APC ISSUE PAPERS

BIG DATA AND SEXUAL SURVEILLANCE

By Nicole Shephard

KEYWORDS: surveillance, gender, sexuality, race, big data, LGBTIQ, women human rights defenders, privacy, anonymity, consent

SUMMARY

Surveillance has historically functioned as an oppressive tool to control women's bodies and is closely related to colonial modes of managing populations. Big data, metadata and the technologies used to collect, store and analyse them are by no means neutral, but come with their own exclusions and biases. This paper highlights the gendered and racialised effects of data practices; outlines the overlapping nature of state, commercial and peer surveillance; and maps the challenges and opportunities women and queers encounter on the nexus between data, surveillance, gender and sexuality. Vulnerable communities as well as sexual rights activists are at heightened risk of data-driven modes of surveillance. In addition to exposing and addressing algorithmic discriminations, feminist data practices oppose the non-consensual collection of data, amplify participatory data projects that empower women and sexual minorities, and protect the data, privacy and anonymity of activists and the communities they work with.

MAIN CONCEPTS

Big data: Vast datasets containing social media data, machine data or transactional data that can be analysed computationally to reveal patterns, trends and predictions about human behaviour and social life.

Dr Nicole Shephard is an independent researcher and writer working on the intersections between gender, sexuality and technology. She holds a PhD in Gender (LSE) and an MSc in International Development (University of Bristol).

APC
ASSOCIATION FOR
PROGRESSIVE
COMMUNICATIONS

Metadata: Information describing other data; in the context of email, telephone calls or instant messaging, metadata includes the sender, receiver, devices, locations, service providers, IP addresses, time, length or size of a message.

Dataveillance: Combines data and surveillance to describe systematic data-based surveillance practices that involve sorting and aggregating large quantities of data to monitor, track and regulate people and populations.

Assemblage: Rather than individual technologies or datasets, this theoretical concept emphasises the intersecting nature of institutions and processes at work in abstracting and reassembling bodies through data, and the fluid socio-technical, economic and political contexts that data and surveillance are embedded in.

KEY FACTS

- Whistleblowers including Edward Snowden,[1] Thomas Drake, Bill Binney[2] and Russell Tice[3] disclosed that contemporary data-driven surveillance practices carried out by the United States National Security Agency (NSA), the United Kingdom Government Communications Headquarters (GCHQ) and their allies globally are ubiquitous and pervasive.
- With state surveillance at the forefront of the public debates it is important to not lose sight of a wider range of sexual surveillance practices historically functioning as a "tool of patriarchy, used to control and restrict women's bodies, speech, and activism."[4]
- The recognition that the data and metadata at stake in surveillance are never neutral[5] is central in relation to gender, sexuality and race.
- Women, people with disabilities, refugees, sexual minorities, people receiving state benefits, or incarcerated populations, among others, can testify to myriad ways in which their privacy has habitually been invaded by surveillance practices in the private as well as in the public sphere.[6]
- Big data is generated in many places – social media, global positioning system (GPS) data, radio frequency identification (RFID) data, health data, financial data or phone records. It extends to sexual and reproductive health, as evidenced by recent research on the spread of menstruation, pregnancy and fertility apps.[7] Experts project 40% annual growth in data generated globally.[8,9]
- Big data has found its way into gender and development discourse, as evidenced by ongoing initiatives that seek to harness data for development[10] and strive to close the gender gap by closing the data gap.[11]
- Large-scale social media data is mined for operative and marketing purposes by corporations.[12] Consent is blurred when social media data is then collected by governments and aggregated with other available databases.
- Women's rights and sexual rights activists do not always hold sufficient technological privilege to mitigate against the risks inherent in their work and to minimise potential adverse effects of surveillance (exposure, harassment, violence) for themselves and vulnerable groups they work with.
- Big data and dataveillance run the risk of algorithmically reproducing past discriminations,[13] creating new exclusions and digital discriminations,[14] and exacerbating epistemic violence that marginalised communities are subject to.[15]

[1] https://edwardsnowden.com/revelations

[2] Radack, J., Drake, T. & Binney, W. (2012). Enemies of the State: What Happens When Telling the Truth about Secret US Government Power Becomes a Crime. Presentation at Chaos Communication Congress. https://media.ccc.de/v/29c3-5338-en-enemies_of_the_state_h264

[3] Democracy Now. (2006, 3 January). National Security Agency Whistleblower Warns Domestic Spying Program Is Sign the U.S. is Decaying Into a "Police State". www.democracynow.org/2006/1/3/exclusive_national_security_agency_whistleblower_warns

[4] Association for Progressive Communications (APC). (2016, August). *Feminist Principles of the Internet 2.0.* www.apc.org/en/pubs/feminist-principles-internet-version-20

[5] Kitchin, R., & Lauriault, T. P. (2014). Towards critical data studies: Charting and unpacking data assemblages and their work. The Programmable City Working Paper 2 SSRN. papers.ssrn.com/sol3/papers.cfm?abstract_id=2474112

[6] Dubrofsky, R. E., & Magnet, S. A. (2015). *Feminist surveillance studies*. Durham: Duke University Press.

[7] Rizk, V., & Othman, D. (2016). Quantifying Fertility and Reproduction through Mobile Apps: A Critical Overview. *Arrow for Change, 22*(1). www.arrow.org.my/wp-content/uploads/2016/08/AFC22.1-2016.pdf

[8] Manyika, J., Chiu, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). *Big data: The next frontier for innovation, competition, and productivity*. www.mckinsey.com/business-functions/business-technology/our-insights/big-data-the-next-frontier-for-innovation

[9] https://e27.co/worlds-data-volume-to-grow-40-per-year-50-times-by-2020-aureus-20150115-2/

[10] UN Global Pulse. (2012). *Big Data for Development: Challenges & Opportunities*. New York: United Nations.

[11] Plan International. (2016). *Counting the Invisible: Using Data to Transform the lives of Girls and Women by 2030*. Woking: Plan International.

[12] Kitchin, R. (2014). *The Data Revolution: Big Data, Open Data, Data Infrastructures & their Consequences*. London: Sage.

[13] Conrad, K. (2009). Surveillance, Gender, and the Virtual Body in the Information Age. *Surveillance & Society*, *6*(4), 380–387.

[14] Lyon, D. (2003). Surveillance as social sorting: computer codes and mobile bodies. In D. Lyon (Ed.), *Surveillance as Social Sorting: Privacy, risk, and digital discrimination* (pp. 13–30). London: Routledge.

[15] Gurumurthy, A., & Chami, N. (2016, 31 May). Data: The New Four-Letter Word for Feminism. *GenderIT.org*. www.genderit.org/articles/data-new-four-letter-word-feminism

INTRODUCTION

This issue paper aims at mapping sexual surveillance by exploring the key issues at stake on the nexus between surveillance, gender and sexuality. Taking the recognition that **surveillance has historically functioned as a "tool of patriarchy, used to control and restrict women's bodies, speech, and activism"**[16] as a starting point, it explores the implications of data-driven modes of surveillance for women and queers. It does so from an intersectional[17] perspective, as gender and sexuality never take place in isolation but interact with race, religion, class, geo-political location, health or bodily diversity.

Big data is generated in many places – social media, global positioning system (GPS) data, radio frequency identification (RFID) data, the internet of things, health data, financial data or phone records are just a few examples of data sources. Some of these data are knowingly created, for example by updating a status, posting an image or writing a tweet. Others are the side product of using services and their features, for example swiping a card or using/carrying a phone.

> *So-called "raw data" is in fact never quite raw. The recognition that the data and metadata at stake in surveillance are never neutral is central in relation to gender, sexuality and race.*

Big data evangelists have boldly declared "the end of theory"[18] as vast quantities of numbers are taken to speak for themselves. Critics have convincingly demonstrated that data, no matter how big, are only ever a representation and a sample,[19] never a population, and that so-called "raw data" is in fact never quite raw.[20] All methods of collecting, organising and analysing data are always cooked up by some***body***. Similarly, "metadata" such as the sender, receiver, devices, locations, time and length of a communication (but not its content), often constructed as unproblematic in terms of the right to privacy, can provide "insight into an individual's behaviour, social relationships, private preferences and identity that go beyond even that conveyed by accessing the content of a private communication."[21] The recognition that **the data and metadata at stake in surveillance are never neutral is central in relation to gender, sexuality, and race**.

This paper aligns itself with a growing body of scholarship taking a feminist approach to surveillance studies, as well as with emerging work that draws attention to the colonial continuities and racial implications of data and surveillance. It explores the relationship between big data and sexual surveillance, as well as the challenges and opportunities that arise for women, queers and their advocates where data and surveillance meet gender and sexuality. Empowering uses of data such as community mapping projects take place alongside efforts to push back against, subvert and resist illegitimate surveillance and its adverse effects.

Ultimately, the paper argues for feminist data practices that are attentive to agency and consent of those involved. Such practices:

- Oppose and resist the non-consensual collection of data.
- Amplify data projects that empower women and sexual minorities.
- Take adequate care of protecting the data, privacy, and anonymity of activists and the communities they engage with.
- Work to expose and level algorithmic discriminations.

[16] Association for Progressive Communications (APC). (2016). *Feminist Principles of the Internet 2.0.* www.apc.org/en/pubs/feminist-principles-internet-version-20

[17] Crenshaw, K. (1989). Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics. *University of Chicago Legal Forum, 1989*, 139–167.

[18] Anderson, C. (2008, 23 June). The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. *Wired*. www.wired.com/2008/06/pb-theory/

[19] Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, *1*(1), 1–12.

[20] Gitelman, L. (2013). *"Raw Data" is an Oxymoron*. (L. Gitelman, Ed.). Cambridge, MA: MIT Press.

[21] United Nations, OHCHR (2014). *The Right to Privacy in the Digital Age*. undocs.org/A/HRC/27/37

BIG DATA <-> SURVEILLANCE

Big data have been defined in many ways. Kitchin[22] characterises them as high in volume and velocity, diverse in variety, exhaustive in scope, fine-grained in resolution, relational, and flexible in scaleability. Recent years have witnessed an exponential growth in data volume, and **experts project 40% annual growth in data generated globally**.[23]

Yet, hyperbolic statements like declaring the death of social scientific analysis[24] because now "n=all"[25] form part of big data myths[26] that have close ties with the spectre of objectivity so well-known to feminist critics of knowledge production. Here it appears in the form that bigger data are considered somehow more true, accurate and objective than other modes of producing knowledge.

Using the example of Twitter data, boyd and Crawford[27] explain succinctly why n, no matter how large, never equals all:

> Twitter does not represent "all people", and it is an error to assume "people" and "Twitter users" are synonymous: they are a very particular sub-set. Neither is the population using Twitter representative of the global population. Nor can we assume that accounts and users are equivalent. Some users have multiple accounts, while some accounts are used by multiple people. Some people never establish an account, and simply access Twitter via the web. Some accounts are "bots" that produce automated content without directly involving a person.

Twitter data illustrates the partiality of n well, but it is worth noting that every trove of data comes with its own exclusions.

Just as big data cannot represent entire populations, aggregating "all" data about a single person never fully represents but partially profiles particular individuals. In addition to never being comprehensive, data are also never neutral. Data are "situated, contingent, relational, and framed, and used contextually to try and achieve certain aims and goals."[28] On Twitter individuals purposefully create (mostly) public data. In other instances, they actively consent to the collection of their data for specific purposes such as advertising or research, and in yet others they are unaware of becoming data points.

> *Consent is blurred when social media data is collected in bulk by governments and aggregated with other available databases.*

Consent is blurred when social media data, however willingly created and shared, is collected in bulk by governments and aggregated with other available databases. Neither are the databases and repositories that hold the data neutral. They are "complex socio-technical systems that are embedded within a larger institutional landscape"[29] that includes research institutions, corporations, as well as government agencies concerned with security, citizenship or public health, and is imbued with power relations.

Scholars have observed quantitative and qualitative shifts in surveillance practices, as **the so-called "war on terror" goes hand in hand with the increased availability of big data**. Lyon[30] notes how data are not only "captured differently, they are also processed, combined, and analysed in new ways. Social media that appeared on the scene at roughly the same time as responses to 9/ 11 boosted the 'surveillance state', are now the source of much data, used not only for commercial but also for 'security' purposes."

The US Customs and Border Protection Agency, for instance, currently proposes collecting information on visitors' social media accounts upon entry to the

[22] Kitchin, R. (2014). Op. cit.

[23] Manyika, J., Chiu, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). *Big data: The next frontier for innovation, competition, and productivity*. www.mckinsey.com/business-functions/business-technology/our-insights/big-data-the-next-frontier-for-innovation

[24] Anderson, C. (2008). Op. cit.

[25] Mayer-Schönberger, V., & Cukier, K. (2013). Op. cit.

[26] boyd, d., & Crawford, K. (2012). Critical Questions for Big Data. *Information, Communication & Society*, *15*(5), 662–679.

[27] Ibid.

[28] Kitchin, R., & Lauriault, T. P. (2014). Towards critical data studies: Charting and unpacking data assemblages and their work. The Programmable City Working Paper 2 SSRN. papers.ssrn.com/sol3/papers.cfm?abstract_id=2474112

[29] Kitchin, R., & Lauriault, T. P. (2014). Op. cit.

[30] Lyon, D. (2014). Surveillance, Snowden, and Big Data: Capacities, consequences, critique. *Big Data & Society*, 1(2), 1–13.

country. The measure aims at providing "greater clarity and visibility to possible nefarious activity and connections" in the assessment of "potential risks to national security and the determination of admissibility."[31]

The concept of "dataveillance" expresses these shifts and describes practices of sorting and aggregating vast datasets to track and regulate populations: "Dataveillance in the present moment is not simply descriptive (monitoring) but also predictive (conjecture) and prescriptive (enactment)."[32]

It arguably differs from targeted surveillance: "Whereas surveillance presumes monitoring for specific purposes, dataveillance entails the continuous tracking of (meta) data for unstated preset purposes."[33]

Targeted surveillance thus requires a suspect to monitor for a purpose, while dataveillance generalises suspicion and algorithmically produces suspects, thus turning the assumption of innocence until proven guilty on its head. Marx[34] details this shift to new modes of surveillance along a staggering 28 dimensions. Most pertinently to a discussion of sexual surveillance are the following:

- New surveillance often lacks **consent**, with higher proportions of involuntary production/collection of data.
- The **location of data and its collectors/analysts** is often remote and less visible.
- After collection the data is **stored remotely and migrated** often.
- The temporality of new surveillance is continuous, omnipresent, and **covers past, present and future occurrences of data**.
- It is **acontextual**.
- **Whole populations** rather than individuals are surveilled.

An understanding that data never emerge in isolation, are always contingent on context, technologies, humans and their algorithms that collect, sort, and analyse them, as well as on the power relations that all of the above are imbued with, is the basis for conceptualising data and its entanglements with surveillance as assemblage.[35]

Haggerty and Ericson[36] theorise a **surveillance assemblage** where information, technology and the human body interact to create a somewhat deterritorialised "data double" to be tracked, commodified, managed and controlled. Kitchin describes a **data assemblage** made up of systems of thought, forms of knowledge, finance, political economy, governmental and legal frameworks, infrastructure, materiality, institutions, places, subjectivities, communities, and markets that all intersect with one another.[37] Such a conceptual approach leaves room for highly localised data practices as well as attention to the ways in which **big data extend to "the global**, through inter-regional and worldwide data sets, data sharing arrangements and infrastructures, and the formulation of protocols, standards and legal frameworks."[38]

Aradau and Blanke see modes of knowledge production, technological devices, institutions, and methods as part of a **data-security assemblage** with stakes in civil liberties, human rights, privacy, and data protection.[39]

> *Thinking of big data and (sexual) surveillance in terms of an assemblage highlights how much more than mere data is at stake, and that the data cannot be thought independently of its wide and rather fluid context.*

[31] Dept. of Homeland Security, U.S. Customs and Border Protection. (2016). Agency Information Collection Activities: Arrival and Departure Record (Forms I-94 and I-94W) and Electronic System for Travel Authorization. www.federalregister.gov/documents/2016/06/23/2016-14848/agency-information-collection-activities-arrival-and-departure-record-forms-i-94-and-i-94w-and#h-11

[32] Raley, R. (2013). Dataveillance and Countervailance. In L. Gitelman (Ed.), *Raw Data is an Oxymoron*. Cambridge, MA: The MIT Press

[33] van Dijck, J. (2014). Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. *Surveillance & Society*, *12*(2), 197–208.

[34] Marx, G. T. (2002). What's new about the "new surveillance"?: Classifying for change and continuity. *Surveillance & Society*, *1*(1), 9–29.

[35] The notion of assemblages is based on the work of philosophers Gilles Deleuze and Félix Guattari, who in *A Thousand Plateaus* (1987, p. 102-103) define it along two axes: "On a first, horizontal, axis, an assemblage comprises two segments, one of content and one of expression. On the one hand it is a machinic assemblage of bodies, of actions and passions, an intermingling of bodies reacting to one another; on the other hand it is a collective assemblage of enunciation, of acts and statements, of incorporeal transformations attributed to bodies. Then on a vertical axis, the assemblage has both territorial sides, or reterritorialized sides, which stabilize it, and cutting edges of deterritorialization, which carry it away."

[36] Haggerty, K. D., & Ericson, R. V. (2000). The surveillant assemblage. *The British Journal of Sociology*, *51*(4), 605–622.

[37] Kitchin, R. (2014). *The Data Revolution: Big Data, Open Data, Data Infrastructures & their Consequences*. London: Sage.

[38] Kitchin, R., & Lauriault, T. P. (2014). Op. cit.

[39] Aradau, C., & Blanke, T. (2015). The (Big) Data-security assemblage: Knowledge and critique. *Big Data & Society*, *2*(2), 1–12.

In a nutshell, thinking of big data and (sexual) surveillance in terms of an assemblage highlights how much more than mere data is at stake, and that the data cannot be thought independently of its wide and rather fluid context. The relationship between computers and humans, as well as discourse around security and big data that justifies ever more surveillance form an integral part of this assemblage,[40] as do everyday practices that produce digital traces and practices of resistance against dataveillance, and efforts to use big data for positive change.

SURVEILLANCE <-> SEXUAL SURVEILLANCE

Gender and sexualities have only recently found their way into the study and discourse of surveillance. This section, however, maps ways in which women and sexual minorities have long been highly visible to technologies of surveillance. The racialised, sexualised and gendered outcomes of sexual surveillance form part of what Lyon describes as mechanisms of social sorting: "Surveillance today sorts people into categories, assigning worth or risk, in ways that have real effects on their life-chances. Deep discrimination occurs, thus making surveillance not merely a matter of personal privacy but of social justice."[41]

Data have historically been used to categorise and manage populations. Big data are but the latest trend in a long tradition of quantification with roots in modernity's fetishisation of taxonomy in the service of its institutional order. Big data and machine learning may take these practices to unprecedented levels, but it is important to not lose sight of the historical continuities the data/surveillance assemblage is embedded in, particularly where gender, race and sexuality are concerned.

GENDER, SEXUALITY AND SURVEILLANCE

Browne describes **surveillance as "a technology of social control**" that produces racial norms and has the power to **"define what is in or out of place."** She argues that this racialised ordering is fluid and contextual but "most often upholds negating strategies that first accompanied European colonial expansion and transatlantic slavery that sought to structure social relations and institutions in ways that privilege whiteness."[42]

Smith notes how contemporary surveillance, so concerned with seeing as much as possible, is equally about "not-seeing" its heteropatriarchally structured colonial legacy.[43] The interdependencies between colonial rule, its particular modes of looking/seeing, collecting, recording and managing, as well as the complex ways in which gender, race, class and sexuality shaped colonialism are well documented.[44,45]

Small highlights the continuities between Antebellum surveillance of slaves and contemporary racial profiling that relies on **surveillance mechanisms disproportionately targeting poor people of colour** who lack the technological privilege to opt out.[46] Post 9/11 modes of profiling non-white people, and Muslims in particular, are unthinkable without considering the racist and orientalist colonial logics long predating 9/11 that enable them.

Tamale furthermore shows how colonial legacies, alongside capitalism and globalisation continue to shape contemporary regimes of sexual surveillance in Africa.[47] She maps a complicated patriarchal landscape where state-supported religion (Christianity and Islam with their emphasis of morality, purity and sin) intersects with cultural taboos and the law to inform and control African sexualities. Sexuality is often regulated by colonially inherited penal codes that retain the criminalisation of adultery, prostitution, abortion, sodomy and elopement while protecting marital rape and allowing for defences based on "mistaken belief" or the victim's supposed immorality.[48] As a result, African sexual rights movements operate in environments where some choose to submit to sexual surveillance

[40] Ibid.

[41] Lyon, D. (2003). Introduction. In D. Lyon (Ed.), *Surveillance as Social Sorting: Privacy, risk, and digital discrimination* (pp. 13–30). London: Routledge.

[42] Browne, S. (2015). *Dark Matters: On the Surveillance of Blackness*. Durham: Duke University Press.

[43] Smith, A. (2015). Not-Seeing: State Surveillance, Settler Colonialism, and Gender Violence. In R. E. Dubrofsky & S. A. Magnet (Eds.), *Feminist Surveillance Studies* (pp. 21–38). Durham: Duke University Press.

[44] Stoler, L. A. (2010). *Carnal Knowledge and imperial Power: Race and the Intimate in Colonial Rule.* Berkeley: University of California Press.

[45] McClintock, A. (1995). *Imperial Leather: Race, Gender and Sexuality in the Colonial Contest*. New York: Routledge.

[46] Small, D. (2014, 8 October). Feeling Some Kind of Way About Surveillance. *Model View Culture*. www.modelviewculture.com/pieces/feeling-some-kind-of-way-about-surveillance

[47] Tamale, S. (2014). Exploring the contours of African sexualities: religion, law and power. *African Human Rights Law Journal*, *14*(1), 150–177.

[48] Ibid.

practices – Tamale cites women agreeing to virginity tests for respectability's sake as an example. Conversely, those who cannot conform (sex workers, queers, rape survivors, or those with HIV amongst others) find themselves further marginalised. The vilification, surveillance and control of the "sexual other" is politically instrumentalised, for example when sexuality is constructed as today's "key moral issue on the continent" to distract from mismanagement and corruption at the root of unemployment and other social problems.[49] While Tamale's analysis relies on the similarities in sexual politics across countries, Hodes'[50] work on the South African context – where law and public policy were largely re-written post-apartheid and abortion has been legally available since 1997 – presents further historical continuities in regimes of sexual surveillance. She finds that the proportion of women seeking illegal and unsafe abortions has barely changed since the implementation of the "Choice Act" legalising abortion in 1997. Women fear social stigma attached to abortion as well as punitive treatment by healthcare professionals, many of whom condemn abortion and only reluctantly comply with the Choice Act.[51] Rather than submitting to breaches of confidentiality and privacy built into the legal provision of abortions (crowded facilities, thin walls) and the culture of disapproval among healthcare professionals, many women evade sexual surveillance and protect their privacy by resorting to risky illegal abortions, thus upholding public "norms of silence and secrecy." [52] Hodes concludes that post-apartheid abortion culture, with abortion legal but publicly condemned, shows significant continuities with the past, when abortion was illegal but privately sanctioned. Not only does the state remain similarly invested in sexual surveillance as during the apartheid era, women continue to seek clandestine abortions in large numbers, and the state faces similar limitations in achieving reproductive control.[53]

In addition to highlighting the gendered and racialised implications of surveillance and its colonial and patriarchal continuities, a **feminist lens extends attention to practices that conventionally have not been studied as surveillance** proper to reveal the asymmetrical exercise of power along gendered, racialised, and sexualised lines. Drone footage, CCTV monitoring, wiretapping or the bulk collection of communications data apply to everyone – if, as Small points out, to diverging effects: "We can apparently kiss old assumptions about privacy goodbye. This is especially true for groups without access to technological privilege, who must also deal with race, ethnicity, class, gender, and lifestyle biases against them."[54]

Harry highlights, for instance, how commonplace the exploitative monitoring of black women by law enforcement, media commentators, and online communities is. Stemming from cultural values long predating surveillance cameras, it continues to extend pervasively to contemporary digital spaces and at times includes white feminists' encounter with black women online.[55]

As relevant as conventional surveillance technologies, however, are public health measures, fertility screenings, birth certificates, social media postings, and other everyday practices that incur data outside of our personal control.[56] Andrejevic argues that the study of surveillance needs to include all those "interests, pressures, prejudices, and agendas" and "forms of control that operate in the name of security, efficiency, risk management, and so on, while simultaneously obscuring the forms of gendered, raced, classed, and sexualised discrimination they advance in the name of an allegedly general interest."[57] The contemporary debate tends to frame this general interest in terms of national security, but the examples discussed here illustrate that public health or population control can be equally complicit.

In a case brought to APC's attention by Bangladeshi NGO BFES,[58] for instance, their breast cancer project faced government pressure to hand over non-anonymised data on affected women they had been in contact with, prompting the NGO to develop a system to safeguard their identity. In an environment where women risk being left by their husbands to avoid liability for medical bills, this data in the wrong hands would have exacerbated the already severe social and familial stigma for women seeking cancer care and information in rural areas.[59]

[49] Ibid.

[50] Hodes, R. (2016). The culture of illegal abortion in South Africa. *Journal of South African Studies*, *42*(1), 79–93.

[51] Ibid.

[52] Ibid.

[53] Ibid.

[54] Small, D. (2014, 8 October). Op. cit.

[55] Harry, S. (2014, 6 October). *Everyone Watches, Nobody Sees: How Black Women Disrupt Surveillance Theory*. www.modelviewculture.com/pieces/everyone-watches-nobody-sees-how-black-women-disrupt-surveillance-theory

[56] Dubrofsky, R. E., & Magnet, S. A. (2015). *Feminist surveillance studies*. Durham: Duke University Press.

[57] Andrejevic, M. (2015). Foreword. In R. E. Dubrofsky & S. A. Magnet (Eds.), *Feminist Surveillance Studies*. Durham: Duke University Press.

[58] www.amadergram.org

[59] Communication between BFES and APC.

Drawing attention to continuities between past and present forms of sexual surveillance, highlighting small-scale and/or analogue sexual surveillance practices that a sole focus on big data might obscure, as well as attending to the ways in which data is partial and processed by potentially sexist, racist, and Islamophobic algorithms at work in dataveillance are pressing feminist tasks.

> *More often than not the norm all data bodies are measured against turns out to be white, male, cisgendered and heterosexual*

The algorithmic methods and categorisations that dataveillance relies on by definition operate by proximity to a norm. More often than not the norm all data bodies are measured against turns out to be white, male, cisgendered, and heterosexual. As Conrad notes, "predictive models fed by surveillance data necessarily reproduce past patterns. They cannot take into effective consideration randomness, 'noise', mutation, parody, or disruption unless those effects coalesce into another pattern."[60]

Dataveillance thus by definition runs risk of **reproducing past discriminations and marginalisations through new modes of algorithmic discrimination**. As public health, development, state security as well as private industries increasingly move online and embrace big data, feminist attention to the data practices involved (as well as to those left behind due to a lack of access) is timely and warranted.

THE BODY UNDER SURVEILLANCE

Monahan[61] conceptualises three overlapping gendered dimensions of surveillance:

- Body discrimination
- Context or use discrimination
- Discrimination by abstraction.

Often entangled in practice, they offer guidance in thinking through sexual surveillance beyond using gender and sex to disaggregate (big) data.

Body discrimination refers to "technologies that simply are not designed with a full range of bodies in mind. These technologies privilege certain bodies – usually male, young, White, and able ones – over others."[62] Voice recognition software that struggles with non-male voices, facial recognition software that struggles with non-white faces, or full body scanners singling out non-normative bodies as suspicious are all instances of body discrimination.

By contrast, **context/use discrimination** refers less to those surveilled than to those doing the surveilling and the context the surveillance takes place in. Reminiscent of Mulvey's "male gaze",[63] it describes the masculinised and remote monitoring of feminised spaces. Monahan notes that "when social contexts are already marked by sexist relations, then surveillance (and other) technologies tend to amplify those tensions and inequalities."[64]

Discrimination by abstraction refers to the ways in which context does not translate into the representation of data, "leaving a disembodied and highly abstract depiction of the world and what matters in it."[65] This poses problems when inequalities are not represented in the data, further obscuring existing discriminations.

Van der Ploeg[66] furthermore describes the "informatisation of the body" that has the capacity to affect embodiment and bodily experience and questions the possibility of a **neat distinction between the body itself and its digital representation**. Along with "the body as data" come numerous ways in which "bodies can be monitored, assessed, analysed, categorised, and, ultimately, managed."[67]

[60] Conrad, K. (2009). Surveillance, Gender, and the Virtual Body in the Information Age. *Surveillance & Society*, *6*(4), 380–387.

[61] Monahan, T. (2009). Dreams of Control at a Distance: Gender, Surveillance, and Social Control. *Cultural Studies <-> Critical Methodologies*, *9*(2), 286–305.

[62] Monahan, T. (2009). Ibid.

[63] Mulvey, L. (1975). Visual Pleasure and Narrative Cinema. *Screen*, *16*(3), 6–18.

[64] Monahan, T. (2009). Op. cit.

[65] Ibid.

[66] van der Ploeg, I. (2003). Biometrics and the body as information. In D. Lyon (Ed.), *Surveillance as Social Sorting* (pp. 57–73). London: Routledge.

[67] van der Ploeg, I. (2012). The body as data in the age of information. In K. Ball, K. D. Haggerty, & D. Lyon (Eds.), *Routledge Handbook of Surveillance Studies* (pp. 176–183). Oxon: Routledge.

> *Facebook "likes" can reveal a user's gender with 93% accuracy, sexuality with 75%-88% accuracy, race with 95% accuracy and relationship status with 67% accuracy.*

Hailed as progressive by some, Facebook's inclusion of more than 50 genders in 2014,[68] for instance, paid little more than lip service to transgender and queer critiques of the previously binary gender options as neither targeted advertising on Facebook nor its application programming interface (API) reflect the users' "custom" gender. Having these options affords those outside the male/female binary the opportunity to see the gender identity they embody represented on screen and thus digitally represent themselves. At the same time, they draw attention to critical questions about the coding of gender (and sexuality) in data. The inclusion of diverse gender options can be read as a step towards overcoming gender binaries, but simultaneously counteracts feminist notions of blurring gendered categorisations and promoting views of gender and sexuality as not fixed. The sorting and managing of bodies, particularly vulnerable or marginalised ones, is always already implicated in gendered, sexualised, and racialised ways of seeing, and by extension, of coding and categorising.

While the "custom" gender option is information the user intentionally shares with friends, a recent study shows how an analysis of Facebook "likes" can reveal a user's gender with 93% accuracy, sexuality with 75%-88% accuracy, race with 95% accuracy and relationship status with 67% accuracy.[69] Big data emerges as a rare instance where ***less*** inclusion may make for a stronger feminist argument, particularly considering that more detailed data goes hand in hand with increased surveillance and further marketisation.

The quantified-self trend provides an instance of bodily surveillance that points to the **fluid boundaries between commercial and lateral modes of surveillance** discussed in the next section of this issue paper, here joined by a form of **"self-surveillance"**. The recording, tracking and sharing of health, fitness and nutrition measures extends to sexual and reproductive health, as evidenced by recent research on the spread of menstruation, pregnancy and fertility apps.

Rizk and Othman conclude that the quantification of women's bodies takes place on an unprecedented scale, with large data sets and metadata collected and shared with the applications' platforms and third parties. The algorithms managing the data remain untransparent, with potential implications for "the creation of new normals, of new standards for reproductive and gynaecological indicators based only on those women who have access to these apps, and those who bother to use them."[70]

In addition to binary categories long under feminist scrutiny such as male/female, non/heterosexual or nature/culture, the data body facilitates additional categorisations, for example known/unknown, wanted/unwanted, normal/abnormal and so on.[71] The body in its virtual iteration has the potential to be re-constituted, controlled, marketised, and quite literally sold to the highest bidder.

Taking place from a distance and often unbeknownst to the user, **these mechanisms have material consequences that extend beyond data back to the material body**, for instance by the means of technological barriers to non-male or non-white bodies, increased (border) policing of genderqueer bodies and those with disabilities, or the formation of normative ideas around health and reproduction that affect women's bodies.

SITES OF SEXUAL SURVEILLANCE

Across the literature related to big data and surveillance, two defining moments emerge which conjointly shape the contemporary landscape.

First, the ongoing securitisation of borders, policing, education, development and many other areas, including everyday life under the banner of the **"war on terror" post 9/11**. Gender, race, and sexuality are implicated in these processes. Initiatives carried out in the name of the "war on terror" have at times assumed a feminist guise, for instance, when women's rights were mobilised to justify the invasion of Afghanistan based on discourse

[68] Following further critique, in 2015 the gender options were additionally expanded by a free text field.

[69] Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *PNAS*, *110*(15), 2–5.

[70] Rizk, V., & Othman, D. (2016). Quantifying Fertility and Reproduction through Mobile Apps: A Critical Overview. *Arrow for Change*, *22*(1). www.arrow.org.my/wp-content/uploads/2016/08/AFC22.1-2016.pdf

[71] van der Ploeg, I. (2012). Op. cit.

around the "saving" of Afghan women from the oppressive Taliban.

And the second are recent disclosures, revealing the scope of pervasive and secretive state-driven dataveillance carried out by the US National Security Agency (NSA), the UK Government Communications Headquarters (GCHQ) and their allies globally. Edward Snowden, whose 2013 disclosures on mass surveillance are the most comprehensive and prominent to date, was not the first to disclose details about mass surveillance, as evidenced by the actions of previous whistleblowers including Thomas Drake, Russ Tice and Bill Binney, nor was he the last. 2014 saw an anonymous source disclose the extent of US tracking of "terror suspects", many of whom appear to have no known connection to terrorism at all.[72] Most recently, former Yahoo employees revealed that the email provider secretly scans all customers' emails for intelligence services and/or law enforcement.[73]

The sheer mass of data gathered to feed a wide range of surveillance programmes suggests that data originally retained by social media, phone or internet providers for commercial purposes, as well as all traces the simple use of online services leave potentially seep into mechanisms of state-driven mass surveillance governed by a "collect it all" logic.

Thus while analytically useful and politically sensible to distinguish between the surveillance practices of states, commercial entities, and individuals, **in practice the boundaries between these categories are fluid**. It is nevertheless important to note that commercial surveillance, for example when large-scale social media datasets are mined for marketing purposes, and lateral surveillance where peers draw on similar mechanisms as states and corporations to monitor one another[74] form part and parcel of dataveillance alongside state surveillance.

Often discussed in relation to gendered algorithmic advertising practices or social media platforms' role in (tackling) technologically mediated violence, the commercial realm of surveillance extends beyond marketing and social media.

While the case of privacy risks on online dating platforms[75] carries obvious implications for sexual surveillance, the example of app-based car-for-hire company Uber shows that sexualised data practices can also be found in seemingly mundane business, such as figuring out when/where to dispatch drivers. Uber's data scientists not only correlated rides to/from prostitution-prone areas with the habitual paydays for benefits recipients, but rebranded the so-called "walk of shame" a "ride of glory" after discovering increased demand of their service based on patterns they associated with one night stands.[76] While Uber published these insights in a (humorous) blog post that was later deleted, they illustrate the potential for sexual, and in this case classed, surveillance in data collected for commercial purposes.

Andrejevic describes instances where peers like family members, friends, or acquaintances keep track of one another as "lateral surveillance."[77] Lateral surveillance involves technologies such as online background screenings, portable cameras, keystroke loggers, spyware, or lie detectors but can equally consist of repeatedly googling someone, intensively following their social media presence, excessive and/or threatening commenting, amongst other mundane pursuits.

> *Sexualised modes of lateral surveillance overlap with online harassment and stalking.*

Sexualised modes of lateral surveillance arise when these activities intrude on women's freedom of expression and right to privacy, or when they overlap with online harassment and stalking. Revenge porn websites or smartphone apps specifically designed to track spouses and family members come to mind foremost as tools for lateral sexual surveillance. However, in addition to those as well as to the surveillance technologies listed above, perpetrators can abuse mainstream apps that use

[72] Scahill, J. & Devereaux, R. (2014, 5 August). Watch Commander. *The Intercept*. https://theintercept.com/2014/08/05/watch-commander/

[73] Menn, J. (2016, 4 October). Exclusive: Yahoo secretly scanned customer emails for U.S. intelligence – sources. *Reuters*. www.reuters.com/article/us-yahoo-nsa-exclusive-idUSKCN1241YT

[74] Andrejevic, M. (2005). The work of watching one another: Lateral surveillance, risk, and governance. *Surveillance & Society*, *2*(4), 479–497.

[75] Reitman, R. (2012, 10 February). Six Heartbreaking Truths about Online Dating Privacy. *EFF*. https://www.eff.org/deeplinks/2012/02/six-heartbreaking-truths-about-online-dating-privacy

[76] Dataconomy (2014, 12 June). Uber: Mapping Prostitution and "The Walk of Shame". www.dataconomy.com/uber-mapping-prostitution-and-the-walk-of-shame/

[77] Andrejevic, M. (2005). Op. cit.

location data, messaging services and even sleep monitoring apps to track women's movement.

Andrejevic furthermore highlights **ways in which governments enrol the public in surveillance practices**, for example when enlisting the public to report anything "suspicious" or "unusual" to authorities. He concludes that the spread of new media, rather than having a democratising effect, has facilitated "the injunction to embrace the strategies of law enforcement and marketing at a micro-level. (…) The result has not been a diminution of either government or corporate surveillance, as evidenced by their converging role in the 'war on terror', but rather their amplification and replication through peer monitoring networks."[78]

The shooting of John Crawford III by US police forces illustrates the interplay of lateral surveillance "with historical discrimination. The police who ultimately ended his life were responding to a report, via citizen surveillance, that he had been observed with a gun."[79] Shot in a supermarket as he was holding a toy gun, he is one of a growing number of young black men killed by US law enforcement. Imagery and videos of these deaths circulate widely, turning them into posthumous spectacles of lateral surveillance. The surveillance footage of his shooting, on the other hand, did not lead to the indictment of the police officer responsible for his death.[80]

The focus of this paper is on data-driven modes of sexual surveillance and its impact on women, sexual minorities, and anyone else who may have a stake in the intersection between bodily integrity, sexuality, sexual rights, freedom of expression, health and the myriad ways in which potentially mineable data traces occur. Therefore, a further important site for sexual surveillance, the private home, intimate relationships, and the family is left underexplored**.**

Women, people of colour, queers and other marginalised groups have always been under sexual surveillance in the private as well as in the public sphere. It is perhaps instructive that increased public and media attention to technologies of surveillance coincides with revelations about the scope and scale of the mass surveillance of ***everyone***, while longstanding gendered and racialised modes of surveillance failed to grasp public imagination to remotely similar extents.

CHALLENGES <-> OPPORTUNITIES

Debates on surveillance rightfully emphasise **the right to privacy,** which **remains a feminist issue in more ways than one.**[81] It has, however, never been equally extended to all. Women, people with disabilities, refugees, sexual minorities, people receiving state benefits, or incarcerated populations, amongst others, can testify to myriad ways in which their privacy has habitually been invaded.[82]

CHALLENGES: DIGITAL EXCLUSIONS, ALGORITHMIC DISCRIMINATIONS, AND AMBIGUOUS VISIBILITIES

Whether any transformative potential of the informatisation of the body can be realised depends on who controls the information: "When the control of a person's information is out of that person's hands, so too is the nature of the potential transformation."[83]

Manovich identifies **three emerging classes in data-driven societies, "those who create data** (both consciously and by leaving digital footprints), those who have the means to **collect it,** and those who have expertise to **analyse it**. The first group includes pretty much everybody in the world who is using the web and/or mobile phones; the second group is smaller; and the third group is much smaller still."[84]

> *The transformative potential of the informatisation of the body depends on who controls the information.*

[78] Ibid.

[79] Harry, S. (2014, 6 October). Op. cit.

[80] Ibid.

[81] The Feminist Principles of the Internet, for instance, advocate the right to privacy and control over one's personal information online and defend the right to anonymity. At the same time, feminism's relationship with privacy is a complicated one that rests on decades' worth of critique of the public/private distinction, not least due to the abuse against women perpetrated in the "privacy" of their homes. The private remains as political as ever.

[82] Dubrofsky, R. E., & Magnet, S. A. (2015). Op. cit.

[83] Conrad, K. (2009). Op. cit.

[84] Manovich, L. (2011). Trending: The Promises and the Challenges of Big Social Data. In M. K. Gold (Ed.), *Debates in the Digital Humanities* (pp. 1–17). Minneapolis: University of Minnesota Press.

The control of data bodies predominantly lies with the latter two groups. Those consist of governments, corporations, research institutions, and international organisations holding the infrastructure to collect and store large amounts of data, and of the subset of analysts who both have the required skills ***and*** access to the data.

Juxtaposed with the gendered and racialised power relations underpinning the surveillance assemblage, inclusion in tech industries, persisting divides in access to the internet,[85,86] and technology-mediated violence against women, this means that the control over sexual surveillance lies outside of the hands of the women and queers, particularly those of colour, it concerns.

Women and queers are at risk of tech-related violence, that is, acts of gender-based violence "committed, abetted or aggravated, in part or fully, by the use of information and communication technologies (ICTs), such as phones, the internet, social media platforms and email."[87]

Tech-related violence "infringes on women's right to self-determination and bodily integrity. It impacts on women's capacity to move freely, without fear of surveillance" and "often what are seen as 'merely virtual' threats soon translate into physical violence."[88] Contextually, the same threats extend to trans* people, intersexed people, and ethnic minority groups.[89]

In terms of big data and sexual surveillance, tech-related violence not only affects the data generated (use of pseudonyms, self-censorship and other blind spots), but also limits access and participation and hinders the freedom of expression online. Tech-related violence is clearly situated within the broader landscape of gender-based violence rather than caused or determined by technologies mediating it. However, it amplifies heteronormative and patriarchal modes of sexual surveillance and **further marginalises women and queers in data, code, and life**. Research on female internet users in Mumbai, for instance, found that young women feel uncomfortable when exposed to male observers in internet cafes and are concerned about lateral surveillance (by family members and peers) of their online activities.[90]

> *A technology in itself may be amoral and not created with discriminatory intent, but can nevertheless have sexist and racist outcomes that reproduce social bias.*

Digital exclusions, be they due to lack of access to the internet or due to the violent silencing of those present, further skew the data in favour of those able to participate freely and safely. Exclusions from big data are exacerbated when joined by algorithmic discriminations and racist technologies. Examples include Google search results returning images of black women with natural hair for "unprofessional hair" or normatively attractive white people for "men" and "women";[91] or the history of colour film technology that until recently was unable to process darker skin tones in good quality.[92] A technology in itself may be amoral and not created with discriminatory intent, but can nevertheless have sexist and racist outcomes that reproduce social bias.

Magnet's work on (the failings of) biometric technologies[93] documents how they fail considerably more often on women, people of colour, and differently abled bodies, all while serving states to track and control marginalised groups such as refugees or recipients of benefits. "Othered bodies" become subject to heightened surveillance, as Magnet and Rodgers' work on full body imaging technology

[85] APC. (2016). *Ending digital exclusion: Why the access divide persists and how to close it*. www.apc.org/en/system/files/APC_EndingDigitalExclusion.pdf

[86] Othman, D. (2016). Access, Legislation, and Online Freedom of Expression: A Data Overview. *Arrow for change*, 22(1), 49-55. www.arrow.org.my/wp-content/uploads/2016/08/AFC22.1-2016.pdf

[87] www.genderit.org/onlinevaw/

[88] Malhotra, N. (2014). *Good Questions on Technology-Related violence*. End violence: Women's rights and safety online, APC. www.apc.org/en/pubs/good-questions-technology-related-violence

[89] Ibid.

[90] Bhattacharjya, M., & Ganesh, M. I. (2011). Negotiating intimacy and harm: Female internet users in Mumbai. In *://EROTICS Sex, Rights and the Internet*. APC. www.apc.org/en/system/files/EROTICS.pdf

[91] Alexander, L. (2016, 8 April). Do Google's 'unprofessional hair' results show it is racist? *The Guardian*. https://www.theguardian.com/technology/2016/apr/08/does-google-unprofessional-hair-results-prove-algorithms-racist-

[92] Caswell, E. (2015, 18 September). Color film was built for white people. Here's what it did to dark skin. *Vox*. www.vox.com/2015/9/18/9348821/photography-race-bias

[93] Magnet, S. A. (2011). *When Biometrics Fail: Gender, Race, and the Technology of Identity*. Durham: Duke University Press.

shows. Particularly the "intersections of transgendered, disabled, fat, religious, female and racialised" bodies are disproportionally affected and singled out for increased surveillance and harassment.[94]

The expectation of transparency (of citizens rather than governments) is often accompanied by "nothing to hide – nothing to fear" narratives, with far-reaching feminist implications. Governed by heteronormative institutions within borders that come with their own racialised and sexualised technologies of control, **how much is safe to reveal and what must remain hidden is not equal for all**. Beauchamp[95] discusses what "nothing to hide" means for trans* people for whom questions of stealth versus visibility take on multiple dimensions. Some negotiate a desire/need to remain hidden with medical and legal records that were never afforded privacy, and others' compliant visibility risks complicity with national security discourse around who can "pass" as a "safe" citizen or traveller (white, conclusively gendered) and who becomes subject to policing. The blame for systemic discrimination is transferred onto the victims; if they suffer negative consequences, it must be because they were hiding what is wrong with them.[96]

The heightened visibility and transparency that accompany increased inclusion in data can pose challenges in terms of sexual surveillance. South African research illustrates that online counter-publics play a significant role in the negotiation of transgender and lesbian identities, sexualities and politics. The authors note, however, that black lesbians, the group facing the most violent response to their sexualities, were reluctant to participate in the research.[97]

Participating in queer or feminist activism online, navigating social media as a member of a sexual minority, particularly when additionally racialised are ambiguous instances of visibility that can come at great cost ranging from involuntary outing to harassment, social stigma, and persecution. Research with Kenyan queers highlights strategies developed to negotiate their online presence: "Nearly all respondents maintain two accounts on Facebook because of high levels of lateral surveillance: a 'straight' account using their real name, where they connect with their family, straight friends and church community; and a queer account under an adopted name where they connect with others in the LGBTQ community. (…) Nevertheless, the two accounts can result in dangerous situations; respondents described being accidentally outed to their families after being tagged under their real name in a photo attending a LGBTQ gathering. Cases of surveillance also include family members actively seeking out queer accounts to expose users."[98]

> *Facebook's insistence on nominal transparency penalises non-normative identities when they disrupt the collection of personal data.*

Such strategies serve the safety of sexual minorities and enable their participation online, but come in conflict with Facebook's real name policy, which has come under critique for its negative impact on trans* and queer individuals and activists. MacAulay and Moldes characterise Facebook's insistence on nominal transparency as penalising non-normative identities when they disrupt the collection of personal data. They found that "Facebook prioritised threats that had a market value, and that profiles that fail to generate useful/marketable data are seen as a greater liability than abusive users."[99]

In conjunction with questionable content-moderation practices,[100] **the politics of real-name policies affect big data** and sexual surveillance in several ways:

- Taking down "infringing" content and profiles leads to further exclusion of already marginalised users.
- Enforcing real names and normative identities reinforces heteronormative logics in social data.

[94] Magnet, S., & Rodgers, T. (2012). Op. cit.

[95] Beauchamp, T. (2009). Artful Concealment and Strategic Visibility: Transgender Bodies and U. S. State Surveillance After 9/11. *Surveillance & Society*, 6(4), 356–366.

[96] Andrejevic, M. (2015). Op. cit.

[97] Prinsloo, J., & McLean N. C. (2011). The internet and sexual identities: Exploring transgender and lesbian use of the internet in South Africa. In *://EROTICS Sex, Rights and the Internet*. APC. www.apc.org/en/system/files/EROTICS.pdf

[98] Ganesh, M. I., Deutch, J., & Schulte, J. (2016). *Privacy, anonymity, visibility: dilemmas in tech use by marginalised communities*. opendocs.ids.ac.uk/opendocs/handle/123456789/12110

[99] Macaulay, M., & Moldes, M. D. (2016). Queen don't compute: reading and casting shade on Facebook's real names policy. *Critical Studies in Media Communication*, *33*(1).

[100] York, J. (2016, 20 September). *Facebook's nudity Ban Affects All Kinds of Users.* Electronic Frontier Foundation. www.eff.org/deeplinks/2016/09/facebooks-nudity-ban-affects-all-kinds-users

- It enables the involuntary outing and direct sexual surveillance of queer individuals.
- It simultaneously makes users more transparent to dataveillance.

Social media data in itself is highly curated – be that due to users' choice, lack of access, self-censorship, or other reasons. Those data are then **marketised** (for example, in targeted advertising), potentially **tracked by peers**, and **aggregated** with other data available to law enforcement and security.

For activists advocating women's and sexual rights online the overlapping nature of these modes of surveillance poses further questions. Imagine for example an unsuspecting activist who has gained a following based on her expertise on topics like abortion, gender-based violence, HIV/AIDS, or LGBT issues, resulting in an increasingly public profile. What she might perceive as successful advocacy work, however, places her wider network as well as those her work concerns at heightened risk.

The surveillance of networks and relationships by adversaries is commonplace, and activists are important nodes in those networks. They become targets themselves, as evidenced for instance by police surveillance of protest movements such as Black Lives Matter[101] or the international trade in surveillance technologies used to track activists,[102] or expose links to other activists and members of the communities they work with. As Ball notes, "when one is exposed, one is exposed to something."[103]

However, under the contemporary surveillance assemblage where lateral, commercial, and state surveillance have become closely entangled, **it is often not intelligible who precisely that is.** What may present as benign online collaboration and exchange amongst feminist activists and those involved in their projects, may in fact be closely watched by hostile governments, groups opposing the cause in question, employers, law enforcement, families or peers of those at risk, and so on. Under these circumstances, the consideration of adequate operational security is crucial to feminist activists to protect themselves as well as those they work with/for from becoming targets.

OPPORTUNITIES: BETWEEN RESISTANCE, SOUSVEILLANCE AND COMMUNITY DATA

The potential in the collection and analysis of large quantities of data has not gone unnoticed by activists, researchers, as well as international organisations. The UN, for instance, run a number of big data projects to support sustainable development and humanitarian action. They note that big data are "no modern panacea for age-old development challenges"[104] and advocate an approach to big data that recognises context as key.

An ongoing project, for instance, works on the real-time monitoring of the implementation of ***Option B+***, a treatment programme to prevent mother-to-child HIV transmission available to HIV positive mothers in Uganda. Pulse Lab Kampala has developed an application that tracks metrics such as number of patients receiving the treatment, number of antenatal care visits, or number of HIV/AIDS cases to monitor the performance of clinics in areas where the programme is in place.[105]

A past project analysed the impact of advocacy movement ***Every Woman Every Child*** using four years worth of tweets and machine learning to identify 14 million relevant tweets about women's and children's health. The analysis showed an increase in the conversation about maternal and child health, as well as demographic trends peaks that could be linked to events internal and external to the campaign.[106]

While such projects are indicative of data-driven opportunities for positive change in terms of sexual and reproductive health, an overly celebratory stance risks masking the epistemic dangers that riddle the data/development terrain. Gurumurthy and Chami liken "data for development"[107] narratives to a "techno-solutionism" that is complicit in epistemic violence and plays into the hands of a neo-liberal capitalist brand of

[101] Ozer, N. (2016, 22 September). *Police Use of Social Media Surveillance Software is Escalating, and Activists Are in the Digital Crosshairs.* ACLU. www.aclu.org/blog/free-future/police-use-social-media-surveillance-software-escalating-and-activists-are-digital

[102] Privacy International (2015, 23 March). *Ethiopia expands surveillance capacity with German tech via Lebanon.* www.privacyinternational.org/?q=node/546

[103] Ball, K. (2009). Information, Communication & Society. *Information, Communication & Society*, *12*(5), 639–657.

[104] UN Global Pulse. (2012). *Big Data for Development: Challenges & Opportunities.*

[105] UN Global Pulse. (2016). *Monitoring in Real Time the Implementation of HIV Mother-to-Child Prevention Programme.* Retrieved from http://www.unglobalpulse.org/projects/monitoring-hiv-mother-child-prevention-programme on 20.09.2016

[106] UN Global Pulse. (2013). *Advocacy Monitoring through Social Data: Womens and Children's Health.* www.unglobalpulse.org/projects/EWEC-social-data-analysis

[107] cf. www.data4sdgs.org/

development.[108] This tension between potential public health benefits of big data and its adverse effects in terms of surveillance is tangible in the UN's recognition that "while metadata can provide benefits", they can also be aggregated to "reveal personal information and … give an insight into an individual's behaviour, social relationships, private preferences and identity."[109]

The remainder of this section turns to civil society-based data projects on gender, sexuality, and sexual rights. **Sousveillance**,[110] i.e. practices of inverse surveillance that aim at watching from below, watching the watchers by observing and recording surveillance practices, teaching and sharing knowledge and technologies to defend vulnerable groups and activists against adverse effects of surveillance, has grown in significance.

Such practices have been spurred by movements of the so-called Arab Spring, Occupy, and Black Lives Matter where citizen-reporting on violent authorities has played a central role, as well as by whistleblowers such as Chelsea Manning or Edward Snowden whose sousveillance of the military and security agencies contributed to wider public awareness of previously hidden wrongdoings. Often taking place on a grassroots scale, **sousveillance practices related to gender and sexuality are emerging**. Where there is power, there is resistance,[111] and efforts to subvert technologies – be that creative resistance against surveillance or efforts to use surveillance technologies for good – deserve amplification in the face of a discursive landscape of terror and national security panics.

Data practices need to be viewed in light of the "agency and reflexivity of individual actors as well as the variable ways in which power and participation are constructed and enacted"[112] rather than algorithmic power alone.

Harry describes how hashtags or street recordings are "ways of using tech to push back against surveillance." Along with other Twitter users and participants in an online community of black women, she used the hashtag **#Yourslipisshowing** "to expose 4chan board members who declared 'war' on black feminists by tracking and attempting to infiltrate their 'ranks'." She notes, however, that such public sousveillance strategies hinge on social capital and relative safety of those involved.[113]

The initiatives cited here differ in context, scale and aims but share participatory and situated approaches to resisting sexual surveillance, fighting its adverse effects, or initiating data practices for positive change.

HarassMap[114] is an initiative founded 2010 in Egypt that maps and calls out sexual harassment by crowdsourcing text messages and online reports of sexual harassment and assault, and mapping the instances online. The community organisers and volunteers behind HarassMap then base communication campaigns and programmes designed to make schools, universities, public places, and workplaces safer for women on the collected data. Initially Cairo-based, the project has resonated internationally, and the Egyptian team has coached organisers in 28 countries, including Algeria, Canada, India, Japan, Kenya, Lebanon, Pakistan, Palestine, South Africa, Syria and the US.[115] It is noteworthy that HarassMap builds on Ushahidi,[116] an open source crowdmapping platform originally developed in Kenya to report instances of post-election violence in 2008.

The **Utunzi Rainbow Security Network** serves as another example of sousveillance by the means of community mapping. The Kenyan collaboration between three LGBTQ organisations seeks to crowdmap violence against queers in Kenya. Individuals as well as organisations can report as well as request assistance when faced with or witnessing human rights violations and other violence abuses based on sexual orientation and/or gender identity and expressions.[117] Utunzi struggled to gain the trust and support of its potential users, as being web-based rather than app-based made it difficult for many potential users to access, due to a lack of trust in an unknown platform, and because the LGBTQ community did not see any tangible benefit of mapping incidents.[118]

[108] Gurumurthy, A., & Chami, N. (2016, 31 May). Data: The New Four-Letter Word for Feminism. *GenderIT.org*. www.genderit.org/articles/data-new-four-letter-word-feminism

[109] undocs.org/en/A/RES/69/166

[110] Mann, S. (2004). "Sousveillance" Inverse Surveillance in Multimedia Imaging. In *Proceedings of the 12th annual ACM international conference on Multimedia* (pp. 620–627).

[111] Foucault, M. (1978). *The History of Sexuality: An Introduction*. New York: Pantheon.

[112] Couldry, N., & Powell, A. (2014). Big Data from the bottom up. *Big Data & Society*, (July-December), 1–5.

[113] Harry, S. (2014, 6 October). Op. cit.

[114] www.harassmap.org/en/what-we-do/

[115] www.harassmap.org/en/what-we-do/around-the-world

[116] www.harassmap.org/en/who-we-are/our-partners

[117] https://utunzi.com/about.php

[118] Ganesh, M.I., Deutch, J., & Schulte, J. (2016). *Privacy, anonymity, visibility: dilemmas in tech use by marginalised communities*. opendocs.ids.ac.uk/opendocs/handle/123456789/12110

In addition to strategic opposition to surveillance systems, and the development of mechanisms to evaluate and contest such systems, Monahan calls for **democratising surveillance practices**.[119] If we are to engage in data practices that are empowering, that have the potential to further sexual rights and improve the quality of life of women and sexual minorities, those practices have to contribute to shifting data power from corporations and governments into the hands of individual women, queers, and their communities.

As the hypothetical example of the somewhat oblivious feminist activist at the end of the previous section suggests, women's rights and sexual rights advocates can face similar challenges as the communities they work with in terms of safety and visibility, with the added danger of placing others at risk. Feminist work thus often makes it necessary to resist surveillance and to protect the privacy and anonymity of marginalised women, queers and activists. Activists, researchers, and journalists working on sexual rights might benefit from closer ties with digital rights movements to improve safety and security in their own activism, to mitigate against the risks of increased visibilities, and to reap the benefits of data activism for sexual rights causes.[120]

A wide range of free open source tools offer ways to protect data, anonymity and privacy and thus resist the surveillance of activist work. The accessible and fairly comprehensive manuals aimed at women, sexual minorities and activists, such as A DIY Guide to Feminist Cybersecurity,[121] Zen and the Art of Making Tech Work for You[122] or the Take Back the Tech! digital safety roadmaps[123] introduce a wide range of safety strategies, practices and tools. While 100% security is an illusion, particularly when faced with powerful state adversaries, using open source encryption, anonymity, and privacy tools go a long way towards keeping feminist and sexual rights activists, their sensitive data, and the vulnerable communities they work with as safe as possible.[124]

CONCLUSION

The picture that emerges on the nexus between big data and sexual surveillance is an ambiguous one. Calls for better representation of women and queers, for reaping data's benefits in terms of development, gender equality, and sexual health, and for better recognition of gender-based and sexual violence have the potential to improve the lives of marginalised groups. They, however, have to be leveraged against concerns about the politics of the collection and analysis of big data, discriminations coded into the collection and algorithmic analysis of data, its colonial legacies, and the complicated politics of visibility that go along with the presence in data.

> *Whether data practices are transformative depends on agency and consent, on how data is collected, by whom, and to what ends.*

Whether data practices are transformative depends on agency and consent, on how data is collected, by whom, and to what ends. Does the data serve the powerful to further marginalise women and queers, or does it empower them to make informed sexual health decisions, resist harmful power structures, exercise their sexual rights, and enjoy bodily integrity? A feminist lens on surveillance contributes depth, perspective, a focus on power relations, and attention to the quirks and outliers so often ignored in big data.

The Feminist Principles of the Internet, developed by APC, call for an internet that empowers women and queers in all their diversity and recognise that the internet is not free-floating but an extension of other social spaces that are often "shaped by patriarchy and heteronormativity."[125] UN resolution A/HRC/32/L.20 on "the promotion, protection and enjoyment of human rights on the Internet" was recently passed as an addition to Article 19 of the Universal Declaration of Human Rights[126] and reaffirms the importance of access to a free

[119] Monahan, T. (2006). Questioning Surveillance and Security. In T. Monahan (Ed.), *Surveillance and Security: Technological Politics and Power in Everyday Life*. New York: Routledge.

[120] Ibid.

[121] tech.safehubcollective.org/cybersecurity

[122] gendersec.tacticaltech.org/wiki/index.php/Complete_manual

[123] https://www.takebackthetech.net/be-safe

[124] See also DATNAV, a guide to working with digital data for human rights researchers, as many of the strategies it lists apply equally to women's and sexual rights work. www.theengineroom.org/wp-content/uploads/2016/09/datnav.pdf

[125] APC. (2016). Op cit.

[126] www.un.org/en/universal-declaration-human-rights

and open internet for all.[127] While this resolution does not oblige governments to provide the infrastructure to bring all citizens online, it represents a welcome affirmation that human rights extend to digital spaces. In contrast to previous resolutions on human rights and the internet, it explicitly stresses the importance of access to information technology and digital literacy for the empowerment of women and girls.

Feminist scholarship and activism on/against surveillance shows **clear continuities between discriminatory practices offline and online**, as well as historical continuities between the patriarchal surveillance and control of women's, non-normative, and racialised bodies and contemporary modes of sexual surveillance. Thinking of sexual surveillance and the data it operates through, no matter how big or small, in terms of situated knowledges[128] places the technologies involved in collecting, storing, and analysing them in their social context.

While **data-driven sexual surveillance** often takes place on the level of abstraction, it **produces embodied consequences and meanings**, as the examples discussed in this paper show. Affording these localised meanings participation and voice in data practices has the potential to drive the data, but also the systems and social relations they are embedded in towards more social justice,[129] not least for women and sexual minorities.

The Feminist Principles of the Internet highlight the importance of "an ethics and politics of consent" that affords women and sexual minorities "informed decisions on what aspects of their public or private lives to share online."[130] The notion of consent in the context of big data and sexual surveillance is complicated by the overlapping nature of lateral, commercial, and state surveillance this paper has outlined. Consent to the collection of particular data for particular purposes (say sexual health services, the use of a social media platform, research, or advocacy) of course does not equal consent to the bulk collection of that same data by government agencies. Data gathered in the course of advocacy work and sexual health projects, sensitive data in the hands of sexual rights activists, or data on activists themselves as well as their wider networks form part and parcel of the data/surveillance assemblage.

While the internet remains an important sphere for the advancement of sexual rights and empowerment of women, advocacy work requires a situated and contextual assessment of (and mitigation against) the risks its data practices may expose activists, women, and sexual minorities to. Community-based open source projects that generate data to be used in the struggle against gender-based and sexual violence and towards the empowerment of women and queers are promising and indicative of democratising data practices.[131]

A feminist practice opposes the non-consensual collection of data and plays its part in preventing and/or subverting the non-consensual use of data already collected.

> *A feminist practice opposes the non-consensual collection of data and plays its part in preventing and/or subverting the non-consensual use of data already collected.*

Given the pervasive yet unaccountable nature of dataveillance practices, the protection of **information privacy and anonymity remain a prerequisite for any transformative use of data**. A feminist practice thus opposes the non-consensual collection of data and plays its part in preventing and/or subverting the non-consensual use of data already collected.

When developing or participating in data practices that aim at furthering social justice goals, leveling algorithmic discriminations, or empowering women and sexual minorities, it is attentive to consent and participation and takes adequate care to safeguard the data of vulnerable groups involved as well as of activists themselves at risk of surveillance.

[127] undocs.org/A/HRC/32/L.20

[128] Haraway, D. J. (1988). Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. *Feminist Studies*, *14*(3), 575–599.

[129] Monahan, T. (2009). Op. cit.

[130] APC. (2016). Op cit.

[131] Monahan, T. (2006). Op. cit.

APC
ASSOCIATION FOR
PROGRESSIVE
COMMUNICATIONS

Internet and ICTs for social justice and development

APC is an international network of civil society organisations founded in1990 dedicated to empowering and supporting people working for peace, human rights, development and protection of the environment, through the strategic use of information and communication technologies (ICTs).

We work to build a world in which all people have easy, equal and affordable access to the creative potential of ICTs to improve their lives and create more democratic and egalitarian societies.

www.apc.org info@apc.org

Graphic Design: MONOCROMO

AMPLIFYCHANGE

THIS ISSUE PAPER FORMS PART OF APC'S EROTICS NETWORK BUILDING IN INDIA, NEPAL AND SRI LANKA,
AND WAS PRODUCED WITH THE SUPPORT OF AMPLIFYCHANGE.

BIG DATA AND SEXUAL SURVEILLANCE
DECEMBER 2016

ISBN 978-92-95102-71-2 APC-201612-WRP-I-EN-DIGITAL-261