Algorithmic Performativity

Emanuel Moss, PhD. – Research Scientist, Intel Labs – emanuel.moss@intel.com

In "Algorithms as Culture", Nick Seaver proposes that algorithms, as sociotechnical systems, be studied both *in* culture and *as* culture and suggests a number of methodological approaches for doing so (Seaver 2017). Algorithms *in* culture take on different meanings and are variably defined from one social context to the next (Dourish 2016). Algorithms *as* culture do not so much act upon culture as some external force, but are also cultural actors that through their operations bring culture into existence in a multiplicity of ways (Mol 2002). Here, I would like to address what happens when applied machine learning researchers build algorithmic representations *of* culture. Algorithmic representations are one of the multiple ways in which algorithms operate alongside others as cultural actors—that is to say, both *in* and *as* culture—and attending to algorithmic representations of culture allows us to focus on how this happens. In the following, I will focus on the performative aspects of algorithms, the ways they enter into and contribute to the enactment of specific cultural practices, and how performativity offers a methodological toolkit for better understanding AI in culture.

Algorithmic representations are always already cultural. They are at once the product of the culturally bound understandings algorithm developers bring to their framing of the projects they engage in, the culturally-contingent practices that structure data collection and classification, the cultural particularity of the phenomena represented as data, and the cultural practices that are enacted through the operation of the algorithmic system. Tracing the practices through which algorithms embed aspects of culture, how those aspects are elided from professional and popular understandings of algorithms *in* culture, and the cultural conventions that support the operation of algorithms *as* culture are key components to understanding the power algorithms hold. This power derives not only from their status as "the manifold consequences of a variety of human practices" (Seaver, 2017, p. 4), or from the condensation of human investments in material objects (Thomas, Nafus and Sherman, 2018), but also from the ways they orchestrate specific *performances* of the culturally-bound understandings they represent (Moss, 2021).

By *performance*, I mean here that algorithms enact the objects of its analysis as *specific kinds of objects*, through the practices of data science and machine learning. Here, I am referring to performativity along the lines of J.L. Austin, in that algorithms produce a pragmatic effect on the world (Austin, 1962), bringing something specific and novel into existence. But they also are performative through the agency they hold to make what Karen Barad calls "agential cuts" (Barad, 1999, 2007); they are an apparatus through which several possible states of the world (or indeterminate meanings in the world) are resolved into a single authoritative state or meaning. For machine learning, this power and authority is enacted through how it comes to affect the phenomena it is built to represent. Recently, the ability of an algorithmic system to “steer a population” in measurable ways through the manipulation of market-like mechanisms has been referred to as *performative power* (Hardt, Jagadeesan and Mendler-Dünner, 2022). While this specific formalization captures, in broad swaths, some aspects of performativity described above, it defines performative power as a measurement of the effects on an algorithmic system has on the phenomena it represents in terms of the data it uses to build that representation. But the performative power algorithmic systems hold extends beyond that which can currently be measured. It includes what Foucault would call *disciplinary* forms of power (Foucault, 1995), what Searle has called *deontic* power (Searle, 2006), and what Barad has referred to as *ethico-onto-epistemlogical* power (Barad, 1999). Each of these forms of power *may* be rendered measurable in terms of the data used to build algorithmic representations, but need not be rendered so in order to deserve consideration for those seeking responsible and accountable deployments of algorithmic technologies.

Take, as an illustrative example, a computer vision model trained to recognize handwritten numerals using the MNIST dataset. This is a common exercise in introductory machine learning courses (Géron, 2017) that uses labeled examples of handwritten numerals compiled by the U.S. National Institute of Standards and Technology (NIST). A fully trained machine learning model that is reasonably accurate at that task will have constructed, across the parameters of each node it uses to accomplish its task, a statistical *representation* of the dataset it has been trained on, reflecting the variances of number-writing archived within the training dataset. The pixels for “1’s” cluster linearly, the pixels for “0’s” cluster annularly, and so on. But while this representation might be taken to be able to stand in for all handwritten Arabic numerals (and in practice it has been effective for doing so), it is only a partial representation of the cultural practices of number-writing. These practices exhibit greater variance than that exhibited within the MNIST dataset.

A closer look at the MNIST dataset reveals that it is comprised of images taken of numerals written on forms filled out by U.S. Census Bureau field staff (National Institute of Standards and Technology, 1994). Census Bureau staff are quite literally professional writers of numerals, and whose professional numeral-writing practices might overlap, but not precisely correspond, with that of the general U.S. population, and even less so with that of other nations. However, through MNIST, the locally-specific numeral-writing practices of Census Bureau workers have come to stand in for the way Arabic numerals *ought* to be written… at least if they are to be recognized by an computer vision system.

It is on this "ought" that the power of algorithmic performativity rests. An algorithmic system, trained on MNIST to represent Arabic numeral-writing practices, creates an expectation of how numbers *are* written, which gives it power to determine how numbers *ought* to be written. Briefly, this is the mechanism through which the various forms of power discussed above are exercised. Imagine someone who is interacting with a handwriting-detection system. Perhaps they are trying to write a check that can be cashed by a mobile banking app or are a medical scribe annotating a patient's chart on a tablet. If that person writes numerals that fail to be detected, they fall under the performative power of the algorithm in that a change is forced upon them through that interaction. That performative power has a disciplinary dimension; as they change their number-writing practices, refining how they inscribe various numerals in an attempt to 'pass' the system and have their writing accurately accepted by the machine, they are re-conditioned to write numbers in ways that align with the representation of writing embodied by the trained algorithmic system. It also has a deontic dimension. Defined as inducements to act in ways contrary to desire, this dimension of performative power pushes against the way a person might want to write—perhaps sloppily in favor of haste (but well enough to be read by other humans), or as an aide-mémoire that needs to be legible only to themself—and causes them to write in a way they do not wish to. Additionally, the performative power of algorithms also has an ethico-onto-epistemological dimension. Of the many ways a numeral *might* be drawn—the number "7" may be drawn with or without a crossbar, a "1" with or without a flag, a "9" with a curving base, etc.—the algorithm makes an "agential cut" that collapses those possibilities into a narrower range of options (Barad, 1999), in effect deciding what such things *are* and *can be* in the world, as well as what their specific ethical stakes consist of. Across all these dimensions, the cultural practices of U.S. Census fieldworkers have come to affect the cultural practices of a far wider group of people.

These dimensions of performative power are, perhaps, small in magnitude for a system built on a dataset like MNIST. But for the increasingly wide set of practices machine learning is currently being developed for, these dimensions might grow quite vast indeed. This requires growing awareness, methodological attention, and robust accountability practices for all involved in the machine learning ecosystem to ensure that this form of power is legible to all involved and—ultimately—contestable by those who hold positions of lesser political, economic, and cultural power in relation to those who develop such technologies.

Works Cited

Austin, J.L. (1962) *How to Do Things With Words: The William James Lectures delivered at Harvard University in 1955*. Oxford, UK: Clarendon Press.

Barad, K. (1999) 'Agential realism: Feminist interventions in understanding scientific practices', in M. Biagioli (ed.) *The Science Studies Reader*. New York: Routledge, pp. 1–11.

Barad, K. (2007) *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Meaning*. Durham, NC: Duke University Press.

Foucault, M. (1995) *Discipline and Punish: The Birth of the Prison*. Vintage Books.

Géron, A. (2017) *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. 1st. Boston, MA: O'Reilly.

Hardt, M., Jagadeesan, M. and Mendler-Dünner, C. (2022) 'Performative Power'. arXiv. Available at: http://arxiv.org/abs/2203.17232 (Accessed: 23 September 2022).

Moss, E. (2021) *The Objective Function: Science and Society in the Age of Machine Intelligence*. Dissertation. CUNY Graduate Center. https://arxiv.org/abs/2209.10418

National Institute of Standards and Technology (1994) *NIST Standard Reference Data Products Catalog*. SP 782. National Institute of Standards, p. 120.

Searle, J.R. (2006) 'Social ontology: Some basic principles', *Anthropological Theory*, 6(1), pp. 12–29. Available at: https://doi.org/10.1177/1463499606061731.

Seaver, N. (2017) 'Algorithms as Culture: Some Tactics for the Ethnography of Algorithmic Systems', *Big Data & Society*, 4(2), p. 205395171773810. Available at: https://doi.org/10.1177/2053951717738104.

Thomas, S.L., Nafus, D. and Sherman, J. (2018) 'Algorithms as fetish: Faith and possibility in algorithmic work', *Big Data & Society*, 5(1), p. 205395171775155. Available at: https://doi.org/10.1177/2053951717751552.