Data Ethics for Non-Ideal Times—Some Notes on the Course
by Anna Lauren Hoffmann

> "…one could say epigrammatically that the best way to bring about the ideal is by recognizing the non-ideal, and that by assuming the ideal or near-ideal, one is only guaranteeing the perpetuation of the non-ideal."
>
> -Charles W. Mills, "'Ideal Theory' as Ideology"

I.

We live in non-ideal times, to say the least.

A global pandemic has upended our entire year, killing hundreds of thousands in the US and interrupting the lives of millions worldwide. It has interrupted our plans and thrown our schedules into disarray, at first tentatively and now—increasingly, it seems—indefinitely. No doubt, few of us planned on taking classes remotely or teaching online. Worse, some of us have been forced to navigate uncertain and unstable living situations. Meanwhile, many of our families have been touched by the virus—either directly in terms of illness or indirectly though job losses, insecurity, and threats of eviction.

And none of us know when or if or how this all ends—or what the world will look like when it does. I feel confident calling that "non-ideal."

In other ways, however, the pandemic has been clarifying. It has, for example, made obvious the already-existing injustices that structure our world. People are getting sick, yes—but not in equal measure. In the US, the fact that Black, Latinx, and Indigenous communities have shouldered a disproportionate share of cases and deaths is, in many ways, simply indicative of the inequalities informed health care pre-COVID. These and other inequities—especially and most pointedly, the violent application of law enforcement—have helped fuel a national movement for racial justice that has, arguably, outstripped even the civil rights movements of the 1960s. From the streets to city councils to the backs of pro basketball jerseys, the pandemic has forced an accounting of how we got here.

It has also thrown a wrench into what was already, inevitably a tumultuous presidential election year in the US. But the coronavirus isn't the source of that tumult; a virus doesn't cause things like white identity politics, partisan polarization, or craven political power-grabs. But it has given those things new toeholds, so that something as wearing a mask has morphed into a frontline cultural conflict—one we see played out in the aisles of Target or Trader Joe's, recorded and

replayed on social and mainstream media alike. Denied their usual consumptive outlets, as Tressie McMillan Cottom describes, "white people are spinning out." [1]

Given that the fault lines that have most profoundly shaped the trajectory of this pandemic are not new, its consequences have not been entirely unpredictable. So why are we still so surprised?

One answer, I think, lies in the ways we have oriented ourselves to the crisis. As a practical matter, we often interpret the world by filtering information through those perspectives and ideals that are more or less ready to hand. Consider, for example, this frank reflection on returning to the classroom during the pandemic, written by Temple University professor Devon Powers:

> "The problem with [back-to-school] guidelines, and the return centered around them, is that they expect us to avoid the very things that bring classrooms and campus to life. They depend on us behaving as perfect rational actors, rather than human beings who are vulnerable and sometimes unsure, especially after so many months of isolation. They place responsibility on us as individuals, letting the leaders and systems that fail to protect us off the hook. And they require us to evaluate encounters in black and white when, just days into this semester, I have already seen many shades of gray." [2]

I have to admit, it stung a bit reading those words. It made me sad to think about the classroom connections and experiences we won't have. But that aside, anyone who has spent any amount of time managing or navigating the intimacy of a college classroom could have predicted exactly what Powers captures here.

And this is hardly an isolated case. Just recently, school officials in Wisconsin had to scramble to address the reality that some parents have knowingly been sending sick kids to school. In response, the officials noted: "Never in a million years did we imagine or think to account for parents deliberately sending their sick or symptomatic child to school."

Never? *Never in a million years?* You don't have to be a parent to recognize that navigating life against the backdrop of a less-than-ideal global pandemic might lead people to make some less-than-ideal decisions. Sick kids being deliberately sent to school might have been easier to anticipate if we started reasoning from a different premise—for example, from the reality that many parents and caretakers are currently desperate and hyper-stressed, overburdened and under-supported, unable to balance childcare with working from home.

This is not to blame some university or other school officials for developing questionable or insufficient plans. They are also often laboring under difficult circumstances exacerbated by a federal response that could be described as, at best, haphazard or, at worst, as actively negligent.

(Though, when the dust settles, there will be plenty of blame to go around.)

We can identify similar trends in reactions to the racial and class-based politics that have informed so much of 2020. The fact that some protests for justice would turn violent is only surprising if you've never breathed tear gas or watched innocent people get snatched up into unmarked vans. Further, the idea that police brutality is real and devastating and common is only surprising if your race and/or class position has spared you regular interaction with police.

Or, the fact that the pandemic has only further proved white racial grievances a danger to our health and safety is unsurprising to those on their receiving end. As McMillan Cottom summarizes: "to some of us, they already were [dangerous]."

II.

At this point you might be saying, "okay, sure—but what does any of this have to do with 'data ethics'?"

The answer, I think, is a lot.

As with the ill-fated pandemic plans and pearl-clutching protest responses referenced above, many of the most prominent approaches to "ethics" also, at times, seem to reason from idealized premises and situations—premises and situations informed by perspectives that don't account for the full (or even a broad) range of messy and less-than-ideal facts of human history and behavior.

Ethics in the Anglo-American analytic tradition, in particular, has long prioritized argumentation over evidence. Take, for example, the work of John Rawls, arguably the most influential political philosopher of the 20th century. Central to his famous theory of justice is a conception of individuals as "free and equal moral persons." Critics have long pointed out that his framing of such persons seems to rely on assumptions about human cognition that exclude many people, including certain people with disabilities. Rawls' response to these critics was to double-down on this exclusion, deliberately putting their concerns aside and arguing that we have to figure out justice—and I am not making this wording up—for "normal cases" first. [3]

Building on this and other examples, philosopher Charles Mills argues that this sort of idealized thinking—known as "ideal theory"—is a consequence of certain dominant strands of institutionalized philosophy, namely strands that prize carefully curated assumptions about

human capacities, relationships, and institutions over messy entanglements with history, ability, and oppression. [4] Ideal theory's assumptions might lend themselves to tidy arguments, but they are often remarkably limited when it comes to thinking about and addressing ethical problems on the ground, in the here and now.

Even worse, Mills argues, given the ways that ideal theory has been (historically and overwhelmingly) dominated by white men with certain class privileges, its curated assumptions about the world are hardly innocuous. Instead, they encode ideas that reflect the non-representative expectations and perspectives of their curators—and these are limited in their ability to address the interests, needs, and demands of the oppressed. For example, arguing from an ideal like Kantian autonomy, where everyone is hypothesized as an inviolable "end" in themselves, ultimately—as philosopher Robin James puts it—"reproduces the patriarchal denial of personhood to women, who in their forced shouldering of reproductive labor are constantly treated as means. Think about it: *somebody* did Kant's laundry." [5]

This is not to say, under certain circumstances, we wouldn't be able to learn something from such work. But it *is* unreasonable to assume that ideas built around on the interests and ideals of economically well-off, able-bodied white men can adequately attend to the interests of people of color, disabled people, women and non-binary people, or the working class. Yet, somehow, that's exactly what centuries of philosophy have asked us to assume.

Reflecting on this situation, Mills asks: "*How in God's name could anybody think that this is the appropriate way to do ethics?*"

Against this, Mills advocates for what he and others—like philosopher Elizabeth Anderson and (in a slightly different way) economist Amartya Sen—call the "non-ideal theory" approach to ethics. Under non-ideal theory, reasoning about ethics starts not from idealized assumptions, but from the actual non-ideal conditions that inform our lives, including (among other things) the facts of domination, oppression, and difference.

An important component of non-ideal theory is, as philosopher Serene Khader describes it, a requirement that we evaluate our ethics, in part, by how they direct our attention. [6] If an ethical proposition starts from some unencumbered assumption about the ideal human and not from actual humans and their capacities, then we should resist that proposition. For example, if an ethical framework starts from the assumption that all people are free and equal autonomous agents while largely ignoring the fact that people have range of abilities, identities, material needs, and social dependencies, then we have little reason to take that framework seriously.

Following the insights of these and other thinkers, our approach to "data ethics"—that is, questions of right and wrong, justice and injustice, in relation to data and its attendant

technologies, like classifications, labels, computation, algorithms, machine learning, and artificial intelligence—will set out from the non-ideal and, in particular, from the facts of domination, oppression, and difference. Accordingly, our discussions will foreground those categories by which difference and exposure to violence and harm are predictably organized, including (but not limited to): race, ethnicity, gender, ability, class, and nationality. It will also identify and critically assess those tools and strategies commonly advanced as ethical responses to difference and oppression, like liberal freedoms, human rights, digital access, social and economic opportunity, and "diversity and inclusion" efforts. We will attend, in particular, to the possibilities and limits of these tools relative to the challenges posed by data science and technology and explore alternatives grounded in domains like critical race theory, feminism, reproductive justice, and critical trans politics.

By framing the class in this way, I do not mean to cordon discussion so as to prevent disagreement. Quite the opposite, in fact. By recognizing that ethical debate is not only about values, but about how we account for certain non-ideal facts about the world, I mean to create more space for robust discussion, passionate engagement, and productive disagreement. But I do insist that our disagreements arise from good-faith engagement with these facts, not in spite of them.

For example, we might have reasonable disagreements about exactly how histories of violence and oppression shape our current world, but we are nonetheless going to start from the recognition that such histories do, in fact, shape our current world. We might also have reasonable disagreements about exactly how categories like race and gender shape people's experiences, but we are nonetheless going to start by recognizing and centering the fact that such categories do, in fact, shape people's experiences.

In terms of data ethics, we might have reasonable disagreements about exactly how domination, oppression, and difference inform the production and application of data science and technology, but we aren't going to bury our heads in the sand and pretend like data science and technology are not informed by these facts. After all, science and technology are part of this world—not apart from it. To think otherwise would be—perhaps now more than ever—wildly irresponsible. It is incumbent on us to recognize that, to paraphrase Mills' critique, an ethics that ignores oppression is an ethics in service of oppression.

III.

To ground our thinking about a non-ideal data ethics, we are going to spend the quarter reading a series of books that marshal a range of perspectives in studying data science and technology's entanglements in the world. We will engage these books earnestly and in good faith, with

particular attention to how they direct us to think about, reason with, and confront data science and technology under non-ideal conditions.

You'll note that these books don't often deploy the term ethics specifically. But this hardly makes them irrelevant for doing ethics. As Mills notes, ethical debates—that is, "debates about right and wrong, justice and injustice"—are not wholly reducible to debates about ethical values. They also entail disputes over how to account for the world, over "competing narratives of what has happened in the past and what is happening right now, alternative descriptive frameworks and interpretations." [7] Or—following philosopher Margaret Urban Walker—our actual, real-world responses to that which we perceive to be unjust or unethical don't issue from some timeless, transcendent value—they are, in the first place, driven by our interpretations of what is salient in a given situation. [8]

Accordingly, these books all highlight and make salient different dimensions of data science and technology—and the worlds within which they are embedded. I chose them, in particular, because they don't offer some idealized version of data science and technology as self-contained or possible outside of the social and political conditions under which they are designed, developed, and deployed. In fact, for all of these authors, doing ethics will mean attending to histories, dynamics, and relationships that exceed any given tool or technology, no matter how sophisticated. At points, you might even find that their data ethics make "data" appear beside the point—and maybe, sometimes, that *is* the point.

To say that our ethics are informed by struggles over how we perceive and account for the world is not, however, to say that our ethics are simply a matter of subjective opinion. Just because we are all situated in particular ways does not mean that everything is or has to be relative. In fact, that's an outcome we actively want to avoid.

But admitting the non-ideal does not preclude us from seeking universal—or at least common—understanding. For example, by centering violent histories of racial or gendered subjugation and abuse we can better see how violence is common tool of domination, deployed in ways that harm people—as Franz Fanon made so clear—physically, mentally, emotionally, and beyond. [9] Working from these actual histories, then, we might conclude that our ethics should profess a common commitment to eliminating—or at least reducing exposure to—structural or systemic forms of violence. Further, we might concede that among those best positioned to teach us how to ethically assess and address the harms and consequences of violence are those with firsthand experience struggling under and against it. Similarly, we might also admit that we have comparatively less to learn from those whose lives largely escape (or actively benefit from) such violence.

Put another way, being critical of our ethics and the ways we account for them does not mean negating the possibility of ethics. Rather, being critical is, as Mayanthi Fernando puts it, "a necessary practice of care for the world, and vital to any project for political justice." [10]

IV.

In terms of data ethics, adopting a non-ideal approach can shift our conversations in profound ways. Take, for example, computer vision systems used for facial recognition. If we start from some idealized assumption about free and equal autonomous agents, the most salient questions will likely revolve around abstract ideals of fairness or consent under more or less ideal conditions. The ethics of such systems will hinge on whether or not idealized agents are, at least hypothetically, treated fairly by the system (i.e., all people are "seen" more or less accurately) and can consent to its use (i.e., its operations are evident and people can opt out). The solutions that fall out of this idealized approach might revolve around prioritizing "inclusive" or more representative datasets (to make them more accurate) and the posting of signs or notices to make people aware of when they are consenting to the use of facial recognition.

That's not to say these solutions could never be useful. But, following the non-ideal approach, we should be skeptical of the degree to which such solutions can address facts of domination, oppression, and violent subjugation.

By contrast, a non-ideal approach will make salient different features of the situation. Instead of asking questions about idealized capacities of idealized agents, we will ask "does this free—or at least reduce—people's exposure to structural or systematic forms of violence?" One immediate consequence of this shift will be a necessary centering of applications and use. In order to assess these systems relative to their capacity for violence, we will have to understand how and where those historically subject to such violence will encounter them—in contexts like law enforcement, employment and education, and securitized borders.

Following non-ideal theory, then, we won't be able to effectively attend to the ethics of facial recognition without at the same time contending with ongoing legacies of state violence, racial profiling, and xenophobic abuse. To separate these things would be to retreat to the very idealized plane we have committed ourselves to resisting. So, while it may be reasonable to say that an idealized facial recognition system developed and deployed in an idealized vacuum is not inherently violent, we're not actually developing and deploying facial recognition systems in vacuums (even if I think we should shoot them into space).

Scholar and writer Virginia Eubanks has made a similar point about the ways data and algorithms are deployed in the regulation and delivery of social services. Most often, social services are fundamentally designed to prevent or root out fraud—that is, they are built on an

ideology that says government services should be limited and that those who need them should be treated with suspicion. The results are systems that are not only difficult to navigate, but actively hostile to those navigating them. But what if we started our ethical reasoning from a recognition of historical and ongoing oppression and, perhaps, a correspoding obligation to recognize and address the fact certain people and groups have been regularly and systematically dispossessed and disadvantaged? As Eubanks argues, we could instead build systems on the idea that people should be able to get access to maximum benefit as easily as possible—but that would necessitate a radically different approach. [11] Here we can see that while the debate between ideal and non-ideal theory, though at times seemingly abstract, has real consequences for what we do in practice.

Additionally, committing to non-ideal theory also means recognizing that the role of articulating and responding to the ethics of any technology should not be ceded to those furthest removed from its worst effects. Today, we are awash in so-called "tech ethics" experts like Tristan Harris and Aza Raskin, whose Center for Humane Technology announces itself as working to save humanity from technology's harmful and dystopic tendencies. They use their Big Tech insider status (Harris, as we are reminded again and again, is a former "design ethicist" at Google) to profess special insight into the ethical stakes of search engines, social media platforms, and social technologies broadly.

Let's not forget that these "prodigal techbros" were, in the first place (and by their own ready and frequent admission!), paid handsomely to help to create the social technologies they now want to save us from. [12] And now they're using their power and position to claim authority over how we define, diagnose, and respond to problems like propaganda, hate, and algorithmic bias. All despite that fact that they themselves are—socially, economically, and otherwise—far removed from their worst effects.

As scholar Pranav Malhotra notes, people like Harris and Raskin largely confine their ethical analyses to questions of design, never grappling with longer-standing facts of inequality, exploitation, and dispossesion. [13] This move is not innocent. By bracketing these non-ideal facts, these "experts" are able to advance an idealized version of tech's ills as confined to design—a rarefied area where they are able to exert maximum authority and influence. For them, admitting the non-ideal into their ethics would be unthinkable, as it would mean surrendering the special insight that keeps them secure at the top of the tech ethics hierarchy. (An ethics in service of oppression, indeed.)

IV.

Unlike other, more nominally "applied" approaches to data ethics, our aim in this class is not to generate some more or less complete set of directives for handling or working with data, writing

and deploying algorithms, or regulating systems. That's not to say such classes don't have their place (though, admittedly, I have little interest in teaching them). And, anyway, there are no such set of directives for us to learn. Different companies have different approaches, internal rules and policies, and expectations for how work is to be done. It's not possible to anticipate precisely what situations and constraints you'll face in your working life.

In lieu of concrete ethical prescriptions, we will instead work together to take some meaningful steps toward cultivating a kind of praxis. That is, we will explore these texts and embrace their lessons with an eye toward how they might reorient us toward data science and technology—and how they might inform how we live, work, and relate to one another. In place of rules, I hope you leave this class with some new knowledge, yes, but also, perhaps, a new sensibility. That is, I hope you leave with a changed sense of how we interpret and assess data technologies that are—as Oscar Gandy has demonstrated—predicated on surveillance, social sorting, and optimization for social, political, and economic gain. [14] Such a sensibility is infinitely more portable than any set of necessarily incomplete prescriptions.

Most importantly, I hope you leave committed to a data ethics that offers, in Lilly Irani's words, more than "minimum viable futures." [14] Following Saba Mahmood, I want you to get better at "juxtaposing the constitutive concepts and practices of one form of life against [those of] another in order to ask a different set of questions, to decenter and rethink the normative frameworks by which we have come to apprehend life." [16]

I'm not going to lie to you: there is some risk here. Grounding our work in the non-ideal means, for the time being, leaving behind the juridical fantasy of tidy rules and buttoned up checklists that other more clinical or analytic approaches to ethics might have to offer. Eschewing certainty and embracing a critical lens is, no doubt, uncomfortable and you're going to feel adrift at times. By abandoning the hermetically sealed surety of ideal theory, we are going to have to learn to live with an ethics informed by, in the words of anti-violence educator and organizer Mia Mingus, "complicated stories of saturated heartbreak and nagging hope, with small victories scattered throughout."

Though it might be difficult, we can at least be confident that we spent less time developing principles for idealized worlds that don't actually exist and more time cultivating our ability to confront actually existing harms and injustices in our actually existing, non-ideal world.

September 30, 2020

References

[1] Tressie McMillan Cottom, "The Meltdown Crisis," *More of a Comment: Bits and Bites from Tressie McMillan Cottom*, July 14, 2020, https://medium.com/more-of-a-comment/the-meltdown-crisis-dda5c3ff51a5

[2] Devon Powers, "A professor's first days back at Temple: 'This pandemic experiment won't work'," *The Philadelphia Inquirer,* August 28, 2020, https://www.inquirer.com/opinion/commentary/temple-reopening-coronavirus-safety-campus-teachers-20200828.html

[3] John Rawls, *Political Liberalism* (New York, NY: Columbia University Press, 1993).

[4] Charles W. Mills, "'Ideal Theory' as Ideology," *Hypatia* 20, no. 3 (2005): 165-184.

[5] Robin James, "The Other Secret Twist: On the Political Philosophy of The Good Place," *Los Angeles Review of Books*, October 11, 2018, https://lareviewofbooks.org/article/secret-twist-political-philosophy-good-place/

[6] Serene J. Khader, *Decolonizing Universalism: A Transnational Feminist Ethic* (New York, NY: Oxford University Press, 2019).

[7] Charles W. Mills, *Black Rights/White Wrongs: The Critique of Racial Liberalism* (New York, NY: Oxford University Press, 2017).

[8] Margaret Urban Walker, *Moral Understandings* (New York, NY: Oxford University Press).

[9] Frantz Fanon, *The Wretched of the Earth*, trans. Richard Philcox (New York, NY: Grove Press, 2004).

[10] Mayanthi Fernando, "Critique as Care," *Critical Times* 2, no. 1 (2019): 13-22.

[11] Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor* (New York, NY: St. Martin's Press, 2018).

[12] Maria Farrell, "The Prodigal Techbro," *The Conversationalist*, March 5, 2020, https://conversationalist.org/2020/03/05/the-prodigal-techbro/

[13] Pranav Malhotra, "*The Social Dilemma* Fails to Tackle the Real Issues in Tech," *Slate*, September 18, 2020, https://slate.com/technology/2020/09/social-dilemma-netflix-technology.html

[14] Oscar H. Gandy, Jr., *The Panoptic Sort: A Political Economy of Personal Information* (Boulder, CO: Westview Press, 1993).

[15] Lilly Irani, *Chasing Innovation: Making Entrepreneurial Citizens in Modern India* (Princeton, NJ: Princeton University Press, 2019).

[16] Saba Mahmood, *Religious Difference in a Secular Age: A Minority Report* (Princeton, NJ: Princeton University Press, 2016).