

A post-truth pandemic?

Taylor Shelton 

Big Data & Society
July–December: 1–6
© The Author(s) 2020
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/2053951720965612
journals.sagepub.com/home/bds



Abstract

As the coronavirus pandemic continues apace in the United States, the dizzying amount of data being generated, analyzed and consumed about the virus has led to calls to proclaim this the first ‘data-driven pandemic’. But at the same time, it seems that this plethora of data has not meant a better grasp on the reality of the pandemic and its effects. Even as we have the potential to digitally track and trace nearly every single individual who has contracted the virus, we have no idea exactly how many people have had the virus, been hospitalized, or died because of it, largely due to a confluence of factors, particularly active obfuscation and mismanagement by public authorities and misinformation spread through social media and right-wing media channels. But beyond these dynamics, there also lies the less nefarious ways that the everyday, subjective practices of data collection, analysis and visualization have the potential to themselves (re)produce these very same dynamics where data is at once valorized and ignored, preeminent and completely useless. That is, the pandemic has revealed only the general inadequacy of our data infrastructures and assemblages to solving pressing social issues, but also the more general shift towards a ‘post-truth’ disposition in contemporary social life. But, as this paper argues, it would be a mistake to see the centrality of data as being somehow the opposite from the larger post-truth apparatus, as the two are instead fundamentally intertwined and co-produced.

Keywords

coronavirus, post-truth, data-driven governance, critical data studies, data visualization, data politics

Each day for the last six months, the Johns Hopkins COVID-19 interactive dashboard has received billions of daily interactions (Perkel, 2020), while governors, mayors, and even the President provide press briefings announcing, as a kind of public ritual, the latest counts of infections and deaths due to coronavirus. Along the way, countless forms of data have been proposed as ways to better understand and counteract the deadly pandemic, from the use of contact tracing apps, social media, and mobile phone trace data that can track people’s everyday movements at extremely granular spatial and temporal scales (Glanz et al., 2020; Watkins et al., 2020) to the hypothesized potential for water quality and fecal sampling of municipal sewers as a way of tracking the virus’ presence in “near-real-time”, even before symptoms begin to appear (Lesté-Lasserre, 2020; Smith, 2020). Together, these phenomena and countless others have lent credence to calls to proclaim this the first “data-driven pandemic” (Geraghty and Frye, 2020; Rocha, 2020).

But at the same time as data is being generated, analyzed, and consumed at dizzying speeds, it seems that our collective grasp on the virus—at least in the United States—is almost no better for it. While testing

regimes have improved considerably since the virus’s initial outbreak in the US in March, it remains true that “the number of infected is close to meaningless” (O’Neil, 2020). So even as we have the potential to digitally track and trace nearly every single individual who has contracted COVID, we have no idea exactly how many people have had the virus, been hospitalized, or died because of it, largely due to a confluence of factors, particularly active obfuscation and mismanagement by public authorities and disinformation spread through social media and right-wing media channels. But beyond these dynamics, there also lies the less nefarious ways that the everyday, subjective practices of data collection, analysis, and visualization have the potential to themselves (re)produce these very same dynamics where data is at once valorized and ignored, preeminent, and completely useless.

Department of Geosciences, Georgia State University, Atlanta, GA, USA

Corresponding author:

Taylor Shelton, Department of Geosciences, Georgia State University,
730 Langdale Hall, Atlanta, GA 30303, USA.
Email: jshelton19@gsu.edu



And so, in taking stock of the broader, lasting implications of the pandemic, it's evident that coronavirus has revealed not only the general inadequacy of our data infrastructures and assemblages, but also the more general shift towards a “post-truth” disposition in contemporary social life, wherein “objective facts are less influential in shaping public opinion than appeals to emotion and personal belief” (Oxford Dictionaries, 2016). But, as I've argued in a recent chapter (Shelton, forthcoming), it would be a mistake to see the centrality of data as being somehow the opposite from the larger post-truth apparatus that leads data and facts to exist on unstable ground. Instead, these two ostensibly opposed dynamics are fundamentally intertwined and co-produced.

Willful obfuscation

Arguably, the most obvious examples of the post-truth pandemic are those where governmental officials are responsible for willfully obfuscating the reality of the virus and its impact on people, and actively intervening to prevent damaging—but still very much factual—information from being released to the public. But the less commented upon dynamic is how these actions remain cloaked in the veneer of being data-driven and scientifically grounded even as the actions and ends to which they are put are anything but.

One noteworthy example is the firing of Florida Department of Health staffer Rebekah Jones, a geographer responsible for managing the state's coronavirus dashboard. Jones was fired after she alleged that her

superiors demanded data be removed from the dashboard, including “data showing that some residents tested positive for the coronavirus in January, even though DeSantis assured residents in March that there was no evidence of community spread,” as well as her being “asked to manually change numbers to wrongly make counties appear to have met metrics for reopening” (Iati, 2020).

When Jones went public with the story of her firing, Florida Governor Ron DeSantis called into question Jones' credentials, saying “She's not a data scientist. She's somebody that's got a degree in journalism, communication and geography,” at once leveraging the social power imbued in “data science” while *also ignoring* the fact that data science is a nebulous, ill-defined field that one need not have a degree in to be a practitioner of. DeSantis similarly accused Jones of “putting data on the portal which the scientists didn't believe was valid data” (Taylor, 2020a), while a spokesperson for the Florida Department of Health argued that Jones' dashboard “aggregates disparate sets of data without considering many of the important guidelines utilized by epidemiologists” (Iati, 2020). By fighting Jones' accusations of mismanagement and obfuscation through the language of data, DeSantis and the state employees who answer to him simultaneously reify both the “post-truth” *and* the “data-driven,” privileging the sanctity of data at the same time as the state's liberal reopening policy and lack of restrictions have made Florida a key epicenter of the coronavirus within the United States.

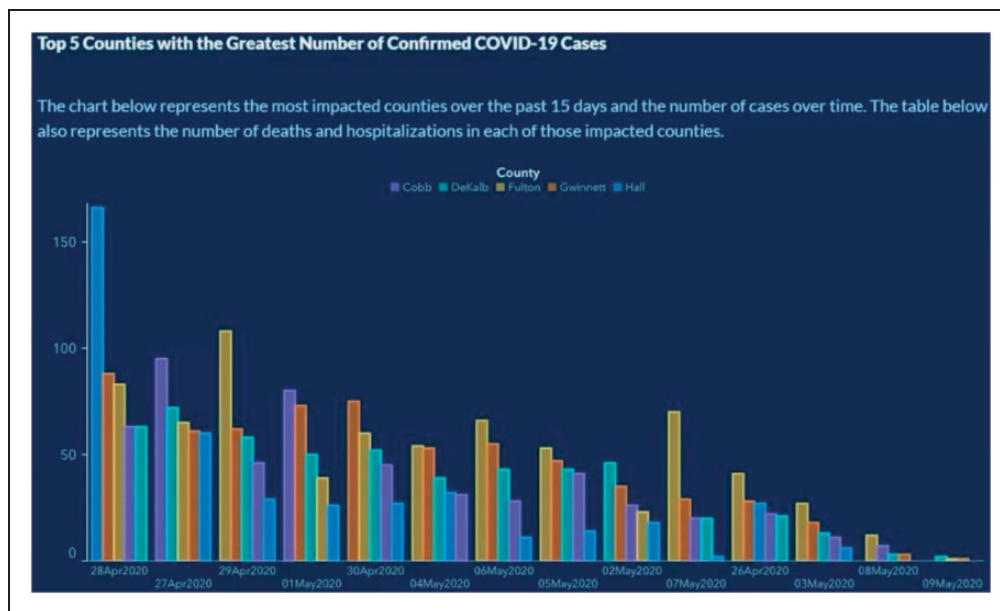


Figure 1. Doctored chart of Coronavirus cases in Georgia showing dates out of order, originally produced by the Georgia Department of Public Health: <https://dph.georgia.gov/covid-19-daily-status-report>.

Similar processes have been at work in Georgia, the first state to reopen back in May. Like Florida, Georgia has been beset by all manner of coronavirus-related mishaps and mishandlings at the hands of Governor Brian Kemp. One of the most notable was the release of the chart seen in Figure 1 back in May, which, at first glance, appears to show a decline in the number of confirmed coronavirus cases in the state's five most-affected counties. Upon closer inspection, the chart shows no such thing, with the dates along the x-axis considerably out of chronological order, as would be conventional or expected for any data visualization where time is a variable.

Given Governor Kemp's attempts to reopen the state as early as possible, the intention behind the manipulated chart was easily imputed. But, according to an article about the controversy in the *Atlanta Journal-Constitution*, Kemp's office and other departments in the state government offered at least three different explanations for the error, blaming it alternately on a user error in sorting the dates, a problem with the software itself, and, ultimately, arguing that the chart wasn't released in error at all, but was meant to simply display a different perspective (Mariano and Trubey, 2020). Ultimately, one state legislator couldn't call it anything but "cuckoo." All the while, Kemp has cloaked himself in the language of being data-driven, having "repeatedly said that data, science and the advice of health officials drive his decision making, including his delays in imposing statewide social-distancing measures to contain the virus' spread" (Judd and Teegardin, 2020).

Social media misinformation

Though the role of government officials in deliberately misleading the public is perhaps the single most important way the response to COVID-19 has been bungled, a more diffuse responsibility lays at the feet of those who have used social media (or other means of mass communication) to spread misinformation and conspiracy theories about the pandemic. The diffusion of this information has been so substantial that the United Nations and World Health Organization have warned of a companion "infodemic" sweeping the globe, noting that fake news "spreads faster and more easily than this virus" (United Nations Department of Global Communications, 2020).

Given that it has now been nearly four years since Donald Trump's election as President of the United States, talk of social media-based misinformation campaigns is commonplace, thanks to the near-mythical role of Russian bots intervening in the 2016 election. But the centrality of social media misinformation in the response to the coronavirus pandemic takes on a

particular salience given the ways social media has long been heralded as a potential means of responding to such crises in a more efficient and effective manner. From the halcyon days of Google Flu Trends in 2010 to the now practically uncountable number of papers declaring social media platforms to be "early warning systems" for disease surveillance, digital trace data is lauded for its promise in making visible threats like coronavirus before more conventional scientific or governmental systems would be able to see them. Nevermind, of course, the countless biases in the data that make such systems less-than-perfect reflections of the material realities they purport to represent (cf. Crawford, 2013; Lazer et al., 2014).

But the pervasiveness of misinformation on social media calls into the question the validity of this data for understanding the actual dynamics of disease transmission during a pandemic like COVID-19. As initial analysis of social media discussions about coronavirus have revealed, as many as half or more of the Twitter accounts discussing coronavirus may be bots, including the majority of the platform's most popular accounts (Hao, 2020). But perhaps more distressingly, the propagation of such misinformation doesn't even require bots, as there are plenty of *actual people* participating as well. As early work by both Stephens (2020) and Gruzd and Mai (2020) has shown, right-wing memes and conspiracy theories about the pandemic are rapidly disseminated from low follower-count users through the tightly-linked (and geographically concentrated) network of right-wing commentators, even in the absence of substantial bot activity. And while the misinformation propagated has surely been enough to convince many citizens that there's no need to take the precautions of staying home, wearing a mask or otherwise working to protect the health and safety of their fellow citizens, such effects would have been minimized if the aforementioned obfuscation by official representatives was not so significant.

The fact that such deliberate misinformation can thrive in the very same medium that's been heralded as a high-tech savior for all manner of social problems points to the dialectical tension between the post-truth and data-driven. Indeed, we might see a post-truth society as the logical outgrowth of the data-driven (or, perhaps more accurately, data-centric) society we've been living in in recent years, where individualized, decontextualized data points serve as the focal point for social and political life. But when these decontextualized data points are foregrounded, we lose the meta-narrative that binds this data together into a comprehensible whole. And in the absence of such a coherent narrative about what's going on and how things got to be this way, each of these ostensibly objective—but not always related—data points lead us

to a place where the contrived narratives of right-wing trolls can come to counteract “truth,” if not overcome it outright.

Everyday contingencies and subjective data practices

Cumulatively, the willful obfuscation of data on coronavirus by government official and the massive spread of misinformation through social media channels have meant that even as data remains abundant, relatively accessible, and very much central to our collective thinking about the pandemic, our experience of and response to the pandemic is also very much shaped by a tendency for “truth” to be put on the back burner. It would be a mistake, however, to see each of these dynamics as being the lone things standing in the way of a truly data-driven, scientifically-informed, and even socially-equitable, response to the COVID-19 crisis. Instead, underlying each and every aspect of coronavirus data is a more pervasive problem that can’t be so easily discarded in the effort to detach ourselves from this larger “post-truth” moment. That is, no matter the context, the production, analysis, visualization, and interpretation of data is a fundamentally (inter-)subjective, power-laden process shaped by any number of social, cultural, political, and economic considerations (cf. Kitchin, 2014). The everyday practice of working with data is one that itself calls into being the foundations of a post-truth society through its messiness and contingency.

As technology writer and founder of the COVID Tracking Project Alexis Madrigal wrote in the early days of the outbreak in the United States:

[t]he point is that every country’s numbers are the result of a specific set of testing and accounting regimes. Everyone is cooking the data, one way or another. And yet, even though these inconsistencies are public and plain, people continue to rely on charts showing different numbers, with no indication that they are not all produced with the same rigor or vigor. (Madrigal, 2020)

With the millions upon millions of individuals looking at maps, charts, and other visualizations of the pandemic on a daily basis, the potential for such a massive amount of information to mislead is significant, even in the absence of an intentional effort to distract or discount (Mooney and Juhász, 2020).

Such is the case of another recent scandal in the visualization of COVID-19 data in Georgia (see Figure 2). While this case shares similarities with the aforementioned chart, the case is not so clear cut an example of willful obfuscation. Instead, the effect of maintaining an overall visual pattern by changing the size of the categories the data is classified into is likely the result of a much more mundane decision about how to classify the data into an easily intuited map. Of course, the problem arises in attempting to use these maps to understand the evolution of the virus over time, where the seemingly apolitical, methodologically-justifiable decision to use a natural breaks classification masks the fact that the number of cases continues to rise considerably each day, with the state being among the worst in the country in terms of sustained outbreaks. Even without the kind of willful malfeasance discussed above, the Georgia map legend points to the ways that everyday decisions made in data analysis can lead to an obscuring of the

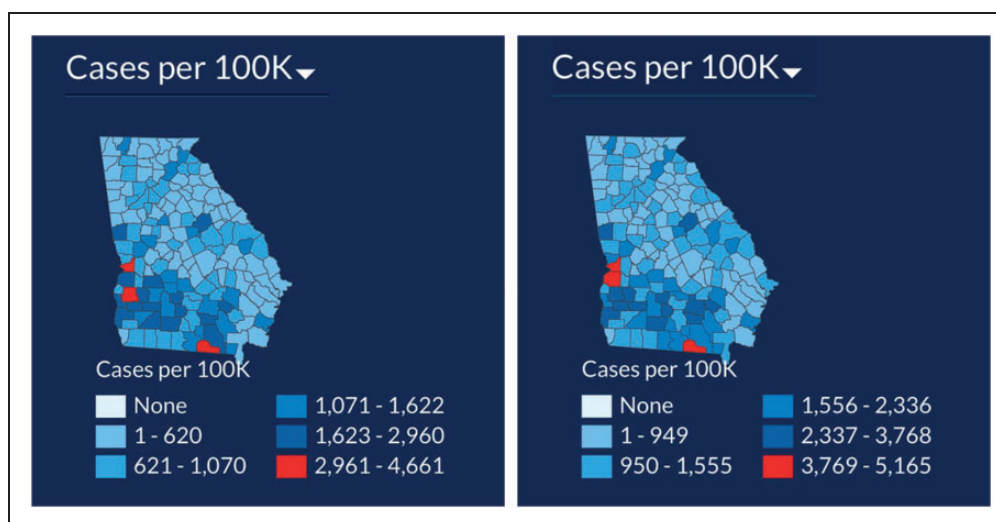


Figure 2. Maps of Coronavirus Cases in Georgia taken from the Georgia Department of Public Health’s COVID-19 Daily Status Report website: <https://dph.georgia.gov/covid-19-daily-status-report> on (L) July 2 and (R) July 17.¹

underlying dynamics, and thus lending credence (again, even if it's unintentional) to the narrative that nothing is actually wrong or worth worrying about.

While these kinds of biases and subjectivities are inherent in all aspects of data production, analysis, and visualization, these processes don't require a uniform adherence to the notion that data provides an unvarnished look at the objective reality of the world, as demonstrated by a slew of what Bowe et al. (2020) call "counter-plots and subaltern maps" of coronavirus, which attempt to provide a more contextualized, critical understanding of the pandemic and its surrounding social conditions. But, by and large, the emphasis on continual representation of whatever data there is, particularly through interactive data dashboards, serves to limit our views of coronavirus to the immediate emergency at hand, obscuring the ways it has been produced through processes operating at a larger scale (both spatially and temporally) (Everts, 2020).

As much as reflecting any underlying "truth" about the realities of COVID's impacts, this slew of rapidly-produced quantitative indicators reflects the inconsistent and fractured regime of epidemiological data collection in the United States, where there are 50 (or more) different ways of tracking and intervening in the virus (Vestal, 2020), thanks in large part to the lack of centralized capacity by a federal government run not only by Donald Trump, but hollowed out after nearly 50 years of neoliberal retrenchment. So, no matter how robust the computational modeling capacity of scientists and how competent and careful are the analysts working to communicate these results to policymakers, there remains an underlying question that this system of fractured federalism produces in the context of a pandemic.

The result is, again, inconsistent data which reveals a tendency to invisibilize already-marginalized groups who are bearing the brunt of this virus (Taylor, 2020b), further limiting the scope of our understanding of the virus and ability to adequately address it. And while it would be impossible to ever produce a truly objective, holistic, and totalizing account of the pandemic, the reality of our broader data infrastructures means that it isn't just a single set of biases that have been dealt with, but rather a near infinite number of combinations of data, methods, and interpretations ripe with contingencies. Appropriately addressing these subjectivities and contingencies can only be done by embracing their existence, *not* pretending that data somehow exist as a salve or panacea for post-truth practices.

Conclusion

The sum of these factors, both in general and in relation to the coronavirus pandemic in particular, has

been a pervasive lack of faith in the state to intervene on behalf of its citizens, but also a lack of faith in data, science, and "truth" more broadly. But these conditions emerge alongside, and arguably even *because of*, the massive investment of social power and cache into data and computation as apolitical, neutral arbiters of truth that cannot be attained through less totalizing methods. But the end result of this widespread investment into data is the fact that the goalposts can always be moved; when bigger (and thus better) data is seen as not only possible, but expected, anything that falls short of such a comprehensive vision or understanding is seen to be lacking, rather than being taken on its own terms. That is, the foundations of a "post-truth" society inhere in data itself. While the lasting impacts of coronavirus on science and the public's connection to data are yet to be determined, these changes have certainly meant that just as we're living through the first data-driven pandemic, we're also living through the first "post-truth pandemic," and it will almost certainly not be the last.

Declaration of conflicting interests

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

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Note

1. This issue was most notably highlighted by Andisheh Nourae in a tweet on July 17 (<https://twitter.com/andishehnourae/status/1284237474831761408>), which has received over 26,000 retweets and nearly 48,000 likes as of July 31. It is worth noting, however, that the state updated its data dashboard on July 28 to address the criticisms of its maps (Trubey, 2020).

ORCID iD

Taylor Shelton  <https://orcid.org/0000-0003-0776-9296>

References

- Bowe E, Simmons E and Mattern S (2020) Learning from lines: Critical COVID data visualizations and the quarantine quotidian. *Big Data & Society* 7(2). <https://doi.org/10.1177/2053951720939236>
- Crawford K (2013) The hidden biases in Big Data. *Harvard Business Review*. 1 April. Available at: <https://hbr.org/2013/04/the-hidden-biases-in-big-data> (accessed 29 September 2020).

- Everts J (2020) The dashboard pandemic. *Dialogues in Human Geography* 10(2): 260–264.
- Geraghty E and Frye C (2020) COVID-19: The First Global Pandemic of the Information Age. *Esri*. 22 June. Available at: <https://storymaps.arcgis.com/stories/a5190c7fd6db422f9a1bab6dac024b99> (accessed 29 September 2020).
- Glanz J, Carey B, Holder J, et al. (2020) Where America Didn't Stay Home Even as the Virus Spread. *The New York Times*. 2 April. Available at: <https://www.nytimes.com/interactive/2020/04/02/us/coronavirus-social-distancing.html> (accessed 29 September 2020).
- 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): <https://doi.org/10.1177/2053951720938405>.
- Hao K (2020) Nearly half of Twitter accounts pushing to reopen America may be bots. *MIT Technology Review*. 21 May. Available at: <https://www.technologyreview.com/2020/05/21/1002105/covid-bot-twitter-accounts-push-to-reopen-america/> (accessed 29 September 2020).
- Iati M (2020) Florida fired its coronavirus data scientist. Now she's publishing the statistics on her own. *Washington Post*. 16 June. Available at: <https://www.washingtonpost.com/nation/2020/06/12/rebekah-jones-florida-coronavirus/> (accessed 29 September 2020).
- Judd A and Teegardin C (2020) Faulty data obscures virus' impact on Georgia. *Atlanta Journal-Constitution*. 16 April. Available at: <https://www.ajc.com/news/faulty-data-obscures-virus-impact-georgia/LhCiI0bVKXOQW9VuEF9OrN/> (accessed 29 September 2020).
- Kitchin R (2014) *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. Thousand Oaks, CA: Sage.
- Lazer D, Kennedy R, King G, et al. (2014) The parable of Google flu: Traps in big data analysis. *Science* 343(6176): 1203–1205.
- Lesté-Lasserre C (2020) Coronavirus found in Paris sewage points to early warning system. *Science*. 21 April. Available at: <https://www.sciencemag.org/news/2020/04/coronavirus-found-paris-sewage-points-early-warning-system> (accessed 29 September 2020).
- Madrigal A (2020) The Official Coronavirus Numbers Are Wrong, and Everyone Knows It. *The Atlantic*. 3 March. Available at: <https://www.theatlantic.com/technology/archive/2020/03/how-many-americans-really-have-coronavirus/607348/> (accessed 29 September 2020).
- Mariano W and Trubey JS (2020) 'It's just cuckoo': State's latest data mishap causes critics to cry foul. *Atlanta Journal-Constitution*. 19 May. Available at: <https://www.ajc.com/news/state-regional-govt-politics/just-cuckoo-state-latest-data-mishap-causes-critics-cry-foul/182PpUvUX9XEF8vO1INVGO/> (accessed 29 September 2020).
- Mooney P and Juhász L (2020) Mapping COVID-19: How web-based maps contribute to the infodemic. *Dialogues in Human Geography* 10(2): 265–270.
- O'Neil C (2020) 10 Reasons to Doubt the Covid-19 Data. *Bloomberg News*. 13 April. Available at: <https://www.bloomberg.com/amp/opinion/articles/2020-04-13/ten-reasons-to-doubt-the-covid-19-data> (accessed 29 September 2020).
- Oxford Dictionaries (2016) Word of the Year 2016. Available at: <https://languages.oup.com/word-of-the-year/2016/> (accessed 29 September 2020).
- Perkel JM (2020) Behind the Johns Hopkins University coronavirus dashboard. *Nature Index*. 7 April. Available at: <https://www.natureindex.com/news-blog/behind-johns-hopkins-university-coronavirus-dashboard> (accessed 29 September 2020).
- Rocha R (2020) The data-driven pandemic: Information sharing with COVID-19 is 'unprecedented'. *CBC News*. 17 March. Available at: <https://www.cbc.ca/news/canada/coronavirus-date-information-sharing-1.5500709> (accessed 29 September 2020).
- Shelton T (Forthcoming) Data-driven governance, post-truth politics. In: Meisterlin L (ed) *Digital Urbanisms*. New York, NY: Columbia Books on Architecture and the City.
- Smith PA (2020) To Spot Future Coronavirus Flare-Ups, Search the Sewers. *Scientific American*. 30 June. Available at: <https://www.scientificamerican.com/article/to-spot-future-coronavirus-flare-ups-search-the-sewers/> (accessed 29 September 2020).
- Stephens M (2020) A geospatial infodemic: Mapping twitter conspiracy theories of COVID-19. *Dialogues in Human Geography* 10(2): 276–281.
- Taylor L (2020a) Ousted manager was told to manipulate COVID-19 data before state's re-opening, she says. *Tampa Bay Times*. 22 May. Available at: <https://www.tampabay.com/news/health/2020/05/22/ousted-manager-was-told-to-manipulate-covid-19-data-before-states-re-opening-she-says/> (accessed 29 September 2020).
- Taylor L (2020b) The price of certainty: How the politics of pandemic data demand an ethics of care. *Big Data & Society* 7(2). <https://doi.org/10.1177/2053951720942539>
- Trubey JS (2020) Georgia revamps virus maps, charts that critics said were confusing. *Atlanta Journal-Constitution*. 28 July. Available at: <https://www.ajc.com/news/coronavirus/georgia-revamps-virus-maps-charts-that-critics-said-were-confusing/YFEBS4VWEJEZZF4VFSV2UVVKF4/> (accessed 29 September 2020).
- United Nations Department of Global Communications (2020) UN tackles 'infodemic' of misinformation and cybercrime in COVID-19 crisis. *United Nations COVID-19 Response*. 31 March. Available at: <https://www.un.org/en/un-coronavirus-communications-team/un-tackling-%E2%80%98infodemic%E2%80%99-misinformation-and-cybercrime-covid-19> (accessed 29 September 2020).
- Vestal C, Fernandez C and Orso A (2020) Public Coronavirus Data Varies Widely Between States. *Pew Charitable Trusts Stateline*. 27 April. Available at: <https://www.pewtrusts.org/en/research-and-analysis/blogs/stateline/2020/04/27/public-coronavirus-data-varies-widely-between-states> (accessed 29 September 2020).
- Watkins D, Holder J, Glanz J, et al. (2020) How the virus won. *The New York Times*. 25 June. Available at: <https://www.nytimes.com/interactive/2020/us/coronavirus-spread.html> (accessed 29 September 2020).