Public Books

FACIAL RECOGNITION IS ONLY THE BEGINNING

1.27.2020

BY JAKE GOLDENFEIN

"Research in Translation" is a series that showcases the most cutting-edge research in artificial intelligence, or AI, technology.

Does the relationship between power and AI mean that all people will be monitored all the time?

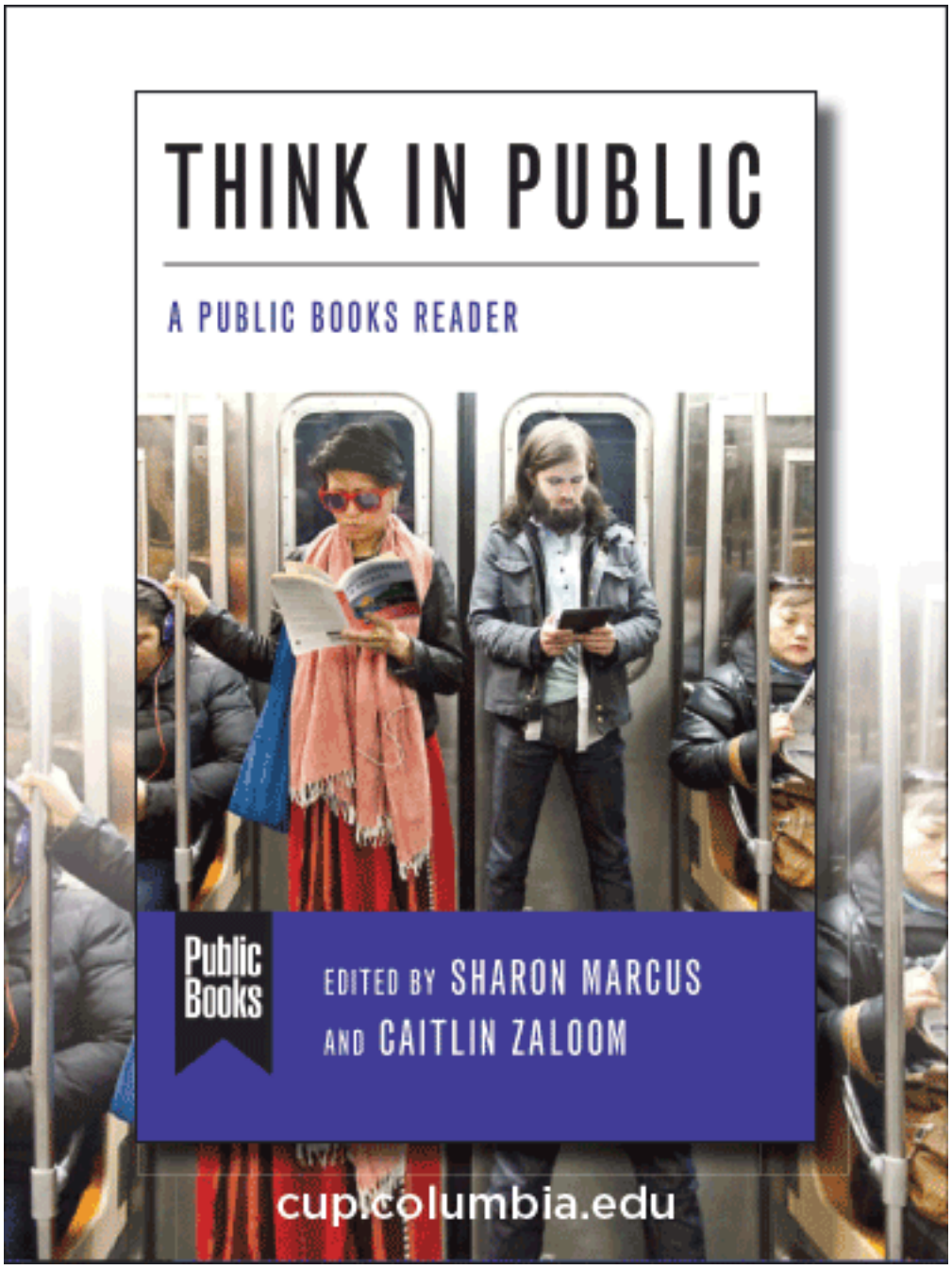
We have become obsessed with automated facial-recognition technology. There is something about the capacity to remotely identify and track a person through space that piques terror within us. But with the continuing development of computer vision and video analytics, facial recognition may ultimately turn out to be one of the more benign applications of camera surveillance. Put another way, facial recognition is just part of a much larger program of surveillance—equipped with higher scale and granularity—associated with cameras, microphones, and the multitude of sensors in modern smart phones.

As these systems become more sophisticated and more ubiquitous, we are likely to experience what a recent ACLU report describes as a "phase transition" from collection-and-storage surveillance to "mass automated real-time monitoring."[1] More importantly, though, as they bring about this change in scale, these tools will also challenge how we understand the world and the people in it.

This degree of surveillance will not emerge simply because the technology is improving. Instead, the technologies that track and evaluate people gain legitimacy because of the influence of the political and commercial entities *interested in making them work*. Those actors have the power to make society *believe* in the value of the technology and the understandings of the world it generates, while at the same time building and selling the systems and apparatuses for that tracking, measurement, and computational analysis.

In other words, beyond any specific attributes of the technology, there is a particular group of actors that benefits from the idea that measuring and computationally analyzing the world affords access to knowledge—and those actors also have the power to enact and

legitimize their knowledge claims. *They* are creating a world where all are monitored and evaluated computationally, because it is in their interests.



THE PROBLEM

The history of technology shows that every new visual technology is believed to offer new access to knowledge and "truth." Photography, for instance, was understood as a way to "see deeply into nature's cabinet."[2] It challenged the "optical unconscious." Walter Benjamin described this access to nature as opening up "the physiognomic aspects of the world of images, which reside in the smallest details, clear and yet hidden enough to have found shelter in daydreams."[3]

Other optical technologies prompted similar narratives. The telescope gave access to celestial knowledge, the microscope to cellular knowledge. X-ray imaging and radiographic measurements of material density afforded a "new sight"[4] that could reveal "hidden existence."[5] For some commentators, the X-ray was the beginning of a new form of "a-visuality," in which sensing technologies "transgress the physical thresholds of an optically available world."[6]

Such historical technologies have shown that our desire to believe that new tools offer new insight is common. But today's situation presents worrisome differences.

When photography rose to prominence, the companies selling film and cameras were not the largest corporations in the world. Today, the companies with the highest commercial valuations in the US are Microsoft, Amazon, and Apple, each with a market capitalization approaching $1 trillion. Like Microsoft, Amazon sells data-science and computer-vision products and services (including to police forces), but it also controls the majority of the data-storage and computing infrastructure in North America. Both companies also work to establish themselves as part of the broader state governance apparatus, with Microsoft to control part of the US Department of Defense data cloud after it won the $10 billion JEDI contract last October (although Amazon is challenging the decision). And while a number of other, smaller companies and start-ups also operate in this space, they typically rely on enterprise-level services and infrastructure provided by the likes of Microsoft (with Azure) and Amazon (with AWS).

While Amazon and Microsoft have become major purveyors of data analytics, companies like Facebook and Google control massive data-collection infrastructures. Together these companies control most of the technological systems responsible for information dissemination and interpersonal exchange.

Some commentators have described the dominance of these companies in terms of "data power,"[7] but their influence could also be characterized in terms of "computational power," "analytics power," and even "epistemological power"—that is, the capacity to validate a knowledge claim or system. Indeed, if the phenomenal advances in artificial intelligence of recent times are the product of massive data collection and the computational power of the cloud, then these private entities control the tools for the production of knowledge—and have an agenda to secure its legitimacy and proliferation.

This weird corporate-state technopolitical alliance also includes the funding of research at universities and other knowledge-generating centers. Together these entities produce power over what it means to know, who is entitled to know, and with what tools. Their outputs are put to work both directly, in surveillance systems and practices, and indirectly, in the policy governing those technologies. That is to say, this same network of interests establishes the parameters of acceptable critique of the technologies it builds and deploys. Indeed, professors and researchers are increasingly funded by these companies, using the money they receive to establish academic data-science schools and institutes at universities, as well as social science and humanities groups that propose ethical and regulatory limitations to the use of these very technologies. These realities must inform the way we begin discussing how to "co-opt AI."

THE CONTEXT

The terrain of the privacy battle is shifting. For a long time, the subject of interest to industry and states alike, as well as to critical scholars, was our interactions with informational environments—sometimes called our "transaction-generated information"[8]—and how that data is collected and processed into predictions and inferences. The combination of machine learning and sensor surveillance, however, extends that tracking regime with real-time measurement and analysis of how you look, sound, and move, and how your body behaves, through a combination of biometrics (measurement of physical and behavioral characteristics), anthropometrics (measurement of body morphology), and physiometrics (measurement of

bodily functions like heart rate, blood pressure, and other physical states).

In this light, the current focus on regulating automated facial recognition, while important and desirable in many respects, might actually be a red herring. For one, if you have a smart phone, you are already identified and subjected to location surveillance via cell towers, GPS, and Bluetooth beacons (like the LinkNYC kiosks around New York City). Second, focusing on facial recognition distracts us from the broader structural question of whether governments or companies should be building massive databases of biometric information in the first place. Third, in the world of machine learning and statistical pattern analysis, you don't need a face to identify a person. Computer-vision systems can also identify individuals through other biometrics, like gait, and not all regulatory proposals (including the ban on facial recognition in San Francisco) deal with that reality. And fourth, these technologies can do much more than simply identify us.

Computer vision and other emerging techniques, like computer listening, afford the capacity to record, computationally analyze, and classify everything about the "real world" that is visible, audible, or otherwise sensible, in real time. This might include our institutional identity—such as our name, image, or unique identifier—but also our clothes and skin color, our age and gender, what we are doing, whom we have been associating with, how we are feeling, how we are likely to behave in the future, and even what type of person we are.

This last technique, sometimes called "personality computing," is being experimented with in multiple domains,[9] including, notoriously, video analytic tools for assessing job candidates. The proliferation of these techniques reflects the belief that sensors and machine learning can give insight into nonsensible traits or characteristics. Beyond "identifying," then, these sensing and machine-learning systems are also capable of "profiling." And we should not be surprised by this. Identity technologies—like automated facial recognition and its precursor, police photography—almost always double as technologies for evaluating people. Historian Jane Caplan reminds us that every technology for verifying identity is inevitably directed at two simultaneous questions: What person is that? And what *type* of person is that?[10]

THE ENTITIES THAT PRODUCE, SELL, AND PROFIT FROM AI TECHNOLOGY HAVE THE CAPACITY TO ESTABLISH AND ENTRENCH THEIR WORLDVIEW AS THE DOMINANT WORLDVIEW.

In the context of "co-opting AI"—and thinking about what it is that AI has been co-opting—this massive increase in the capacity to process and evaluate information captured from physical space is part of a set of particular trends.

First, physical spaces are becoming *cyber-physical spaces*. Computer vision is especially

meaningful for its ability to translate the "real world," and the people within it, into numbers for statistical analysis and automated decision-making. These cyber-physical systems can be constantly and pervasively monitored but, more importantly, computationally interpreted and understood. In the cyber-physical world, the computational sentinels that manage access to public spaces, sporting events, shops, public transport, and whatever else will have a great deal of information on which to base their decisions, including who you are, what kind of person you are, and what you are doing.[11]

Second, as we grow more accustomed to and accepting of those automated decisions, our human visual understanding of the world is supplemented, challenged, and perhaps even negated by computational ways of sensing and knowing. As sensors measure the world and the people within it, techniques of pattern matching and statistical analysis—in particular, deep learning and neural networks—are used to produce classifications, predictions, insights, and knowledge. For those systems, the greater the quantity of data that can be sensed (irrespective of how tangential it may seem), the finer the ability to classify and produce insight. The more data points, the more sensitive the pattern recognition. But this is not an approximation of human vision—it works by clustering data points (i.e., numerical measurements for red, blue, and green values per pixel) into high-dimensional statistical maps that humans cannot navigate.

The dramatic power of these techniques has provoked some (controversial) researchers to argue that our faces and bodies contain more information than can be registered and processed by the human brain.[12] The suggestion is that *only through computation* can the excesses of the physical world, and of the people within it, be perceived, processed, and understood. In other words, there is a belief that data science can know people better than they can know themselves, representing a shift from a logic of "approximation" to a logic of "revelation" in the application of these techniques to people.

Well before computer vision and video analytics became as prominent as they are today, media theorist and philosopher Vilém Flusser described this type of belief in "technical images" as the result of humans forgetting they created images in order to orient themselves in the world, and instead using images to build it.[13] Computer vision clearly expresses this potential not only to gain insight into and understand the world, but also to rewrite it. In this reality, knowledge becomes probabilistic infill, with only a few entities offering the tools to compute it.

In other work, I have described this epistemological position, and methods of applying techniques from the physical sciences to the understanding of people, especially in the realm of machine learning, as a type of "computational empiricism."[14] In my new book, I describe how privacy and information laws have, over time, tried and failed to address this technological narrative that representations afford a more reliable account of a person's identity than a person can provide for themselves.[15] Computational empiricism emerges as a knowledge-logic when these profiling techniques are extended into analyzing people, actions, and events in the "real world," premised on the idea that nonvisible truths reside only in what can be sensed technologically and evaluated statistically.

THE PROPER TARGET

The proliferation of AI as a way of understanding people and the world is as much a product of structural power as it is of, for instance, the technology's success as a knowledge-generating tool. The entities that produce, sell, and profit from this technology have the capacity to establish and entrench their worldview as the dominant worldview, thus making the use of AI systems appear like common sense.

In the Western context, however, most critiques of these systems focus on the ways in which *particular applications* have discriminatory outcomes, instead of on whether they are undesirable expressions of power, or whether they generate objectionable social transformations. For instance, researchers have rightly pointed out the problematic ideologies behind using facial analysis to evaluate nonvisible personality traits, connecting contemporary efforts to dubious historical practices of statistical criminology.

Indeed, shortly after the invention of the daguerreotype, criminal portraits were subjected to statistical analyses—to identify the average or mean criminal facial morphology—by Francis Galton, the godfather of statistical "correlation." Galton was also the founder of British eugenics, the discredited school of thought that posits that an individual's inherent worth and future are determined by biological characteristics.

These ideas unquestionably live on in contemporary face-analysis projects. While Galton's experiments failed to reveal the face of criminality, since at least 2012, those 19th-century experiments have been replicated, with neural networks being used to predict intelligence and personality—and, more recently, sexuality and criminality[16]—from images and video of people. The data scientists involved in this field typically try to separate themselves from the history of social Darwinism by claiming algorithmic objectivity (i.e., claiming the data is speaking for itself). But others correctly argue that such claims to technological objectivity are, in fact, simply attempts to whitewash and repurpose disproven ideologies of eugenics for a new age.[17]

"Laundering" ideology using claims of technology's neutrality may be problematic. But this critique—which focuses on specific questionable applications of AI technologies like personality computation or facial recognition—does not tackle the bigger issue. There is a broader agenda to be addressed, one that substantiates these automated decisions and classifications as legitimate or desirable, *whether or not* they are accurate or correct. This is the agenda of the very entities that benefit from that technology's proliferation, and that are building our emerging cyber-physical society—and it is *big* business.

The private entities building these tools have the scope and influence to legitimize the knowledge claims their technologies produce, and the political power to implement their systems. We have to ask what interests and worldview this configuration serves, what the goals of cyber-physical society are now and ought to be in the future, and how might these technologies be used and understood by society for purposes not determined by big business.

This article was commissioned by Mona Sloane.

---

1. Jay Stanley, *The Dawn of Robot Surveillance: AI, Video Analytics, and Privacy* (American Civil Liberties Union, 2019), p. 5.
2. Roberta McGrath, *Seeing Her Sex: Medical Archives and the Female Body* (Manchester University Press, 2002), p. 114.

3. Walter Benjamin, “A Short History of Photography,” translated from the German by Phil Patton, in *Classic Essays on Photography*, edited by Alan Trachtenberg (Leete Island, 1980).
4. “The New Sight and the New Photography,” *The Photogram*, 1898.
5. Mary Warner Marien, *Photography: A Cultural History*, 2nd. ed. (Laurence King, 2006), p. 216.
6. Akira Mizuta Lippit, *Atomic Light (Shadow Optics)* (University of Minnesota Press, 2005), p. 44; quoted by Susan Schuppli, “Atmospheric Correction,” in *On the Verge of Photography: Imaging Beyond Representation*, edited by Daniel Rubinstein, Johnny Golding, and Andy Fisher (ARTicle, 2013), p. 23.
7. Orla Lynskey, “Grappling with ‘Data Power’: Normative Nudges from Data Protection and Privacy,” *Theoretical Inquiries in Law,* vol. 20, no. 1 (2019), p. 189.
8. Oscar H. Gandy Jr., “Statistical Surveillance: Remote Sensing in the Digital Age,” in *Routledge Handbook of Surveillance Studies*, edited by Kirstie Ball, Kevin D. Haggerty, and David Lyon (Routledge, 2012).
9. Julio C. S. Jacques Junior et al., “First Impressions: A Survey on Computer Vision–Based Apparent Personality Trait Analysis,” arXiv:1804.08046, v1 [cs.CV], April 21, 2018.
10. Jane Caplan, ‘“This or That Particular Person’: Protocols of Identification in Nineteenth-Century Europe,” in *Documenting Individual Identity: The Development of State Practices in the Modern World*, edited by Caplan and John Torpey (Princeton University Press, 2001).
11. Oscar H. Gandy Jr., “Engaging Rational Discrimination: Exploring Reasons for Placing Regulatory Constraints on Decision Support Systems,” *Ethics and Information Technology*, vol. 12, no. 1 (2008), p. 29.
12. See, e.g., Yilun Wang and Michal Kosinski, “Deep Neural Networks Are More Accurate Than Humans at Detecting Sexual Orientation from Facial Images,” *Journal of Personality and Social Psychology*, vol. 114, no. 2 (2018), p. 246.
13. Vilém Flusser, *Towards a Philosophy of Photography*, translated from the German by Anthony Mathews (Reaktion, 2000).
14. Jake Goldenfein, “The Profiling Potential of Computer Vision and the Challenge of Computational Empiricism” (paper presented at the Conference on Fairness, Accountability, and Transparency, Atlanta, Georgia, January 29–31, 2019). Note the term “computational empiricism” was used previously by Paul Humphreys in “Computational Empiricism,” in *Topics in the Foundation of Statistics*, edited by Bas C. van Fraassen (Springer, 1997), p. 119, to describe changing scientific methods associated with the adoption of powerful instrumentation and powerful computational devices.
15. Jake Goldenfein, *Monitoring Laws: Profiling and Identity in the World State* (Cambridge University Press, 2019).
16. Yilun Wang and Michal Kosinski, “Deep Neural Networks Are More Accurate Than Humans at Detecting Sexual Orientation from Facial Images,” *Journal of Personality and Social Psychology*, vol. 114, no. 2 (2018), p. 246; Xiaolin Wu and Xi Zhang, “Automated Inference on Criminality Using Face Images,” arXiv, November 13, 2016.
17. Blaise Agüera y Arcas, Margaret Mitchell, and Alexander Todorov, “Physiognomy’s New Clothes,” Medium, May 6, 2017.

Featured image: *Facial Recognition Concept*. Source: Piqsels

---

#ARTIFICIAL INTELLIGENCE #CO-OPTING AI #DATA #DESIGN #JUSTICE #POLICY #PRIVACY #RESEARCH IN TRANSLATION #SURVEILLANCE #TECHNOLOGY