

## Chinese Computing and Computing China as Global Knowledge Production

Angela Xiao Wu

New York University

angelaxwu@nyu.edu

### Abstract

As data-driven technologies and business models pervade on a global scale, China's enormous digital economy often signals its dominating power by dint of data extraction. Complicating this view, this critical commentary focuses on knowledge production, an important dimension for examining the ways in which postsocialist China transpires in global political economy in the age of Big Data analytics. First, I show how Chinese commercial surveillance analytics profits from legitimation lent by the West-centric hierarchical academe. Then, I move to transnational academic repurposing of Big Data from China, which becomes increasingly common. Such social research tends to yield specters of China that are untethered to the lived realities of those whose data are taken. Drawing on decolonial thinking and feminist care ethics, this commentary concludes by urging social scientists to "stay with the trouble," making China "legible" in their computing of Chinese Big Data.

### Introduction

On February 6 after 9 p.m., the word about Li Wenliang's critical condition gathered attention from various parties and the online information situation was heating up. At 3:48 a.m. February 7, Wuhan Central Hospital confirmed via its official Weibo account: Li Wenliang

Wu, Angela Xiao (2020). Chinese Computing and Computing China as Global Knowledge Production. *Catalyst: Feminism, Theory, Technoscience*, 6(2), 1–12.

<http://www.catalystjournal.org> | ISSN: 2380-3312

© Angela Xiao Wu, 2020 | Licensed to the Catalyst Project under a Creative Commons Attribution Non-Commercial No Derivatives license

passed away at 2:58 a.m. Hereby we report on the situation.

Thus begins the red-letterhead “Major Net Intelligence Special Report” by Womin Technology, a company founded in Beijing in 2007. Ambiguously addressed to “Relevant Central Authorities,” the report was submitted mere hours after the death of Dr. Li Wenliang and leaked to the internet hours after. Dr. Li, whose mask-clad visage has since become the face of the pandemic, blew the whistle on a new virus on 30 December 2019. In early January 2020, after being detained and muzzled by local police, he returned to treating patients and contracted the disease, now known as COVID-19.

With registered capital of 132.6 million RMB (US\$19m), Womin was among the first movers in China’s extremely profitable commercial industry of online opinion monitoring. By scraping, aggregating, repackaging, and analyzing online content and activity data, these companies provide analytics and visualizations that purportedly inform decisions ranging from crisis and public relations management tactics to both platform and state-ordered censorship. Industry clientele spreads across public and private sectors. Womin, for example, mainly serves financial sector clients and chooses to offer free intelligence to the government during extraordinary moments. “We have nothing to hide with our work,” Womin’s CEO told the *Washington Post* after the above-quoted report was leaked (Shih, 2020).<sup>1</sup>

In this essay, rather than focusing on the global knowledge production *about* computing in/from China (e.g., Milan & Treré, 2019), I use Womin and its leaked report as a point of departure from which to analyze computing in/from China as an *integral part* of global knowledge production. First, I discuss how Chinese computing such as Womin’s surveillance analytics profits from legitimation lent by the West-centric hierarchical academe. Then, I move to transnational academic computing of Big Data supplied by Chinese surveillance tech, which, I illustrate, often yields specters of China as “portables” that are untethered to the lived realities of those whose data are taken.

As the capitalist expansion of data extraction “makes us all subjects to capital in new distinctive ways,” write Couldry and Mejias (2019), the United States and China, home to the two largest digital economies, have emerged as the world’s dual poles of “colonial power” (pp. 337, 339). Complicating this view, I call for attention to the entanglements of the predominantly anglophone production of world knowledge in the silencing and contortion of Chinese people, who are being subject to increasingly comprehensive politico-cultural governance in state-

managed neoliberalism.<sup>2</sup> My analysis adds a layer to our understanding of how, in the age of Big Data analytics, postsocialist China transpires in global political economy.<sup>3</sup> Drawing on the feminist literature on care, I end by urging social scientists working with Chinese Big Data to be willing to “stay with the trouble” (Haraway, 2016), to bring into their computation a conscious commitment to also unravel the “relationships that set the political, economic, and geopolitical conditions of knowledge-making, world-making, [and] forgetting” (Murphy, 2015, p. 732).

## Computing in China: Academic Legitimation of Surveillance Analytics

Womin’s four-page report presents data analytics in two figures. The first is a line graph that plots over time the numbers of mentions of “Li Wenliang” on China’s social media platforms such as WeChat and Weibo, mobile news apps, online forums, webpages, videos, as well as “foreign social media” and “foreign mainstream media.” These mentions peaked at 11 p.m. on 6 February, supposedly as “the word about Li Wenliang’s critical condition” spread. To illustrate the “heat,” the report uses Weibo as an example: among just the top fifty accounts on the platform, their content featuring Dr. Li’s name had drawn 1.005 billion visits. The report calls this a “screen flooding-level event.” “Screen flooding” (*shuaping*) is often used to describe spamming and the appeal of celebrity scandals and publicity stunts.



Figure 1 Heat of attention (Womin’s leaked report)

The second figure is titled “Computing Emotion, Measuring the Heart of the

People (*renxin*)". Based on sentiment analysis of the supposedly same dataset, it shows the percentage and over-time fluctuation of five types of sentiment: optimism, anger, sadness, disgust, and fear. "Prior to 11 p.m. 6 February, sadness reigned," explains the report, "Afterwards anger increased, which requires urgent attention."

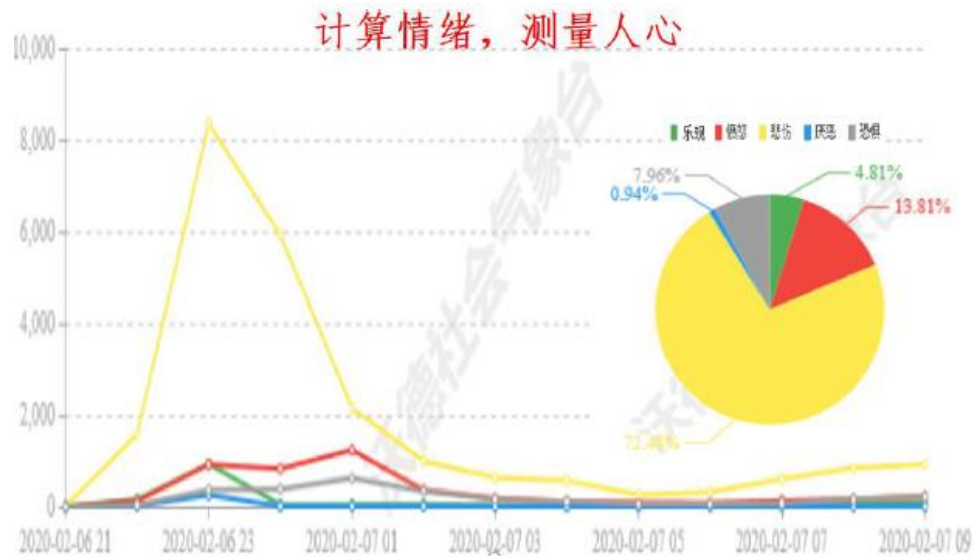


Figure 2 Computing Emotion (Womin's leaked report)

"The event of Li Wenliang's death is still fermenting online," the report continues, recommending the government take measures "in order to dissolve the diffusion of negative emotions, especially anger." The top three measures were (1) to affirm Dr. Li's contribution because doing so "meets the expectation of most internet users"; (2) to promise that relevant parties will be held accountable but delay the main action "till the epidemic is under control"; and (3) to forcefully remove and rectify rumors, countering the malicious intent of foreign entities who fomented them.

As the coronavirus rampaged through the country and people mourned for the deaths and sacrifices of Dr. Li and so many others, Womin's leaked report aroused much ire on the Chinese internet: the company's eagerness to advise on censorship was "shitty." That aside, its analytics, as some had remarked with disbelief, appeared no less shitty. No information on data collection and categorization was given, the unit of analysis unclear, and even the Y axes of the two figures came unspecified. The relations between data results and recommended measures were described as, at best, disjointed.

This judgment, however, missed the point; the computing expertise Womin offers its clients is never academic. Rather, it operates in “spaces between fields” (Eyal, 2013). People Online Research Institute (2018), a subsidiary of the mouthpiece *People’s Daily*, considers the industry an optimal collaboration between government, commercial software development, media, and research and education (“*guan shang mei jiao*”). In online opinion monitoring and other emerging data-driven sectors, computing is marketed through exploiting connections to multiple established institutional fields. To begin, the Chinese government plays a formative role. The government buys, sells, and invests in online expression monitoring, an industry that dovetails nicely with the dual goals of postsocialist China’s AI policies: economic development and political stability (Chen, 2019).<sup>4</sup> While all levels of government and state-owned enterprises procure its products, state media set up offshoot monitoring firms that use inherited political authority to gain a competitive edge. Womin’s “free intelligence” report represents a typical strategy to nurture relationships with government agencies.

But when it comes to “Chinese digital surveillance,” it is necessary to clarify that the authoritarian state is not the sole force at play. Often overshadowed is the legitimation work that academic actors perform for these surveillance tech companies. When expertise, such as online opinion analytics, is hard to assess with clear and open standards, academic credentials and liaisons may serve as key indicators of competence. This is all too clear to Womin and its thousands of rivals. In the leaked report, in sharp contrast to the hurried, cryptic analytics is a two-and-a-half-page introduction on the company itself. “Womin Technology,” it begins, “was founded by multiple experienced professionals in the disciplines of philosophy, sociology, computer science, law, and communication.” What follows are bios of the seven key figures behind the 150-person firm. Two stand out for their connections to state media and the military. Four teach at top universities, including a prominent Princeton sociologist who serves as Womin’s chief academic advisor. Even those not in academia obtained graduate degrees from top Chinese and Hong Kong universities. Also touted are collaborations with various research institutes and publications in academic journals and conferences, with international ones particularly highlighted. Importantly, whereas the Chinese government and media organizations are bound by the sovereign territory, its academics operate as part of a hierarchical transnational network. Working closely with Chinese universities, the online opinion monitoring industry

especially craves linkages to (Euro-)American academia, handing out not only money, but also data, in exchange for domestic legitimation.

## Academic Computing of Chinese Big Data: China as Portables

China became a land of massive datasets largely due to its postsocialist conditions, which include the preponderance of monopolistic platforms in close collaboration with the state, laxer market regulation and labor protection, the absence of public organizing about data privacy (as with everything else), and an elephantine digital world thanks to long-term government infrastructural investments. While Chinese Big Data have been stealthily moving across borders to fuel Silicon Valley-centric innovations by feeding the neural networks of startups, here I focus on their destination within academic computing, which distinguishes itself by being open, authoritative, cumulative, and (nominally) non-profit driven.

In an age when online behavioral trace data support a growing academic ecosystem, access to such data from China has become an enviable asset for scholarly publication.<sup>5</sup> It comes as no surprise that China's online opinion monitoring industry is a major supplier, touting data volumes easily in the billions, like those in Womin's report. This industry-academia collaboration poses grueling ethical and political questions, but I draw attention to what seems to be an epistemological transfiguration that stems from this data appropriation: How is China known through such data? How does China, so computed, fit in the elaborate dispersed knowledge machine that is the transnational field of social sciences?

China has long been notorious for lacking reliable quantitative data, be they public opinion surveys or economic statistics. In her overview about the unprecedented explosion of digital trace data from China, political scientist Daniela Stockmann (2018) hopes that, combining data science with China-specific knowledge, scholarship will "visualize the complex and dynamic nature of Chinese society and...allow this vision to better travel beyond the China field" (p. 407). While concurring with her advocacy, I am less optimistic given the institutional circumstances that determine actual practice.

Inquiries of non-Western societies have always occupied a marginal corner in disciplinary social sciences, where constructing generalizing theory takes primacy over engaging empirical particulars.<sup>6</sup> Their marginality is only worsening. In

American universities, government funds and private philanthropy, once a lifeline for developing “indigenous knowledge” of foreign territories, has been dwindling since the end of the Cold War. As jobs become fast-paced and competitive, in these disciplines acquiring area expertise is discouraged by graduate mentors (Stevens et al., 2018). Adding to this is the embrace of STEM talent thanks to the recent trend of “computational social science.” In short, in the US-dominated social sciences, the influx of Chinese Big Data meets the continued dilution of China expertise. Meanwhile, Big Data social research within China, as Stockmann (2018) also notes, steers clear of “politically sensitive topics” and concentrates on narrow “application” problems.

Against this backdrop, Chinese online opinion monitoring companies, whose business is to intervene in the Chinese web ecology, also step in to supply its data for academic research transnationally.<sup>7</sup> As their provision services get streamlined and standardized, the presumptions and tinkering intrinsic to their data mining and processing are further veiled. Sentiment analysis deployed by Womin (Figure 2), for example, draws on opaque analytics to label “emotion,” which becomes ready-made data points. Notably, as the most used Chinese Big Data, platform trace data and data scraped from other venues usually fall in the confines of the online behavior of individual user accounts or the semantic content of sprawling, anchorless textual fields. Lacking direct measures for gender, income, ethnicity, citizenship (*hukou*), these data, despite their abundance, are “unfriendly” to the investigation of systemic inequality; they also tend to obscure organizational practices of content production and censorship at work.

In its extreme forms, social science computing of such data not only flattens out the intimate knowledge of those whose traces are captured and the historical and institutional conditions they live through. It also actively occludes Chinese heterogeneities in attempts to justify these data as making up proxies for generic social constructs (e.g., “public opinion” and “social capital,” typically emerging from Euro-American contexts).<sup>8</sup> It is not difficult to see that research both in- and outside China has this tendency; both employ Chinese Big Data to expand the intelligible, portable plane of abstracted human behavior. As such, the new spectral China begins frequenting conferences, journals, and PowerPoint slides. While critical scholarship may readily call out the Cold War ideology behind popular portrayals of the Chinese state as monolithically totalitarian or its people as unthinking masses, to confront this new China requires different vocabulary and analytical lenses.

To make sense of the multilayered extraction (i.e., legitimating surveillance tech and using surveillance data) at work, I believe decolonial thinking offers a way out. This direction faces at least two challenges. First, China studies, which consolidated during the Cold War, had undertaken active expurgation of colonialism (and in turn decolonial analytics) to fit America's larger geopolitical project: China was never really colonized; it belonged to a guarded sphere (Barlow, 1997). Second and more substantially, academic computing of Chinese Big Data largely falls outside of what scholars call the "longue durée of datafication" that is underwritten by the characterization and management of the territory-specific Other (Raval, 2019). Indeed, there may be such a continuum between some current data science initiatives on American campuses and the data-intensive projects dating back to philological studies formulated in Cold War-era Area Studies, or further back to those integral to colonial governance (Asif, 2019). But the employment of Chinese Big Data by disciplinary social sciences at a broader scale is of a different nature. Stemming from a confluence of transnational forces and its own institutional history, it leads to a spectral China voided of sociocultural tapestry, or its "Chinese-ness."

To decolonize this assembly of knowledge production, I argue, presupposes a commitment to make China truly legible in Big Data; that is, making legible in data analytics what it means to be a person—being tracked on the Chinese Internet—in Chinese society at this time. Adopting a feminist perspective of care in the production of scholarship helps unsettle the calcifying facade of China as a data source. Approaching China as "matters of care" (de la Bellacasa, 2011) is more than acquiring "indigenous knowledge" or "area expertise." Care as affectively charged attentiveness can be embroiled in agendas of orientalism, colonialism, neoliberalism, and right-wing politics (Hobart & Kneese, 2020; Murphy, 2015). What distinguishes the feminist ethics of care is its emphasis on open-endedness, responsiveness, and, crucially, the reflexivity on power differentials (Martin et al., 2015).

Viewed in this lens, social scientists repurposing Chinese Big Data must stay accountable to the politics and privilege involved in their work, which necessitates that they restore and confront the contexts of datafication. Further, Chinese people are distributed across variegated situations of vulnerabilities by local political, economic, and ideological structures—from China's widening wealth gap, its rural-urban divide instituting manifold exploitation, its resurging patriarchy sanctioned by the government, to the operations of Chinese media system and various state apparatuses, corporate capital, platform technologies.

Using Chinese Big Data, analyses should aspire to reveal circumstances, murmurs, and life chances that these interlocking forms of subjection intend to obscure. But to approximate this goal, social scientists must be willing to step outside these data, painstakingly weave together contextual knowledge, and also strive to develop a standpoint reflective of Chinese cultural, social, and political realities (see Harding, 2005). This entails “a politics of reckoning with a world already violated” (Murphy, 2015, p. 732). Only then can quantitative abstraction, while inevitably flattening lived experience, be wielded to challenge power (also see D’Ignazio & Klein, 2020).

Under his final Weibo post, the living memorial of Dr. Li Wenliang is still unfolding. As of early June 2020, fellow Chinese left well over a million posts to celebrate his life and reflect on their own quotidian experience. The volume surely is immense, but all may just vanish in the milliseconds of computing.

## Acknowledgements

This essay draws from a larger project supported by an Early Career Fellowship from the Henry Luce Foundation/ACLS Program in China Studies. I am grateful to Sareeta Amrute, Julie Yujie Chen, as well as two anonymous reviewers and the editors of *Catalyst* for their help in improving this essay.

## Notes

<sup>1</sup> Womin’s leaked report and a partial translation by Josh Rudolph can be found at <https://chinadigitaltimes.net/2020/02/translation-recommendations-for-handling-netizen-reaction-to-li-wenliangs-death/>

<sup>2</sup> Beltrán (this issue) also discusses the tendency of STS scholarship to privilege technoscientific practice over geopolitics, histories, and economic structures, running the risk of “leaving out particular differences in the name of connecting across others” to build macro-scale theoretical frameworks.

<sup>3</sup> See McElroy (this issue) for Romania’s different postsocialist pathway, where the quest of “becoming Western” leads to full-sail embrace of Siliconization.

<sup>4</sup> In comparison, US AI policies focus on national and economic security.

<sup>5</sup> See Okune’s (this issue) study of developing qualitative, ethnographic data infrastructures in Kenya. While constant interviews make certain Kenyan populations “feel over-researched,” due to the “unobtrusive” nature of data

mining, in China the immense production lines for behavioral trace data eschew people's awareness.

<sup>6</sup> Disciplinary social sciences refer to those dominated by quantitative methods, such as economics, communication, information science, political science, and sociology.

<sup>7</sup> In addition to social science computing Chinese trace data, it is worth noting that its appropriation of platform-generated trace data begs questions about epistemic assimilation in service of platform power. Similar concerns arise in regard to appropriating data captured by essentially administrative enterprises as such Womin. See Wu & Taneja, 2020.

<sup>8</sup> I refrain from citing specific studies and deliberately keep the description generic to invite readers to reflect on the literature in breadth.

## References

Asif, M. A. (2019, November 20). Technologies of power: From area studies to data sciences. *Spheres: Journal for Digital Cultures*, 5. <https://spheres-journal.org/contribution/technologies-of-power-from-area-studies-to-data-sciences/>

Barlow, T. E. (1997). Colonialism's career in postwar China studies. In T. E. Barlow (Ed.), *Formations of colonial modernity in East Asia* (pp. 373–411). Duke University Press.

Beltrán, H. (2020). Hacking difference: Coding expertise and belonging from the South. *Catalyst: Feminism, Theory, Technoscience*, 6(2).

Chen, W. (2019). Now I know my ABCs: U.S.-China policy on AI, Big Data, and cloud computing. East-West Center. <https://www.eastwestcenter.org/publications/now-i-know-my-abcs-us-china-policy-ai-big-data-and-cloud-computing>

Couldry, N., & Mejias, U. A. (2019). Data colonialism: Rethinking big data's relation to the contemporary subject. *Television & New Media*, 20(4), 336–349.

de la Bellacasa, M. P. (2011). Matters of care in technoscience: Assembling neglected things. *Social Studies of Science*, 41(1), 85–106.

D'Ignazio, C., & Klein, L. F. (2020). *Data feminism*. MIT Press.

Eyal, G. (2013). Spaces between fields. In P. S. Gorski (Ed.), *Bourdieu and historical analysis* (pp. 158–182). Duke University Press.

Haraway, D. J. (2016). *Staying with the trouble: Making kin in the Chthulucene*. Duke

University Press.

Harding, S. (2005). Negotiating with the positivist legacy: New social justice movements and a standpoint politics of method. In G. Steinmetz (Ed.), *The politics of method in the human sciences: Positivism and its epistemological others* (pp. 346–365). Duke University Press.

Hobart, H. J. K., & Kneese, T. (2020). Radical care. *Social Text*, 38(1), 1–16.

Martin, A., Myers, N., & Viseu, A. (2015). The politics of care in technoscience. *Social Studies of Science*, 45(5), 625–641.

McElroy, E. (2020). Corrupting Techno-normativity in postsocialist Romania. *Catalyst: Feminism, Theory, Technoscience*, 6(2).

Milan, S., & Treré, E. (2019). Big data from the South(s): Beyond data universalism. *Television & New Media*, 20(4), 319–335.

Murphy, M. (2015). Unsettling care: Troubling transnational itineraries of care in feminist health practices. *Social Studies of Science*, 45(5), 717–737.

Okune, A. (2020). Open ethnographic archives as collaborative ethnography. *Catalyst: Feminism, Theory, Technoscience*, 6(2).

People Online Research Institute. (2018, January 24). On the state and prospect of our online opinion industry. *People Online*.  
<http://media.people.com.cn/n1/2018/0124/c416772-29784789.html>

Raval, N. (2019, November 20). An agenda for decolonizing data science. *Spheres: Journal for Digital Cultures*, 5. <https://spheres-journal.org/contribution/an-agenda-for-decolonizing-data-science/>

Shih, G. (2020, February 8). As anger explodes over doctor's death, Beijing firm gives Chinese government advice. *The Washington Post*.  
[https://www.washingtonpost.com/world/as-anger-explodes-over-doctors-death-beijing-firm-gives-chinese-government-advice/2020/02/08/6899e3e6-4a52-11ea-9164-d3154ad8a5cd\\_story.html](https://www.washingtonpost.com/world/as-anger-explodes-over-doctors-death-beijing-firm-gives-chinese-government-advice/2020/02/08/6899e3e6-4a52-11ea-9164-d3154ad8a5cd_story.html)

Stevens, M., Miller-Idriss, C., & Shami, S. (2018). *Seeing the world: How US universities make knowledge in a global era*. Princeton University Press.

Stockmann, D. (2018). Toward area-smart data science: Critical questions for working with big data from China. *Policy & Internet*, 10(4), 393–414.

Wu, A. X., & Taneja, H. (2020). Platform enclosure of human behavior and its measurement: Using behavioral trace data against platform episteme. *New Media & Society*. OnlineFirst. <https://doi.org/10.1177/1461444820933547>

## Author Bio

**Angela Xiao Wu** is an assistant professor in the Department of Media, Culture, and Communication at New York University. Her past research has examined cultures and practices of data analytics, media use and media (infra)structures, and Chinese information technology. She is currently working on a book that explores how public culture takes shape when systems thinking informs its conception and governance.