CULTURAL ANTHROPOLOGY

Openings and Retrospectives

EXPLORATION OR ALGORITHM? The Undone Science Before the Algorithms

DAWN NAFUS
Intel Corporation
http://orcid.org/0000-0002-2976-6662

The societal harms that many algorithmic systems have caused—and could cause in the future—are widely known to both scholars and the wider public (see Gray et al. 2016; Irani 2015; Pasquale 2015; Barocas and Selbst 2016; O'Neil 2016). In 2017, there was a veritable bumper crop of ethics statements outlining broad principles for ethical algorithm development, put in place to mitigate those harms. They came from engineering professional associations like the Institute of Electrical and Electronics Engineers, university-based institutions such as New York University's AI Now, and companies including my own. Many of these statements argued that the people most affected by algorithms' design ought to have a seat at the table in their making. While I could not agree more heartily with this view, and indeed work in an industrial R&D lab to ensure that this happens, the claim raises a further, much trickier question. What new knowledge is *not* created precisely because algorithms inspire preoccupations with automation? What roads remain not taken, even with a people's seat at the table?

CULTURAL ANTHROPOLOGY, Vol. 33, Issue 3, pp. 368–374, ISSN 0886-7356, online ISSN 1548-1360. *DOI: 10.14506/ca33.3.03*

Algorithms are difficult to untangle from automation in a very particular sense. Well before we get to questions of job loss through automation, algorithms have to be integrated into a deeper system of sensors, processors, communication infrastructures, software layers, and so forth. They don't just run on any old machine. The skills required to integrate an algorithm into a system (or optimize a system around an algorithm) are elaborate, and the task is full of uncertainties. Building the right team and assembling or creating the right set of tools is a remarkable undertaking in and of itself. Systems, not just algorithms, distinguish between cats and things that are not cats, or between a sleep disturbance and awakening, and these are expensive to develop and maintain. Such systems do not make sense to build unless these distinctions need to be made over and over again, by a machine. This particular aspect of automation means that the cost of getting it wrong—of creating a parsing with little economic or social value—can be huge. While anthropologists might sometimes see the world of technology as fast-changing, the reality is that machine execution relies on unwieldy, often brittle infrastructures that are anything but (see also Taylor 2016).

The economics of these systems incentivize certainty in design, which means that the road to algorithm development is littered with many alternative ways of parsing data that were not taken up, but that in fact could have even more immediate human value than any eventual algorithm. I will share two examples from my own work—work that did not end in algorithm development but very well could have—to explain what I mean. My first example comes from a collaboration with the Atlas of Caregiving project, a nonprofit organization that uses methods developed in the Quantified Self movement to teach family caregivers how to use data for critically reflecting on their situation. These techniques are very much a "data dialect" (Churchill 2017) in that they are not techniques anyone will teach you in a standard data science class, but are born out of everyday experiences of people figuring out which ways of collecting and parsing data best support critical reflection.

In the pilot study (Mehta and Nafus 2016), the value of data to the participant was unpredictable. Some got no value at all, other than the satisfaction of sharing their story with others. Others got practical value. For example, one participant saw her sleep data and decided that of all the complexities and difficulties of caregiving, her lack of sleep (due to the need to support her care recipient at night) constituted the biggest problem. Looking at her data helped her decide it was time to hire someone for nighttime care. Another used a graph of ambient noise levels to solidify her sense of how badly nearby construction affected her mother, an Alzheimer's sufferer. Others still found meaning in their "care map," a research technique where we asked people to draw a social network diagram of who takes care of whom in their family. The social network data was not news—they themselves had drawn it. Yet of all the fancy sensing gadgets and data visualization techniques we tried, this form of manual computation—the adding together of all the people involved and the time spent—had the most profound effect of all, as it yielded a visual form that enabled participants to deepen their appreciation for the ways that others contributed. We encountered one caregiver who found numbers and technology hugely intimidating, but who nevertheless sat down with us to go through it all and in the process started formulating opinions about how data should be cleaned. So much for data science as an esoteric skill only relevant to the privileged.

The second example comes from a collaboration with Bay Area environmental justice groups and the Drexel University–based Fair Tech Collective, which explored how air-quality data could be made more usable for residents of a high-pollution area. High-grade air-quality monitors that sensed the air in real time had long been operating, but few people were using the data. We wondered if wearable technology, which sensed bodies at roughly the same rate, would give the data a different valence: more meaning, perhaps more urgency. Is it possible, we asked, to find a time-based relationship between air pollution and cardiovascular health? If so, what would it look like for residents to meaningfully participate in the research design and analysis? Would the results change residents' ability to advocate for change, or would it increase their burdens in some way?

The pilot study had only ten participants, but we nevertheless ended up with 1.2 million data points. This dataset was modest in size by computer science standards, but it created a lot of potential pathways. We sat down with participating individuals to have a look at the data, and we found pathways that we could not have from afar: moments when travel out of the high-pollution area appeared to improve blood oxygen, or moments when questions emerged about why heart rate seemed to spike at the same time of day—was it a pattern in pollution or an infuriating television show? Having one data scientist decide which is the signal and which is the noise would not do, but working out how to do it otherwise was itself an experiment in how social relations can be formed around a dataset this complex. All of this involved putting together a fairly significant infrastructure and set of work practices for accessing data and exploring it with participants. In a sense we built, on a small scale, a setup analogous to what my engineering colleagues do when they build multimodal sensing systems, but with more interpreters situated closely to the data at hand and fewer algorithms.

In both of these examples, algorithm design could have been the end goal. We could have sought an algorithm that took signals identified in the Atlas of Caregiving pilot and inferred situations where stress was most likely to build up. Key stakeholders of that project indeed advocated that we move in that direction. Perhaps an algorithmic system might consistently point out the areas when health and social systems fail caregivers, thus generating stress, as opposed to making caregivers individually responsible for their own stress relief every time that stress is inferred. We could also speculate about an algorithm that learns how to define and identify, in high resolution, the extent of cumulative cardiovascular health effects that air pollution causes. In the right hands, such a tool could make a powerful case for public health action. Those are, in my view, reasonably promising scenarios for algorithm design. I will not venture a guess as to how feasible they are, but I do want to point out the value that gets lost once we race toward them.

In these projects, there was a good deal of undone science that got done precisely because the goal was not specifically to end in algorithm design. Clues about the sources of stress or illness were surfaced. In science and technology studies, the concept of undone science points to the choices made about which research questions are asked and which go underinvestigated, such as the many unasked questions in environmental health (Frickel et al. 2010). Even if the new knowledge we were creating was social and cultural, not necessarily scientific, this notion encourages us to think about how critique can take the form of knowledge production that opens up or elaborates a particular line of inquiry, rather than simply identifying problems with current technical systems. Both these projects pointed to undone science that needed doing. They surfaced alternative lines of inquiry by appropriating datasets that were originally designed to fit very different categories and by giving the subjects of that data the opportunity to reframe and reconsider its meaning.

The undone science will not be done, however, if fairness advocates limit their sense of inclusion to matters of algorithm design. By the time anyone is at the stage of algorithm design, the data that an algorithm would work on is usually well understood, the categories to which they are believed to refer have been narrowed, and investments in other aspects of the system have already been made. Assumptions might be flawed, but there is some consensus about the purpose. When inclusion is limited to algorithm design, important adjustments around the edges can be made. A self-driving car can be made to recognize objects on roads different from the roads found in Palo Alto, and facial recognition software can be made to recognize faces other than the white male ones on which they were first tested. However, the more fundamental question of what the data signals, and what human value those signals have, can only be asked in the most limited ways.

The recent calls for broader participation in algorithm development, then, are not wrong as a first attempt to better situate algorithms in more equitable social relations. They are all but toothless, however, without a rejoinder. There must also be a call for broader participation in data *explorations* that do not necessarily end in algorithm development directly. Such explorations could shape the consensus formation that happens before algorithm development is even in the cards and could inform new technical directions as they emerge. Such explorations should also be thought of as socially useful end goals in and of themselves. In both of my examples, a reflexive, participatory approach to data opened up areas of direct value to the participants themselves. Unexpected forms of value became available precisely because we had lingered in the data together and were not narrowly thinking about data as the thing that was on its way to training an algorithm to categorize something. These were not forms of value that required automation: no one needs to be told twice that they really ought to hire someone to assist with night care. Value emerged instead through human interaction supported by a technical infrastructure, one that is fundamentally not the stuff of automation.

Patterns in data that are numerically calculable and deeply valuable to the very people who generated them are often left behind because they require slowness—a slowness not attributable to the fragility and complexity of infrastructures necessary for automation, but rooted in the intellectual patience necessary for real meaning to emerge through lived, rather than imagined, human experience. Exploration cannot be short-circuited by speculating on data's meaning in a conference room somewhere, even with "domain experts" (as nontechnical people are called in data science circles) present. The instances of travel that brought oxygen into someone's bloodstream, or the construction sounds that distressed an Alzheimer's sufferer and her daughter, will never be found in an armchair version of data science. They are found in dialogue. Such dialogues do require meaningful investment in labor and infrastructure, but of a very different kind.

It is possible to ask about the extent to which everyday life should or should not be algorithmically encoded while also pointing out those places where radically different technical directions (and distribution of social and technical resources) would prove beneficial. Alternative paths might end in a different kind of algorithmic system, or they might end in the production of new knowledge not necessary to repeat. Good data scientists, of course, explore how a dataset came to be in the first place before attempting to write algorithms for it. Anthropology can take this practice further, by doing that in a skillful way *with* the people to whom it refers. This is a seemingly obvious move that often surprises the data scientists I meet, but it has also frequently been welcome.

While I have emphasized the practical and societal value of tracing out routes not elaborated in large-scale technical systems, I also see scholarly value in doing things in this way. These instances of engagement with data and with people constitute rich ethnographic moments. The prospect of proceeding along these lines raises a further question: what would it be like to have our own vernaculars in data—our own data dialects—attuned to the theoretical and methodological commitments that many anthropologists share? Indeed, others have started to raise this issue (Knox and Nafus 2018), and so perhaps the question is not entirely speculative. It is nevertheless one that is worth our consideration. I remain curious about the social worlds that could be made when anthropologists and the people we study go beyond the fraught politics of getting a seat at the table—whether that seat is taken through public critique or internal advocacy—and start taking a crack at building those worlds ourselves.

REFERENCES

Barocas, Solon, and Andrew D. Selbst
2016 "Big Data's Disparate Impact." *California Law Review* 104, no. 3: 671–732. https://doi.org/10.15779/Z38BG31.

Churchill, Elizabeth
2017 "The Ethnographic Lens: Perspectives and Opportunities for New Data Dialects." Perspectives, EPIC website, September 26. https://www.epicpeople.org/ethnographic-lens.

Frickel, Scott, Sahra Gibbon, Jeff Howard, Joanna Kempner, Gwen Ottinger, and David J. Hess
2010 "Undone Science: Charting Social Movement and Civil Society Challenges to Research Agenda Setting." *Science, Technology, and Human Values* 35, no. 4: 444–73. https://doi.org/10.1177/0162243909345836.

Gray, Mary L., Siddharth Suri, Syed Shoaib Ali, and Deepti Kulkarni
2016 "The Crowd is a Collaborative Network." In *Proceedings of the Nineteenth ACM Conference on Computer-Supported Cooperative Work and Social Computing*, 134–47. https://doi.org/10.1145/2818048.2819942.

Irani, Lilly
2015 "Justice for 'Data Janitors.'" *Public Books*, January 15. http://www.publicbooks.org/nonfiction/justice-for-data-janitors.

Knox, Hannah, and Dawn Nafus, eds.
2018 *Ethnography for a Data-Saturated World*. Manchester, UK: Manchester University Press.
Mehta, Rajiv, and Dawn Nafus
2016 "Atlas of Caregiving Pilot Study Report." https://atlasofcaregiving.com/core-research/pilot-study.
O'Neil, Cathy
2016 *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown Publishing.
Pasquale, Frank
2015 *The Black Box Society: The Secret Algorithms that Control Money and Information*. Cambridge, Mass.: Harvard University Press.
Taylor, Alex
2016 "Data, (Bio)Sensing, and (Other-)Worldly Stories from the Cycle Routes of London." In *Quantified: Biosensing Technologies in Everyday Life*, edited by Dawn Nafus, 189–209. Cambridge, Mass.: MIT Press.