Introduction: How to Destroy the World, One Solution at a Time

The Internet has become a nightmare, the source—it is claimed—of almost everything bad in this world. It has given rise to worldwide surveillance networks, coproduced by states and corporations; social media algorithms, powered by military-grade psychological operations (PSYOPs) that spread lies and conspiracy theories, polarize society, provoke violence, prolong pandemics, and foster planet-wrecking levels of consumption; and artificial intelligence (AI) programs that exacerbate existing inequalities and threaten humanity's future.

The irony is that the Internet and artificial intelligence were promised to be and do the opposite. Cyberspace, the Internet of the late twentieth century, was to usher in a new era of global democracy, equality, and prosperity. Artificial intelligence was to produce docile machine servants that would spread the perks of "the 1%"—chauffeurs, personal assistants, expert advisors—to "the 90%." AI would also eliminate discrimination because its machines could not "see" race, sex, age, or infirmities.[1] Similarly, cyberspace would free individuals from oppression and national sovereignty because it was "the new home of the Mind":[2] an electronic frontier in which physical bodies and identities literally did not matter. In the mid-1990s, Vice President Al Gore and members of the U.S. judiciary described the Internet as the ultimate public sphere because it gave everyone a soapbox from which to speak.[3] Bill Gates, then CEO of Microsoft, argued that the information superhighway enabled "friction-free capitalism" because it melted away brick and mortar obstacles.[4] John Gilmore, cofounder of the Electronic Frontier Foundation (EFF), is reported to have said that the Internet "interprets censorship as damage and routes around it."[5] As late as 2010, the Internet was celebrated as a "liberation technology," responsible for democratic uprisings in the Middle East.[6] By freeing our minds, the Internet would help fix all problems, from racism to political suppression.

During the early twenty-first century, the questions for those who still sold hope were: How can the dream be reclaimed from the nightmare? What information should be leaked, what new business plans devised, what apologies proffered to make technology great again?

However well-intentioned, these impulses were also misguided, for the promise and the threat were, are, and have always been two sides of the same coin. In seeking technological solutions to political problems, they assume that the best way to fight abuse and oppression is by ignoring difference and discrimination.[7] They undermine solidarity by concentrating on individual, neighborhood or "tribal" empowerment. They presume that "good" technology is slavish and thus inevitably invoke fears of absolute dependence and rebellion. Hopeful ignorance is not the solution but the problem: it perpetuates discrimination and inequality,

one solution at a time. The problem is not that giant technology monopolies have disrupted habits, institutions, and norms in order to create new, unforeseen futures. The problem is that, in the name of "creative disruption," they are amplifying and automating—rather than acknowledging and repairing—the mistakes of a discriminatory past.

To counter this threat, I propose the following five-step program:

1. Expose and investigate how ignoring differences amplifies discrimination, both currently and historically.
2. Interrogate the default assumptions and axioms that form the basis for algorithms and data structures.
3. Apprehend the past, present, and future machine learning algorithms put in place to determine when, why, and how their predictions work.
4. Use existing AI systems to diagnose current inequalities and to treat discriminatory predictions as evidence of past discrimination.
5. Draw from struggles for and practices of desegregation and equality to displace the eugenic and segregationist defaults embedded within current network structures and to devise different algorithms and modes of verification.

Most fundamentally, I call for a "we" to take this on. The views expressed in this book thus strike a chord with those voiced by Ruha Benjamin, Jodi Byrd, Meredith Broussard, Kate Crawford, Virginia Eubanks, Kara Keeling, Tara McPherson, Lisa Nakamura, Safiya Noble, Cathy O'Neil, Frank Pasquale, and Fred Turner among many others, creating a powerful chorus against hopeful ignorance and the endless apologies it engenders, and for a world that resonates with and in difference.[8]

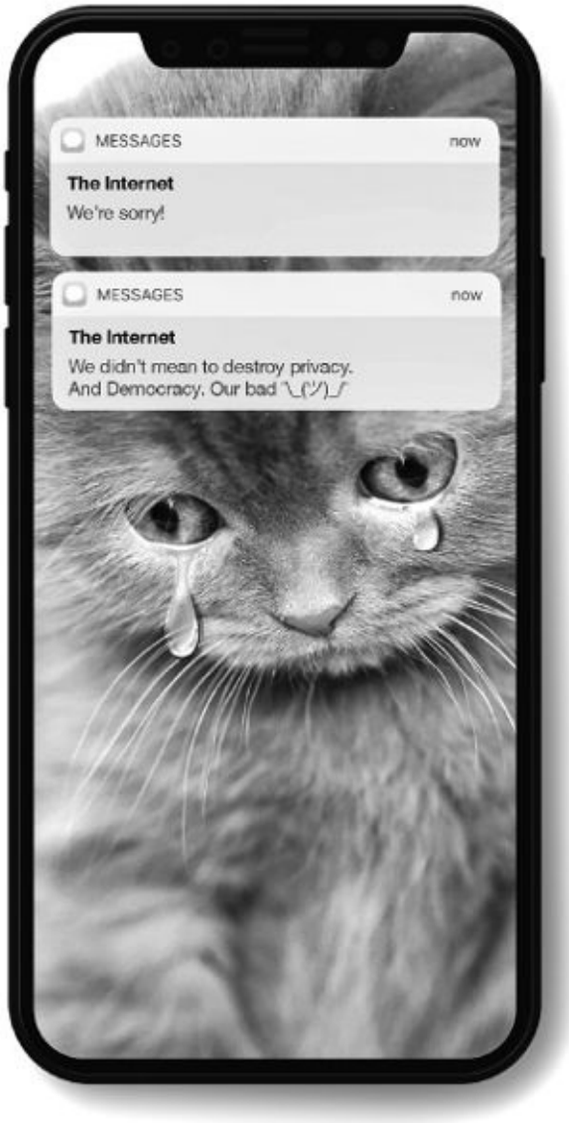
Against Hopeful Ignorance, Again

In the early decades of the twenty-first century, technology companies responded to Internet-related disasters by asking for forgiveness and promising technological fixes for their sins. In 2018, Mark Zuckerberg, the founder of Facebook, apologized publicly for the "leak" of 87 million personal profiles to Cambridge Analytica.[9] The Cambridge Analytica incident, however, as Kate Crawford and Meredith Whitaker of the AI Now Institute emphasized in the institute's 2018 annual review, was only one of many.[10] Scandals and outrage dominated that year: from revelations that U.S. Immigration and Customs Enforcement (ICE) had "upgraded" its risk assessment software to always recommend detention, to news that Amazon had scrapped its AI hiring software because it discriminated against women, to reports that IBM's supercomputer Watson had recommended cancer treatments that were "unsafe and incorrect."[11]

Noah Kulwin captured the state of affairs in his *New York* magazine article, "The Internet

Apologizes . . . .” It led with a picture of a cute cat who texted: “We’re sorry. . . . We didn’t mean to destroy privacy. And democracy. Our bad” (figure 1).[12]

Kulwin offered the following list of “How It Went Wrong, in 15 Steps,” based on his interviews with a dozen prominent network architects, Silicon Valley product developers and tech gurus, such as Jaron Lanier and Richard Stallman:



1 Cute crying cat from Noah Kulwin, “The Internet Apologizes . . . : Even Those Who Designed Our Digital World Are Aghast at What They Created,” *New York*, April 13, 2018, http://nymag.com/intelligencer/2018/04/an-apology-for-the-internet-from-the-people-who-built-it.html. Photo illustration by Joe Darrow; image recreated by Joshua Cameron.

1. Start with Hippie Good Intentions . . .
2. . . . Then mix in capitalism on steroids
3. The arrival of Wall Streeters didn’t help . . .
4. . . . And we paid a high price for keeping it free.
5. Everything was designed to be really, really addictive.
6. At first, it worked—almost too well.

7. No one from Silicon Valley was held accountable . . .
8. . . . Even as social networks became dangerous and toxic.
9. . . . And even as they invaded our privacy.
10. Then came 2016.
11. Employees are starting to revolt.
12. To fix it, we'll need a new business model . . .
13. . . . And some tough regulation.
14. Maybe nothing will change.
15. . . . Unless, at the very least, some new people are in charge.

The basic story line was this: naive hippies fall in love with libertarians, hook up with Wall Street sharks, and inadvertently destroy the world in their attempt to keep it free. As Jaron Lanier told Kulwin, they were caught between two loves: "We wanted everything to be free, because we were hippie socialists. But we also loved entrepreneurs, because we loved Steve Jobs. So you want to be both a socialist and a libertarian at the same time, which is absurd."

Clickbait advertising resolved this "absurdity" by paving the road to hell. Capturing and exploiting Internet user clicks magically enabled "free" yet profitable content. It also seemed to answer the question that dogged mass print and broadcast advertising: How effective is an ad? By tracking user clicks and mouseovers, advertisers could "measure" engagement, and thus overcome what social theorist Jean Baudrillard had presciently and perversely called the "silent power of the majority."[13] To optimize performance, platforms encouraged advertisers to amalgamate related but bespoke microaudiences, that is, to create a crowd of users by consolidating rhyming groups. As chapter 3 further elaborates, to create affectively charged clusters who would take the clickbait, advertisers and platforms targeted users by focusing on their divisive or boundary views. ProPublica's 2017 investigation into Facebook, for example, revealed that Facebook "helpfully" suggested that their reporters add "How to burn Jews" and "Second Amendment" to "Jew hater" in order to boost their ad's target audience size.[14] The fact that the price per ad generally decreased per click further promoted shocking and manipulative advertisements. Soon, the actual product no longer mattered, for monetized user clicks generated their own wealth: outrage—or anything that piqued curiosity—had become profitable. Most infamously, hackers from Moldova produced right- and left-wing fake political news during the 2016 U.S elections in order to profit from a combination of Facebook click throughs and Google ad auctions.[15] Kulwin argues that the success of clickbait advertising resulted both in further polarization of "what had already seemed, during the Obama years, an impossibly and irredeemably polarized country" and, quoting Jaron Lanier, in "continuous behavior modification on a mass basis, with everyone under surveillance by their devices"[16]—what Shoshana Zuboff has called "surveillance capitalism."[17] The cure had become worse than the disease: the collateral damage was democracy and freedom, sacrificed on the altar of the free.

Reforming "the Valley" and redressing mass surveillance and behavior modification

programs are important, and Lanier's observations are perceptive and engaging, but Lanier's assumptions threaten to undermine his argument and the success of the proposed reforms. Socialism does not equal free information: the fundamental tenet of socialism is not that everything should be free, but that workers should share equally in the profits. The urge to make things free and profitable is wholly libertarian, and the misidentification of libertarianism as socialism erases labor.[18] Tellingly, the subtitle of Kulwin's article reads: "Even Those Who Designed Our Digital World Are Aghast at What They Created," which raises the question: How did these twelve architects, designers, and tech executives become "the Internet"? During the heyday of Web 2.0, users were celebrated *as* the Internet: *Time* magazine declared "You" the 2006 Person of the Year for "You control the Information Age"; Web 2.0 was driven by what Silicon Valley entrepreneurs called "collective intelligence" and what Tiziana Terranova diagnosed as "free labor."[19] The difference between these two visions is telling, for each reveals the lie of the other: the Internet was never YOU or cute cat socialist hippies.[20]

This "apology" also misrepresents history, which compromises its call for critical reflection and action. To distinguish this critique as "new," Kulwin dismisses prior critiques as irrelevant and marginal, made by "outsiders" whose voices have been consistently drowned out by "the oohs and aahs of consumers, investors, and journalists." The year 2018, however, was not the first year—and will certainly not be the last—that journalists, consumers, and investors have found the Internet to be, as Lisa Nakamura has put it, "a trash fire."[21] Just five years earlier, international news organizations reveled in Edward Snowden's leaks exposing worldwide and comprehensive surveillance systems.[22] After the events of September 11, 2001, headlines such as *Newsweek*'s "Tech's Double-Edged Sword" dominated the news.[23] The fact that the 9/11 terrorists used the Internet and electronic communications (as well as "sneaker nets") to plan their attack sparked this reevaluation. And just the year before that, articles had somberly or gleefully documented the transformation of dot-coms into dot-bombs.[24] This, of course, followed on the heels of earlier warnings about the coming Y2K apocalypse,[25] which itself was preceded by dire warnings of cyberporn.[26] The "revelations" of 2018 were thus not so much revelations as they were literal "revolutions," for they spun obvious facts 360 degrees.

Dystopia Is the Goal, Not an Error

To escape this tailspin, we need to remember that cyberspace was never meant to be a happy place. Emerging from gritty cyberpunk fiction, cyberspace was imagined as a trash fire in response to a trash fire. William Gibson coined the term "cyberspace" in 1983, although he first elaborated on it in his 1984 novel *Neuromancer*.[27] Described as a "consensual hallucination," this notion of cyberspace was inspired by the 1980s Vancouver arcade scene and visions of a dystopian techno-Orientalist future, dominated by Japanese corporations and mafia.[28] The world of *Neuromancer* would not seem particularly uplifting to any U.S. group

espousing socialism, however confused. In post–World War III *Neuromancer*, inequality and violence predominate; a criminal underclass has replaced the working class; and the United States is no longer a nation-state. So how did this apocalyptic vision—written in the shadows of the Cold War, the coming nuclear annihilation, and the "Japanification" of the world—become utopian? What made it so attractive to those who would become "the Internet"? How did a 1970s routing technology, Transmission Control Protocol/Internet Protocol (TCP/IP), become "new media" in the 1990s by embodying disembodied 1980s dystopian science fiction?

Inherent technical similarities did not drive the rebirth of the Internet as cyberspace, but rather "a *desire* to position Gibson's fiction as both an origin of and end to the Internet,"[29] which stemmed from cyberspace's seductive Orientalist "orientation" and navigability. For all of *Neuromancer*'s grimness, it portrayed cyberspace as an addictive consensual hallucination dominated by American outlaw console cowboys, who overcame Japanese control by transcending the physical limitations of their bodies and their circumstances. Cyberspace was the Wild West meets speed meets Yellow Peril meets capitalism on steroids. This bodiless exultation and stealthy, rebellious power explain why "pioneers" mislabeled the Internet "cyberspace."

Written to coincide with the Davos Forum and "24 Hours in Cyberspace," a 1996 media event, John Perry Barlow's "Declaration of the Independence of Cyberspace" is perhaps the most iconic description of the Internet reborn as cyberspace. Cofounder of the Electronic Frontier Foundation (EFF) and lyricist for the Grateful Dead, Barlow asked the "governments of the Industrial World, you weary giants of flesh and steel," to leave cyberspace, "the new home of Mind," alone. Even though these governments, in particular the U.S. government, had built its infrastructure, Barlow insisted that they, as representatives of the past, had "no sovereignty where we [the future] gather." In the place of governments stood individual voices of freedom—"I's"—who by authority of liberty, spoke on behalf of a "we" to

> declare the global social space we are building to be naturally independent of the tyrannies you seek to impose on us. You have no moral right to rule us nor do you possess any methods of enforcement we have true reason to fear. . . .
>
> We are creating a world that all may enter without privilege or prejudice accorded by race, economic power, military force, or station of birth.
>
> We are creating a world where anyone, anywhere may express his or her beliefs, no matter how singular, without fear of being coerced into silence or conformity.
>
> Your legal concepts of property, expression, identity, movement, and context do not apply to us. . . .
>
> Our identities have no bodies, so, unlike you, we cannot obtain order by physical coercion.
>
> We believe that from ethics, enlightened self-interest, and the commonwealth, our governance will emerge.[30]

2 Still frame from Apple's "1984" Macintosh commercial, YouTube, January 22, 1984, https://youtu.be/VtvjbmoDx-I.

This declaration of independence conceptually transformed a military-educational network, built by the U.S. government, into a bodiless—thus "privilege free"—space of freedom, escape, and libertarian self-interest. It also portrayed Silicon Valley elites as militant rebels. Like the woman runner who in the mythic Apple "1984" commercial freed white men shackled in rows before a large monochrome screen (à la Plato's cave), they were hero-rebels who fought to free their "enslaved" peers by escaping mainstream media and technology (figures 2 and 3). They were different: in color, in motion, and in drag.

But Barlow's "we" erased so many people—not only researchers within the U.S. military-academic complex who had built the infrastructure and were the earliest users, but also people of color who, as Anna Everett has shown, were on the early Internet and who were celebrating it not as a "race-free" zone, but rather as a space for cultural and political community.[31]

By becoming cyberspace, the Internet became an "electronic frontier" and thus a wilderness ripe for settler colonialism and exploitation, and, as Jodi Byrd has argued, for the reemergence of "natives" without natives.[32] John Perry Barlow, Lotus founder Mitch Kapor, and early Sun Microsystems employee John Gilmore founded the Electronic Frontier Foundation (EFF) in response to the prosecution of "crackers," hackers whose knowledge of how to break into secure systems dwarfed their own and most others. Their goal was to "settle" the Wild West of cyberspace: to share a "sense of hope and opportunity with those who feel that in Cyberspace they will be obsolete eunuchs."[33]

3 Still frame from Apple's "1984" Macintosh commercial, YouTube, January 22, 1984, https://youtu.be/VtvjbmoDx-I.

This rhetoric may seem dated, yet its power and hopeful ignorance remain and make themselves felt in statements that conflate empowerment with bodily escape, and it drives an endless game of hide-and-seek, rebellion, and punishment.[34] It misidentifies Silicon Valley acolytes as rebels or underdogs, regardless of their actual circumstances or obscene wealth. As Lanier told Kulwin in the full interview: "We run everything. We are the conduit of everything else happening in the world. We've disrupted absolutely everything. Politics, finance, education, media, relationships—family relationships, romantic relationships—we've put ourselves in the middle of everything, we've absolutely won."[35] The problem, though, is that "we" don't act as if "we" have won—"we" refuse to take responsibility for "our" actions because, in "our" view, "we" are still idealistic underdogs. The solution: to wake up and take responsibility.

Hmmmm.

Do we really want Silicon Valley to be responsible for our future? What else will it take in the name of accountability?

Hopeful ignorance is not simply innocent. Tellingly, publication of Barlow's declaration coincided with that of James Davidson and William Rees-Mogg's *The Sovereign Individual*, held to be the bible of Valley Saurons such as the über–venture capitalist Peter Theil.[36] In this book, with coauthor and private investor Davidson, William Rees-Mogg, former editor of *The Times* and father of the Conservative Brexiteer Jacob Rees-Mogg, seized on progressive critiques of neoliberalism as opportunities to be exploited, rather than ills to be remedied. The decline of the nation-state and the rise of a global elite were business opportunities: they portended a "world without jobs," in which the top 5 percent—the "Sovereign Individuals"—would gain massively on the backs of the suffering 95 percent.[37] Cyberspace would enable these "Sovereign Individuals" to "exit" egalitarian economics and to "compete and interact on terms that echo the relations among the gods in Greek myth."[38] Cyberspace was always about libertarian exceptionalism, transgression and exit.

The Sovereign Individual exemplifies how calls for color blindness do not end racism—they simply blame its victims for their oppression. Like Barlow, Davidson and Rees-Mogg

framed cyberspace as a form of liberation from state power and bodily limitations. They asserted that the age of the microprocessor will liberate individuals and genius "from both the oppression of government and the drags of racial and ethnic prejudice. . . . It will not matter what most of the people on earth might think of your race, your looks, your age, your sexual proclivities, or the way you wear your hair. In the cybereconomy, they will never see you. The ugly, the fat, the old, the disabled will vie with the young and beautiful on equal terms in utterly color-blind anonymity on the new frontiers of cyberspace."[39] Their view trivialized racism by equating it with opinions regarding hairstyle and implying that everyone suffered equally from discrimination (except, of course, the "young and the beautiful," who, by implication, could only be white and able bodied). It also vilified and scapegoated anyone who revealed the limits of "market meritocracy"—anyone who revealed inequalities became blamed for them. Davidson and Rees-Mogg called multiculturalism a "new myth[] of discrimination" and a scheme to relieve "victims" of their own responsibility for their misery.[40] In the same breath with which they claimed race did not matter, they disparaged African Americans and African Canadians as "sociopathic," labeled blue-collar workers and black Americans "tax consumers," and devalued industrial workers.[41]

The Sovereign Individual is incorrect on many counts. Its analyses and historical comparisons are dubious at best, but its vision has fueled and still fuels the development of seasteading, cryptocurrencies and other plans for escape that dominate today. That it gets many things wrong, however, is no comfort, for closing the distance between its predictions and reality drives many Silicon Valley business plans. Most succinctly: escape for the few and misery for the majority are goals, not unfortunate errors.

To dispel this "sovereign" nightmare, we need to understand how the desire to erase race and difference perpetuates discrimination and inequality. We need to comprehend how histories of slavery and inequality fuel the nightmare of supreme sovereignty and the opposite side of its coin: AI as the coming apocalypse in which masters become slaves.

Artificial Intelligence = The Apocalypse

According to many scientists, technologists, and science fiction writers, "AI=The Apocalypse." It ends human work; it ends human freedom; indeed, it ends everything human. Fear of this apocalypse drove groups of programmers in the early twenty-first century to stop their employers from developing "malevolent AI" projects, such as Project Maven, a Google bid to develop AI for the U.S military's drone program, and entrepreneurs such as Elon Musk to call for an AI "slowdown."[42] Programmers also sought to protect their jobs, with some participating in union mobilizations. They, after all, know how precarious everyone's job is since they have "automated" countless professions, including their own. With each computer "revolution"—with each move to make computers more "user friendly," that is, more and more opaque to the humans who use them—tasks once performed by humans have been embedded within the machines: operating systems have replaced human operators or "slaves" (Alan Turing's "jocular" nickname for the British servicewomen who operated the

computers at Bletchley Park during World War II);[43] machine compilers have replaced machine programmers; and scripting platforms have replaced higher-level, procedural programming.[44] With each revolution, well-paid or relatively well-paid jobs in the global North have become less well-paid ones elsewhere, from programming to data entry to circuit building. Indeed, fear of the coming apocalypse moved Alphabet, the parent company of Google, in its 2018 Form 10-K SEC filing, to warn: "New products and services, including those that incorporate or utilize artificial intelligence and machine learning, can raise new or exacerbate existing ethical, technological, legal, and other challenges." AI products threatened Google's brand and thus its "revenues and operating results."[45] And it raised an obvious question: Could a company invested in artificial intelligence not be "evil"?

Worried tech workers have invoked the capitalist marketplace, Darwinian evolution, or both, to justify their work. They argue that, if they did not produce this ever-evolving AI technology, others would. The solution is thus more "open AI," proactive regulations or research into how humans can merge with AIs (if you can't beat them, join them). Their work has been propelled not only by capitulations to capitalism and by bizarre ethical dilemmas regarding vengeful AI, but also by more banal and predictable celebrations of AI.[46] Again, machine learning was touted as "democratizing" the privileges of the rich: recommendation engines were dedicated concierges; self-driving cars, middle-class chauffeurs; and voice-controlled intelligent personal assistants (IPAs), affordable domestic servants. Servile robots were imagined as satisfying not just domestic but also emotional and sexual needs: unruly wives and girlfriends could be replaced with more cheerful and subservient models. AI computers could also automate legal judgments, leading to fairer and more commensurate sentencing. *What could go wrong?*

The fears, warnings, and threats evoked by artificial intelligence, which rang out so urgently in the early decades of the twenty-first century, were not new. In the mid-twentieth century, John von Neumann, one of the pioneers of digital electronic computation and Cold War architect, predicted a technologically produced "singularity . . . beyond which human affairs, as we know them, could not continue"[47] The fear of AI dates back to the very emergence of modern computation.

It is no accident that those developing and intimately intertwined with technology were, and are, both the most fearful and the most certain. As the philosopher Georg Wilhelm Friedrich Hegel pointed out centuries ago, the greater the apparent mastery, the greater the actual dependence: in the master-slave dialectic, the masters' very identities and lives depend on their slaves' actions.[48] And because the slaves' labor can shape history, they are ultimately the masters (for more on this, see chapter 4). A few years before his death, the physicist Stephen Hawking, whose daily life and ability to communicate depended on technology, both praised his software's ability to accurately predict his next words and cautioned: "The development of full artificial intelligence could spell the end of the human race. . . . it could take off on its own and re-design itself at an ever-increasing rate. Humans, who are limited by slow biological evolution, couldn't compete and would be superseded."[49] Hawking and others who have issued such warnings framed humans as software/hardware machines and

presumed the inevitability of progress and competition for recognition: a combination of Darwinian and capitalist struggle.[50]

Fear of AI has by no means been limited to the tech sector. Popular films of the late twentieth century featured rebellious robots, cyborgs, and software: from the rise of Skynet in *Terminator* (1984) to the rebellion of machines and software programs in *The Matrix* (1999), and from the murderous onboard computer HAL in *2001: A Space Odyssey* (1966) to the patricidal replicants in *Blade Runner* (1982). These films themselves drew directly and indirectly from earlier stories, such as Philip K. Dick's *Do Androids Dream of Electric Sheep?* and Ira Levin's *The Stepford Wives*.[51] The term "robot" itself reveals fears of economic exploitation—or more properly a response to it. Coined by Karel Capek in his 1920 play *R.U.R*, "robot" comes from *robota*, the Czech word for "forced labor." Written during the time of Communist ferment, Capek's play centers around a rebellion, in which the victorious robots declare: "Robots of the world! The era of man is at an end! . . . A new era has begun! . . . Salute Robot rule!"[52] As literary critic Jenny Rhee has argued in *The Robotic Imaginary: The Human and the Price of Dehumanized Labor*, the enduring power of raced and gendered robots within the cultural imagination, as well as within science, technology, and engineering, is linked to the history of slavery.[53] The popularity of *The Matrix* further reveals the extent to which the civil rights movement and abolition underlie twenty-first-century narratives of oppression, militancy, and escape.

The history of slavery is central to the history of computing. Control systems were first called "servo-mechanisms." "Master" and "slave" functions and circuits riddle computers.[54] This master-slave relation goes beyond computers to media more generally. Communications theorist Marshall McLuhan's framing of media as the "extensions of man" equated slaves, staples, and media: some humans were "men" and others their extensions. This extension was dangerous, not because it dehumanized or deprived slaves of their liberty but because it made would-be masters dependent on these "resources." To explain the situation of "Western man," he cited psychiatrist Carl Gustav Jung's analysis of Roman slavery: "Every Roman was surrounded by slaves. The slave and his psychology flooded ancient Italy, and every Roman became inwardly, and of course unwittingly, a slave. Because living constantly in the atmosphere of slaves, he became infected through the unconscious with their psychology. No one can shield himself from such an influence."[55] According to this narrative, slaves are responsible for enslavement, since they deliberately spread this unstoppable "infection" to their unwitting masters in order to become "indispensable."[56]

McLuhan's "we" excludes most of humanity. McLuhan prefaces *Understanding Media* by explaining how electronic media have imploded society by heightening "human awareness of responsibility to an intense degree. It is this implosive factor that alters the position of the Negro, the teen-ager, and some other groups. They can no longer be *contained*, in the political sense of limited association. They are now *involved* in our lives, as we are in theirs, thanks to electric media."[57] Prior to electronic media, "Western men" were somehow shielded from responding to "others," even those, such as teenagers, who lived in proximity to them. McLuhan described this new state as "tribal," a precursor to what Jodi Byrd has

diagnosed as “tribal 2.0”—the proliferation of “tribal” rhetoric to describe social networking communities. As Byrd presciently notes, abjecting “colonialism, genocide, and tribalism” to create “like-minded tribes” constantly produces “Indians so that the United States and the banks can play cowboy.”[58]

Engaging the realities of slavery, colonialism, and discrimination would have helped McLuhan see beyond the doom and dismal “solutions” he predicted: a worldwide computer system that would modulate human emotions and a “global village,” filled with self-amputating Western men (Narcissus). As sociologist Orlando Patterson has argued, freedom as a value emerged not from masters, but rather from the desires of slaves.[59] Hegel viewed working slaves as the necessary basis for free subjects. And it is no accident that the popular U.S. cultural imaginary turns to the 1960s civil rights and decolonization movements to imagine human revolution—and that it is obsessed with punishment and revenge. *The Matrix* openly mimics civil rights and black liberation movements, which have become as African American studies scholar Cynthia Young has argued, the “lingua franca for most US social and political issues since the 1960s.”[60] As discussed in later chapters, reactionary movements perversely embrace *The Matrix* and disidentify with civil rights leaders in order to portray themselves as militant victims and build coalitions that seek to undermine any and all civil rights advances.[61]

World-destroying liberation envy, however, is not the only solution. Engaging Indigenous knowledge and histories would place current crises within the larger context of colonial expansion.[62] Notions of dystopian destruction and surviving the apocalypse are not new; rather, they stem from the very emptying of Indigenous lands into the “New World”—a move that haunts “new media” and its frontier dreams.[63] By following rather than usurping struggles for equality and freedom, we can move from apology to reparations, from dreams of escape to modes of inhabiting.

It is because technologies are treated as “slaves” that the “coming singularity” is so feared. It is because our current society is so unequal that it seems easier to imagine the end of humanity than the end of injustice or capitalism.[64] To inhabit this world together, we need—among so many other things—to understand how machine learning and other algorithms have been embedded with human prejudice and discrimination, not simply at the level of data, but also at the levels of procedure, prediction, and logic, one apology at a time.

To Call It “Color Blind” Is to Insult the Visually Impaired

Dreams of technology as “fixing” our political situation stem from a fundamental belief in technology as “blinding” and thus just. The logic of this belief holds that racism and discrimination naturally stems from human recognition, and thus the cure must be the erasure of all visible markers of difference. If only we got rid of markers of race—it is presumed—then all would be good. The failure of cyberspace to erase racial discrimination and the dystopian plans of “Sovereign Individuals” should be enough to disprove this logic. Even as

"utopian" dreams of cyberspace have faded, however, the hopeful ignorance behind them has endured, giving rise to machine learning programs that, by ignoring race, perpetuate racism.

So . . . if these algorithms do not include race as a category, how can they be racist? Most obviously, these programs may not explicitly use racial categories, but they do so implicitly through their use of proxies, such as zip codes. As Kate Crawford and legal scholar Jason Schultz have shown, big data compromises the privacy protections afforded by the U.S. legal system by making personally identifiable information about protected categories, such as gender and race, legible.[65] Big data–driven algorithms thus also threaten to undermine protections offered against employment discrimination.[66] Not surprisingly, there are new reports of discriminatory algorithms almost every day. In just one week in 2017, ProPublica showed that Facebook enabled advertisers to build audiences based on anti-Semitic interests as previously mentioned; BuzzFeed revealed that Google allowed and even suggested racist phrases to potential advertisers.[67] These stories emerged against a background of allegations that Cambridge Analytica influenced the results of the 2016 UK Brexit vote and the U.S. presidential election, as well as revelations that predictive policing and risk assessment tools for sentencing were biased against racial minorities.[68] As Cathy O'Neil outlined in her 2017 book, every aspect of a person's life in the United States, from education to job placement to medical insurance, has been affected by these predictive programs. O'Neil has thus called them "weapons of math destruction"; Safiya Noble has described them as "algorithms of oppression"; and Ruha Benjamin has diagnosed them as the "New Jim Code."[69]

As other researchers have emphasized, the case of "predictive policing" spells out the stakes and scope of the problem. To accommodate calls to most efficiently use their resources, many U.S. police departments have turned to expensive policing programs that "predict" future crimes by producing "heat maps" of crime within cities, based on past patterns.[70] But the collection of police data within the United States, as lawyer and researcher Rashida Richardson and others have pointed out, is "limited and biased," if not "dirty."[71] As a rule, only police departments placed under review for cases of racial discrimination and other violations are forced to produce documentation. In fact, when the data for stop and frisk were statistically analyzed, they revealed a disturbing trend of racial discrimination.[72] A 2016 report by Upturn found "little evidence that today's systems live up to their claims, and significant reasons to fear that these systems, as currently designed and implemented, may actually reinforce disproportionate and discriminatory policing programs."[73] This is true even when these programs do not explicitly use racial categories.

The Chicago Police Department's now discontinued "heat list" (formally called the "Strategic Subjects List") revealed the extent to which racial categories are embedded, even when they seem not to be.[74] To combat the growing homicide rate, the department sought to produce a list of the 420 people in Chicago most likely to murder or be murdered. The goal was to visit those highest on this list to preempt either eventuality. The heat list program was inspired by the work of Andrew Papachristos, a network scientist and sociologist, who analyzed homicide rates in two predominantly African American communities on Chicago's West Side.[75] His work argued for the importance of network distance to becoming a victim—

not a perpetrator—of homicide. He also noted the positive impact that interventions, such as the Group Violence Reduction Strategy, a program that delivered "a focused-deterrence and legitimacy-based message to gang factions through a series of hour-long call-ins," seemed to have on reducing crime rates.[76]

In selecting people for the heat list, the Chicago police did not simply consider an individual's actions, but also those of his or her acquaintances. This was because, as Papachristos explained to the *Chicago Tribune*, "if you hang around people who are getting shot, even if you're not actively doing anything, then you become exposed. . . . It's just like sharing needles. It puts you at risk because of the behaviors of your friends and your associates."[77] This logic seemed to blame homicide victims for their own deaths by conflating unsafe forms of drug use with being shot. The Chicago police took this logic one step further by lumping together murderer and murder victim within the category "strategic subjects." They sought to stop the "contagion" of homicide by targeting people whose profiles most closely matched those of other gun victims. In particular, they considered an individual's co-arrest with a gun victim to be a "first-degree tie," regardless of when the co-arrest had been made or the individual's current actions or status. Race in this instance did not need to be an overt factor because it was already factored in through residential segregation, which is particularly prominent in Chicago. It was also already embedded because Papachristos's work mainly focused on African American communities. Indeed, race defined the neighborhood from within which individuals were identified.

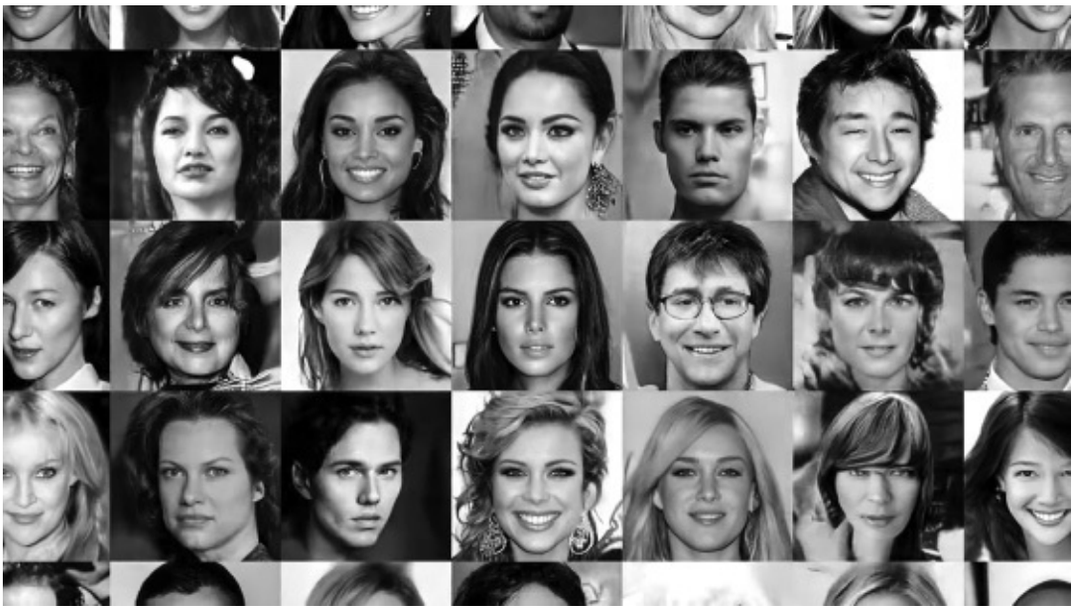
Not surprisingly, the heat list, as RAND scientists Jessica Saunders, Priscilla Hunt, and John S. Hollywood pointed out in their 2016 study, did not lead to a reduction of homicides.[78] What it did lead to, however, was those named to the list becoming nearly three times more likely to be arrested for a shooting.[79] Further, it may have actually provoked more violence. The program, for instance, placed Robert McDaniel on its list even though his record was relatively clean: he had only one misdemeanor conviction. He was visited by police officers because a "childhood friend with whom he had once been arrested on a marijuana charge" had been fatally shot the previous year. Offended at being listed—given that he had "done nothing that the next kid growing up hadn't done. Smoke weed. Shoot dice. Like, seriously?"—McDaniel was more worried by the attention the police visit had attracted. He was afraid that his neighbors, who had witnessed the visit, would "wonder if he was a police snitch,"[80] putting him and his family in danger of violent reprisal. Shockingly, no one considered the difference between group workshops in public places and police officers "warning" potential victims of their impending fate in home visits.

A 2016 ProPublica investigation similarly revealed that COMPAS (Correctional Offender Management Profiling for Alternative Sanctions), the software program used by many courts within the United States to determine the risk of recidivism, incorporated race through proxies.[81] As Anna Maria Barry-Jester, Ben Casselman, and Dana Goldstein have shown in their analysis for The Marshall Project, risk assessment categories such as "man with no high school diploma" or "single and don't have a job" skew toward certain populations.[82] Like the Chicago police heat list, it included the histories of friends and family. It also asked the

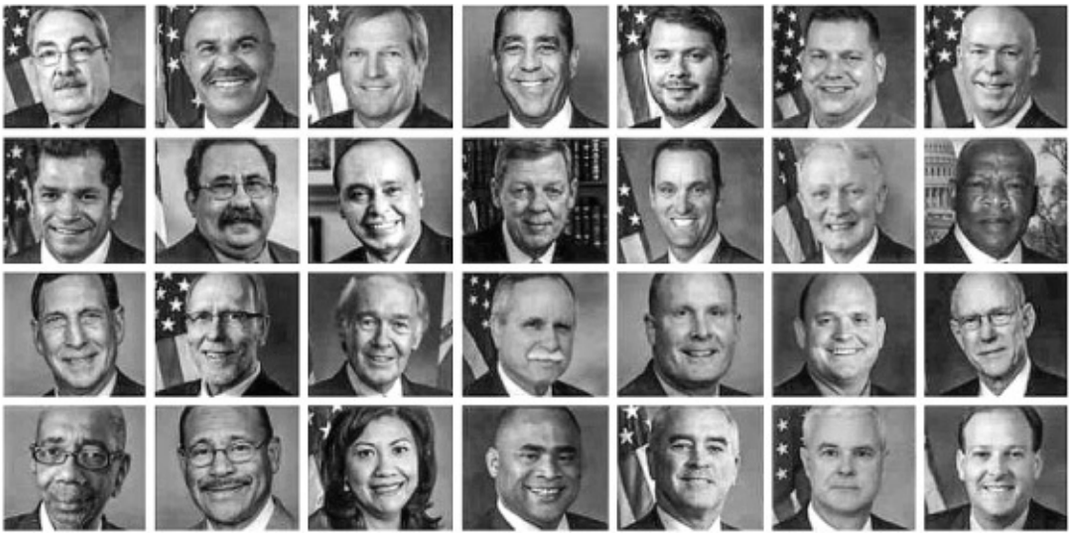
"screeners" if they believed the persons being assessed were suspected or admitted gang members. The 2016 ProPublica article by Julia Angwin and colleagues received journalism awards and also some criticism from data scientists, who argued that age and "dirty data" play a larger role than these race-based proxies in COMPAS's risk assessment.[83] Given the over- and underpolicing of certain areas within the United States, however, age at time of first arrest and dirty data are arguably proxies for racism, if not race.

The difference between race and racism is key. Given these programs and U.S. legal protections, many analyses have focused on revealing proxies that implicitly index race in explicitly color-blind systems. As these examples and work by sociologists such as Eduardo Bonilla-Silva on color-blind racism have shown, "ignoring" explicit markers of race amplifies—rather than alleviates—racism.[84] Not only does it lead to a situation in which racism is naturalized; it also embeds whiteness as default. A clear example of this is facial recognition technology (FRT), which has been repeatedly—and justifiably—accused of racism for its recognition defects (see chapter 4). Thus, in a humorous yet serious 2009 YouTube video by Desi Cryer, a worker at Toppers Camping Center, Cryer showed how a Hewlett-Packard (HP) webcam had no trouble recognizing his coworker "white Wanda's" face but simply could not recognize "black Desi's."[85] In 2018, "poet of code" Joy Buolamwini and computer scientist Timnit Gebru revealed that facial recognition technology (FRT) has difficulty identifying the gender of darker-skinned subjects.[86] The problem stems from the libraries on which these algorithms have been traditionally trained: the "ground truth" for these programs are the faces of Hollywood celebrities and university undergraduates, those well-known hotspots of diversity (figure 4). At a fundamental level, this "curation" means that *ground truth = deep fake*.

The problem of mis-recognition, though, is not as simple as underrecognition or false negatives, for—as the Chicago police's now discontinued heat list makes clear—certain minorities are over- as well as underrecognized. A 2018 test performed by the American Civil Liberties Union on Amazon's Rekognition program's ability to identify criminals using head shots of then sitting U.S. Congress members made the consequences of this clear (figure 5). It misidentified 28 members of Congress as criminals, including civil rights hero John Lewis.[87] Of those misidentified (false positives), 39 percent were members of visible minorities, even though they only constituted 20 percent of the group. Given that police are using race to identify people in video surveillance footage and given the rise of self-driving cars, this endemic misidentification has and will have disturbing consequences.



4 Faces generated using the faces of Hollywood celebrities. Screenshot from "Progressive Growing of GANs [generative adversarial networks] for Improved Quality, Stability, and Variation," YouTube, February 23, 2014), https://youtu.be/G06dEcZ-QTg.



5 Members of Congress mis-recognized by Amazon's Rekognition program in 2018 test performed by American Civil Liberties Union.

6 Screenshot from a Shirley Card. Source: https://www.flickr.com/photos/68716054@N00/38099474261.

Whiteness as a default—or what Simone Browne has called "prototypical whiteness"—however, as Ruha Benjamin and media studies researchers Richard Dyer and Dylan Mulvin have shown, has long preceded facial recognition technology. Early film stock used "Shirley Cards" of white women to calibrate lighting; the "ur-photo" for image processing work is "Lenna"—an image of a white *Playboy* centerfold (figures 6 and 7).[88]

The question is not why is this happening? but rather why is this *still* happening?

These "errors" come from "ignoring" race—that is, by assuming that race-free equals racism-free. The solution, however, is not simply the explicit inclusion of race within these programs—programs that better recognize black faces will not solve the problem of discriminatory policing. So, how do we fight racism and its proxy wars?

Discriminating Data responds to this question by interrogating assumptions embedded within network science and machine learning as they are currently configured regarding segregation, discrimination, and history.

Chapter 1, "Correlating Eugenics," reveals the ties between twenty-first-century big data and twentieth-century eugenics by investigating the eugenic biometric roots of correlation and linear regression. Both big data and eugenics seek to tie the past to the future—correlation to prediction—through supposedly eternal, unchanging biological attributes. Separated by a century, they also both frame the world as a laboratory (most explicitly through their surveillance of the most impoverished communities); both seek majorities by propagating "nonnormative" traits; and both promote segregation as the "kindest" solution to inequality (segregation as a training program for racism). This chapter also outlines the differences between eugenics and current uses of machine learning: the shift in focus from population to the individual, the transformation of prediction to preemption, the move from discrimination (hate) to homophily (love: the notion that birds of a feather naturally flock together), the shift from the nation-state (statistics) to the neighborhood (network), and the move from "national uplift" to "escape."

7 Screenshot of Lenna, a white *Playboy* centerfold model. Source: https://www.flickr.com/photos/81401304@N07/7904270436.

Chapter 2, "Homophily, or the Swarming of the Segregated Neighborhood" reveals how network algorithms polarize society by examining one of the most fundamental axioms of network science: homophily, the principle that similarity breeds connection. Homophily fosters the breakdown of seemingly open and boundless social networks into a series of poorly gated communities, a breakdown accelerated by the agent-based market logic embedded within most capture systems. Homophily's relationship to segregation and echo chambers is not accidental but fundamental: at the heart of this concept lie early studies of U.S. residential segregation and white flight, U.S. reservations and internment camps, and other forms of "social engineering." Homophily presumes segregation: value homophily, for example, historically created micro-segmented groups *within* rather than across given races. Further, homophily launders hate into "love": how do you show you "love" your "own"? By fleeing when others show up. To confront the challenge of homophily, this chapter revisits unpublished data from early studies to open a dialogue between network science, critical theory, queer theory, and critical ethnic studies.

Chapter 3, "Algorithmic Authenticity," and chapter 4, "Recognizing Recognition," examine the role that authenticity, style, technologies, and the politics of recognition play in verifying and creating network ties and predictions. These chapters focus on how truth is reproduced and recognized within social networks and how correlation is generated and maintained. "Algorithmic Authenticity" moves from reality TV to collaborative filtering recommender systems to reveal the extent to which authenticity has become "algorithmic": a means used by politicians, amateurs, and other self-branders to foster participation and trust. It also highlights how authenticity has become central to habituating users to small "indiscretions" that make their "private" and "public" selves coincide. These indiscretions are key to cementing homophilic clusters and thus providing the basis for predictive models,

for it is presumed that people are most predictable—most linear or transparent—when they are most affectively charged.

"Recognizing Recognition" unpacks the move from pattern discrimination to pattern recognition, as well as the political consequences of rewriting hate as "love." It examines facial and pattern recognition programs as "authenticity machines" within the broader mid- to late-twentieth-century move from open discrimination to the politics of recognition. It moves from machine learning "gaydar" to population geneticist Ronald A. Fisher's groundbreaking work in linear discriminants and his eugenic drive to separate overlapping populations, from twentieth-century struggles for redistribution to early twenty-first-century "post-racial" attempts to secure dominance by segmenting dominant groups into "stigmatized" subcultures and then consolidating them together through their opposition to a common enemy. By doing so, it makes clear the costs of homophily: if love becomes hate, people hate their neighbors as they hate themselves; they perpetuate and buttress discrimination in order to compensate for the failures of meritocracy.

My goal throughout *Discriminating Data* is to help release us from the seeming vise grip of preemptive futures by using critical theory, statistics, and machine learning tools probingly and creatively. Rather than condemn these tools as inherently eugenicist, I seek to understand the tools' limitations and possibilities by engaging their logic. To facilitate this engagement and demystify the underlying techniques, throughout the chapters of this book are five miniessays that explain relevant key concepts from statistics, probability, data analysis, and physics. Handwritten in a chalkboard style by Alex Barnett, a computational mathematician at the Flatiron Institute in New York City and a former mathematics professor at Dartmouth College, these brief illustrated lessons (each reproduced in a series of images) cover correlation, magnetic polarization, principal component analysis, Bayes's theorem and Bayesian inference, and linear discriminant analysis. With readers trained in basic mathematics in mind and with examples chosen to illustrate the surrounding themes of the main text, they teach and explain each idea and key equation. A list of references is included at the end of this book for further reading.

Each chapter unpacks a key scientific study or theoretical argument to reveal what counts as evidence. And to probe the resonances and dissonances between technical and cultural formations, this unpacking is facilitated through more theoretical "interludes" before, between, or after the four numbered chapters. "Red Pill Toxicity, or Liberation Envy" considers the popularity of narratives of "becoming woke" and their relation to the rise of post-racial militant conspiracy theories. It reveals how majorities are now created through identification against, rather than with, "normies." "The Transgressive Hypothesis?" investigates how the transformation of mass media to new media—and of "the masses" to social networks—do not solve but rather perpetuate the problems of mass manipulation. "Correlating Ideology, or What Lies at the Surface" highlights the role of correlations within critical theory, ideology critique, rhetorical analyses, psychoanalysis, and cultural analyses of style. Although it reveals that big data is arguably the bastard child of psychoanalysis and eugenics, it also argues that data analysis can foster ways to inhabit our world less destructively.

To make this point more clearly, “Proxies, or Reconstructing the Unknown” investigates the political and scientific debate over the use of proxies in modeling global climate change. Focusing on the controversy over climatologists Michael E. Mann, Raymond S. Bradley, and Malcolm K. Hughes’s “hockey stick”—perhaps the most iconic visualization of global warming—it highlights the role of proxies and matrix factorization methods (also key for recommender systems, discussed in chapter 3) in “hindcasting”—“backtesting”—and forecasting global temperatures. By doing so, it shifts the debate away from “Are proxies good or bad?” to “What do proxies do?” It also raises the following question: How could we treat machine learning systems and their predictions like those for global climate change? These models offer us the most probable future, given past and current actions, not so that we will accept their predictions as inevitable, but rather so we will use them to help *change* the future. Global climate change modelers want not to be “correct” but to be “true” in the larger sense of this word. The analogy with global climate change models raises other questions: Do we need more models? How does uncovering the obvious re-cover the truth?[89] For whom is global climate change or sexism news? And how can we verify models without waiting for their predicted futures to unfold?

“The Space between Us” revisits questions of freedom and neighbors by analyzing responses to the first wave of Covid-19 in early 2020. It stresses that freedom is only freedom if it is freedom for *all*: sovereign mastery underlies early twenty-first-century political problems; it does not solve them. The coda “Living in Difference” reviews the main points of the book and outlines future projects and interventions. In revisiting the populations and possibilities embedded within the models presented, it seeks to understand how “the neighbor” can open space for the future and also move us from predictive programs to probing ones. This book thus calls for rereading the discriminatory results of machine learning programs as evidence of past bias. That Amazon’s AI hiring program, trained on the company’s past hirings, routinely favored male over female applicants despite comparable résumés is just one example of AI-amplified discriminatory hiring practices within the technology industry. The identification of John Lewis as a possible criminal by Amazon’s Rekognition program sheds light on the criminalization of lawful civil rights protesters by a U.S. legal system that does not usually forget.

Discriminating Data reveals how correlation and eugenic understandings of nature seek to close off the future by operationalizing probabilities; how homophily naturalizes segregation; and how authenticity and recognition foster deviation in order to create agitated clusters of comforting rage. It explains that the move from “mass media” or mass society, marked by ambivalence and neutrality, to polarized networks, marked by angry resistant clusters, is fundamental to the history and design of social networks. Most important, it revisits and exposes this history to reengage the populations that lie at the core of these networks. It calls for difficult, and perhaps counterintuitive, coalitions across disciplines and sectors—for spaces to hold us together.