HRC/39/29 (3 August 2018); UN General Assembly, 'Interim Report of the Special Rapporteur on freedom of religion or belief, Ahmed Shaheed: Freedom of thought' A/76/380 (5 October 2021); EU Fundamental Rights Agency, *Getting the Future Right: Artificial Intelligence and Fundamental Rights* (2020), at 61.
[41] Ibid, at §67.
[42] Ibid. at §73
[43] Ibid at, §69.

emerging digital technologies when employed by particular actors in specific contexts. For example, a growing body of research now exists which shows the potential adverse human rights outcomes when digital technologies (individually and in combination) are deployed in sectors such as law enforcement,[44] education,[45] social care,[46] welfare,[47] and migration.[48] Much of this research focuses on deployment within the public sector in close partnership with private actors, for example, through public-private partnerships. Such analysis has enabled human rights researchers in academia and practice to highlight the potential effects of technologies on all human rights, individually and collectively, depending on why state and private actors design, develop and deploy such technologies in contexts in which human rights are already often at risk. For example, the deployment of new and emerging digital technologies may disproportionality affect people already in positions of marginalization through the targeting of new and emerging digital technologies at what are deemed 'expensive' public services.[49] The introduction of technologies in such contexts can therefore amplify existing forms of discrimination and experiences of marginalisation, surveillance and distrust in public services.[50]

Research into the human rights effects of new and emerging digital technologies plays a pivotal role in justifying the engagement of national, regional and international human rights systems and thus harnessing the accompanying standards and norms as well as institutions with investigative, monitoring, standard-setting, accountability and remedial mandates.[51] Squarely characterising the design, development and deployment of new and emerging digital technologies as generative of human rights issues is therefore critical in responding to any pushback that often attends to perceived expansions of human rights systems into new areas.[52]

---

[44] Pete Fussey and Daragh Murray (n36).
[45] Human Rights Council, 'Artificial intelligence and privacy, and children's privacy: Report of the Special Rapporteur on the right to privacy, Joseph A. Cannataci', A/HRC/46/37 (25 January 2021); UN Human Rights Council (right to education) (n31).
[46] Crowther & McGregor, Macaskill, Report of the Independent Expert on the enjoyment of all human rights by older persons (n6).
[47] UN General Assembly (Extreme poverty and human rights) (n22).
[48] Mijente et al (n24); Petra Molnar, 'Technology at the Margins: The Human Rights Impacts of AI in Migration Management' 8(2) *Cambridge Journal of International Law* (2019); Dimitri Van Den Meerssche, 'Virtual Borders: International Law and the Elusive Inequalities of Algorithmic Association', 33(1) *European Journal of International Law* 171 (2022).
[49] Kate Crawford and Jason Schulz, 'AI Systems as State Actors' 119 *Columbia Law Review* 1941 (2019).
[50] Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police and Punish the Poor* (2018).
[51] For a discussion of the varying functions of international human rights bodies, *see,* Laurence Helfer, 'Populism and International Human Rights Law Institutions: A Survival Guide' in Gerald Neuman (ed), *Human Rights in a Time of Populism: Challenges and Responses* (2020) at 218.
[52] Hurst Hannum, *Rescuing Human Rights: a Radically Moderate Approach* (2019) (discussing arguments against expansion).

Framing new and emerging digital technologies as a human rights problem thus provides a language of harm[53] as well as a language of governance,[54] theoretically at least constructing certain parameters to decisions on if and how to design, develop and deployment new and emerging technologies, including how to assess their potential impact, monitoring their use and account for, and remedy any resulting human rights violations.[55]

B. Expanding human rights research to examine the structural drivers enabling human rights harm

The previous section charted dominant approaches to human rights research which analyse the human rights effects of new and emerging digital technologies. As noted, this research has been central to engaging the international human rights system as well as providing the foundation for research on the applicability of international human rights standards and norms to the governance of new and emerging digital technologies. However, human rights standards and norms alone are unlikely to be able to effectively protect human rights. In this regard, understanding how and why these technologies have come to play such a central role in areas of life that affect people's enjoyment of their human rights is also a critical component to governing their design, development and deployment and to preventing future harm. While analysis of structural enablers is not a new approach in human rights research, it is only beginning to emerge in relation to new and emerging digital technologies in contrast to research in related disciplines and fields that has long located technologies within their social, economic and political context as socio-technical assemblages.[56]

In this regard, some researchers in fields such as critical data studies cast doubt over the utility of the human rights framework to addressing the impact of new and emerging digital technologies on society due to its perceived focus on individuals.[57] Researchers have

[53] See evelyn douek, 'The Limits of International Law in Content Moderation' 6 *UC Irvine Journal of International, Transnational, and Comparative Law* 37 (2021) (challenging this presentation).
[54] Barrie Sander, 'Democratic Disruption in the Age of Social Media: Between Marketized and Structural Conceptions of Human Rights Law' 32 *European Journal of International Law* 159 (2021), at 192.
[55] McGregor, Ng and Murray (n12) (setting out the role of international human rights law principles in governing algorithms).
[56] *See,* footnote 12.
[57] Institute for the Future of Work, *Equality Task Force: Mind the gap: How to fill the equality and AI accountability gap in an automated world* (2020) at 46; Currie et al (n17) at 14 (arguing that 'work in data justice is a challenge to dominant discourses, this time those in the guise of fairness, privacy and the protection of individual rights, which have tended to characterise broader discussions of the social impact of data-driven technologies).

highlighted concerns about how technological systems can emphasise individualism[58] both in terms of personalisation as well as through placing the burden on individuals to manage their data, for example, through consent-based models. Where human rights frameworks are also understood as focused on the individual and individualised harm, researchers have critiqued such a framing as overlooking the scale of harm produced by the use of these technologies[59] and the need to address the structural power of actors involved in their design, development and deployment.[60] Such assessments align with wider critiques of human rights. For example, Dominique Clément argues that framing a grievance as a human rights grievance denies the possibility of examining 'systemic social problem[s]' and inevitability directs resolution to 'legal avenues'.[61]

While one way of interpreting the human rights framework, it can first underplay the importance of capturing individual harm and harm to groups in positions of vulnerability, both in affording individuals whose rights have been violated a remedy and to making claims of technological harm less abstract, which can facilitate the denial or underplaying of harm by powerful actors. Second, it can overlook the evolution of the human rights framework from one which traditionally may have leaned towards a 'violation-violator-remedy' model,[62] but has long focused on analysing the systems and structures that can enable violations of economic, social and cultural rights moving beyond what Allison Corkery refers to as 'events-based methods' to methods to identify and analyse 'complex causal chains'.[63] Further, the human rights system increasingly assumes a wider systems lens as a means of preventing human rights violations[64] and in guaranteeing non-repetition as part of the right to an adequate and effective remedy.[65] In this regard, to guarantee non-repetition of a violation typically

[58] Lina Dencik, Arne Hintz and Jonathan Cable, 'Towards Data Justice: Bridging anti-surveillance and social justice activism' in Didier Bigo, Engin Isin, and Evelyn Ruppert (eds), *Data Politics: World, Subjects, Rights* (2019) at 179.
[59] Linnet Taylor, 'What is Data Justice?' *Big Data & Society* (2017).
[60] Institute for the Future of Work (n57), at 46; evelyn douek (n53) at 51.
[61] Dominique Clément (2018) 'Human rights or social justice? The problem of rights inflation', 22 *International Journal of Human Rights* 155 (2018), at 160.
[62] Madeline Baer, 'The human right to water and sanitation: champions and challengers in the fight for new rights acceptance' in Alison Brysk and Michael Stohl, *Expanding Human Rights: 21st Century Norms and Governance* 94 (2017), at 98.
[63] Allison Corkery, 'Methodological choices in human rights research are political, not just technical' *Open Global Rights* (21 December 2017) https://www.openglobalrights.org/methodological-choices-in-human-rights-research-are-political-not-just-technical/?lang=English
[64] However, *see,* Galit Sarfaty, 'Can Big Data Revolutionise International Human Rights Law?' 39 *University of Pennsylvania Journal of International Law* 73 (2017) at 75 (arguing that international human rights law traditionally 'underemphasiz[ed] the prevention of human rights violations').
[65] Alexander Mayer-Rieckh, 'Guarantees of Non-Recurrence: An Approximation' 39 *Human Rights Quarterly* 416 (2017).

requires an assessment of what led to the violation in the first place and thus aligns closely with the construction of laws, policies and practices to prevent human rights violations from recurring.

As noted by Barrie Sander in relation to social media, how human rights law is 'understood' also shapes the approach taken. Thus, if human rights are conceived of as manifestations of 'abstract individualism' they may not offer a route to addressing the structural enablers of human rights harm in the design, development and deployment of new and emerging digital technologies.[66] However, if human rights are viewed as 'structural conceptions [which] tend to adopt more systemic perspectives that strive to take into account imbalances of power in the social media ecosystem as well as the effects of state and platform practices on the social media environment as a whole',[67] they are more likely to attend to the causes of human rights harm.

Equally, within the field of new and emerging digital technologies, the structural enablers of human rights harm produced through the design, development and deployment of new and emerging digital technologies have not featured as centrally within existing research. However, the human rights framework provides scope to incorporate a more structural lens.[68] For example, the human rights law test of legality, necessity and proportionality not only enables assessments of whether an individual's rights have been violated but also provides a means by which to interrogate the underlying justifications for introducing new and emerging digital technologies. This process can expose the evidence-base relied upon to make such a decision, including whether other less invasive technological and non-technological means were considered. This is particularly important as the underlying evidence-base is not neutral. For example, as Linnet Taylor demonstrates in her research on smart cities, evidence drawn from data analytics may point towards very different policy interventions than qualitative evidence from people's experiences.[69]

Focus on the justification for the introduction of technology can also locate it within wider policies which themselves may be the drivers for technological uptake. For example, where the

[66] Barrie Sander (n53), at 188.
[67] Ibid, at 162.
[68] Dror and Shany (n12) at 1253 (discussing the increasingly structural focus of General Assembly and Human Rights Council resolutions).
[69] Linnet Taylor, 'The taming of chaos: Optimal cities and the state of the art in urban systems research' 58 *Urban Studies* 3196 (2021), at 3198.

application of this framework reveals the deployment of surveillance-based technologies to securitise borders, the problem definition widens from the human rights impact of surveillance technologies to the human rights impact of securitised border policies, thus expanding the governance approaches required to effectively protect human rights.[70] In a similar vein, research on private actors such as major technology companies has focused on the structural enablers of human rights harm such as the business models of major technology companies rather than the impact of specific types of technologies, such as algorithms.[71]

At the same time, a turn to analysing the structural causes and enablers of human rights harm remains in the early stages of development, and therefore consideration of the methodological approaches to such studies is timely. As already noted, researchers in other disciplines and fields have long researched the place and role of technologies from such perspectives, highlighting the importance of locating new and emerging digital technologies within a 'full socio-technical assemblage'.[72] Nick Couldry notes that this not only requires 'a close understanding of such assemblages and their role in shaping social processes and institutions' but also 'approaches to undertaking complexity that, by also addressing questions of value, help us to grasp the larger social order, and associated forms of hierarchy and exclusion, being built through processes of datafication'.[73] Discussing algorithms, Rob Kitchin argues that to not place technologies within their social and political context is 'like considering a law without reference to the debate for its introduction, legal institutions, infrastructures such as courts, implementers such as the police, and the operating and business practices of the legal profession. It also risks fetishising the algorithm and code at the expense of the rest of the assemblage'.[74] While not writing on human rights specifically, engagement with this research highlights the need for, but complexity of, a layered approach to analysing the human rights impact of new and emerging digital technologies when integrated into existing systems and contexts.

[70] Lorna McGregor and Petra Molnar, *Digital Border Governance: A Human Rights Based Approach* (forthcoming); Ayelet Schachar, 'Instruments of Invasion: The Global Dispersing of Rights-Restricting Migration Policies' 110 *California Law Review* 967 (2022).
[71] Amnesty International, *Surveillance Giants: how the business model of Google and Facebook threatens human rights* (2019) (as an example of a human rights focusing on business models as a structural issue).
[72] Kitchin (n11); Iliadis and Russo (n11).
[73] Couldry (n11).
[74] Kitchin (n11).

Similarly, the adoption of a wider lens on the reasons for technological uptake also makes visible the structural power[75] and political economy which may be driving states in particular to adopt technologies where private actors present them as 'solutions' to challenges they are facing.[76] For example, writing in the field of data justice, Linnet Taylor points to the demand created for public-private sector partnerships by states seeking to gain insights through data analytics but with low capacity to do so without working with the private sector. She observes that 'markets are a central factor in establishing and amplifying power asymmetries to do so with digital data, and that new strategies and framings may be needed that can address the public-private interface as an important site for determining wither data technologies serve us or control us'.[77] This wider lens again points to the importance of human rights research which goes beyond a focus on the human rights effects of specific technologies to a more structural examination of how new and emerging digital technologies enter the public sector, which can arise as a result of private actors presenting – and even creating – technologies as purportedly simple, cost effective and efficient solutions to complex public sector challenges. Private actors may therefore not only respond to states' tendering processes but may create a political economy for technological uptake.

Scholarship within fields such as critical data studies can therefore offer important methodological insights for human rights researchers seeking to understand how new and emerging digital technologies have come to be embedded in key areas of concern for human rights. Even where researchers within critical data studies are not themselves convinced of a human rights framing of the problem or the contribution of the human rights framework to dealing with the problems arising from new and emerging digital technologies, value exists in cross-field dialogue, engagement and challenge particularly as it can 'create new, rather than merely parallel, conversations and outcomes'[78] in addition to intellectual and practical friction which may generate fresh insights about how to deal with the challenges posed by new and emerging digital technologies.

[75] Safiya Umoja Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism* (2018).
[76] Kate Crawford, *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence* (2021), Conclusion.
[77] Linnet Taylor (n59).
[78] Tom McLeish and Veronica Strang, *Leading interdisciplinary research: transforming the academic landscape'* Leadership Foundation for Higher Education (2014), at 3.

Combined with growing research articulating the nature, severity and extent of human rights harm presented by the design, development and deployment of new and emerging digital technologies, deeper research accounting for the motivations and political economy driving the ubiquitousness of such technologies in so many areas of life provides a thicker view of the problem to be governed and a lens through which to assess the adequacy and effectiveness of existing and proposed governance models as well as the development of alternatives. Such research is therefore foundational to the prevention of human rights violations in the first place as well as access to remedies and accountability where harm occurs.

4. Multi-Method Contributions to Human Rights Research on the Structural Causes and Consequences of the Design, Development and Deployment of New and Emerging Digital Technologies

Research on the causes and consequences of the design, development and deployment of new and emerging digital technologies remains in the early stages, despite the significant progress that has been made in the last decade. In this part of the chapter, I first argue that as human rights research develops, much greater prioritisation is needed of research led by individuals and groups affected by the design, development and deployment of new and emerging digital technologies.

I then turn to examine the methods that can facilitate knowledge creation in this area. In this regard, this part of the chapter examines how the harnessing of multi-methods can facilitate the assemblage of knowledge on the view that 'knowledge is power' and 'the first step'[79] in processes of accountability, redress, and prevention. Many methods and methodological configurations can serve this objective.[80] In the space of this chapter, I critically assess the role of new methods to research elite actors and technological methods to counter opacity constructed by states and private sector actors before turning to analyse the role and importance of traditional human rights methods in this space which can be at risk of dismissal due to the fetishisation of technology.[81]

[79] Hannah Bloch-Wehba, 'Access to Algorithms', 88 *Fordham Law Review* 1265 (2020) at 1295
[80] Hepp et al (n10) at 7.
[81] *See,* Currie et al (n17).

To date, existing human rights research, with some exceptions, has not had a dominant empirical base.[82] Rather, researchers have adopted a range of approaches. For example, research may apply a human rights analysis to academic studies or policy presentations of specific technologies.[83] It may also draw on the empirical research of researchers in related fields, such as the intersecting field of critical data studies, as well as the work of investigative journalists.[84] While this methodological approach has highlighted the extent of potential human rights harm, in this part of the chapter, I argue that the development of a more extensive body of dedicated empirical human rights research can deepen knowledge of the human rights implications of new and emerging digital technologies, both from the perspective of individuals and groups whose human rights are compromised by their use and elite actors involved in their design, development and deployment.

A. Prioritising research led by individuals and groups affected by the design development of new and emerging digital technologies

A human rights-based approach to research on new and emerging digital technologies necessitates the foregrounding of the participation, voices and experiences of people most and disproportionately affected by technological use.[85] While civic participation in research on new and emerging digital technologies through methods such as citizen juries or 'mini-publics' constitutes an important participatory methodology,[86] it is distinct from research specifically led by, and focused on, people who have experienced human rights harm through the use of new and emerging digital technologies in their lives, often as a continuum of disproportionate surveillance and monitoring practices, for example, through the over-policing of particular communities or through interaction with welfare services.

---

[82] Fussey and Murray (n36).
[83] *See,* for example, McGregor, Murray and Ng (n12) at 336 (analysing from a human rights perspective a study published at Stanford University into whether deep neural networks could predict sexual orientation from facial images and concluding that such an exercise would be prohibited by international human rights law).
[84] *See,* for example, Jones (n12), section 2.2 (citing New York Times, Politico, JAMIA Open, EDRi, Nature, Just Security blog, MIT Technology Review for support for arguments made on the 'key human rights and ethical challenges posed by AI).
[85] *See,* for example, A Human Rights Based Approach: PANEL Principles', Scottish Human Rights Commission,
available at https://www.scottishhumanrights.com/media/1409/shrc_hrba_leaflet.pdf
[86] *See,* for example, Aidan Peppin, 'The Role of Trust: Findings from Citizens' Juries on the Good Governance of Data in Pandemics' *The Ada Lovelace Institute* (28 July 2022) at 45 (discussing methodology).

Some researchers have conducted qualitative interviews with individuals and groups subject to the use of new and emerging digital technologies as well as community groups in areas such as border governance[87] and social security.[88] However to date, the experiences of people most affected by new and emerging digital technologies have been relatively marginalised in human rights research on new and emerging digital technologies. Further, despite arguments by actors such as the Special Rapporteur on the rights of persons with disabilities for the development of research 'in active partnership and co-production with organizations of persons with disabilities … To investigate the uses of artificial intelligence and their impacts on the rights of persons with disabilities … [and] explore ways of rectifying those impacts and avoiding them in the future',[89] co-design and co-production of research led by communities disproportionately affected by the use of new and emerging digital technologies has been insufficiently prioritised.

These gaps further highlight the power asymmetries in the design, development and deployment of new and emerging digital technologies with the result that knowledge and understanding of how such technologies should be governed and the design of effective monitoring, oversight, accountability and remedial processes is diminished. For example, the Nuffield Foundation and Leverhulme Centre have emphasised the importance of research led by individuals and groups whose rights have been affected by the use of new and emerging digital technologies in noting that 'without understanding the perspectives of various different groups in society, we risk making trade-offs that favour the values and needs of the majority at the expense of minorities'.[90] Understanding the direct experience of individuals and groups affected is also critical to building support for governance frameworks that focus on the protection of human rights.[91] This is particularly important in an area where it is easy to dismiss the impact of technology as purely technical or to minimise the human rights impact of their employment. For example, it can be easy to dismiss the role of machine learning risk assessments within decision-making processes by characterising them as a digital form of evidence without understanding the influential role their outcomes can have on human

[87] For example, *see,* Petra Molnar, *Artificial Borders: AI, Surveillance, and Border Tech Experiments* (The New Press 2024).
[88] Virginia Eubanks (n47).
[89] Human Rights Council, Rights of persons with disabilities (n6), at §83.
[90] Jess Whittlestone, Rune Nyrup, Anna Alexandrova, Kanta Dihal, Stephen Cave, *Ethical and Societal Implications of Algorithms, Data, and Artificial Intelligence: A Roadmap for Research* (2019) Ada Lovelace Institute and the Leverhulme Centre for the Future of Intelligence, at 20.
[91] European Union, Fundamental Rights Agency, *10 keys to effectively communicating human rights* (26 September 2018).

decision-makers should they defer to the algorithm's findings without critically interrogating how the risk assessment works.

As discussed above, the foregrounding of the voices and experiences of individuals and groups whose rights have been compromised through the deployment of new and emerging digital technologies can challenge claims about the role of technologies in society based on data analytics and can also reshape the way in which the challenges with technologies are framed and the governance approaches that therefore flow from it. In particular, the experiences of people most affected are a central means by which to situate technology within its social, political and economic contexts as well as connect it to ongoing and historical patterns of human rights harm. For example, Seeta Peña Gangadharan and Jedrzej Niklas conducted research with civil society groups at a point in time at which much of the dominant legal and policy discourse focused on the impact of technical systems in the abstract. Their empirical research pointed to the need to widen analysis from 'technical analyses, especially fairness, accountability, and transparency studies, … [to] a deeper exploration of potential negative externalities of automated systems.'[92] They argued that rather than examine 'data violence' in isolation, data violence should be considered 'alongside structural violence and racist search engines alongside racism perpetuated by social institutions'.[93]

Accordingly, much greater investment through funding but also politically and through meaningfully having a 'seat the table' is needed in research led by individuals and communities most affected. However, as highlighted by a range of commentators, this must not result in 'participation-washing'[94] but rather achieve a 'shift to recognise the value of different types of expertise, and seeing marginalised people's experiences and knowledge as legitimate, and equal'.[95]

B. Cautionary inclusion of new methods including technological

[92] Seeta Peña Gangadharan and Jedrzej Niklas, 'Decentering technology in discourse on discrimination' 22 *Information, Communication & Society* 882 (2019), at 896.
[93] Ibid.
[94] *See,* for example, Mona Sloane, 'Participation-washing could be the next dangerous fad in machine learning' *MIT Technology Review* (25 August 2020).
[95] Jef Ausloos, Alexandra Giannopoulou and Jill Toh, Marginalised Communities in Democratic Governance' in Ada Lovelace, *Rethinking Data and Rebalancing Digital Power*, at 75.

When faced with opacity, human rights researchers and researchers in related fields have developed creative methods to unearth the risks posed by certain new and emerging digital technologies as well as to examine their enabling power structures. Some of these methods emphasise the role of sources which are not typically the focus of human rights research but which can provide insights into the practices of elite actors holding power and knowledge into technological design, development and deployment. For example, such sources may include trade literature, 'attending trade fairs',[96] 'municipal contracts'[97] and 'documenting the biographies of key actors and the histories of projects'.[98] In addition, as the uptake of many technologies deployed in areas of public life of key concern to human rights may enter the public sector from public-private partnerships, the 'monitor[ing of] state and local procurement processes' can offer a lens through which to document and study not only the technology but the way in which it is promoted by private actors and justified by state actors.[99]

Often through multi-method approaches, researchers have also designed research that incorporates new and emerging digital technologies themselves. Some researchers have used data scraping techniques from publicly available websites, to compile information about government use of algorithms as part of a multi-methods approach. For example, researchers at the Data Justice Lab in Cardiff 'used search engines to scrape UK government sites and media sites' as part of an exercise 'to map developments across the UK'.[100] Other organisations, such as the American Civil Liberties Union (ACLU) have 'built a face data base and search tool using 25,000 publicly available arrest photos. [It] … searched that database against public photos of every current member of the [US] House and Senate' using the Amazon facial recognition system, Rekognition.[101] In doing so, it found that the search 'incorrectly matched 28 Members of Congress, identifying them as other people who have been arrested for a crime'; these Members of Congress were 'disproportionately people of color'.[102]

[96] Kitchin (n11), at 25.
[97] *See also,* Tamara Nopper, '"Power Structure" is not Abstract Phrase: On methods for studying elites' *Data & Society Points* (22 February 2023); Jenna Burrell, 'A Guide to Investigating the Datafied State through Documents: Gleaning knowledge from documents involves more than just reading words on a page' *Data & Society Points* (9 March 2023).
[98] Kitchin (n11) at 25.
[99] Michele Gilman, 'Poverty Lawgorithms: A Poverty Lawyer's Guide to Fighting Automated Decision-Making Harms on Low-Income Communities' *Data & Society* (2020) at 7.
[100] Lina Dencik, Arne Hintz, Joanna Redden and Harry Warne (n23), at 18.
[101] Jacob Snow, 'Amazon's Face Recognition Falsely Matched 28 Members of Congress with Mugshots' *American Civil Liberties Union* (26 July 2018).
[102] Ibid.

The EU Fundamental Rights Agency has also employed synthetic data (artificially created data) to conduct 'a simulation of a feedback loop in the area of predictive policing', explaining that a 'feedback loop occurs when predictions made by a system influence the data that are used to update the same system', meaning that '[a]ny bias in algorithms can therefore potentially be reinforced over time and exacerbated'.[103] In relation to migration, researchers also 'co-opt' similar technologies employed by states and private actors for the purpose of surveillance and to 'prevent unauthorized border crossings' such as 'satellite imagery, automatic identification system vessel tracking data, and geo-spatial mapping … to document and seek accountability for the violation of migrant' rights … The focus is not only on the events that occurred, but also the information infrastructure and surveillance technology that governments had at their disposal at the time of the event'.[104] Thus, in some instances, human rights researchers employ new and emerging digital technologies as part of wider methods aimed at countering opacity within socio-technical systems.

These types of methods assist in making technological uptake and its impact on human rights visible and thus offers new ways to document human rights harm, especially when combined with other research methods. At the same time, much greater debate and discussion is needed into the ethics of particular technological methods. Questions most obviously arise where researchers use personal data, even if publicly available, to test technological systems. In theory, synthetic data, may overcome such challenges. However, complex issues can also arise with synthetic data, particularly if it is originally based on real-world data, including personal data, which could lead to re-identification.[105]

Moreover, even when part of multi-method approaches, care still has to be taken to avoid the overall undermining of the strength of traditional research methods such as qualitative interviews through an expectation of the inclusion of evidence or analysis generated through technology. As noted by Molly Land and Jay Aronson access to 'satellite imagery, statistical

[103] European Union Fundamental Rights Agency, *Bias in Algorithms: Artificial Intelligence and Discrimination Getting the Future Right: Artificial Intelligence and Fundamental Rights* (2020) at 8.
[104] Daniel Ghezelbash, 'Technology and countersurveillance: holding governments accountable for refugee externalization policies', Globalizations (2022) at 7 (discussing the Forensic Oceanography project Charles Heller and Lorenzo Pezzani who are part of the Forensic Architecture Team at Goldsmiths, University of London which employs multi-methods to investigate and 'reconstruct' specific incidents involving migrants attempting to cross the Mediterranean Sea - Forensic Oceanography, *The Nivin Case: Migrants' resistance to Italy's strategy of privatized push-back* (2019).
[105] Michal Gal and Orla Lynskey, 'Synthetic Data: Legal Implications of the Data-Generation Revolution' *LSE Working Paper* (forthcoming 109 *Iowa Law Review* (2023)) at 19, 30.

methodology, and sophisticated data analysis techniques' is power-based and uneven[106] and a methodological prioritisation of such tools can 'divert resources' and 'neglect the more traditional advocacy and grassroots mobilization strategies that are necessary to generate the political will required for social change.'[107] Further, it can result in buy-in by human rights researchers to the proposition that technology best supports evidence-based decision making which is one way in which techno-solutionism can become embedded in contexts in which human rights are put at risk, as discussed in the first part of this chapter.[108] For example, Galit Sarfaty observes that '[t]he use of big data in the international human rights field is part of an increasing focus on quantification in global governance'.[109] She argues that such an emphasis 'can reify power inequalities, lead to a concentration of decision-making power in the hands of technical experts, and risk distorting or obscuring the phenomena they are meant to take account of',[110] particularly by 'abstract[ing] away the individual and the local'.[111]

Accordingly, while new and emerging digital technologies can offer innovative methodological tools to human rights research seeking to document the human rights impact of the use of such technologies, particularly in countering opacity, they still require problematisation and critical assessment around whether they should be used.

C. Emphasising the continuing relevance of traditional human rights research methods

Notwithstanding the important contributions of new methods to human rights research in this area, traditional qualitative human rights research methods have a critical role to play within a multi-methods approach. These methods both avoid the fetishisation of new and emerging digital technologies through over-emphasis of their distinctiveness from other human rights challenges as well underscore the embeddedness of new and emerging technologies within

[106] Molly Land and Jay Aronson, 'Part II: Technology and Human Rights Enforcement', in Molly Land and Jay Aronson (eds), *New Technologies for Human Rights Law and Practice* (2018) 125 at 127 (discussing 'broad disparities in distribution')
[107] Molly Land and Jay Aronson, 'The Promise and Peril of Human Rights Technology' in Land and Aronson ibid at 13.
[108] Evgeny Morozov, *To Save Everything Click Here: The Folly of Technological Solutionism* (2014).
[109] Sarfaty (n6), at 93.
[110] Ibid, at 93
[111] Ibid, at 94.

wider social contexts, thus constituting socio-technical assemblages.[112] However, as discussed in this section, they also encounter limitations, again underscoring the contribution of a multi-methods approach to knowledge creation.

Qualitative interviews with central governmental and state actors at different points in the technological lifecycle, from commissioning to development to procurement to repurposing as well as policy formation and oversight can offer insights into how actors directly developing, using, overseeing or monitoring the use of new and emerging digital technologies experience them in practice.[113] However, such research may only be possible for networked researchers who are able to access state and particularly corporate actors.[114] Similarly, freedom of information requests can reveal information about the procurement of specific technologies. However, researchers have documented challenges in successfully making such requests.[115] For example, Hannah Bloch-Wehba points to the exemptions incorporated within freedom of information legislation which can prevent disclosure, such as assertions of proprietary interest, '"trade secrets and commercial or financial information obtained from a person and privileged or confidential"'.[116]

Further, within scholarship on human rights research methods, the convening of meetings, often labelled as 'expert', is rarely referred to as a research method. Yet, expert meetings typically with academia, civil society and international organisations constitute a common part of the research process, particularly if policy-orientated. These convenings can serve as a vehicle for knowledge-exchange; problem diagnosis; analysis of human rights protection gaps, for example through a failure to implement or operationalise human rights standards and

[112] Stine Lomborg, Lina Dencik and Hallvard Moe, 'Methods for datafication, datafication of methods: Introduction to the Special Issue' 35 *European Journal of Communication* 203 (2020) at 207 (discussing the role of innovative methods in this field while avoiding 'algorithmic fetishism' and recognising that '[n]ew methods do not emerge in a vacuum – there are historical continuities and predecessors, and old methods are still relevant,' particularly with regard to 'the situated, contextual aspects of data as a way to understand dynamics of power'.)

[113] European Union Fundamental Rights Agency, *Getting the Future Right: Artificial Intelligence and Fundamental Rights* (2020), Foreword and at 6.

[114] Gavin Sullivan, *The Law of the List: UN Counterterrorism Sanctions and the Politics of Global Security* (2020), at 34.

[115] Lina Dencik, Arne Hintz, Joanna Redden and Harry Warne, (n23) at 14-15.

[116] Hannah Bloch-Wehba (n79) at 1300; Algorithmic Transparency for the Smart City Robert Brauneis & Ellen P. Goodman, Algorithmic Transparency for the Smart City' 20 *Yale Journal of Law and Technology* 103 (2018) at 110 (discussing 'impediments to making government use of big data prediction transparent [as]: (1) the absence of appropriate record generation practices around algorithmic processes; (2) insufficient government insistence on appropriate disclosure practices; and (3) the assertion of trade secrecy or other confidential privileges by government contractors.')

norms; to collaboratively develop research agendas; and to issue recommendations to states, corporate actors and international organisations. They can be distinguished from multistakeholder meetings which typically include states and corporate actors although these types of meetings can also provide a means of knowledge production and policy development. The distinction derives from the function of expert meetings to critically assess and consolidate existing human rights knowledge among human rights researchers in academia and practice, although notably they can sometimes suffer from their own power imbalances if representatives of groups disproportionately affected by the deployment of new and emerging digital technologies are not part of such meetings, thus highlighting a participatory gap. When presented as based on group analysis, expert meetings can be perceived as more authoritative to state and corporate actors, thus enhancing the possibilities for influencing policy and practice.

In an emerging field of inquiry for human rights such as new and emerging digital technologies, the convening of expert meetings between scholars and practitioners has provided a means to assemblage knowledge and analyse the human rights impact of the design, development and deployment of such technologies and in doing so, to produce new forms of knowledge and insight into human rights protection. Such convenings are particularly important given the layers of opacity surrounding technological practices. While not a human rights focused workshop, Archie Drake et al reflect on the role of expert workshops in offering 'an unusually broad experience-based insight into the practical implications of AI in law, drawing on the views of practitioner professionals as well as those of researchers. It therefore addressed the significant con-temporary UK AI policy agenda from the fresh perspective of a group with some practical experience of relevant harms and associated reasons for mistrust.'[117] Thus, while not always labelled as a research method, expert meetings have played a central role in human rights research in this area. This points to a gap in scholarship on human rights methodology which would benefit from greater reflection and insight in order to strengthen the methodological rigour by which such meetings are convened and to critically assess their role in the generation of new human rights knowledge.

[117] Archie Drake, Perry Keller, Irene Pietropaoli, Anuj Puri, Spyros Maniatis, Joe Tomlinson, Jack Maxwell, Pete Fussey, Claudia Pagliari, Hannah Smethurst, Lilian Edwards & Sir William Blair (2022), 'Legal contestation of artificial intelligence-related decision-making in the United Kingdom: reflections for policy', 36 *International Review of Law, Computers & Technology*, 251 (2022), at 267.

Equally, while a multi-method approach can help to chip away at layers of opacity and construct new knowledge both of the human rights effects of new and emerging digital technologies and the reasons why, it will not be a panacea. In this regard, particularly in engagement with and the design of governance models, it is also important for human rights researchers to systematically document what is not known – about why and how states and private actors harness new and emerging digital technologies with resulting human rights harm, including through each (partial) failure to uncover information, for example, through rejected interview and freedom of information requests.

5. Future Research Agendas and Methods: The role of forecasting

As discussed in this chapter, human rights research on new and emerging digital technologies to date has largely been reactive and retrospective, drawing on empirical work in other fields and disciplines.[118] While somewhat inevitable for a new field, particularly where technological systems have been steeped in opacity, the question arises whether a retrospective approach suffices or whether human rights research should also start to look forward through the forecasting of future technological developments and thus future potential manifestations of human rights harm. In this regard, it could be argued that the scale, sophistication and speed at which new and emerging digital technologies are being developed calls for the development of anticipatory methodology to forecast the types of technologies likely to be developed and deployed and their potential human rights impact.

Forecasting should be distinguished from existential research focused on general purpose artificial intelligence, for example, which may never come into fruition, but particularly as it is often led by major technological actors, researchers express concern that it detracts from current societal challenges presented by the place of technologies in our lives.[119] By contrast, forecasting is a methodological approach to look into the near future to counter reactive tendencies with approaches to governance of new and emerging digital technologies, in favour of a more anticipatory case. The development of methodological frameworks to forecast emerging digital technologies would also provide space for early deliberative and participatory

[118] Drake et al ibid at 255.
[119] *See,* Access Now, RightsCon Costa Rica, 'In Conversation with Timnit Gebru and Melissa Chan' (2023).

methods,[120] particularly with individuals and groups most likely to be subject or affected by such technologies, to feed into decisions about whether and how to introduce such technologies.

Nevertheless, forecasting new and emerging digital technologies is itself a challenging task methodologically. Adding analysis of potential human rights impact adds layers of complexity both substantively and due to the lack of existing forums within which such analysis takes place. It could also carry similar risks to existentialism by shifting resources (financial but also legal, social and political) to a focus on the future (even if the near future) instead of concentrating on addressing current-day human rights challenges and constructing adequate and effective governance frameworks. Further, it could be argued that a separate approach is not needed where governance frameworks incorporate anticipatory dimensions to encompass future forms of technology or socio-technical assemblages without concretely defining what they might look like.[121]

Accordingly, critical questions require exploration into the contributions and risks of widening human rights research to encompass technological forecasting. The questions are not only normative but also methodological as the process by which technological development is forecast may shape the value in such undertakings. In this regard, some researchers have commissioned 'landscape overview[s] on 'future technologies'.[122] Other routes to forecasting could include dialogue across disciplines that traditionally have not worked together within the human rights field, particularly as many technologies which later become commercially viable start within academia as well as expert meetings as discussed in the previous section. However, as most technologies are concentrated within a small number of global companies, such methods may only operate at the fringes. Thus, the viability of dialogue and interaction between academia and practice, particularly with industry would also be important to explore, including by identifying possible limitations, such as proprietary interest and anti-trust legislation which may limit the openness of such dialogue, thereby potentially inhibiting such exercises. Accordingly, this section highlights the value in considering the role of forecasting as part of

[120] Arne Hintz, Lina Dencik, Joanna Redden, Emiliano Treré, Jess Brand and Harry Warne, 'Civic Participation in the Datafied Society: Towards Democratic Auditing?' *Data Justice Lab* (2022) at 21; Ada Lovelace Institution, 'The Citizens' Biometric Council: Recommendations and findings of a public deliberation on biometrics technology, policy and governance (March 2021).
[121] For a discussion on anticipatory regulation, *see,* Sofia Ranchordas, 'Future Proofing Legislation for the Digital Age' University of Groningen Faculty of Law Research Paper No. 36/2019.
[122] Ada Lovelace Institute, *Rethinking data and rebalancing digital power* (2022), at 90.

future agenda building while calling for careful consideration of the viability of such exercises, particularly in light of current power configurations.

6. Conclusion

Despite the proliferation of human rights research into new and emerging digital technologies, knowledge and understanding of the human rights impacts of such technological design, development and deployment and the structural conditions and motivations leading to their uptake remains in the early stages. In this regard, this chapter has both recognised the barriers presented by multiple layers of opacity to understanding the motivations and effects of technological uptake but has also emphasised the contribution of interdisciplinary research and interdisciplinary interaction in chipping away at such opacity to produce new forms of knowledge. In this regard, it has underscored the role – as well as limitations – of multi-method approaches and the need for a much greater accent on empirical research documenting, and led by, individuals and groups directly affected by technological use, particularly in areas in which human rights may already be at risk. Equally, the chapter recognises that much of the current focus remains on existing human rights challenges which are exacerbated by ongoing technological roll-out in contexts in which dedicated regulation remains incomplete and embryonic as does the engagement of existing governance frameworks. As such, the focus of research as well as advocacy, policy and litigation continues to centre on the present. However, the chapter ends by asking whether any space exists to develop methods to forecast future technological developments, contributions and risks to human rights as a form of prevention or whether the political economy and infrastructure of power prohibits such exercises in any meaningful way, with the accompanying potential for distraction from today's challenges.