Geography Faculty Publications   Geography

11-9-2020

Viral Data

Agnieszka Leszczynski
Western University, Canada

Matthew Zook
University of Kentucky, zook@uky.edu

Repository Citation

Viral Data

Viral Data

Agnieszka Leszczynski[1] ⓘD **and Matthew Zook[2]** ⓘD

Abstract

We are experiencing a historical moment characterized by unprecedented conditions of virality: a viral pandemic, the viral diffusion of misinformation and conspiracy theories, the viral momentum of ongoing Hong Kong protests, and the viral spread of #BlackLivesMatter demonstrations and related efforts to defund policing. These co-articulations of crises, traumas, and virality both implicate and are implicated by big data practices occurring in a present that is pervasively mediated by data materialities, deeply rooted dataist ideologies that entrench processes of datafication as granting objective access to truth and attendant practices of tracking, data analytics, algorithmic prediction, and data-driven targeting of individuals and communities. This collection of papers explores how data (and their absences) is figuring in the making of the discourses, lived realities, and systemic inequalities of the uneven impacts of the coronavirus pandemic.

This article is a part of special theme on Viral Data. To see a full list of all articles in this special theme, please click here: https://journals.sagepub.com/page/bds/collections/viraldata

We are experiencing a historical moment characterized by unprecedented conditions of virality: a viral pandemic, the viral diffusion of misinformation, and conspiracy theories including the idea that 5G data networks are spreading COVID-19, the viral momentum of ongoing Hong Kong protests, and the viral spread of #BlackLivesMatter demonstrations and related efforts to defund policing. These viralities designate both epidemiological phenomena, as in the case of the novel coronavirus, and also capture the ways in which digital networks and data serve as vectors for the rapid, decentralized transmission, diffusion, and circulation of ideas, ideology, mis/information, and social movements (Kumar, 2015; Marwick, 2013; Tufekci, 2017; Varis and Blommaert, 2015). These viralities are deeply interconnected, concomitant with the layering of political, ecological, health, and economic crises and traumas, many rooted in systemic racism and white supremacy made most recently visible in the killings of, in the United States, George Floyd, Ahmaud Arbery, and Breonna Taylor, as well as far too many others.

These co-articulations of crises, traumas, and virality both implicate and are implicated by Big Data practices occurring in a present that is pervasively mediated by data materialities, deeply rooted dataist ideologies that entrench processes of datafication as granting objective access to truth (Van Dijck, 2014), and attendant practices of tracking, data analytics, algorithmic prediction, and data-driven targeting of individuals and communities. As the COVID-19 pandemic unfolded, it quickly became apparent that data—as well as their absences—is figuring in the making of the discourses, lived realities, and systemic inequalities of the uneven impacts of the coronavirus. Also evident are intensifications and accelerations of datafication

[1]Geography, Western University, London, Canada
[2]Geography, University of Kentucky, Lexington, USA

Corresponding author:
Matthew Zook, Geography, University of Kentucky, Lexington, KY 40506, USA.
Email: zook@uky.edu

which invoke COVID-19 as an *ex post* justification for the deepening entrenchment of Big Data in processes of regulation, governance, surveillance, and subjectification. *Big Data & Society's* status as the flagship interdisciplinary journal for original work prefiguring the implications of Big Data practices for societies motivated us to invite contributions[1] exploring the connections between Big Data and COVID-19 across the social sciences and humanities, which are brought together here under the banner of "Viral Data".

In naming this symposium *Viral Data*,[2] we mobilize "virality" in multiple senses of the term to signal and explore emergent configurations of Big Data, datafication, and analytics in the context of COVID-19: as a rapid pivot of big tech toward the viral pandemic as a new opportunity to apply troves of data and data-driven practices (Cinnamon, 2020); the viral diffusion of data solutionism across jurisdictions; COVID-19 (mis)information "going viral" over social media channels; and widespread gaps in available data that, much like a virus of the epidemiological variety, are often hidden and difficult to address yet also demand solutions (a vaccine, alternative data epistemologies, visualizations, and praxes). The contributions to this symposium engage the "virality" of data vis-à-vis the COVID-19 pandemic across this spectrum of meanings, often marshaling "viral" in more than one sense of the term.

Under this wide range of meanings, contributors to the symposium offer a first engagement with COVID-19 Big Data practices as implicated in challenges to, as well as reimaginations and reconfigurations of, larger social, cultural, political, and economic processes. Maalsen and Dowling (2020), Sandvik (2020), Milne and Costa (2020), and Milan (2020) detail the drive for data-centric techno-solutionist interventions motivated by COVID-19's crises, and the discursive power of framing the coronavirus pandemic *as* crisis. As Maalsen and Dowling (2020) identify, the spectrum of lockdown and shelter-in-place orders around the world has emboldened incursions of smart technologies into the ever-more intimate spheres of the home under the guises of facilitating work productivity, schooling, and maintaining social connection to the outer world. In so doing, these technologies accelerate the datafication of the home, intimacy, and what has historically been considered the private sphere in ways that weaken the rights of tenants, workers, and citizens. Focused on the British context, Milne and Costa (2020) examine how the pandemic is disrupting and dislocating digital health data and futures. They outline how techno-revolutionary imaginaries start with an envisioned future from which to reverse-engineer a present moment of crisis/opportunity for datafication while simultaneously mobilizing data-driven interventions

in the health field as templates for actualizing these futures.

Sandvik (2020) outlines a similar recourse to crisis in the way that *Smittestopp*—a specifically Norwegian attempt to "appify" the coronavirus response (Datta, 2020) via a poorly designed COVID-19 contact tracing app—leveraged cultural senses of responsibility to the collective as well as high levels of trust in government institutions to secure social buy-in for this intervention. Ultimately the insufficient privacy protections and questionable social benefits resulted in the government canceling the Smittestopp app. The *Smittestopp* experience of data-driven intervention in the context of crisis is not unique: as Milan (2020) argues, the urgency toward the quantitative datafication of COVID-19 and its impacts is not necessarily met with capabilities or readiness to take advantage of datasets. Drawing on the work of decolonial scholars, Milan emphasizes the importance of thinking from the margins, particularly when designing responses that enable communities to retain agency and autonomy over their data rather than being abstracted into data flows via techno-solutionist interventions. Elsewhere, drawing on a comparison of the Estonian and Finnish experience in applying mobile Big Data to social mobility during the pandemic, Poom et al. (2020) highlight the challenges of designing and maintaining institutional frameworks that allow meaningful access to mobile data given its geographically contingent, spatially uneven, and potentially privacy-eroding nature. Noting the strong spatial elements to the pandemic—contact tracing, mobility monitoring, exposure notification, and quarantine confinement—Poom et al. (2020) argue for the importance of designing flexible institutional arrangements that enable researchers to have better (yet secured) access to mobile Big Data during future crises. These concerns of access are further complicated by issues of data's selective availability, where in many jurisdictions, such as the United Kingdom and the United States, "policy inputs" and priorities "determine the data available" (Taylor, 2020, pp. x; see also D'Ignazio and Klein, 2020).

In addition to the contingencies of data availability, many of the contributions to this symposium elucidate the importance of attending to data's glaring absences. These concerns revolve around data availability about the uneven impacts of the coronavirus on those most structurally susceptible to the multi-dimensional impacts of COVID-19: migrants and refugees, gig laborers, frontline and essential workers, carers, Black and minority ethnic (BAME) and Indigenous communities, and the socioeconomically less well advantaged (D'Ignazio and Klein, 2020; Pelizza, 2020; Taylor, 2020). As Pelizza (2020) argues, these data absences are actively underwritten by the viral

circulation of virulently racist (and patently false) misinformation about the inherent immunity of BAME communities to COVID-19, which has informed underreporting and even exclusion of these communities from data flows that shape policy, healthcare interventions, and the distribution of resources. Viral misinformation is likewise explored by Gruzd and Mai (2020), who chronicle how a single hashtag—*#FilmYourHospital*—spawned a conservative-fueled social media conspiracy theory that COVID-19 was a left-wing hoax. Using social network analysis, Gruzd and Mai (2020) demonstrate the difficulty of countering the spread of misinformation, particularly where its diffusion maps onto ideological cleavages that can be readily leveraged under current social media practices and regulation by operatives seeking political gain.

Focusing on the viral proliferation of data visualizations as a means of evidencing and parsing the pandemic, Bowe, Simmons and Mattern (2020) further detail the ways politics fuel selective representations that have circumscribed interpretations of COVID-19's spatial and social impacts. Yet, as they write, alongside official data visualizations generated by the state, a viral burgeoning of counter-plots and subaltern mappings of the pandemic has also emerged to document the quarantine quotidian. When read together, both these official and counter-visualizations capture and render the visible multiple scales of the pandemic and its material effects. Drawing on a number of examples, Bowe, Simmons and Mattern (2020) position these counter-plots as means to invite critical reflection about the sourcing and visualization of COVID-19, including questions about which kinds of subjects are and are not registering in Big Data flows. Taylor (2020) likewise argues that a critical COVID-19 data praxis must divest itself from a preoccupation with absences of data as merely "gaps" to be filled through the collection of more and more data. Taylor (2020) argues that such a response only further dehumanizes life (see also Raji, 2020). Instead, Taylor (2020) advocates for a feminist ethics of care to re-embody COVID-19 in the personhood of those rendered both knowable and unknown in systems of datafication. D'Ignazio and Klein's (2020) intervention makes a similar call for a re-embodiment of data to both make labor and subjects visible. In extending their seven principles for a feminist data science to COVID-19, they profile the work being done by movements such as Data for Black Lives to rectify data absences and to produce counter-data for actionable and more equitable data practices and interventions.

On 25 May 2020, George Floyd was killed in the United States by a Minneapolis police officer, spurring demonstrations of grief, outrage, and solidarity with Black Lives and mobilizations for restorative justice around the world. The focus of this *Viral Data* theme on co-articulations of Big Data and COVID-19 may now read as off-centered and out of step with these more recent events. While this is an artifact of timing rather than of selective focus, pervasive data inequalities, the disproportionate impacts of COVID-19 on BAME and Indigenous communities, and police brutality perpetrated against persons of color are co-implicated, rooted in anti-Blackness, structural racism, and white supremacy. As Barrington Walker (in CBC News, 2020, n.p.) has recently elucidated:

> we're seeing the connection between longer histories of socio-economic marginalization, the impoverishment the neglect of [B]lack communities [the United States and Canada] and how that connects to the historical line that you can draw from those older histories to the condition of [B]lack communities today.

And data, both big and small, has played profoundly important roles in this systemic marginalization and impoverishment.

A small sampling from the history of datafication in the social sciences includes the racist classification of cranial measurements used to "prove" the superiority of white persons (Gould and Gold, 1996), the use of U. S. neighborhood racial composition to power data visualizations justifying redlining policies to block African-American access to home mortgages (Aaronson et al., 2019 [2017]), and claims that inherited traits tied to race determine intelligence and socio-economic status (Fischer et al., 1996). More recently, within the context of Big Data, researchers have analyzed how predictive crime mapping legitimizes and expands racialized policing policies (Jefferson, 2018), how bias in existing datasets used for artificial intelligence training produces facial recognition algorithms that have difficulty identifying non-white faces (Buolamwini and Gebru, 2018), and how Big Data algorithms incorrectly predicted that African-American parolees were at higher risk of reoffending (Angwin et al., 2016). Returning to the COVID-19 pandemic, it is impossible to ignore the direct outcomes of racialized societies and data, including "racial disparities in exposure to [toxic] pollutants" associated with, and which exacerbate, co-morbidities that make Black Americans more likely to die of COVID-19 (Washington, 2020: 241; Yancy, 2020). Moreover, as Bowe, Simmons and Mattern (2020) note, this also manifests in some COVID-19 data collection and curation practices that disassociate variables of race from cases obscuring the impact of the pandemic on Black and other racially marginalized communities.

We recognize that while these issues are among the most pressing concerns at the intersections of Big Data

and society, they have been under-explored in the pages of *Big Data & Society*. While the journal has always been open to all scholarly submissions on themes related to Big Data, we will undertake to solicit submissions that explore and analyze systemic inequalities rooted in and co-concomitant with structural and systemic racism. We welcome contributions that do so through either stand-alone papers or proposals for special themes in response to calls by the journal which are announced twice a year. The next deadline for special theme proposals is 15 September 2020 (see: http://big datasoc.blogspot.com/2020/06/call-for-special-theme-proposals-for.html). We are further committed to pro-actively seeking research from, and working with, emerging scholars.

ORCID iDs

Agnieszka Leszczynski ![ORCID] https://orcid.org/0000-0002-5167-0499
Matthew Zook ![ORCID] https://orcid.org/0000-0002-6034-3262

Notes

1. We openly acknowledge and critique the lack of scholars of color or scholars focused on race and Indigeneity and data in this symposium. When we began to identify contributors in mid-April, we identified racial inequality and anti-Blackness as a priority for interventions and invited a number of BAME and Indigenous scholars specializing in these issues to participate. However, given the challenges of the pandemic, we recognize that some potential contributors were already overcommitted. This became an even a greater issue after the killing of George Floyd as the demands of the moment are for scholarly engagement in praxis rather than academic commentary. We recognize that these demands are being placed disproportionately on, and being met by, Black, Indigenous, and minority ethnic scholars.
2. We thank Evelyn Ruppert, Founding Editor of *Big Data & Society*, for suggesting the "Viral Data" title for this symposium.

References

Aaronson D, Hartley D and Mazumder B (2019 [2017]) *The effects of the 1930s HOLC "Redlining" Maps*. Federal Reserve Bank of Chicago Working Paper WP 2017-2. Available at: https://www.chicagofed.org/publications/working-papers/2017/wp2017-12 (accessed 20 July 2020).

Angwin J, Larson J, Mattu S, et al. (2016) Machine Bias There's software used across the country to predict future criminals. And it's biased against blacks. *ProPublica*, 23 May. Available at: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing (accessed 20 July 2020).

Bowe E, Simmons, E and Mattern S (2020) Learning from lines: Critical COVID data visualizations and the quarantine quotidian. *Big data & society*, 7(2), 2053951720939236.

Buolamwini, J. & Gebru, T. Gender shades: intersectional accuracy disparities in commercial gender classification. 2018. *Proceedings of Machine Learning Research* Vol 81: pp. 1–15. Vol 81 of PoMLR comes from the Conference on Fairness, Accountability and Transparency, 23-24 February 2018, New York, NY, USA.

Butler OE (1993) *Parable of the Sower*. New York: Grand Central Publishing.

CBC News (2020) More people ready to be 'actively anti-racist' in light of George Floyd's death, Waterloo activist says. *CBC News*, 03 June. Available at: https://www.cbc.ca/news/canada/kitchener-waterloo/george-floyd-death-unrest-racism-waterloo-region-1.5596645 (accessed 21 June 2020).

Cinnamon J (2020) Platform philanthropy, 'public value', and the COVID019 pandemic moment. *Dialogues in Human Geography 10*(2): 242–245.

Datta A (2020) Self(ie) governance: Technologies of intimate surveillance in India under COVID-19. *Dialogues in Human Geography 10*(2): 234–237.

D'Ignazio C and Klein LF (2020) Seven intersectional feminist principles for equitable and actionable COVID-19 data. *Big Data & Society 7*(2): 2053951720942544.

Fischer CS, Hout M, Jankowski MS, et al. (1996) *Inequality by Design: Cracking the Bell Curve Myth*. Princeton, NJ: Princeton University Press.

Gould SJ and Gold SJ (1996) *The Mismeasure of Man*. New York: W.W. Norton & Company.

Gruzd A and Mai P (2020) Going viral: How a single tweet spawned a COVID-19 conspiracy theory on Twitter. *Big Data & Society, 7*(2), 2053951720938405.

Jefferson BJ (2018) Predictable policing: Predictive crime mapping and geographies of policing and race. *Annals of the American Association of Geographers 108*(1): 1–16.

Kumar S (2015) Contagious memes, viral videos and subversive parody: The grammar of contention on the Indian web. *International Communication Gazette 77*(3): 232–247.

Maalsen S and Dowling R (2020) Covid-19 and the accelerating smart home. *Big Data & Society 7*(2): 2053951720938073.

Marwick A (2013) Memes. *Contexts 12*(4): 12–13.

Milan S (2020) Techno-solutionism and the standard human in the making of the COVID-19 pandemic. *Big Data & Society 7*(2): 2053951720966781.

Milne R and Costa A (2020) Disruption and dislocation in post-COVID futures for digital health. *Big Data & Society 7*(2): 2053951720949567.

Pelizza, A. (2020). "No disease for the others": How COVID-19 data can enact new and old alterities. *Big Data & Society* 7(2), 2053951720942542.

Poom A, Järv O, Zook M, and Toivonen T (2020) COVID-19 is spatial: Ensuring that mobile Big Data is used for social good. *Big Data & Society* 7(2): 2053951720952088.

Raji ID (2020) The discomfort of death counts: Mourning through the distorted lens of reported COVID-19 death data. *Patterns 1*: 100066.

Sandvik KB (2020) "Smittestopp": If you want your freedom back, download now. *Big Data & Society* 7(2): 2053951720939985.

Taylor L (2020) The price of certainty: How the politics of pandemic data demand an ethics of care. *Big Data & Society* 7(2): 2053951720942539.

Tufekci Z (2017) *Twitter and Tear Gas: The Power and Fragility of Networked Protest*. New Haven/London: Yale University Press.

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

Varis P and Blommaert J (2015) Conviviality and collectives on social media: Virality, memes, and new social structures. *Multilingual Margins: A Journal of Multilingualism from the Periphery* 2(1): 31–45.

Washington HA (2020) How environmental racism is fuelling the coronavirus epidemic. *Nature 581*(7808): 241.

Yancy CW (2020) COVID-19 and African Americans. *JAMA 323*(19): 1891–1892.