Commentaries for Special Forum

Algorithmic ethnography, during and after COVID-19

Angèle Christin

Stanford University, USA; Institut d'Etudes Avancées and Sorbonne Université, France

Abstract

Social scientists are increasingly turning to digital interactions as a primary source of qualitative data. Online activities in turn typically take place on algorithmically mediated platforms, which shape what people do and say in crucial ways. Here, I offer a toolkit for what I call algorithmic ethnography, that is, the ethnographic study of how computational systems structure online activities. First, scholars need to follow the data and take into consideration the tracking strategies, monetization systems, and business models of the platforms where online interactions unfold. Second, ethnographers should focus on the details of algorithmic sorting, since platforms typically have more content than they can display and thus rely on algorithmic procedures to personalize their pages. Third, ethnographers should include metrics in their fieldwork and study their effects on interactions, hierarchies, and representations. Together, these angles afford a fine-grained understanding of the computational texture of online exchanges.

Keywords

Algorithms, data, ethnography, metrics, platforms

Following the closing of offices, schools, borders, restaurants, bars, stadiums, museums, and other social institutions to curb the spread of COVID-19 in 2020, social scientists have begun to reflect on what these changes meant for the future of research. In the words of Social Science Research Council (SSRC) President Alondra Nelson (2020),

> how do we *do* social research at a time when, for the foreseeable future, borders are closing, global cooperation is yielding to widespread mistrust, and necessary public health accommodations such as "social distancing" create hurdles for both human connection and research?

Qualitative scholars are particularly affected by these changes. What does it mean to do fieldwork if there is no field to work on? Facing this difficult question, qualitative researchers have begun to think about possible ways to continue doing research. Through online seminars and panels—with titles ranging from "Research, Interrupted" to "What—the Field?"—scholars are articulating new methods for ethnographic and qualitative research. Unsurprisingly, these methods entail identifying, gathering, and analyzing digital data.

At a time when most of our social, professional, and intimate interactions take place online, such a digital turn makes perfect sense. Yet it is important to acknowledge that such a turn is also far from new. As

Corresponding author:
Angèle Christin, Department of Communication, Stanford University, Stanford, CA 94305, USA.
Email: angelec@stanford.edu

anthropologist Tom Boellstorff argued in 2008, "humans have always been virtual" (Boellstorff, 2008; Proctor, 2020). Virtual and digital ethnography is now an established tradition, with classics and a research canon—admittedly an interdisciplinary and flexible one (Boellstorff et al., 2012; Coleman, 2014; Glatt, 2020; Hine, 2015; Knox & Nafus, 2018; Mattern, 2020; Murthy, 2008). In some cases, scholars fully embrace the virtual medium and conduct all their fieldwork online. In other cases, ethnographers recommend gathering digital data in addition to in-person observations and interviews. Regardless, digital ethnographers pay close attention to the same handful of questions: What is the texture of online interactions? How do software and platforms shape exchanges and representations? What is the position and role of the ethnographer in online communities?

Going a step further, I would argue that qualitative scholars analyzing digital data also need to engage in what I call *algorithmic ethnography*, that is, the ethnographic study of the computational systems enabling and shaping online interactions (Christin, 2020b). Adopting the lens of algorithmic ethnography entails paying close attention to the role of algorithms in structuring the back and front end of the digital platforms that increasingly mediate digital exchanges. The concept of "platform" is notoriously slippery (Gillespie, 2010; van Dijck, Poell, & de Waal, 2018) but broadly encompasses all the digital infrastructures that function as technological, economic, and cultural intermediaries between online users, including social media platforms (Facebook, Twitter, YouTube, Instagram, TikTok), e-commerce and online labor markets (Amazon, Upwork, Uber), streaming platforms (Netflix, Hulu, Spotify, Apple Music, Amazon Prime Video), as well as communication and collaboration platforms (Zoom, Skype, Slack, Teams). Specifically, ethnographers analyzing digital interactions need to consider three interconnected questions: the kind of data being collected about online users, the role of sorting algorithms in ordering the content users see on the platform, and the effects of metrics on people's interactions and representations. Here, I briefly discuss each of these points and their implications.

First, ethnographers should *follow the data*. Digital platforms are not neutral intermediaries:

They are for-profit companies that make money by monetizing the data provided by online users. Platforms typically draw on one or several of the following business models: free access and targeted behavioral advertising (e.g. "pay with your data" through online advertising), subscription models (where people pay to use from the platform, per unit or per period), and percentage-based models (where the platform takes a percentage or royalty on every sale). Depending on their business model, platforms will collect different kinds of data about their users; they will also embed specific "nudges" in their design in order to maximize behaviors that are profitable to them. For instance, if the product is free but users' eyeballs are monetized through targeted ads, platforms will primarily focus on and optimize online engagement (time spent, content provided, etc.). Furthermore, depending on their position and audience, they will either seek to expand absolute numbers of users or focus on attracting "valuable" segments (with specific socioeconomic characteristics) that are worth more to advertisers. If platforms primarily rely on subscriptions, they will nudge users toward buying more expensive plans, tracking their engagement in order to offer the nudge at the right moment. In contrast, when the business model of platforms is primarily based on percentages, their incentive is to maximize the total volume of sales, thus investing primarily in recommendation algorithms that will nudge users toward exchanging more. Such business models—and the tracking apparatuses that go with them—profoundly shape the architecture and design of digital platforms, which in turn provide the built environment within which online interactions take place (Wu & Taneja, 2020). As such, they need to be taken seriously by digital ethnographers.

Second, ethnographers need to *pay attention to the details of algorithmic sorting*. All platforms host much more content than they can display. Consequently, they carefully select what they will show to users at any given moment. This selection process operates through the features of each platform's "feed" and homepage, as well as through recommendation and sorting algorithms that automatically identify and display relevant content to individual users, based on platform-level rules but

also personalized items, including past behavior. Online users in turn internalize such sorting mechanisms, developing what Taina Bucher (2016) calls "algorithmic imaginaries," or representations about how the algorithm works. Based on these algorithmic imaginaries, and in order to be visible to other users, people adapt their behavior on the platform. In other words, the architecture of digital platforms has performative effects. Such effects in turn need to be part of what digital ethnographers analyze in order to do justice to the material they gather online.

Third, digital ethnographers should *consider metrics in their fieldwork*. In addition to data tracking and algorithmic sorting, metrics play a key role on most platforms. Be it rankings, time engaged, stars, likes, hearts, smiling emojis, and reviews, platforms rely on an arsenal of metrics in order to entice and assess user engagement—and, in the process, to feed the sorting algorithms mentioned above. Not only are metrics often visibly displayed, they also mirror and reinforce existing hierarchies and status orders. Drawing on Erving Goffman's work, Alice Marwick and danah boyd (2011) wrote that people experience a form of "context collapse" on social media platforms: They cannot be sure of who (family, friends, acquaintances, "fans," etc.) will view the content they post online. Consequently, for platform users, metrics and comments are essential indicators signaling the feelings and feedback of their heterogeneous audiences. Through metrics, groups can confirm their identity and "groupness," create and maintain symbolic boundaries with outsiders, and signal fine-grained hierarchies within collectives. Hence, digital ethnographers need to incorporate these metrics in their fieldwork (Christin, 2020a). This entails keeping track of likes and stars through screenshots; scrolling down the comments, even if they become offensive; and asking people how they interpret metrics, and why, in follow-up interviews.

The history of ethnographic classics is replete with failed travels, missed encounters, unplanned failures, and unforeseen refusals. Such failures are part and parcel of the research process. In seeking to adapt their project to circumstances that are not (only) of their making, ethnographers have to find ways to turn unfortunate constraints into promising avenues for new research. This essay provides a few suggestions for turning the constraints of confinement into opportunities for analyzing our rich online social lives, during and after COVID-19.

References

Boellstorff, T. (2008). *Coming of age in second life: An anthropologist explores the virtually human*. Princeton University Press.

Boellstorff, T., Nardi, B., Pearce, C., & Taylor, T. L. (2012). *Ethnography and virtual worlds: A handbook of method*. Princeton University Press.

Bucher, T. (2016). The algorithmic imaginary: Exploring the ordinary affects of Facebook algorithms. *Information, Communication & Society*, *20*(1), 30–44.

Christin, A. (2020a). *Metrics at work: Journalism and the contested meaning of algorithms*. Princeton University Press.

Christin, A. (2020b). The ethnographer and the algorithm: Beyond the black box. *Theory & Society*. Advance online publication. https://doi.org/10.1007/s11186-020-09411-3

Coleman, E. G. (2014). *Hacker, hoaxer, whistleblower, spy. The many faces of anonymous*. Verso.

Gillespie, T. (2010). The politics of "Platforms.*" New Media & Society*, *12*(3), 347–364.

Glatt, Z. (2020). *LSE digital ethnography collective reading list*. https://zoeglatt.com/wp-content/uploads/2020/04/LSE-Digital-Ethnography-Collective-Reading-List-April-2020.pdf

Hine, C. (2015). *Ethnography for the Internet*. Bloomsbury.

Knox, H., & Nafus, D. (Eds.) (2018). *Ethnography for a data-saturated world*. Manchester University Press.

Marwick, A. E, & boyd, d. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society*, *13*(1), 114–133.

Mattern, S. (2020). Toolkit: Digital ethnography. *Designing Methodologies*. http://www.wordsinspace.net/designingmethods/spring2018/2018/01/07/toolkit-digital-ethnography/

Murthy, D. (2008). Digital ethnography: An examination of the use of new technologies for social research. *Sociology*, *42*, 837–855.

Nelson, A. (2020, April 23). Society after pandemic. *Items*. https://items.ssrc.org/covid-19-and-the-social-sciences/society-after-pandemic/

Proctor, D. (2020, March 25). So you want to "Do" digital ethnography? *The Geek Anthropologist*. https://thegeekanthropologist.com/2020/03/25/so-you-want-to-do-digital-ethnography/

van Dijck, J., Poell, T., & de Waal, M. (2018). *The platform society: Public values in a connected world*. Oxford Scholarship Online.

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

Author biography

Angèle Christin is an assistant professor in the Department of Communication at Stanford University. She is the author of *Metrics at Work: Journalism and the Contested Meaning of Algorithms* (Princeton University Press, 2020). She is currently a Visiting Researcher with the Social Media Collective at Microsoft Research New England.