

An archive of contestation: Digital surveillance and human-machine recompositions of resistance

Big Data & Society
July–September: 1–17
© The Author(s) 2026
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20539517261467353
journals.sagepub.com/home/bds



Emma Mc Cluskey¹, Claudia Aradau² , Tobias Blanke³ , Thais Lobo⁴
and Tommaso Venturini⁵

Abstract

This article proposes to make sense of the ‘atmosphere of failure’ diagnosed by key players in the mobilisation against mass surveillance in the United Kingdom. Firstly, by entangling oral history interviews and computational analyses of NGO websites, we recompose an archive of contestation that is not temporally limited to the Snowden ‘moment’. We advance the notion of a ‘discordant machine’ to render the generativity of discordances and surprises in entangling qualitative and digital methods. Secondly, we argue that the campaigns against surveillance did not lead to a shared vocabulary of critique and a collective practice of contestation. We show that contesting digital state surveillance was mostly expressed through the conduits of privacy and transparency, while more collective democratic claims were connected instead to ‘elsewheres’, such as democratisation in Global South states. Although the Snowden disclosures were a key moment for academic scholarship, our analysis shows that they played little role in mobilisations against mass surveillance. References to oversight were limited to the immediate post-Snowden moment. Finally, corporate surveillance remained a feature of NGO advocacy but was represented mainly at a distance from state surveillance. Our analysis suggests that the challenge for mobilisations against digital surveillance is not only to hold together its complexities but also to articulate its antidemocratic effects.

Keywords

Digital surveillance, democracy, civil society, critique, mobilisation, resistance

Introduction

On 15 June 2017, a blog post on the Don’t Spy on Us (2017) coalition website announced that it had ended. Set up in the wake of the 2013 Snowden disclosures of mass surveillance by intelligence, police and security agencies, the coalition brought together civil society organisations calling for accountability and a stop to mass surveillance in the United Kingdom. The executive committee of the Don’t Spy on Us coalition consisted of six organisations: Article 19, Big Brother Watch, English PEN, Liberty, Open Rights Group and Privacy International. The formal coalition disbanded in the wake of the passing of the UK’s Investigatory Powers Act in 2016. It had campaigned to amend the law but ultimately failed to bring about substantive changes.

This atmosphere of failure to stop the legalisation of mass surveillance became evident to us as we researched civil society responses within the framework of intelligence oversight. By 2020, when we were conducting interviews,

mass surveillance was largely viewed by our civil society interlocutors as ‘the ship that sailed’. The story we repeatedly heard was one of failure, dejection and the normalisation of mass surveillance (Interviews with civil society members A, B, C, D).² As the Investigatory Powers Act – also known as ‘Snoopers’ Charter’ – passed into law, one expert labelled the debate around mass surveillance ‘now effectively closed’ (Danezis, 2016).

¹Criminology, University of Westminster, London, UK

²War Studies, King’s College London, London, UK

³University of Amsterdam, Amsterdam, The Netherlands

⁴War Studies, King’s College London, London, UK

⁵University of Geneva, Geneva, Switzerland

Corresponding author:

Tobias Blanke, University of Amsterdam, Postbus 94550, 1090 GN Amsterdam, Netherlands.
Email: t.blanke@uva.nl



What struck us from the interviews was the rather peripheral role of the Snowden disclosures, mentioned only in relation to the changing media landscape that they brought into being (Interviews with civil society members A, B). Most civil society interlocutors, however, began their stories about contesting surveillance some years prior, often in relation to older iterations of surveillance legislation, such as the 2000 Regulation of Investigatory Powers Act, or even the informal gatherings of activists in pubs in 1990s London, which concerned debates about encryption.

As these stories of resisting surveillance were not centred around the Snowden disclosures, we became curious about how this failure came to be. The framing of our story, therefore, took place within this overall narrative arc of disillusionment regarding mass surveillance and the bill's passage into law, following several years of contestation by civil society organisations, with few identifiable 'wins' at the end of the process.³ Across the social sciences, scholars have already drawn attention to the limitations of practices of oversight – even broadly understood to encompass contestation – to eliminate or even significantly reduce mass surveillance (Brown et al., 2015; Brown and Korff, 2009; Kniep et al., 2024; Tréguer, 2023). They saw the increased legalisation of practices of digital surveillance alongside the proceduralisation of oversight as ultimately legitimising state violence and producing non-accountability (Bigo et al., 2024). How could we account for this atmosphere of failure? Could we trace shifts in discourse that would help us to better understand how the scandals of mass surveillance fizzled out and how the practice ultimately became normalised?

To answer these questions, we took an entangled methodological approach, combining different qualitative and digital methods by moving back and forth between them.⁴ The article is an invitation to entangle methods, not to develop a more comprehensive analysis of a topic or issue but to show how these combinations enable the production of 'discordances' that challenge rather than complement findings. Despite the Don't Spy on Us campaign and civil society mobilisation, we argue that a shared vocabulary for contesting mass surveillance ultimately did not crystallise, and that many asymmetries remained in the terms and targets of contestation.

Mobilising 'oral history' (Passerini, 2002, 2017), we conducted interviews with key actors involved in contesting state surveillance in the United Kingdom and asked them how they remembered events playing out, both within parliamentary legislative struggles and in public advocacy and litigation. To understand the 'atmosphere of failure', we drew on a repertoire of digital methods to analyse the public-facing online material from NGOs at the forefront of the campaign against the legalisation of mass surveillance. Digital methods were formulated as techniques that repurposed the affordances of social media (Rogers, 2013; Venturini et al., 2018: 4198). Distinct from approaches

that see the internet as an alternative sphere for traditional methodologies such as online surveys, digital methods interrogate the native data architectures and algorithmic logics inherent to online platforms, creatively adapting these mechanisms for social and political inquiry (Venturini et al., 2018). Since their initial focus on social media devices, digital methods have expanded to address other digital devices that can be repurposed to answer social and political questions. As we were interested in the vocabularies and trajectories of contestation, we also supplemented the traditional repertoire of digital methods, focused on 'methods of the medium' (Rogers, 2013: 1), with topic modelling, a method used for mapping themes across large datasets.

To trace key themes in the NGO discourses, we used the Wayback Machine to retrieve the websites of the seven organisations involved in the Don't Spy on Us campaign and the mass surveillance litigation before the European Court of Human Rights (Big Brother Watch and Others v the UK, 2018, 2021). While most of the academic literature on the contestation of mass surveillance has focused on the Snowden 'moment', we have analysed the contestation of mass surveillance over time through the websites of the seven pre-selected NGOs.

To explore these changes in the contestation of mass surveillance, the article is structured in four sections. Firstly, we discuss how the literature that has analysed mobilisations and contestations of mass surveillance has tended to focus on a particular moment in time – the 'Snowden' moment. Secondly, we introduce our oral histories of key players in the civil society mobilisations to draw out key themes as a point of departure for investigating the evolution of the NGOs' public-facing discourse. Thirdly, we unpack how we have traced these topics in the online material using digital methods and developed the notion of 'discordance machines' for multi-method inquiries. Lastly, we offer an analysis of the vocabularies and themes around (mass) surveillance, which helps us crystallise our argument about the limitations that the archive of contestation reveals and account for the perceived failure that our interlocutors spoke about.

Contesting mass surveillance: the centrality of the 'Snowden moment'

The 2013 Snowden disclosures of mass surveillance by intelligence agencies appeared to revitalise the contestation of surveillance and drew attention to new digital forms. The language of Big Data and datafication helped scholars attend to the specificities of large-scale or mass surveillance. Conversations about surveillance post-Snowden have taken place in various scholarly interstices, spanning international relations, law, surveillance studies and sociology more broadly. Much of the literature has centred on the

‘Snowden moment’ and has focused on parliamentary inquiries, hearings, new legislation and strategic litigation to contest mass surveillance in the wake of the disclosures (Aradau, 2017; Aradau and Mc Cluskey, 2022; Bauman et al., 2014; Dencik and Cable, 2017; Dencik et al., 2016; Gros et al., 2017; Hintz and Brown, 2017; Ochoa et al., 2021; Stevens and Allen-Robertson, 2021; Tréguer, 2017). Many scholars have also shown the limitations of these contestations, as mass surveillance could be challenged only to a limited extent; it was often legalised and subjected to ‘safeguards’ rather than being banned or reduced. While much of the focus was on state surveillance, there was also increasing attention to corporate surveillance.

Some of the earliest conversations within surveillance studies and adjacent disciplines emerged from a normative perspective, asking what anti-surveillance resistance should look like, and how it could be more effective (Bennett, 2010; Fernandez and Huey, 2009; Monahan, 2006). Whilst these discussions were dominated by individualised acts of resistance to surveillance in everyday life (Gilliom and Monahan, 2012; Huysmans, 2014; Marx, 2014), some dialogues within legal studies post-Snowden opened a space for thinking about collective challenges to surveillance. Gandy (2021) notes that the literature has focused heavily on techno-legal responses related to the development and use of encryption, as well as on policy advocacy around privacy and data protection. However, more recently, some scholars have argued that digital surveillance should instead be framed in relation to collective questions of social justice.

Against the so-called atomisation of individuals, who are urged to take measures to protect their personal data, a more collective approach understands data as the outcome of shared relations (Dencik et al., 2016; Ruppert et al., 2015). In this regards, Elspeth Guild (2019) speaks of the ‘data citizen’ to anchor the different subjectivities produced by data surveillance within a rights framework as a means of resisting mass surveillance. Privacy has also been re-framed, moving away from narrow understandings related to data and information to link the concept of bodily integrity through the notion of ‘physical privacy’ (Hariharan, 2023). A separate line of critique has addressed the socio-materialities of digital surveillance, including questions of Big Data, machine learning and the infrastructures enabling the extraction and exchange of data (e.g. Andrejevic, 2019; Aradau and Blanke, 2015; Bauman et al., 2014; Lyon, 2014, 2015). These interdisciplinary debates have often drawn on science and technology studies to bring discussions of the entangled social, technical and political dimensions of Big Data and machine learning algorithms to critical studies on security and surveillance.

Another dimension of these scholarly and public conversations concerns the connections between state and corporate surveillance. The existing literature points to the

ambivalent role of big tech companies in enabling and opposing digital surveillance. In her study of the constellation of actors formed to resist the UK’s Investigatory Powers Act of 2016, Amy Stevens (2024: 90) casts these corporations as part of an ‘assemblage of resistance’, whereby the precarious and transient alliance of big tech companies with civil liberties and digital rights organisations momentarily coalesced to stop the passing of the law. Stevens and Allen-Robertson (2021) demonstrate how, in the UK state surveillance debate, justifications for encryption became intertwined with human rights discourses to divert public attention away from big tech companies’ own practices of corporate surveillance. Felix Tréguer (2019) has mapped the evolution of the tech industry within a ‘security assemblage’ on a cooperation-resistance spectrum, shrewdly demonstrating how the immediate overt resistance to mass surveillance post-Snowden was accompanied by routine ‘backstage’ cooperation in the mining of metadata for use by security professionals. He argues that, after 2015 and under the banner of ‘data governance’, such cooperation has increased through measures to police propaganda and counter ‘hate speech’ (Tréguer, 2019: 152).

The increasing merging of surveillance, self-exposure on social media, extraction of personal data through sensors and other devices, and ‘intelligence’ by governmental agencies has raised questions about the target of criticism. Some scholars have unpacked the conditions for this ‘normalisation of surveillance’ (Dencik and Cable, 2017; Eley and Rampton, 2020; Lyon, 2001), for instance, through the role of social scientists in ‘building societal resilience’ towards surveillance-based security technologies under various research funding streams in Europe (Hayes et al., 2009; Jones, 2017; Larsson, 2020; Martin-Mazé, 2020; Martin-Mazé and Perret, 2021). The ‘normalisation’ of surveillance also unfolds through the reproduction of the ‘myth of digital surveillance as active and individualised, a Big Brother vision of surveillance that ignores or erases how automated surveillance and algorithmic evaluation of human beings can reinforce and exacerbate structural inequalities’ (Kelly, 2022). Other scholars have examined the narrowing of democratic vocabularies in challenges to surveillance, reduced to concepts such as ‘compliance’ (Kuldova, 2022), ‘transparency’ (Viola and Laidler, 2022) and ‘[re-building] trust’ (Mc Cluskey and Aradau 2023), which neuter critique along technical lines. Reappropriating and hyper-visibilising the normalisation of surveillance through aesthetic and performance art has also been conceived as a modality of resistance, albeit from a position of vulnerability (Hogue, 2024; Huysmans, 2014).

By orienting our analysis towards transformations over time rather than at a specific ‘moment’, we aim to shed light on how contestations of surveillance changed over time. Instead of labelling actors as ‘resisting’ or ‘cooperating’ with the expansion of surveillance, we look at ways in

which they are connected and linked in the digital realm, and how these constellations ebb and flow.

Oral histories of challenging surveillance

For our oral history interviews, we selected seven NGOs, which played a crucial role in resisting and critiquing mass surveillance. Of the six NGOs that formed the executive committee of the Don't Spy on Us coalition, Privacy International, Liberty and Big Brother Watch were the main players in bringing forward litigation, with the former two forming larger transnational constellations of NGOs. Big Brother Watch was established with the aim of resisting state surveillance and was also notable as the only NGO operating in the political space of the centre-right (Interview with civil society member A). Open Rights Group was the only NGO at the time to focus solely on digital rights issues and to advocate for greater public engagement (Interview with civil society member E). Both Article 19 and English PEN were deemed significant because 'freedom of expression' played an important role in the litigation procedures (Big Brother Watch and Others v the UK, Grand Chamber Judgement, 2021). English PEN had never litigated against the UK government before and had traditionally viewed privacy advocates as being antagonistic to freedom of expression. The seventh NGO, Amnesty International, was not a member of the Don't Spy on Us executive committee. Nonetheless, Amnesty International was identified by other actors as crucial for the symbolic capital and resources it brought to both the litigation and the parliamentary struggle, even though it was initially said to be hesitant about getting involved (Interviews with civil society members F, D, C).

The interviews took place over approximately 16 months, from January 2020 to May 2021, as part of a broader pan-European project on oversight and accountability of transnational data sharing by state intelligence agencies. Though we spoke to around twenty key practitioners who were directly involved in attempts to reform intelligence oversight in the wake of the Snowden disclosures, this article draws primarily on seven interviews with civil society activists most directly involved in the Don't Spy on Us coalition. The interviews took the form of oral histories, a method focused on eliciting memories (Passerini, 2017; Thompson, 2002). We asked participants to recount their recollections of the advocacy campaigns against mass surveillance, which were often intertwined with involvement in legislative struggles and strategic litigation.

For historian Paul Thompson, the merit of oral history is that it allows for a democratisation of history through a pluralisation of voices and perspectives. Borrowing from the simple definition provided by the Oxford English Dictionary, he builds on oral history as 'tape-recorded historical information drawn from the speaker's personal knowledge', but emphasises the importance of a

transdisciplinary and thus socially conscious, approach in the (re)telling of the past (Thompson, 2000: xi). Framing oral history as a distinctly political project with the aim of enacting concrete political interventions, Thompson points to the capacity of oral narratives, as opposed to just archival documents, to allow for 'the original multiplicity of stand-points to be recreated' (xx). Passerini (2017: 86–91), however, moves beyond the use of oral sources as merely factual to bring oral history closer to anthropology, considering the method to be primarily concerned with capturing 'an expression and representation of culture', also able to allow analyses of a subject's multifarious relations to institutional power. For Passerini, therefore, a type of reflexivity is necessarily built into oral history as an analytical tool. An interlocutor's reaction may force us to question our habits of thought and the way in which we have formulated questions.

Applying thematic analysis to the interview data, we identified recurring topics and situated these within broader frames related to the normalisation of mass surveillance. We began framing these topics as a series of binaries: state surveillance versus corporate surveillance, general democracy versus singular democracy and resistance versus oversight. Although these topics overlapped and intersected in our interlocutors' recollections, the binary structure enables us to trace relations and connections in the web archives. The final category, actors, was intended to map relations between entities that critiqued surveillance, looking to discover the extent and form of coalition-building among surveillance campaigners.

The first binary pointed to a tension that repeatedly emerged in interviews between state and corporate surveillance and how to articulate arguments about the intrusive capabilities of the state versus big tech. Actors often shifted from one to another to deflect or attribute responsibility. The mobilisation against the Investigatory Powers Act took place at a time when intelligence agency spokespersons publicly attacked big tech for becoming the 'command and control centre' for terrorists (Hannigan cited in Quinn et al., 2014). Yet, the role of big tech as a mediator and enabler of state surveillance was partly eclipsed in public discourse about mass surveillance.

In some NGO narratives, this ambivalence was expressed as a tension how to frame arguments in order to gain the most traction and how to engage certain 'publics' they attempted to reach through the Don't Spy on Us campaign:

These powers that the Home Office keep telling us are being used to fight terrorists are being used by the taxman.....That [message] cuts through with a section of the public who don't like terrorists and don't like paedophiles but nonetheless do think that the state uses its powers towards them (Interview with civil society member A).

The blurring between state and corporate surveillance was also articulated along the lines of the 'revolving

door' between state and private actors (Interviews with civil society members A and B), and the limits in challenging the materiality of surveillance architecture, which could not be reduced to solely state or solely corporate (Interviews with civil society members G and B).

The second binary captured the distinction between more individualised and more generalised democratic claims ('singular' vs. 'general' democracy). We worked with a heuristic distinction between rights held in common, such as in discourses of civil liberties, versus more individualised rights, such as privacy. However, we wanted to know whether this distinction was operational or whether any distinction between different dimensions of democracy could be observed over time or across campaigns. In our interviews, the exasperation at being forced into narrower frames of critique was evident in how many of our civil society interlocutors discussed the legislative struggle against the Investigatory Powers Act. Frustration lay in the parameters of the debate already being set by the security agencies, with civil society 'essentially tagging along' (Interview with civil society member E), or expressing resignation that 'the winds of the time were what they were', constraining and limiting critique to narrow, procedural interventions (Interview with civil society member D). As one of our interlocutors explained, 'We did not talk about how problematic it was to have mass surveillance per se. We just continued to have the same arguments ... hiding behind the idea that "we need more safeguards".' (Interview with civil society member E). Although this was a strong feature in memories of the legislative struggle, we were interested in whether the public campaigns simultaneously transformed their vocabularies for critiquing resistance, as these 'winds' were changing and the terms of debate were increasingly being set by the agencies and the UK government.

The final binary we identified was between 'oversight' and 'resistance'. This emerged from a distinction between more or less formalised practices of contestation and the way in which accountability was enacted in different spheres. This was particularly relevant to us as the Don't Spy on Us coalition advanced six principles of reform for UK surveillance laws, including oversight. The six principles were: No surveillance without suspicion; transparent laws, not secret laws; judicial, not political, authorisation; effective democratic oversight; the right to redress; and a secure web for all (Don't Spy On Us, 2016). We were particularly interested in how the idea of oversight was taken up (or not) by our civil society interlocutors. What was their relationship to 'oversight'? Did they see themselves as part of the landscape of oversight?

Taking our cue from Passerini (2017) and paying attention to the silences and discrepancies in our oral material, it was notable how little 'oversight' was mentioned by our interviewees, despite it being one of the six principles for reform. For the most part, when oversight was invoked, it was in relation to formal structures of external oversight, with

comparisons made to earlier (pre-Investigatory Powers Act) commissioners who were framed as being more open and transparent than the current independent executive oversight body (Interviews with civil society members E, F, G). Comparisons were also made with other oversight systems in different jurisdictions, with Germany and the United States described as 'better' than the United Kingdom (Interviews with civil society members F, G).

Many interlocutors hesitated when we asked if they considered themselves part of the 'oversight landscape'. Instead, some preferred to label themselves as 'activists' or 'campaigners' (Interviews with civil society members A, C, E) and saw what they were doing more in terms of organising and resisting surveillance from the outside (Interview with civil society member D). When pressed, one of our participants framed the work of reading legislation line by line, challenging parts of it and engaging with MPs as oversight (Interview with civil society member D), but the work of public advocacy was not conceived of in these terms.

This absence of spontaneous invocation of 'oversight' was important in prompting us to reformulate how we approached the social and political space of these civil society actors. We included it as a topic to trace whether this absence could also be observed in the web archives, and positioned it next to a broader, more informal notion of 'resistance', which included mechanisms of holding to account such as whistleblowing and campaigning that did not fit into the rather tightly demarcated conception of oversight put forward by our civil society interlocutors (but see Kniep et al., 2024).

Lastly, we included 'actors' to interrogate the connections between each NGO and, for example, different activists, journalists and experts. The aim of highlighting these topics was to trace their presence in the NGO webpages. How did mobilisation against surveillance play out on these websites? How did topics shift over time? We also wanted to illuminate the transversal lines and connections between the actors.

Entangling this oral history with digital methods allowed for the machine to be thought of as an 'interpretive ethnographer' of sorts. In a way analogous to how our assumptions about countering surveillance can be challenged or reinterpreted in our encounters with interlocutors, so too can digital methods help rewrite the categories through which we structured our understandings of contesting surveillance. Munk et al. (2022) have developed the notion of a 'thick machine' to render the ability of machine learning algorithms to explicate rather than explain: to do the interpretive work of making sense of cultural texts and pointing the researcher in the direction of interesting areas of Geertzian 'thick description'. 'Thick' carries here an ambiguity, as these scholars note that it is often on the occasions in which the machine *fails* that post-hoc interpretability becomes interesting, as 'failure' suggests ambiguity and layers of meaning; the anthropological 'moment of clarity' (Munk et al., 2022: 12).

In our entangled methods approach, it was not simply ‘failure’ but also unexpected and discordant dimensions and connections that required us to revisit our initial framings of surveillance, resistance and oversight. Discordance and surprise were crucial to our interpretive endeavour. As the following section discusses, what we had thought important was absent from the computational results, with the category of oversight completely overwritten and transformed into a concept concerned with censorship, books and culture.

Data sprint as methodological event: tracing counter-surveillance in the web archive

For our interdisciplinary methodological experiment, we drew on the format of the ‘data sprint’, an intensive research and coding workshop where participants work together on a research question (Venturini et al., 2018). Our data sprint was a two-day event that brought together researchers working on surveillance and security, qualitative methods, digital methods, Artificial Intelligence and archives to analyse the websites of the seven NGOs engaged in contesting mass surveillance. The data sprint enabled prior insights from the oral history method to shape how subsequent methods were approached and refined.

The WayBack Machine (Figure 1), the digital library founded by the Internet Archive to preserve versions of websites, was a means of studying the history of anti-surveillance campaigning. Using the tool’s calendar function, the NGOs’ archived pages were qualitatively assessed to verify the most relevant page for each website at specific periods; that is, the page that provided the best description of the NGOs’ activities and campaigns. When the homepage did not offer such a description on a given date, more suitable page sections were selected (e.g. ‘news’, ‘campaigns’ and ‘blog’). Finally, we had to account for changes to URL addresses, as some organisations had changed their domain names over time (e.g. libertyhumanrights.org.uk/ vs. liberty-human-rights.org.uk).

With the seed URLs, we employed Link Ripper (Digital Methods Initiative, 2007) to harvest archived webpages and their timestamps. The Internet Archive does not sample websites equally.⁵ Its crawl priority is higher for sites that are more likely to change frequently, such as websites of larger organisations, news pages, blogs and other dynamic content. This imbalance means that smaller NGOs – those with fewer resources and less dynamic content – risk being underweighted in computational processing. To redress this imbalance, we decided to sample up to four-page versions per month from the Link Ripper output, if these were available. Otherwise, we selected the maximum number of available page versions, which can also be read as an indication of activity from the perspective of the Wayback Machine.

Figure 2 shows the number of websites by NGO, from each website’s date of origin – the earliest being Amnesty in October 1996 – until June 2022, when the collection was conducted. As international NGOs like Amnesty have different sites depending on their operations worldwide, we chose Amnesty International UK.

We then downloaded the HTML pages and extracted their texts. The Internet Archive generally does not capture dynamic JavaScript elements, resulting in simplified HTML. To extract the text from the HTML, we employed `readtext`, which skips non-textual content, such as script tags.⁶ `readtext` is written in R, which has been our programming language throughout. This approach was manually verified, and we removed most of the non-textual content. While not perfect, we worked with the affordances of the Wayback machine. In the next step, we de-duplicated the data by removing documents with more than 90% similarity. This further addressed activity-biased sampling and created a dataset of unique content production. The de-duplication employed the `TextReuseCorpus` tool (Mullen and Li, 2026), which uses Jaccard similarity on 7-grams. The choice of 7-grams is long enough to avoid noise but short enough to capture overlapping passages.

We standardised the websites further by removing empty texts, lower-casing, deleting non-Unicode characters and merging NGO names into single tokens (i.e. Article 19 becomes `Article_19`). We also removed non-relevant words or characters, such as whitespaces and stopwords, and employed Google’s 10,000 most common English words⁷ list to filter out misspelt words and non-English ones. This resulted in a corpus of over 2700 websites, comprising over 19 million words, which represents the transformation of civil society public campaigns. During the data sprint, the new dataset aided our discussions and experimentations to explore the intersections and differences of the insights from the interviews, topic modelling and network visualisation.

To trace the topics we were interested in across time, we used topic modelling, which can help investigate and summarise developments and transformations over time by identifying underlying topics within large collections of text, without relying on predefined structures (e.g. Blanke and Wilson, 2017; Guldi, 2019). We chose Latent Dirichlet Allocation (LDA) topic modelling (Jelodar et al., 2019), which assumes that each document is associated with multiple topics (k) with certain probabilities, and that each topic consists of several predefined keywords. Interpreting topic models can, however, resemble ‘reading tea leaves’ (Chang et al., 2009), because the topics that emerge are often abstract and ambiguous, requiring human interpretation. To counteract these limitations, we used the topics identified through the oral history interviews. Therefore, we also employed a modified version of topic modelling, known as ‘seeded’ or ‘guided’ topic modelling. Rather than opposing human and machine or working with

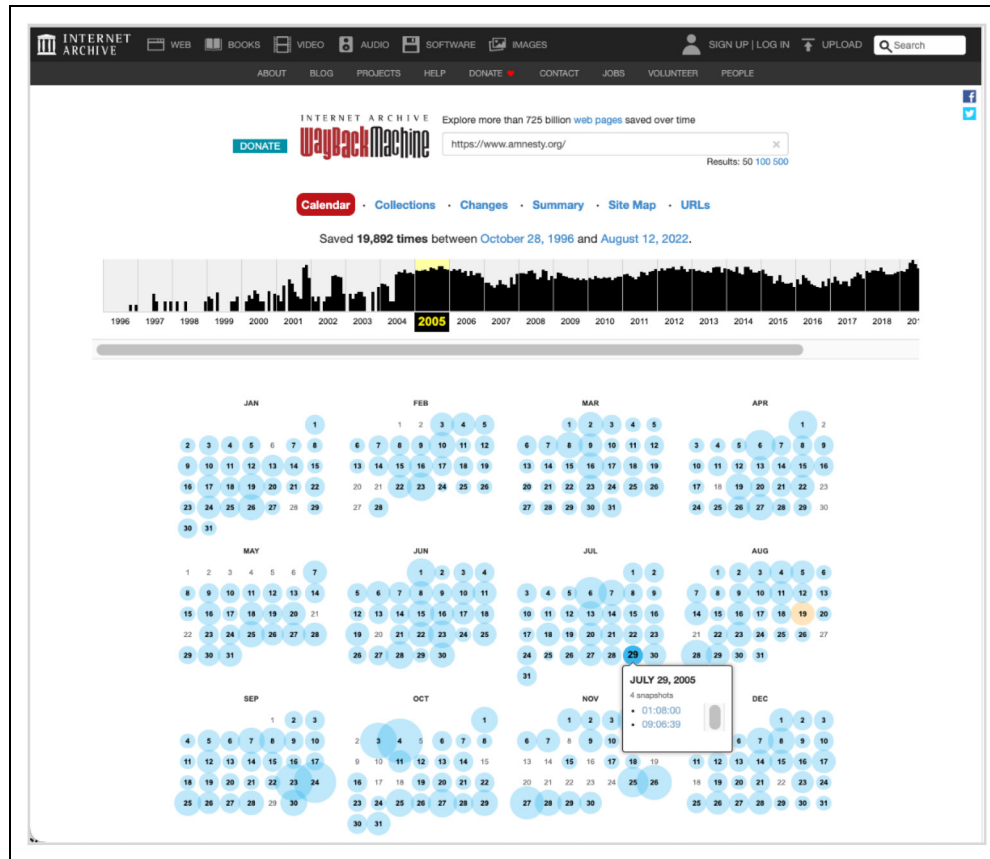


Figure 1. Internet Archive's WayBack Machine view of Amnesty's website and timestamps indicating changes.

the machine as an explainer, seeded topic modelling entangles human and machine interpretations. Meaning and interpretations change, create surprises and bring up unthought-of discordances.

In standard LDA, terms that frequently co-occur are grouped into topics and, after several iterations, the model reaches convergence. By contrast, in seeded LDA, topics are not just derived from the document collection but are influenced by predefined 'seed' keywords, which can be read as introducing a 'bias' or an interpretive orientation in the model (see also Jagarlamudi et al., 2012). R's *seededlda* package guides topic models with 'seed words' derived from our oral history interviews. For example, the topic of state surveillance was seeded with words such as state, security and police. Table 1 shows all the initial topics and keywords that we assumed to find in the websites before running the topic modelling.

Based on the interviews, we expected phrases such as 'mass surveillance' to be present. Yet, we quickly found that they were not important on the websites, so we decided to use keywords only. The seeded topic model then grouped documents where these words occurred, while still discovering new, related terms. This is a probabilistic process in which the seeds act as an orientation rather than a strict

rule. If seed words do not appear on the websites, they have no effect. Therefore, topics may drift, mix or include unexpected words, which is what we hoped to achieve.

Table 1 shows that our 'seeded' topics differentiate between state and corporate surveillance, general and singular democracy, resistance and oversight, while adding actors as an independent topic. As the previous section showed, the seven topics were derived from themes in the oral history interviews. Distinctions such as state versus corporate surveillance or general versus singular democracy directly reflected tensions the interviewees articulated. Actors encompassed publics, media, experts and institutions. Keywords were then selected through an iterative process that combined interview vocabulary, domain knowledge and unseeded LDA runs on the corpus. Reading sample documents allowed us to verify whether the terms were used in the expected sense. During the data sprint, we juxtaposed our own expectations for keywords from the interviews and scholarship on mass surveillance with what the machine brought together.

Reviewing the results of our iterations, we created a new predefined dictionary of seed keywords to be allocated to topics, and a seeded topic model was created from the seeds in Table 1. Running this model during the data sprint

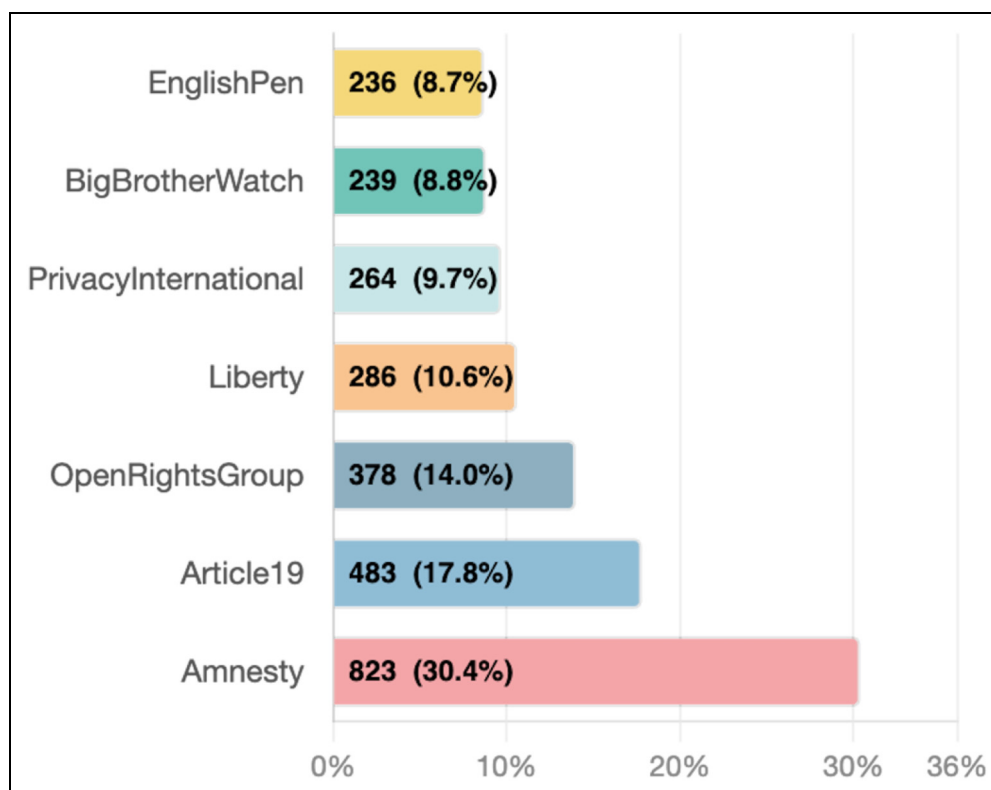


Figure 2. Number of NGO pages harvested with WayBack machine.

delivered the final topic model in Figure 3. Its topics combine keywords from the seeds with new and partly unexpected ones.

Figure 4 is the first visualisation that resulted from the data sprint. For each archived webpage, the model estimates

Table 1. Associated keywords for seeded topic models.

State surveillance	<i>State, mass surveillance, agency, agencies, home office, data, information, hack*, security, police, bulk data, interception, track*, database, spy*</i>
corporate_surveillance	<i>facebook, big tech, data, information, track, twitter, google, microsoft, amazon</i>
general_democracy	<i>democracy, rights, digital rights, civil liberties, freedom of expression, free speech, discrimination, chilling effect]</i>
singular_democracy	<i>privacy, transparency, safeguard*, trust, data protection, rights, democracy]</i>
actors	<i>public, press, media, journalist*, activist*, expert, experts, snowden, parliament, committee, commissioner, congress, government, wikileaks</i>
resistance	<i>whistleblow*, campaign*, scandal, petition, lobby*, report, media, press, court, leak*</i>
oversight	<i>scrutiny, oversight, snoop*, act, bill, tribunal, court, commissioner, committee, oversee, control</i>

The asterisk matches any ending.

the probability that it reflects each seeded topic from Table 1, weighted by the density and co-occurrence of the associated keywords across the text of the website rather than simple keyword counts. These per-page probabilities are then averaged across all pages belonging to each NGO, yielding a single score per NGO per topic, thus allowing systematic comparison of topic prominence across organisations. It is notable, for example, that the topic ‘singular_democracy’ is more prominent in some NGOs such as Amnesty and Liberty compared to English PEN. In fact, many topics are limited in their importance to one or two NGOs. Our definition of an ‘oversight’ topic, for example, features strongly only with English PEN. As we show further down, this is based on a complete reconfiguration of the meaning of oversight.

What does the ebb and flow of topics over time tell us about the atmosphere of failure with which we started the article? In Figure 5, we visualised the mean topic distribution per year. Some topics appear to have up-and-down spikes, while others have more meandering contours. Again, ‘oversight’ is limited, being mainly relevant in the immediate post-Snowden year. ‘Singular democracy’ appears almost throughout but has distinct peaks in importance in the early 2000s and around 2020–21. The ‘actors’ topic seems to disappear after 2017–18, while ‘general democracy’ is also limited temporally.

state_surveillance corporate_surveillance general_democracy singular_democracy			
[1,]	"data"	"data"	"rights"
[2,]	"information"	"information"	"discrimination"
[3,]	"police"	"google"	"democracy"
[4,]	"security"	"microsoft"	"islands"
[5,]	"state"	"track"	"human"
[6,]	"database"	"amazon"	"republic"
[7,]	"spy"	"org"	"international"
[8,]	"agencies"	"copyright"	"new"
[9,]	"agency"	"digital"	"guinea"
[10,]	"tracking"	"can"	"south"
[11,]	"tracks"	"people"	"news"
[12,]	"tracker"	"open"	"congo"
[13,]	"hack"	"comments"	"justice"
[14,]	"track"	"posted"	"palestinian"
[15,]	"spyware"	"internet"	"korea"
[16,]	"tracked"	"blog"	"french"
[17,]	"hacker"	"online"	"samoa"
[18,]	"surveillance"	"like"	"uk"
[19,]	"privacy"	"uk"	"country"
[20,]	"uk"	"one"	"sudan"
actors	resistance	oversight	
[1,]	"government"	"campaigns"	"bill"
[2,]	"press"	"report"	"act"
[3,]	"public"	"campaign"	"control"
[4,]	"media"	"press"	"court"
[5,]	"journalists"	"media"	"committee"
[6,]	"parliament"	"court"	"commissioner"
[7,]	"committee"	"petition"	"tribunal"
[8,]	"journalist"	"lobby"	"literature"
[9,]	"activists"	"read"	"translation"
[10,]	"commissioner"	"english"	"prize"
[11,]	"experts"	"us"	"book"
[12,]	"expert"	"expression"	"english"
[13,]	"congress"	"freedom"	"books"
[14,]	"article"	"cookies"	"david"
[15,]	"freedom"	"work"	"international"
[16,]	"expression"	"latest"	"writing"
[17,]	"africa"	"writers"	"festival"
[18,]	"asia"	"events"	"world"
[19,]	"east"	"website"	"poetry"
[20,]	"read"	"translation"	"writers"

Figure 3. Topic modelling results based on seed topics and keywords.

Seeded topic modelling does not offer an explanation for these changes over time, but it produces surprises and discordances. Rather than the ‘thick machine’ that Munk and colleagues described, we propose to speak of a ‘discordant machine’. In his engagement with John Dewey’s logic of inquiry, anthropologist Paul Rabinow explains the centrality of discordance: ‘Thinking, then, is a situated practice of active inquiry, the role and goal of which is to initiate a motion that results in a movement from a discordant situation to a less discordant situation’ (Rabinow, 2002: 137). Therefore, a ‘discordant machine’ sets an inquiry in motion, working with existing discordances without

aiming to flatten, surpass or otherwise erase them. To work with these discordances, we performed one more entanglement by returning to the method of textual analysis, as we will discuss in the final section.

As our inquiry started from the Don’t Spy on US coalition, we were interested in the connections that sustained the campaign. During the data sprint, we also developed a hyperlink analysis (Rogers, 2019) to make visible the underlying network of affiliations – both among the Don’t Spy on Us coalition and with external entities they were connected to – as well as shared digital infrastructures that were not immediately apparent in the textual content

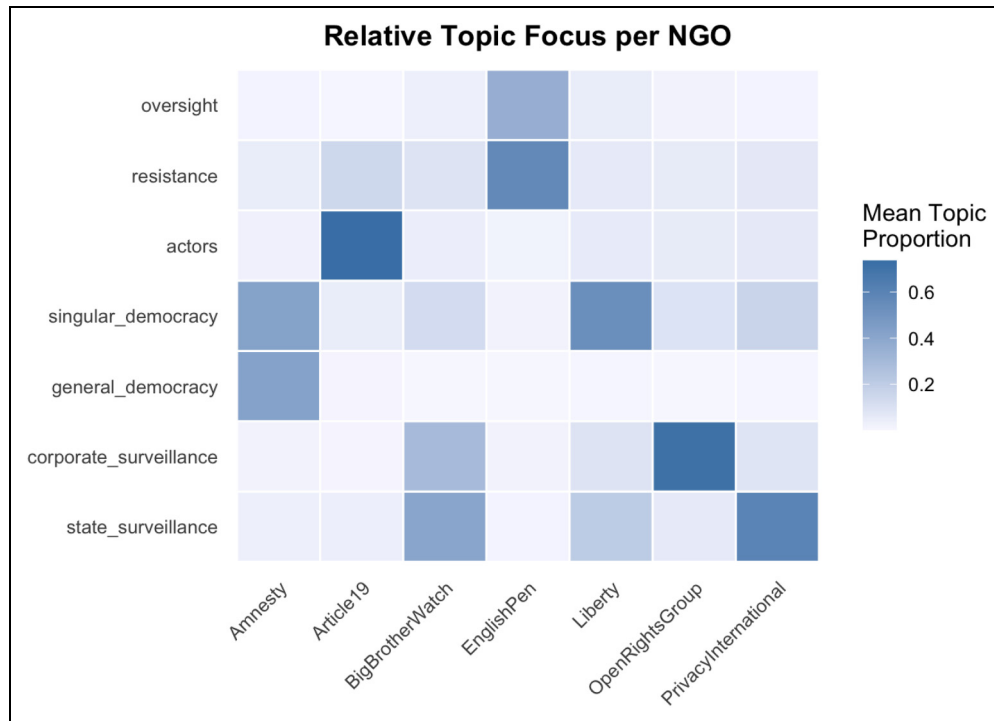


Figure 4. Relative topic focus per NGO.

of the websites. Here, connections are understood through common elements between the various websites, namely URLs. We extracted the URLs in the downloaded pages and reduced them to their domain name. Internal hyperlinks

(i.e. pointing to a page within the website) were not considered in the analysis. Then, URLs linked by at least two NGOs or at least six times by the same NGO were part of the 11 categories in Table 2.

Links under ‘infrastructure’ and N/A labels were not considered in the analysis. We used the remaining 135 domain names and their connections to create a weighted hyperlink network using Gephi (Bastian et al., 2009). The nodes in the network in Figure 6 represent the websites cited by the seven NGOs we analysed (shown as the larger white nodes). They are coloured according to their category and sized according to their degree (the number of connections they have). In order to reveal which domains are cited by whom, we arranged the nodes of the NGOs on a circle, fixed their positions and then let the other nodes take their places between them according to their connections. To do so, we used a force-directed layout in which connections work as springs or rubber bands, attracting the nodes that they connect (Jacomy et al., 2014). As a consequence, in Figure 6, domain names are situated between the NGOs that cite them (Venturini et al., 2021).

The network of hyperlinks shows that three of our actors are particularly well connected through the hyperlinks featured on their websites: Big Brother Watch, Privacy International and Open Rights Group. The connections between them and other NGOs are secured predominantly by social media platforms such as Twitter, Facebook, YouTube as well as by traditional media

Table 2. Types of extracted URLs in the retrieved NGOs’ pages.

NGO	Non-governmental organisations and related campaigns/projects and coalitions initiatives; funding charities; not-for-profit/social enterprises; policy research/advocacy-driven organisations
Media	News outlets, social media platforms and crowdfunding/petitions/events websites
Infrastructure	Content-delivery infrastructures such as advertisement engines, firewall applications, website development, URL shortening services, among others
Blog	Blogs and portfolios/brand-content websites
Government	Governmental bodies
National authority	National non-UK-based oversight bodies
Local authority	UK-based oversight bodies
International organisation	Intergovernmental organisations; international industry associations
University	Higher education entities as universities, institutes and departments, as well as academic publications
Corporation	Private companies
N/A	URL not accessible or redirecting to an unrelated website

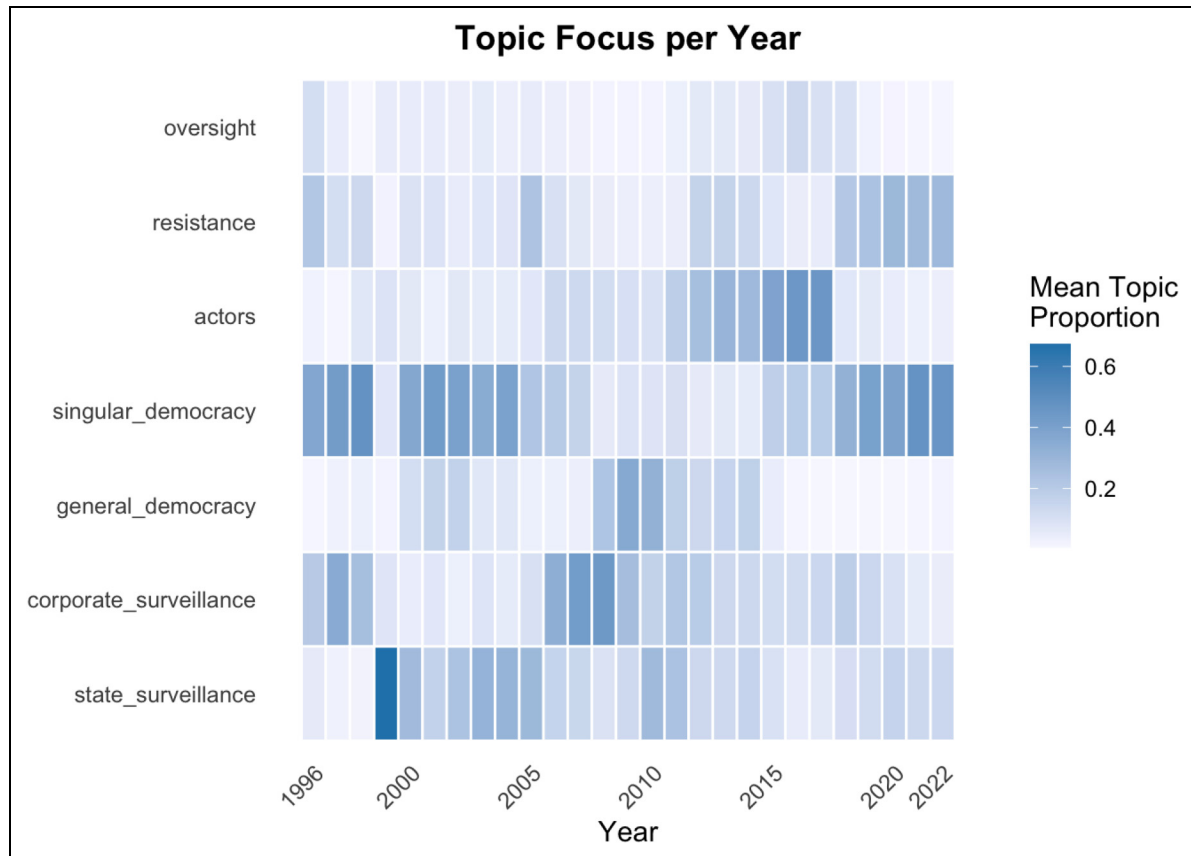


Figure 5. Topic focus per year.

such as the BBC and The Guardian. Overall, URLs under the labels ‘NGO’ and ‘media’ account for 56% of the extracted links, highlighting the dominance of such actors on the websites. By contrast, links to UK-based oversight bodies (7%), intergovernmental organisations (4%) and private companies (4%) are much less mentioned, suggesting a preference for the former as authoritative sources.

The Do Not Spy on Us coalition website appears in the centre-right, more closely connected to Open Rights Group and Big Brother Watch. From its positioning in the network, it can also be inferred that the website is closely connected to pages that also link to URLs of the UK and European parliaments and the European Commission, indicating how publications about the campaign have been intertwined with discussions of parliamentary debates. Amnesty, Article 19, English PEN and Liberty link to fewer external URLs, with many of them being campaign websites or different language versions of their own sites.

The coalition was, therefore, rather asymmetric. We also discovered many asymmetries in the topics we expected to find across the websites. In the final section, we show how these discordances – both machine-generated and produced

through our methodological entanglements – can shed new light on the contestations of mass surveillance in the United Kingdom.

Recomposing contestation

Our keywords were organised into binaries that aimed to render different forms of contestations and their vocabularies. We therefore performed one more methodological entanglement and returned to the archive. With traditional keyword searches, we aimed to understand the discordance between our keyword selection and the topic modelling about general versus singular democracy, state versus corporate surveillance and oversight versus resistance.

Undifferentiated democracy

The seeded topic modelling shows that there are differences between ‘singular’ and ‘general’ democracy, but that general democracy appears to a much lesser extent. Rather than being invoked for the purposes of counter-surveillance, the ‘democracy’ that speaks to the rights of the people is connected to particular countries and international campaigns, centred mainly around struggles in

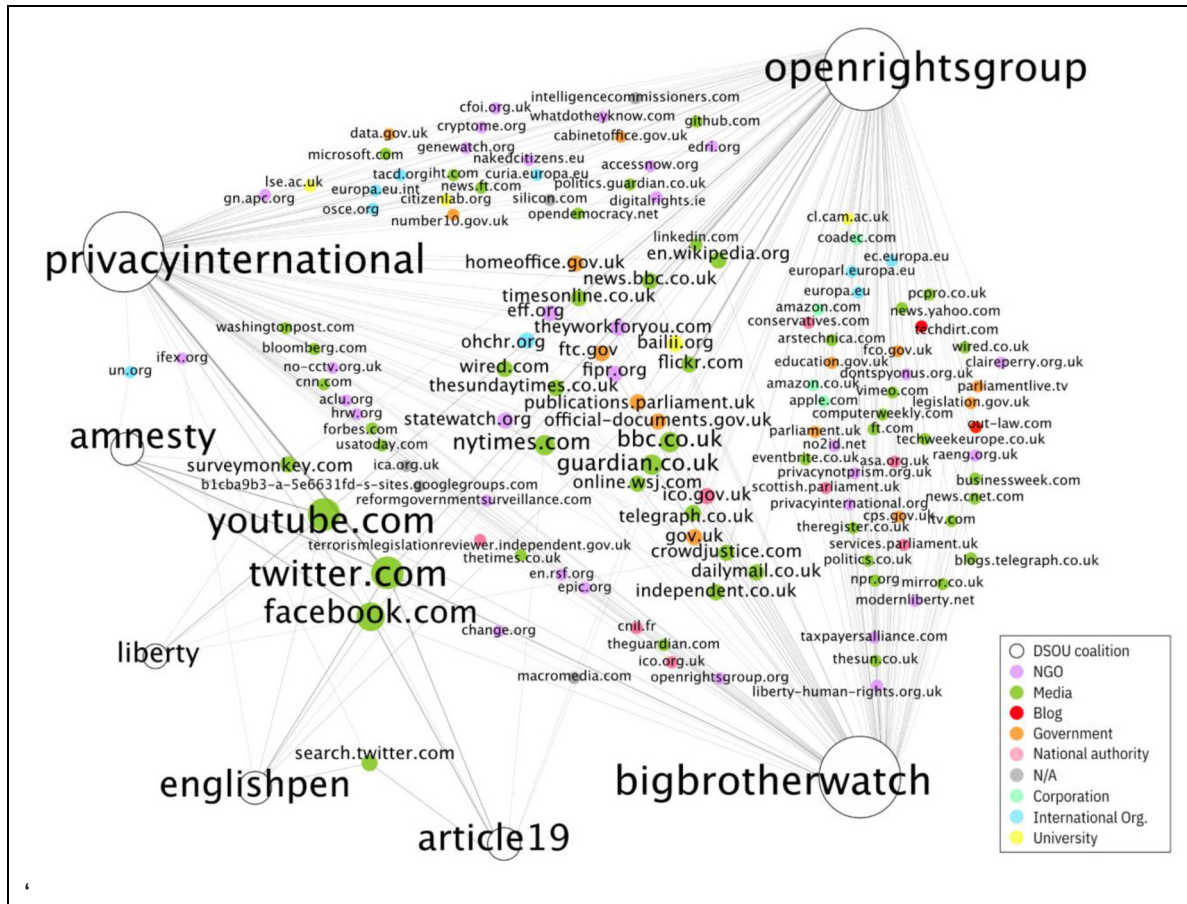


Figure 6. Network of hyperlinks in NGOs' webpages.

the Global South – such as ‘palestinian’, ‘congo’, ‘samoa’ and ‘korea’ as shown in Figure 6. A discordance emerges between what we have called general and singular democracy, but one which plays out geopolitically rather than in relation to surveillance. When democracy is connected with privacy, transparency, trust and rights, its addressee is the liberal subject in the Global North. This is the subject who is interpellated to ‘donate’, ‘read’, ‘contact’, ‘join’, ‘find’, ‘take’ and be ‘involved’ through watching webinars, etc.; all directive verbs which act to enrol citizens in actions of supporting NGOs.

The topic of general democracy has also been largely re-composed by the topic model. Besides ‘rights’, ‘discrimination’ and ‘democracy’, none of the words we had chosen in Table 1 featured in the final topic model in Figure 3. This absence of ‘general democracy’ was also present in the interviews. Interlocutors did not use the language of democracy or ‘human rights’ to speak about their practices, except when explicitly asked to do so. The distinction between general and singular democracy was ultimately not one relevant to mass surveillance, as general democracy was connected to actions in the Global South and singular

democracy to the Global North. Thus, we repeatedly see the topics of privacy, transparency and data protection.

Entangling the topic models with reading individual websites, it became clear that only Big Brother Watch, Privacy International and Open Rights Group oriented their discursive practices around countering surveillance and upholding digital rights – a pattern mirrored in the network analysis, which showed these NGOs sharing more domains and being tightly linked to the Don’t Spy on Us site. If we ‘zoom in’ to the beginning of June 2013, immediately after the Snowden disclosures, not much had changed on English PEN’s website, which continued to defend ‘the freedom to write and read in the UK and internationally’ (2013/06/01).⁸ Amnesty’s main page also did not change, and its coverage remained focused on human rights violations internationally, while its main campaigns in the immediate post-Snowden moment focused on Roma discrimination and producing a ‘panic button app’ to turn activists’ phones into an alert system’ (2013/09/11). It was only in 2015 that Amnesty’s main page focused on surveillance, with a petition to the Obama administration not to punish the NSA whistleblower (2015/06/02). At the same time, Article 19 led with two articles

on the surveillance state (2015/07/01). This focus was, however, short-lived, as both Amnesty's and Article 19's websites returned to headlining more traditionally focused campaigns around the world in 2016.

A focus on rights violations and violence abroad forms a continuous thread across the decades. From the late 1990s, when the first snapshots of Amnesty and Article 19 were archived by the Internet Archive, we see Amnesty focusing on international campaigns such as the arms trade in Burma (2005/01/05), on actions against repressive states around the world, from Zimbabwe and Russia to Iraq and Nigeria, as well as the USA's and UK's involvement in causing 'lawlessness' in Iraq (2003/04/20), and a global campaign against internet repression (2007/06/07). Article 19 was similarly focused 'outwards' during this time, highlighting campaigns in Nepal, Azerbaijan, Macedonia and Kosovo (2005/08/01). This was also the case for English PEN, though this organisation's webpages were available only from 2007 onwards.

The NGOs that had campaigned against surveillance pre-2013 continued to do so. In the 1990s, Privacy International defined itself as 'a human rights group formed in 1990 as a watchdog on surveillance by governments and corporations' (1999/01/25). Its websites throughout the 2000s focused on surveillance, data protection and ID cards. After the Snowden disclosures, Privacy International immediately reported on the NSA's PRISM surveillance programme and, in subsequent months, on the Five Eyes and the international trade in surveillance technology. For Big Brother Watch, which from its first archived website in 2010 had fashioned itself as countering the 'nanny state', we can see a shift towards arguments for surveillance transparency post-Snowden (2013/12/03). Open Rights Group, established in 2007 with digital rights and civil liberties at the core of its mission (2007/07/03), ran several campaigns centred on digital rights and democracy in the following years (2009/04/14). After Snowden, we see Open Rights Group start using the term 'mass surveillance' interchangeably with digital surveillance (2013/12/03) and campaign on its proportionality and necessity (2013/12/03). Liberty occupies an in-between space, campaigning around the 'Snoopers' Charter' as part of a broader effort against anti-terrorism legislation, which fits with its historical mandate and strategic orientations to defend 'the whole spirit of British freedom'.

Differential surveillance

In our topic modelling, both state and corporate surveillance are recurring topics. However, corporate surveillance is mainly associated with companies and linked to subjects encouraged to like, blog and post, whereas state surveillance is connected with 'security', 'police', 'agencies' and 'spy[ing]'. The only overlapping keywords are 'data', 'information', 'track[ing]' and 'UK'.

The limited overlap between corporate and state surveillance also became evident through a close reading of a selection of archived websites. State surveillance was represented as a problem of the 'surveillance state', often implicitly a liberal state like the UK or other EU member states. Mobilising against the surveillance state featured claims to privacy, data protection, transparency and lawfulness. These were the claims of what we have called 'singular democracy'. The repressive state remained 'elsewhere', spatially differentiated from 'home' and was the target of democratisation, while the surveillance state was approached through existing legal frameworks of privacy, data protection and transparency. This distinction also highlights the dichotomy between targeted and mass surveillance, as reproduced by NGOs in their campaigns. While targeted surveillance was framed as acceptable in a liberal society, mass surveillance was too general in its reach and therefore unacceptable. This framing does not allow for a critique of surveillance to include a critique of repression. Instead, this type of critique frames surveillance as sometimes necessary for 'the common good' and securing the United Kingdom against threats (see also Aradau and Mc Cluskey, 2022).

A close reading of the websites offers further nuance here. Privacy International's pages intertwine state and corporate surveillance. This is especially true from around 2013 (2013/12/01) and in 2017, when the organisation mapped companies producing surveillance and spyware internationally and called for regulation of biometrics exports. Big Brother Watch also mentioned the companies around the same time, but focused more on transparency. However, there isn't a joint campaign around state and corporate surveillance, highlighting the difficulty of undoing the framing of big tech as human rights actors (see Stevens, 2024) or as facilitating terrorist communication (Hannigan, cited in Quinn et al., 2014), as opposed to acting as part of the state surveillance assemblage. Against the collective action to counter state surveillance, discourses around corporate surveillance envisage different modes of engagement – for instance, through individual actions by a liberal subject who can use encryption and other more 'protective' technologies.

Absent oversight

Although oversight and various forms of regulation have been central to debates about digital surveillance and featured in the campaign's six claims, 'oversight' is rarely invoked on NGO websites. Among the seven NGOs, oversight is most often used by Privacy International. Oversight is about institutions mandated to monitor information systems or serve a data protection role, such as the European Data Protection Supervisor. At times, oversight and auditing are used interchangeably, as in references to Booz Allen Hamilton's role in providing oversight of the

SWIFT surveillance programme (2007/01/05). Open Rights Group highlighted the lack of oversight for new forms of digital data collection by private companies (2009/06/18). Big Brother Watch also used oversight to indicate its absence (2013/01/16). Formulations of legal, judicial or ‘proper’ oversight remain ambiguous, only gesturing to the role of law and legal authorities. Indeed, as our topic modelling shows, oversight is also limited in time. In the early 2000s, only Privacy International mentioned oversight. It was only in the 2010s that Big Brother Watch, Open Rights Group and Privacy International increasingly turned to a critique of the lack of surveillance oversight, intelligence oversight or biometrics oversight.

Oversight, however, does not become a shared category and remains ambiguous. The 2021 scandal around the NSO Group’s Pegasus spyware and global provision of surveillance technologies led PI and Amnesty International to ask for ‘greater oversight, accountability and remedy of corporate structures that have been reported to contribute to government surveillance of individuals, including human rights defenders’ (2021/06/03). ‘Oversight’ also appears on the website, citing company measures such as Facebook’s new ‘oversight board’ (2019/10/02).

Thus, our initial surprise at seeing books, translations, festivals and writers grouped under ‘oversight’ becomes less surprising. Oversight was recomposed computationally, generating discordance between our assumptions about oversight and patterns in the archive. The make-up of oversight we thought was important was ultimately absent, as oversight was recomposed through keywords related to ‘writing’ and ‘writers’. As Figure 5 suggests, oversight has little significance, so the model fills oversight with other words to turn it into a ‘topic’, which can be balanced against the other topics. Indeed, in our interviews, civil society actors did not speak about themselves as oversight actors unless explicitly invited to do so. This required an extension and reframing of oversight through struggle, as Kniep et al. (2024) have noted.

With oversight associated with particular institutions, submissions of written evidence and parliamentary inquiries, the term ‘resistance’ on the websites was, by contrast, more present and less fragmented. Resistance was connected to rendering practices visible, with a strong orientation towards investigative reporting. In this vein, we can conceive of contestations of mass surveillance more in the language of campaigning, making transparent and enabling the circulation of particular discourses of safeguards.

Conclusion

This article has proposed to make sense of the ‘atmosphere of failure’ that emerged in the wake of the Investigatory Powers Act in the United Kingdom among NGOs that had been active in countering it. As campaigns and mobilisations against mass surveillance came to a close, and

legislation was passed in the United Kingdom and elsewhere, scholars and activists tried to understand how we had got there. By entangling qualitative and digital methods, we could unfold a somewhat different story about mass surveillance, contestation and accountability. We did not see a sudden, widespread mobilisation of civil society against state surveillance in the immediate post-Snowden moment. Nor did we see a complete appropriation of the terms of debate by state intelligence and security agencies. Instead, we get a picture that is much more nuanced. The Snowden ‘moment’ was much less significant than many analyses assume, with several of the main NGOs involved in the broader legislative struggle around the Investigatory Powers Act barely engaging in public advocacy on their webpages, even immediately after the whistleblower’s disclosures.

Though the oral histories pointed to our civil society interlocutors being frustrated by being forced into the narrow framing of the terms of debate set by security actors, the web archives revealed that the prevalent vocabularies of democratic claims were privacy, transparency and data protection. ‘Oversight’ was a limited part of the websites’ discursive repertoire. Instead, a particular notion of resisting mass surveillance focused on investigating secret state actions and rendering them visible. We, the public of liberal citizens in the Global North, are invited to sign petitions, watch webinars or donate money, but are never fully enrolled in the struggles.

The recompositions of surveillance contestations were made possible through the methodological entanglement of qualitative and digital methods, including web archives, topic modelling and network visualisation. We proposed to approach multi-method inquiries with digital methods through the notion of a ‘discordant machine’. Methodological discordances led us to question the binaries through which we approached the archives of contestation: general/singular democracy, corporate/state surveillance and oversight/resistance. As these binaries became undone, other binaries emerged along geopolitical lines of the Global North/Global South. The surveillance state in the United Kingdom and Europe was distanced from the ‘repressive’ state, which was always over there, far away and unrelated to our lives in a liberal democracy. Similarly, the surveillance state is also somewhat distanced from our interactions with big tech companies, who are responsible for ‘tracking’, and appear in an ambiguous relation with state surveillance – as both deploying less accountable surveillance tech and not being as dangerous as the state, as they are allegedly more accountable to their customers.

While the methods we used within the framework of data sprint cannot offer an exhaustive or definite answer to the ‘atmosphere of failure’, they shed light on some of the limitations of contesting digital surveillance. The contestations did not lead to a shared vocabulary and critique of digital

surveillance. Furthermore, the network visualisation revealed little cross-referencing among coalition members' publications and limited links to oversight bodies, with platform capitalism infrastructures as the primary points of connection. Ultimately, our analysis suggests that the challenge for mobilisations against digital surveillance is not only to hold together its complexities but also to articulate its anti-democratic effects more explicitly.

Acknowledgements

We are grateful to the anonymous reviewers and the Big Data & Society editors for their comments and suggestions that helped clarify and sharpen our methodology and visualisations. Special thanks to Ibtehal Hussain for her contribution to the data sprint and for conversations during the GUARDINT project.

ORCID iDs

Claudia Aradau  <https://orcid.org/0000-0003-2010-9006>

Tobias Blanke  <https://orcid.org/0000-0003-0956-987X>

Thais Lobo  <https://orcid.org/0009-0009-5794-8672>

Ethical considerations and informed consent statements

The project received ethics approval at King's College London [LRS-18/19-11787]. All participants consented to being interviewed.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Economic and Social Research Council through the Open Research Area grant GUARDINT 'Intelligence and Oversight Networks. Who guards the guardians?' [grant number ES/S015132/1]. Tobias Blanke's contribution was supported by the H2020 European Research Council-funded project Deep Culture – Living with Difference in the Age of Deep Learning (grant number 101141330).

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data availability statement

The datasets from the study are not publicly available due to privacy restrictions but are available from the first author on reasonable request.

Notes

1. We use digital and mass surveillance interchangeably as mass surveillance was enabled by the spread of digital devices and big data.
2. To protect the anonymity of interviewees, any identifying information has been concealed to reduce the possibility of

deductive disclosure. Interviewees are simply listed with alpha-betised identifiers.

3. Several civil society interlocutors identified one not insignificant 'win' as the judicial 'double lock' mechanism, whereby approval by a judicial commissioner within the Investigatory Powers Commission Office, in addition to the relevant authorities, is required before the most intrusive surveillance powers can be used.
4. Tommaso Venturini defines digital methods as qualitative given their interest 'not only in exploiting the data generated by digital technology but also in understanding how these technologies make collective existence measurable and computable' (2024).
5. See <https://help.archive.org/help/wayback-machine-general-information/>.
6. On the the jaccard_similarity comparison function, see <https://github.com/ropensci/textreue>.
7. More details can be found at <https://github.com/first20hours/google-10000-english/blob/master/google-10000-english.txt>.
8. We reference the websites by the date (YYYY/MM/DD) on which they were scraped by the WayBack Machine.

References

- Andrejevic M (2019) Automating surveillance. *Surveillance & Society* 17(1/2): 7–13.
- Aradau C (2017) Assembling (non)knowledge: Security, law, and surveillance in a digital world. *International Political Sociology* 11(4): 327–342.
- Aradau C and Blanke T (2015) The (Big) Data-security assemblage: Knowledge and critique. *Big Data & Society* 2(2). <https://doi.org/10.1177/2053951715609066>.
- Aradau C and Mc Cluskey E (2022) Making digital surveillance unacceptable? Security, democracy and the political sociology of disputes. *International Political Sociology* 16(1): 1–19.
- Bastian M, Heymann S and Jacomy M (2009) Gephi: An open source software for exploring and manipulating networks. *Proceedings of the International AAAI Conference on Web and Social Media* 3(1): 361–362.
- Bauman Z, Bigo D, Esteves P, et al. (2014) After Snowden: Rethinking the impact of surveillance. *International Political Sociology* 8(2): 121–144.
- Bennett CJ (2010) *The Privacy Advocates: Resisting the Spread of Surveillance*. The MIT Press.
- Big Brother Watch and Others v The UK (2018) Applications nos. 58170/13, 62322/14 and 24960/15). Available at: <https://hudoc.echr.coe.int/> (accessed 3 November).
- Big Brother Watch and Others v the UK (2021) Applications Nos. 58170/13, 62322/14 and 24960/15). Grand Chamber Judgement. European Court of Human Rights. Available at: <https://www.bailii.org/eu/cases/ECHR/2021/439.html> (accessed 25 June 2022).
- Bigo D, Mc Cluskey E and Tréguer F (2024) *Intelligence Oversight in Times of Transnational Impunity: Who Will Watch the Watchers?* Taylor & Francis.

- Blanke T and Wilson J (2017) Identifying epochs in text archives. In: *2017 IEEE International Conference on Big Data (Big Data)*, pp.2219–2224.
- Brown I, Halperin MH, Hayes B, et al. (2015) Towards multilateral standards for surveillance reform. In: Miller RA (ed) *Privacy and Power: A Transatlantic Dialogue in the Shadow of the NSA-Affair*. Cambridge University Press, 461–491.
- Brown I and Korff D (2009) Terrorism and the proportionality of internet surveillance. *European Journal of Criminology* 6(2): 119–134.
- Chang J, Gerrish S, Wang C, et al. (2009) Reading tea leaves: How humans interpret topic models. *Advances in Neural Information Processing Systems* 22: 288–296.
- Danezis G (2016) IP Draft Bill. Available at: <https://conspicuouschatter.wordpress.com/2015/11/05/uk-draft-ip-bill-the-last-policy-discussion-about-surveillance-before-the-mass-gagging/> (accessed 5 June 2022).
- Dencik L and Cable J (2017) The advent of surveillance realism: Public opinion and activist responses to the snowden leaks. *International Journal of Communication* 11: 763–781.
- Dencik L, Hintz A and Cable J (2016) Towards data justice. *Big Data & Society* 3(2): 1–12.
- Digital Methods Initiative (2007) Internet Archive Wayback Machine Link Ripper. Available at: <https://tools.digitalmethods.net/beta/InternetArchiveWaybackMachineLinkRipper/> (retrieved 17 March 2025).
- Don't Spy on Us (2016) Don't spy on us: Our campaign. Available at: <https://www.dontspyonus.org.uk/our-campaign> (accessed 25 May 2026).
- Don't Spy on Us (2017) DSOU coalition has ended. Available at: <https://archive-2017.dontspyonus.org.uk/blog/2017/06/15/dsou-coalition-to-end/index.html> (accessed 25 May 2026).
- Eley L and Rampton B (2020) Everyday surveillance, Goffman and unfocused interaction. *Surveillance and Society* 18(2): 199–215.
- Fernandez LA and Huey L (2009) Is resistance futile? Thoughts on resisting surveillance. *Surveillance & Society* 6(3): 199–202.
- Gandy OH Jr (2021) *The Panoptic Sort: A Political Economy of Personal Information*. Oxford: Oxford University Press.
- Gilliom J and Monahan T (2012) Everyday resistance. In: *Routledge Handbook of Surveillance Studies*. Routledge, pp.405–411.
- Gros V, De Goede M and İşleyen B (2017) The snowden files made public: A material politics of contesting surveillance. *International Political Sociology* 11(1): 73–89.
- Guild E (2019) Data rights: Claiming privacy rights through international institutions. In: Bigo D, Isin E and Ruppert E (eds) *Data Politics*. Routledge, 267–284.
- Guldi J (2019) Parliament's debates about infrastructure: An exercise in using dynamic topic models to synthesize historical change. *Technology and Culture* 60(1): 1–33.
- Hariharan JC (2023) *The Protection of Physical Privacy under English Law*. PhD Thesis. University College London, UK.
- Hayes B, Rowlands M and Buxton N (2009) *NeoConOpticon: the EU Security-Industrial Complex*. Transnational Institute.
- Hintz A and Brown I (2017) Enabling digital citizenship? The reshaping of surveillance policy after Snowden. *International Journal of Communication* 11: 782–801.
- Hogue S (2024) Individual vulnerability and collective resistance under surveillance: Claiming the right to existence against discriminatory suspicion. *International Political Sociology* 18(3). <https://doi.org/10.1093/ips/olae023>.
- Huysmans J (2014) *Security Unbound: Enacting Democratic Limits*. Routledge.
- Jacomy M, Venturini T, Heymann S, et al. (2014) Forceatlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PLoS ONE* 9(6): e98679.
- Jagarlamudi J, Daumé H III, et al. (2012) Incorporating lexical priors into topic models. In: *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, pp.204–213.
- Jelodar H, Wang Y, Yuan C, et al. (2019) Latent Dirichlet allocation (LDA) and topic modeling: Models, applications, a survey. *Multimedia Tools and Applications* 78: 15169–15211.
- Jones C (2017) *Market Forces: The Development of the EU Security-Industrial Complex*. TNI and Statewatch.
- Kelly N (2022) Facing surveillance: Personified surveillance, algorithmic injustice, and the myth of big brother in post-snowden popular culture. *Surveillance & Society* 20(2): 172–185.
- Kniep R, Ewert L, Reyes BL, et al. (2024) Towards democratic intelligence oversight: Limits, practices, struggles. *Review of International Studies* 50(1): 209–229.
- Kuldova TØ (2022) *Compliance-Industrial Complex: The Operating System of a Pre-Crime Society*. Springer Nature.
- Larsson S (2020) The civil paradox: Swedish arms production and export and the role of emerging security technologies. *International Journal of Migration and Border Studies* 6(1–2): 26–51.
- Lyon D (2001) *Surveillance Society: Monitoring Everyday Life*. Open University Press.
- Lyon D (2014) Surveillance, Snowden, and big data: Capacities, consequences, critique. *Big Data & Society* 1(2): 2053951714541861.
- Lyon D (2015) *Surveillance After Snowden*. Polity.
- Martin-Mazé M (2020) The power elite of security research in Europe: From competitiveness and external stability to data-veillance and societal security. *International Journal of Migration and Border Studies* 6(1–2): 52–73.
- Martin-Mazé M and Perret S (2021) Designs of borders: Security, critique, and the machines. *European Journal of International Security* 6(3): 278–300.
- Marx GT (2014) Security and surveillance contests: Resistance and counter-resistance. In: Balzacq T (ed) *Contesting Security: Strategies and Logics*. Routledge, 15–28.
- Mc Cluskey E and Aradau C (2023) From abuse to trust and back again: Intelligence scandals and the quest for oversight. In: Bigo D, Mc Cluskey E and Treguer F (eds) *Intelligence*

- Oversight in Times of Transnational Impunity: Who Will Watch the Watchers?* Routledge, 130–151.
- Monahan T (2006) Counter-surveillance as political intervention? *Social Semiotics* 16(4): 515–534.
- Mullen L and Li Y (2026) Textreuse: Detect text reuse and document similarity. Available at: <https://docs.ropensci.org/textreuse/> (accessed 4 May).
- Munk AK, Olesen AG and Jacomy M (2022) The Thick Machine: Anthropological AI between explanation and explication. *Big Data & Society* 9(1): 20539517211069891.
- Ochoa CS, Gadinger F and Yildiz T (2021) Surveillance under dispute: Conceptualising narrative legitimisation politics. *European Journal of International Security* 6(2): 210–232.
- Passerini L (2002) Work ideology and consensus under Italian fascism. In: *The Oral History Reader*. Routledge, 67–76.
- Passerini L (2017) *Memory and Utopia: The Primacy of Inter-Subjectivity*. Routledge.
- Quinn B, Ball J and Rushe D (2014) GCHQ chief accuses US tech giants of becoming terrorists' 'networks of choice'. *The Guardian*. Available at: <https://www.theguardian.com/uk-news/2014/nov/03/privacy-gchq-spying-robert-hannigan> (accessed 04 May 2023).
- Rabinow P (2002) Midst anthropology's problems. *Cultural Anthropology* 17(2): 135–149.
- Rogers R (2013) *Digital Methods*. The MIT Press.
- Rogers R (2019) *Doing Digital Methods*. Sage.
- Ruppert E, Harvey P, Lury C, et al. (2015) *Socialising Big Data: From Concept to Practice*. CRESC, The University of Manchester and the Open University.
- Stevens A (2024) *Resisting State Surveillance in the Digital Age: Precarious Coalitions, Contested Knowledge, and Diverse Opposition to Mass-Surveillance in the UK*. Taylor & Francis.
- Stevens A and Allen-Robertson J (2021) Encrypting human rights: The intertwining of resistant voices in the UK state surveillance debate. *Big Data & Society* 8(1): 2053951720985304.
- Thompson P (2000) *Voice of the Past: Oral History*. 3rd ed. Oxford University Press.
- Thompson P (2002) The voice of the past: Oral history. In *The Oral History Reader*. Routledge, 35–42.
- Tréguer F (2017) Intelligence reform and the Snowden paradox: The case of France. *Media and Communication* 5(1): 17–28.
- Tréguer F (2019) Seeing like big tech: Security assemblages, technology, and the future of state bureaucracy. In: Bigo D, Isin E and Ruppert E (eds) *Data Politics*. Routledge, 145–164.
- Tréguer F (2023) From radical contention to deference. Intelligence Oversight in Times of Transnational Impunity. In: Bigo D, Mc Cluskey E and Tréguer F (eds) *Intelligence Oversight in Times of Transnational Impunity: Who Will Watch the Watchers?* Routledge, 15–69.
- Venturini T (2024) Quali-quantitative (or non-binary) methods. In: Felt U and Irwin A (eds) *Elgar Encyclopedia of Science and Technology Studies*. Edward Elgar Publishing, 176–185.
- Venturini T, Bounegru L, Gray J, et al. (2018) A reality check (list) for digital methods. *New media & Society* 20(11): 4195–4217.
- Venturini T, Jacomy M and Jensen P (2021) What do we see when we look at networks: Visual network analysis, relational ambiguity, and force-directed layouts. *Big Data & Society* 8(1): 20539517211018488.
- Venturini T, Munk A, Meunier A, et al. (2018) Data-sprinting: A public approach to digital research. In: Lury C (ed) *Routledge Handbook of Interdisciplinary Research Methods*. Routledge, 158–163.
- Viola LA and Laidler P (2022) *Trust and Transparency in an Age of Surveillance*. Routledge.