



People Not Code: The Case for a Digital Civil Society Observatory

Rachel Coldicutt, Careful Trouble

Abstract

Socio-technical change happens at a very different pace to legislative and regulatory activity. This is partly due to bureaucratic and institutional norms and partly due to the difficulty of gathering timely and useful evidence about the social and other impacts of technologies. To date in the UK, legislative and regulatory attention has tended to focus on socially prevalent, or fully emerged, harms; as such, mitigations in critical areas including child online safety, data protection, and financial fraud have not been introduced until the scale of harm has reached a crisis point. Deploying more anticipatory methods would reduce the number of wide-spread harms; doing so requires a reliable evidence base of emergent societal impacts. This essay makes the case for Digital Civil Society Observatory, an evidence-gathering body that would sit alongside the AI Safety Institute and the Turing Institute. This would draw on the empirical knowledge and expertise of the broad field of civil society to anticipate, understand, and mitigate the ongoing societal impacts of technologies and ensure that innovation delivers public benefit and a stronger society, strengthening democratic outcomes and bringing the range and complexity of public interest and experience to greater prominence in the technology policy and governance landscape.

Keywords: anticipatory regulation; participation; community; social intelligence; civil society

Introduction

The societal impacts of digital and data-driven technologies do not play out in uniform or entirely predictable ways. This has presented a recurring Pandora's Box-type problem for successive UK governments, whose tendency to 'wait and see' has been repeatedly shocked by the apparently sudden prevalence of successive waves of digital technologies and their associated harms. From social networks to smartphones to generative AI, the next big thing has often hidden in plain sight, appearing to unfold 'slowly, then all at once',¹ sheltered from government intervention by the sway of the corporate lobby and the shade of economic expectations.

This 'pro-innovation'² attitude has resulted in an ex-post regulatory approach, focused on articulating risk rather than protecting rights. The Online Safety Act (2023) – which covers

a range of illegal and harmful 'content and activity'³ – is a case in point: more than two decades after search engines and social media became commonplace, the UK's regulatory response remains a work in progress, even though concerns about child safety and content standards have been shared among digital professionals since at least the early 2000s.⁴

Predicting which harms might affect which groups of people, which shared goods and infrastructures are needed, and applying appropriate mitigations and solutions in a timely and useful manner, is a collective endeavour. Current legislative assumptions operate upon the principle that a digitally mediated harm must be apparent to or be experienced by a significant proportion of the general population before it can be acted upon,⁵ but such a basis makes effective regulation almost impossible. As in other areas of technology forecasting, noting and acting on 'weak signals' – the early indicators of 'strategic phenomena ... that are distant to the perceiver's frame of reference'⁶ – is essential for mitigating harms as they arise: this essay argues that the best way to perceive and effectively act on these early indicators is by creating a Digital Civil Society Observatory.

The Digital Civil Society Observatory

The creation of a Digital Civil Society Observatory as a UK non-departmental public body would ensure that public experience is more fairly represented in all aspects of digital policymaking. The traditional role of civil society organisations in technology policy and governance is as expert contributors who engage with consultations and policy proposals and advocate for the specific needs of their beneficiaries; for politicians, these same groups can also offer a useful source of community connection, putting the names and faces of real people to issues that may otherwise seem abstract. However, civil society also plays a much under-valued role in sensing, understanding, and responding to early indicators of societal changes and challenges as they emerge.

The Digital Civil Society Observatory (DCSO) would draw on the empirical knowledge and expertise of the broad field of civil society to anticipate, understand, and mitigate the ongoing societal impacts of technologies and ensure that innovation delivers public benefit and a stronger society. This would strengthen democratic outcomes and bring the range and complexity of public interest and experience to greater prominence in the technology policy and governance landscape.

¹ With thanks to Deb Chachra for this observation. See also, Steve Levy, 'How Hemingway Gradually—Then Suddenly—Defined the Zeitgeist', *Wired* (8 July 2022).

² Department for Science, Innovation and Technology, 'Policy paper: A Pro-Innovation Approach to AI Regulation', Gov.UK, (updated 3 August 2023), (accessed on 24 March 2024)

³ [Online Safety Act 2023](#), Gov.UK (25 October 2023), (accessed on 2 April 2024)

⁴ For instance: 'Overwhelmingly, the debate over the regulation of young people's use of the Internet is oriented towards the avoidance of harm.' Sonia

Livingstone and Moira Bovill, 'Families and the Internet: An Observational Study of Children and Young People's Internet Use', *London School of Political and Economic Science* (2001), p. 38 (accessed 2 April 2024).

⁵ Online Safety Act 2023, Gov.UK, Chapter 7, Section 154; see also Centre for Data, Ethics, and Innovation, 'Public attitudes to data and AI: Tracker survey (Wave 3)', Gov.UK (6 December 2023; last updated 12 February 2024), (accessed on 2 April 2024),

⁶ Barbara L. van Veen, J.Roland Ort, 'Unifying weak signals definitions to improve construct understanding', *Futures* (Vol. 134, 2021), 102837.

The DCSO would join the AI Safety Institute, a technology evaluation body, and the Alan Turing Institute, the national research institution for data science and AI, as the third in a triumvirate of national research and monitoring bodies that promote responsible innovation. It is vital that public interest and the social impacts of technology are placed on an equal footing with the technical and academic perspectives offered by these other bodies, which both receive £100m in direct government funding.⁷ Funding for the DCSO could be sponsored by DSIT and drawn in equal parts from the Treasury and the National Lottery Community Fund, reflecting its purpose as a representative body for UK civil society.

The functions of the DCSO would encompass foresight studies and social horizon scanning; synthesising ongoing research and insight from across civil society; and delivering in-depth qualitative research into emergent sociotechnical issues. Its outputs would include social impact assessments of technologies, products, and services; policy proposals that promote public benefit and illustrate what ‘digital good’ looks like; training and dissemination to share insight and best practice with regulators, industry, and public services. A significant proportion of any funding received would be distributed to civil society researchers, advocates, and organisations to fund and support ongoing horizon scanning activity.

Mirroring the cross-cutting nature of the Digital Regulation Cooperation Forum,⁸ the DCSO would operate beyond technological and market silos to understand and analyse the sociotechnical impacts of technologies as they are experienced by people and communities, and bring together socially representative, cross-sector expertise in the analysis of issues as they arise.

The public interest gap

Public interest is currently under-represented in the melee of stakeholders that surround UK technology policy. As members of the public, our engagement with technology policymaking does not have parity with other interested stakeholders. While we have a degree of influence over some of the specific technologies we use, our methods for creating change – exerting consumer power, voting, and campaigning – are all indirect routes that have a slow, and unpredictable, rate of return.

A rapid review into public participation in AI governance by the Ada Lovelace Institute found that:

*the public want to have a meaningful say on decisions that affect their everyday lives, not just to be consulted. ...people expect their views, in the full spectrum of their diversity, to be taken as seriously as those of other stakeholders, including in legislative and oversight processes.*⁹

However, the same review also found that public consultation tends to be restricted to a small number of formats, including polling and citizens’ juries. The focus of such consultations is, of course, framed by those who design the process: people who participate in polling answer the questions that have been written by the polling company and briefed by the client; participants in citizens’ juries work together to consider a “clear question with defined scope”.¹⁰ While such findings are useful to policymakers and others, they are incomplete, and form a very particular view of public attitudes. They show what happens when decision-makers shape public engagement, rather than what happens when people shape technology.

The weaving of public opinion into governance models that are run by state or corporate actors is important, but it should not be mistaken for being entirely representative public concerns or an accurate indicator of emerging harms. The act of convening, of questioning, of sharing evidence – particularly on emergent topics where there may be low-levels of public knowledge and experience – risks distorting outcomes, and it is not a replacement for the informal horizon scanning and sensemaking that takes place in communities. Indeed, the risk of such methods being misused or manipulated could become non-trivial, particularly for ethically contentious uses of technology, including justice, welfare, education and defence.¹¹ One example of this misalignment is broad public support for the use of AI in policing, in spite of a growing body of evidence that AI can actually reduce both the accuracy and effectiveness of policing.¹²

In the areas defined by the Department of Science, Innovation and Technology (DSIT) as comprising ‘critical tech’,¹³ corporate, technical, and academic perspectives have, since 2019, had significantly higher levels of cut-through than those of civil society organisations.

⁷ Department for Science, Innovation and Technology, ‘Policy paper: Introducing the AI Safety Institute’, Gov.UK (updated 17 January 2024), (accessed 24 March 2024); ‘£100m Investment in The Alan Turing Institute Announced’, UK Research and Innovation (7 March 2024), (accessed 2 April 2024).

⁸ Competition and Markets Authority, Information Commissioner’s Office, Ofcom, Financial Conduct Authority, ‘Collection: The Digital Regulation Cooperation Forum’, Gov.UK (updated 28 March 2023), (access 24 March 2024)

⁹ Anna Colom, ‘Meaningful Public Participation and AI: Lessons and Visions for the Way Forward’, Ada Lovelace Institute (1 February 2024), (accessed on 5 March 2024)

¹⁰ ‘Democratising decisions about technology: A toolkit’, Royal Society of the Arts (2019), (accessed on 5 March 2024)

¹¹ On framing effects, see M. M. Bechtel, J. Hainmueller, D. Hangartner, M. Helbling, ‘Reality Bites: The Limits of Framing Effects for Salient and Contested Policy Issues’, *Political Science Research and Methods* (2015), 3(3): 683-695.

¹² Centre for Data Ethics and Innovation, ‘Research and Analysis: Public Attitudes to Data and AI: Tracker survey (Wave 3)’, Gov.UK (updated 12 February 2024), (accessed 2 April 2024); weapons scanning on the New York City Subway system is shown to produce 85% false positives, Felipe de la Holz, ‘NYC Has Tried AI Weapons Scanners Before. The Result: Tons of False Positives’, *Hellgate* (2 April 2024), (accessed 2 April 2024) and Kristian Lum and William Isaac, ‘To predict and serve?’, *Significance* (Royal Statistical Society, 7 October 2016), (accessed 2 April 2024).

¹³ Department of Science, Innovation and Technology, ‘Policy paper: The UK Science and Technology Framework’, Gov.UK (updated 9 February 2024), (accessed 2 April 2024)

For instance, in the run-up to the 2023 AI Safety Summit at Bletchley Park, DSIT ministers are on record as having taken part in policy discussions with more than 50 corporate AI and consulting firms;¹⁴ meanwhile no UK-headquartered civil society organisations were reported as attending the Summit and only a handful were publicly consulted in the weeks prior to the event.¹⁵ More generally, business interests are further bolstered by a number of effective trade bodies and by larger technology companies' considerable investment in public affairs and influence; the significantly higher levels of resource available to corporate stakeholders also means that they have a greater opportunity to frame and influence media narratives.¹⁶

The influence of UK technology research organisations on technology policymaking is also significant. Measuring influence is, of course, an inexact science, but recent government levels of investment in AI research indicate that the area is rising in priority. In addition to the recent additional £100m investment in the Alan Turing Institute, since January 2024 UKRI has committed another £100m to 'transformative' AI innovations; meanwhile the Advanced Research and Innovation Agency (ARIA), the UK's 'moon-shot agency' which exists to 'unlock scientific breakthroughs for everyone's benefit' is backed by £800m.¹⁷ Further to this, if we are to take representation at the 2023 AI Safety Summit as one measure of a sector's influence on policymaking, UK technology research organisations accounted for ten of the 120 global organisations in attendance.¹⁸

Moreover, the relative lack of cut-through for UK digital civil society in policy and influence has almost certainly been exacerbated by its comparatively low levels of funding compared to equivalent organisations in the US and Europe, where multi-funder networks exist to support digital civil society activity. No such funder groups currently exist or fund activity in the UK.¹⁹

The role of civil society

Relational ethics asks that for any solution that we seek, the starting point be the individuals and groups that are impacted the most

Abeba Birhane, 'Algorithmic Injustice: A Relational Ethics Approach'²⁰

Civil society is the term used to describe the space 'outside the family, market and the state'²¹ that includes, but is not limited to, community groups, social movements, communities of interest, faith groups, formally instituted charities, and labour unions. It is an expansive term that includes established third-sector bodies and spontaneous moments of collective action, described by the Inquiry into Civil Society as a state of action rather than a type of organisation, a mode of behaviour that includes 'all of us ... when we act not for profit nor because the law requires us to, but out of love or anger or creativity, or principle'.²²

Civil society creates both physical and digital places for us to gather with people with whom we share experiences, interests, beliefs, or opinions. These spaces range in scale from local and informal settings such as playgroups, sports clubs, and book groups to online support groups and fan forums; from national charities offering advice and advocacy to global networks of researchers and campaigners. As such, civil society brings together many of the people and places that help us navigate uncertainty and times of need – offering practical help, advice, or reassurance – so we can make better sense of the present and future.

Just as there are place-based civil society groups and organisations, there are also digital civil society networks and organisations. These range from the informal to the well-established, from the focussed to the general. Much of the richness and complexity of life is reflected here.

¹⁴ Joseph Bambridge, Laurie Clarke and Vincent Manancourt, 'Who Said the UK Can't Do Chips?', *Politico Pro Morning Technology UK* (28 March 2024), (accessed 2 April 2024, with thanks to Gavin Freeguard).

¹⁵ Research organisations were well-represented in the UK's cohort, including the Ada Lovelace Institute, Alan Turing Institute, Oxford Internet Institute, and the British Academy. For a full list of attendees see Cristina Criddle, Madhumita Murgia, Anna Gross and Yuan Yang, 'How Sunak's Bletchley Park Summit Aims to Shape Global AI Safety', *Financial Times* (27 October 2023).

¹⁶ Influential UK membership bodies include but are not limited to: the Start-Up Coalition, TechUK and BCS, the Chartered Institute for IT and does not include incubators or VC networks. Exact expenditure figures are not available for technology public affairs spending in the UK; in Brussels, as the EU AI Act, Digital Services Act, and Digital Markets Act were in progress, the Corporate Europe Observatory reports that 'big tech' firms spent €113m a year on lobbying activities – this does not include public policy analysts and public affairs activity outside of the official lobby. 'Lobbying Power of Amazon, Google and Co. Continues to Grow', *Corporate Europe Observatory* (08 September 2023), (accessed 26 March 2024).

¹⁷ '£100m boost in AI Research will Propel Transformative Innovations', *UK Research and Innovation* (6 February 2024), (accessed 26 March 2024); 'Frequently Asked Questions', ARIA, <https://www.aria.org.uk/faqs/> (accessed 26 March 2024).

¹⁸ Department for Science, Innovation and Technology, 'Guidance: AI Safety

Summit: Confirmed Attendees (governments and organisations)', *Gov.UK* (updated 31 October 2023), (accessed 2 April 2024).

¹⁹ Ahead of the 2023 AI Safety Summit, VP Harris announced 'a New Funders Initiative to Advance AI in the Public Interest... Ten leading foundations are announcing they have collectively committed more than \$200 million in funding toward initiatives to advance the priorities laid out by the Vice President, and are forming a funders network to coordinate new philanthropic giving', 'Fact Sheet: Vice President Harris Announces New U.S. Initiatives to Advance the Safe and Responsible Use of Artificial Intelligence', *The American Presidency Project* (1 November 2023), (accessed 2 April 2024). On a much smaller scale, the European AI and Society Fund is backed by a consortium of funders who have supported €7.5 million worth of activity in the EU since 2020, see *European AI and Society Fund*, (accessed 2 April 2024).

²⁰ Abeba Birhane, 'Algorithmic Injustice: A Relational Ethics Approach', *Patterns* (Vol. 2, Issue 2, 2021).

²¹ World Economic Forum, 'The Future Role of Civil Society' (January 2013), p. 8, (accessed 25 March 2024)

²² The Independent Inquiry into Civil Society, 'The Story of Our Times: Shifting Power, Bridging Divides, Transforming Society' (November 2018), (accessed 25 March 2024).

At the grassroots end, digital civil society is home to groups as diverse as campaigning parents concerned for their children's welfare; race and equality campaigners protesting facial recognition technology; and volunteers who teach skills in community centres. There are also more formal organisations that work at regional, national, and international levels representing the interests of specific communities or demographic groups, for instance, Digital Defend Me 'exists to protect children's rights to privacy and family life'; Data, Tech, and Black Communities 'work(s) with and for Black communities in the UK to make technology a more positive influence in the lives of Black people'; Chayn supports 'survivors of abuse, across borders'. Another type of digital civil society organisation brings a specific area of domain expertise to cross-cutting issues: for instance, Foxglove specialises in strategic litigation order to 'make tech fair for everyone'; Who Targets Me create transparency around political advertising to support democratic outcomes; Glitch's mission is to 'end abuse online [and] champion digital citizenship'. There are also larger, more established, civil society organisations that existed before the digital society came into being that have expanded their remit to include the ways technologies affect people's rights; these include the Trades Union Congress, Liberty, and Amnesty International.

Weak signals and early indicators

As well as providing support networks, advocacy, and services, civil society activities offer a significant source of empirical knowledge around emerging trends and outcomes. This data is not always captured in spreadsheets or databases; it often stays close to communities and is used to inform day-to-day interventions and social relations. It might appear as the handwritten list of items needed by a local foodbank, an anecdote exchanged between two youth workers, the list of concerns that a faith leader keeps watch on, sometimes filed at the back of their mind. These ephemeral, sometimes mundane-seeming observations inform day-to-day decision making and relationship building, but often disappear in the ether of daily interactions.

In isolation, these early indicators might appear trivial, but they are often early indicators of a change in the balance of how life is lived. This kind of real-time empirical knowledge is often shared informally, in conversations on WhatsApp or over cups of tea, and tends to be used to inform immediate

action well before (if ever) it appears in a well-structured data set. In the field of foresight – which is often used as an input into technology forecasting – such indicators might be monitored as early strategic indicators, or 'weak signals', gathered to 'mitigate surprise'.²³ This kind of monitoring is relatively common in both military and corporate contexts, where anticipation of emerging threats and trends is considered a strategic asset;²⁴ digital harms, however, are often left to become socially prevalent before interventions are considered. As policy researcher Nanjira Sambuli notes:

*[civil society] organisations have long been a bellwether for the risks and harms that... digital fervour imposes. However, civil society is woefully under-resourced and even undermined in an ecosystem that prioritises digital innovation and quantifiable metrics, whether or not they are appropriate in varied contexts.*²⁵

Stafford Scott, founder and Director of Tottenham Rights and Guest Professor at Forensic Architecture, is a community activist and race equality specialist²⁶ who brought the Gangs Violence Matrix to the notice of Amnesty International. The Gangs Violence Matrix was a predictive analytics system used by the Metropolitan Police to monitor the behaviour of and impose sanctions on young men who were perceived to be at risk of gang membership.²⁷ Amnesty's investigation of the programme – which led to its censure by the Information Commissioner's Office (ICO) in 2018 and subsequent closure in October 2022 – exposed the extent to which young Black men were wrongly criminalised by the system.

Identification of a person as a 'gang nominal' on the Matrix, based on profiling information that included data points such as their nicknames and social-media habits, resulted in enforcement by what the ICO referred to as 'the "Al Capone Approach"', meaning general action was taken against groups of possible gang members, regardless of their individual risk level. Possible sanctions included 'prison licence conditions, increased stop and search, TV licensing, parking enforcement, truancy, benefits sanctions, housing action (including eviction) and immigration enforcement'.²⁸ Amnesty International found that inclusion on the Matrix was unduly weighted towards young Black men, who accounted for 78% of those listed, when 'in reality black people are responsible for just 27% of serious youth crime'.²⁹ Moreover, 64% of those recorded on the Matrix were found to be low risk, and 100 entries were for children under the age of 16.³⁰

²³ For a discussion on the definition of weak signals, see Barbara L. van Veen, J. Roland Ort, 'Unifying Weak Signals Definitions to Improve Construct Understanding', *Futures* (Vol. 134, 2021), 102837, (accessed 26 March 2024)

²⁴ Dewey Murdick, 'Foresight and Understanding from Scientific Exposition (FUSE)', *Office of the Director of National Intelligence* (24 July 2012), (accessed 2 April 2024); Scott Smith and Madeline Ashby, *How to Future* (Kogan Page Inspire, London, 2020).

²⁵ Nanjira Sambuli, 'On the Patient Capital Needed from Philanthropy in Tech', *WINGS*, (accessed 5 January 2022).

²⁶ Stafford Scott, *Forensic Architecture*, (accessed 2 April 2024).

²⁷ 'Gangs Violence Matrix', *Metropolitan Police*, (accessed 2 April 2024)

²⁸ Elizabeth Denham, 'Enforcement Notice', *Information Commissioner's Office* (13 November 2018), (accessed 2 April 2024).

²⁹ 'Trapped in the Matrix: Secrecy, Stigma, and Bias in the Met's Gangs Database', *Amnesty International* (May 2018), www.amnesty.org.uk/gangs (accessed on 6 March 2024)

³⁰ 'Elizabeth Denham, 'Enforcement Notice'.

Speaking about his work in a podcast interview, Scott relayed advice given to him by racial justice campaigner Darcus Howe who said:

“This isn’t about leading people and leading communities, it’s about listening to them, hearing their needs and expectations and then finding ways to facilitate and enable that to happen. So the key thing I think is about how we listen and hear what’s happening in our communities more. And listening means really deep listening. Ambulance people when doing the triage... they don’t rush to the people who are making the most noise, even though they’re in pain and in need, you sometimes have to go to those who are making the least noise because they’re the ones who are more likely to have a fatal crisis within that incident.”³¹

Scott’s ability to listen to young men in his community who had been wrongly associated with gangs meant that he was able to seek redress on their behalf and, ultimately, bring about the closure of a biased and faulty system long before it would have otherwise come to general public attention. In doing so, Scott interrupted what Uphol Ehsan et al. have termed the ‘algorithmic imprint’ of the Gangs Violence Matrix, in which the after-effects of algorithmic decisions have a lingering half-life that ‘can be inferred at the infrastructural, social, and individual levels’ for many years to come.³²

Informal gatherings of people who have experienced a sociotechnical harm are another important source of sociotechnical intelligence. For instance, in the case of the Post Office Horizon IT Scandal, the Justice for Subpostmasters Alliance formed,

with a meeting of 30 victims in 2009 and over the intervening years, the determination of the group was solid and at meetings of the victims of Post Office’s brutality, people who had run businesses often in the heart of a community, met to offer support to others and confirm their resolve to expose the real truth no matter how long it took. It was a long slow laborious process for the JFSA to eventually get Post Office into court in 2017 in a group litigation action by over 550 mainly ex Subpostmaster.³³

At the time of writing, some fifteen years later, the Subpostmasters’ case is subject to a public statutory inquiry; having received evidence from 189 affected people, the inquiry is now seeking submissions from ‘all who may have been affected, including current or former sub-postmasters, family members, friends, community members, or other

members of the public who may have reflections to share’. Inquiry Chair Sir Wyn Williams has reflected on this process, saying:

I know that many more people have been affected either directly or indirectly. Submissions I have received on the subject of compensation also tell me that the impacts on many people’s lives continue to unfold, often with very grave consequences. These stories must be brought into the light.³⁴

Victims of both the Gang Violence Matrix and the Post Office Horizon IT Scandal were wrongly accused of criminal activity; in both cases, community action was needed for harms experienced by individuals to become more widely observable and actionable. And in both cases, affected individuals from less powerful groups – which included many victims who were minoritised due to factors including race and ethnicity, social status, and age – were forced to operate outside of the market and the state, within the space known as civil society, to create the conditions necessary to seek justice from powerful and opaque institutions. The lack of public-facing infrastructure to gather these weak signals and support redress means that a combination of luck and serendipity is currently needed to surface harms of this nature and scale; without a Stafford Scott or the Justice for Subpostmasters Alliance, it is unknowable whether these injustices would ever have been revealed.

Governance as a process of social negotiation

Writing in 1988, Mary Warnock introduced her final report on the ‘social, legal and ethical implications of ... human assisted reproduction’³⁵ with this assessment of legislation:

In recognising that there should be limits, people are bearing witness to a moral idea of society. But in our pluralistic society it is not to be expected that any one set of principles can be enunciated to be completely accepted by everyone ... What is legally permissible may be thought of as the minimum requirement for a tolerable society.³⁶

Warnock grapples with the moral responsibility of setting limits and acknowledges that, while the function of the law is to establish a ‘minimum requirement’, understanding what good looks like is more relative and relational, a matter of social negotiation rather than a hard and fast set of rules. Legislation is a vital part of a society we can ‘praise and admire ... with our conscience clear’,³⁷ but it also exists in the context of personal, moral, and professional judgements.

³¹ Sally Warren and Stafford Scott, ‘A Dialogue of Equals: Stafford Scott on Community Activism’, *The King’s Fund*, (accessed 26 March 2024)

³² Uphol Esan, Ranjit Singh, Jacob Metcalf, Mark O. Riedl, ‘The Algorithmic Imprint’, *AcM FAccT* (2022), (accessed 2 April 2024).

³³ ‘About Us’, *Justice for Subpostmasters Alliance Website*, (accessed 26 March 2024).

³⁴ ‘Inquiry Launches “In Your Own Words” to Hear Impacts of the Scandal’, *Post Office Horizon IT Inquiry* (13 March 2024), (accessed 2 April 2024).

³⁵ ‘M. Warnock, ‘Report of the Committee of Inquiry into Human Fertilisation and Embryology’, *Department of Health* (London, 1988), p. iv.

³⁶ ‘Warnock, pp. 2-3

³⁷ Warnock, p. 3

In 2024 the challenge of setting norms for emerging technologies in a pluralistic society is no simpler.

Demographic and economic shifts in the UK have seen pluralism turn into social divisions, while the proliferation of digital and data-driven technologies has created numerous new challenges and opportunities. Data-driven decisions, digitised public services, social platforms, the development of artificial intelligence, and the disintermediation of news and media are just some of the many technological changes around which societal norms are still developing. As such, what good looks like in a digital society, and how that good is stewarded into being, remains a work in progress. Understanding how tools such as generative AI will unfold in, influence, and harm discrete communities and micro-publics will help to promote social cohesion and produce effective mitigations; waiting for significant impacts to emerge at a general population level will simply surface and create new axes of division.

Hard systems and human beings

Gathering and analysing the unfolding sociotechnical impacts of is also an essential adjunct to the technical auditing and evaluations being conducted by the AI Safety Institute. The Oxford Dictionary of Business and Management describes a sociotechnical system as ‘involving the interaction of hard systems and human beings, in ways that cannot be separated or are thought to be inappropriate to separate’.³⁸ This interaction is complex, systemic, and not always reproducible, as different humans experience different outcomes in different contexts.

At the time of writing, government policy recognises the need for ‘sociotechnical infrastructure’ to understand and mitigate the harms created by emerging technologies. However, this infrastructure is currently situated in the AI Safety Institute, which takes a ‘technically grounded’ approach to understanding societal harms through evaluations of ‘usage data and incident reporting’, with a focus on ‘the direct impact of advanced AI systems’.³⁹

Documents released by the DSIT show that the AI Safety Institute approach starts with technical audits to extrapolate diverse societal impacts on ‘individuals and society’; these are based on outcomes that might be apparent to the Institute’s small staff, ‘a team of technical staff, including researchers, engineers, and behavioural/social scientists’.⁴⁰ The lack of representative empirical inputs to this process shows a strictly ‘hard systems’ view of sociotechnical change that does not fully account for the experiences of people and communities

who are not accurately reflected in extant data. Moreover, it is deeply unlikely that the numerous circumstances of ordinary people can be either guessed at or accurately modelled by the ‘researchers, engineers, and behavioural/social scientists’ tasked with undertaking evaluation.

Such a ‘technically grounded’ approach will also overlook second- and third-order social and political outcomes for general purpose technologies that have been created for adaptability, and will not account for the long-term ‘algorithmic imprint’ effects of data-driven decisions. This model assumes that societal shifts such as workforce displacement and democratic engagement can be accurately forecast and monitored through methods that originate in technical quality assurance. It is, at best, a tool for telling half of the sociotechnical story.

People and context are vital inputs for situating the past, current, and future societal impacts of technologies. The richness of social experience conveyed by civil society is a necessary complement to the evaluation-focussed approach of the AI Safety Institute, and one that will become more important as the rate of sociotechnical change and adaptation increases.

Conclusion

The pattern of the last twenty-five years of regulatory progress – in which harms created by a small number of technologies and technology companies have been allowed to emerge and flourish – is not sustainable in a thriving democracy. The increased adoption of artificial intelligence will intensify the technologies we already have and create new social dilemmas. Anticipating and mitigating more of these harms before they fully emerge will enable more effective innovation and strengthen social bonds. Rather than acting on Mark Zuckerberg’s injunction to ‘move fast and break things’, the Digital Civil Society Observatory will build a rolling evidence base, empower more people to live good lives with technologies, and build the UK’s capacity for anticipatory governance.

While cultivating moon-shots is both exciting and inspirational, it is also a risky undertaking that offers an uncertain yield.⁴¹ The ability to mitigate surprise while enabling more socially equitable outcomes is just as, if not more, important for an innovative state. Putting people before code, and building a Digital Civil Society Observatory, is a concrete step to ensuring public-interest innovation can deliver for everyone in the UK.

³⁸ Jonathan Law, *Oxford Dictionary of Business and Management*, 5th edition (Oxford, Oxford University Press, 2009).

³⁹ Department for Science, Innovation and Technology, ‘[Policy paper: Introducing the AI Safety Institute](#)’, Gov.UK (updated 17 January 2024), (accessed 24 March 2024); AI Safety Institute, ‘[Notice: AI Safety Institute Approach to Evaluations](#)’,

Gov.UK (9 February 2024), (accessed 24 March 2024).

⁴⁰ AI Safety Institute, ‘[Notice: AI Safety Institute Approach to Evaluations](#)’ Sir Patrick Vallance and Dame Nancy Rothwell, ‘[Principles for science and technology moon-shots to achieve by 2030](#)’, Council for Science and Technology (25 June 2020), (accessed on 2 April 2024).

References

- Amnesty International, 'Trapped in the Matrix: Secrecy, Stigma, and Bias in the Met's Gangs Database', (May 2018), www.amnesty.org.uk/gangs (accessed on 6 March 2024).
- ARIA, '[Frequently Asked Questions](#)', (accessed 26 March 2024).
- Bambridge, J., Clarke, L. and Manancourt, V. '[Who Said the UK Can't Do Chips?](#)', *Politico Pro Morning Technology UK* (28 March 2024), (accessed 2 April 2024).
- Bechtel, M. M., Hainmueller, J., Hangartner, D., Helbling, M. 'Reality Bites: The Limits of Framing Effects for Salient and Contested Policy Issues', *Political Science Research and Methods* (2015); 3(3): 683-695.
- Centre for Data, Ethics, and Innovation, '[Public attitudes to data and AI: Tracker survey \(Wave 3\)](#)', *Gov.UK* (6 December 2023; last updated 12 February 2024), (accessed on 2 April 2024).
- Birhane, A., '[Algorithmic Injustice: A Relational Ethics Approach](#)', *Patterns* (Vol. 2, Issue 2, 2021).
- Colom, A. '[Meaningful Public Participation and AI: Lessons and Visions for the Way Forward](#)', *Ada Lovelace Institute* (1 February 2024), (accessed on 5 March 2024).
- Competition and Markets Authority, Information Commissioner's Office, Ofcom, Financial Conduct Authority, '[Collection: The Digital Regulation Cooperation Forum](#)', *Gov.UK* (updated 28 March 2023), (accessed 24 March 2024).
- Criddle, C., Murgia, M., Gross, A. and Yang, Y. 'How Sunak's Bletchley Park Summit Aims to Shape Global AI Safety', *Financial Times* (27 October 2023).
- Denham, E. '[Enforcement Notice](#)', *Information Commissioner's Office* (13 November 2018), (accessed 2 April 2024).
- Department for Science, Innovation and Technology, '[Guidance: AI Safety Summit: Confirmed Attendees \(governments and organisations\)](#)', *Gov.UK* (updated 31 October 2023), (accessed 2 April 2024).
- Department for Science, Innovation and Technology, '[Policy paper: A Pro-Innovation Approach to AI Regulation](#)', *Gov.UK* (updated 3 August 2023), (accessed on 24 March 2024).
- Department for Science, Innovation and Technology, '[Policy paper: Introducing the AI Safety Institute](#)', *Gov.UK* (updated 17 January 2024), (accessed 24 March 2024).
- Esan, U., Singh, R., Metcalf, J., Riedl, M.O. '[The Algorithmic Imprint](#)', *AcM FAccT* (2022), (accessed 2 April 2024).
- [European AI and Society Fund](#), (accessed 2 April 2024).
- Holz, F. de la, 'NYC Has Tried AI Weapons Scanners Before. The Result: Tons of False Positives', *Hellgate* (2 April 2024), (accessed 2 April 2024).
- Independent Inquiry into Civil Society, '[The Story of Our Times: Shifting Power, Bridging Divides, Transforming Society](#)' (November 2018), (accessed 25 March 2024).
- [Justice for Subpostmasters Alliance](#), (accessed 26 March 2024).
- '[Inquiry Launches "In Your Own Words" to Hear Impacts of the Scandal](#)', *Post Office Horizon IT Inquiry* (13 March 2024), (accessed 2 April 2024).
- Law, J. *Oxford Dictionary of Business and Management*, 5th edition (Oxford, Oxford University Press, 2009).
- Levy, S. 'How Hemingway Gradually—Then Suddenly—Defined the Zeitgeist', *Wired* (8 July 2022).
- Livingstone, S. and Bovill, M. '[Families and the Internet: An Observational Study of Children and Young People's Internet Use](#)', *London School of Political and Economic Science* (2001), p. 38, (accessed 2 April 2024).
- Lum, K. and Isaac, W. '[To predict and serve?](#)', *Significance* (Royal Statistical Society, 7 October 2016), (accessed 2 April 2024).
- Murdick, D. '[Foresight and Understanding from Scientific Exposition \(FUSE\)](#)', *Office of the Director of National Intelligence* (24 July 2012), (accessed 2 April 2024).
- Metropolitan Police, '[Gangs Violence Matrix](#)', (accessed 2 April 2024).
- [Online Safety Act 2023](#), *Gov.UK* (25 October 2023), (accessed on 2 April 2024).
- Royal Society of the Arts, '[Democratising decisions about technology: A toolkit](#)', (2019), (accessed on 5 March 2024).
- Smith, S. and Ashby, M., *How to Future* (Kogan Page Inspire, London, 2020).
- Sambuli, N. '[On the Patient Capital Needed from Philanthropy in Tech.](#)', *WINGS*, (accessed 5 January 2022).
- Scott, S. [Forensic Architecture](#), (accessed 2 April 2024).
- UK Research and Innovation, '[£100m Investment in The Alan Turing Institute Announced](#)', (accessed 2 April 2024).
- UK Research and Innovation, '[£100m Boost in AI Research will Propel Transformative Innovations](#)', (6 February 2024), (accessed 26 March 2024).
- Vallance, P. and Rothwell, N. '[Principles for science and technology moon-shots to achieve by 2030](#)', *Council for Science and Technology* (25 June 2020), (accessed on 2 April 2024).
- Veen, B. L. van and Ortt, J. R. '[Unifying weak signals definitions to improve construct understanding](#)', *Futures* (Vol. 134, 2021), 102837.

Warnock, M. 'Report of the Committee of Inquiry into Human Fertilisation and Embryology', Department of Health (London, 1988), p. iv.

Warren, S. and Scott, S. '[A Dialogue of Equals: Stafford Scott on Community Activism](#)', *The King's Fund*, (accessed 26 March 2024).

World Economic Forum, '[The Future Role of Civil Society](#)' (January 2013), p. 8, (accessed 25 March 2024).

The British Academy is the UK's national academy for the humanities and social sciences. We mobilise these disciplines to understand the world and shape a brighter future.

From artificial intelligence to climate change, from building prosperity to improving well-being – today's complex challenges can only be resolved by deepening our insight into people, cultures and societies.

We invest in researchers and projects across the UK and overseas, engage the public with fresh thinking and debates, and bring together scholars, government, business and civil society to influence policy for the benefit of everyone.

The British Academy
10-11 Carlton House Terrace
London SW1Y 5AH

Registered charity no. 233176

thebritishacademy.ac.uk
Twitter: @BritishAcademy_
Facebook: TheBritishAcademy

© The author. This is an open access publication licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License.

To cite this report:
Coldicutt, R. (2024). *People Not Code: The Case for a Digital Civil Society Observatory*, The British Academy.

doi.org/10.5871/digital-society/9780856726910.001

ISBN 978-0-85672-691-0

Published September 2024