Race and Robots

J. Jesse Ramírez

➡ *For additional information about this article*

Race and Robots

J. Jesse Ramírez

Race after Technology: Abolitionist Tools for the New Jim Code. **By Ruha Benjamin. New York: Polity, 2019. 286 pages. $64.95 (cloth). $19.95 (paper).**

The Robotic Imaginary: The Human and the Price of Dehumanized Labor. **By Jennifer Rhee. Minneapolis: University of Minnesota Press, 2018. 240 pages. $108.00 (cloth). $27.00 (paper).**

Surrogate Humanity: Race, Robots, and the Politics of Technological Futures. **By Neda Atanasoski and Kalindi Vora. Durham, NC: Duke University Press, 2019. 256 pages. $94.95 (cloth). $24.95 (paper).**

A neglected motivation behind the white supremacist terrorist attack in El Paso, Texas, on August 3, 2019, was the perpetrator's racialized anxieties about automation. The bleak future predicted in the killer's online manifesto reads like a scene from the dystopian film *Elysium* (2013): working-class whites will not be able to reap the benefits of the new automatic technologies because they will be overrun by poor, unemployed, government-dependent Latinxs. This is certainly not the future promised by today's dominant automation discourse, the business science fiction of the "second machine age" and "rise of the robots"; nor is it the future affirmed by "accelerationists" and other leftist thinkers who claim that full automation—combined with universal basic income—will usher in socialism or "luxury communism." But the El Paso terrorist's combination of the myth of white genocide and speculations about the radical automation of work is not as peculiar as it seems. For robots, as products of US history and culture, are cast from a substance that is simultaneously more immaterial and more real than their sensors and actuators: race. Recent work at the intersections of critical race and ethnic studies, decolonial science and technology studies (STS), critical code studies, feminist science studies, and literary and film studies suggests that what is at stake in automation is the technical reproduction of the US racial formation.

There have been two major waves of automation discourse in the twentieth-century United States: the Depression-era debates over mechanization (1929–40) and the hopes and anxieties of the short American Century (1945–73), when the term *automation* was coined. (One could add briefer moments in the 1980s and early 1990s.) Amy Sue Bix's cultural history of automation, which is weighted toward the 1930s; David F. Noble's social history of numerical control; Shoshana Zuboff's study of the computerization of pulp mills and office work; Ruth Schwartz Cowan's work on household technology; and Venus Green's book on race and the Bell System's switch to direct dial are indispensable resources for understanding these periods.[1] But now, in the wake of the Great Recession of 2008, we are in a third era. While previous conjunctures coalesced around Fordist mass production and visions of the automatic factory, ours is the era of collaborative robots and AI, IBM's Watson and DeepMind's AlphaGo, Siri and Alexa, self-driving cars and military drones. These and other automated objects and systems are allegedly encroaching on all labor that is repetitive and predictable, be it manual or mental, and bleeding into the infrastructures of social life. Capitalism's utopians claim that the coming automation wave will lift all boats, raise the standard of living, and free us from drudgery, while the dystopians foresee a jobs apocalypse. The public appears to agree with the latter. A recent poll by the Pew Research Center indicates that 82 percent of Americans believe that robots will take over "much of the work currently done by humans" within the next thirty years, while 76 percent think this transformation will "likely" increase class inequality.[2]

Pew's language reveals the centrality of the human ("work done by humans") in the contemporary automation debate. This is the same figure shown on the covers of two of the most widely cited works of business science fiction, Erik Brynjolfsson and Andrew McAfee's *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies* (2014) and Martin Ford's *Rise of the Robots: Technology and the Threat of a Jobless Futures* (2015).[3] Both book covers depict abstract, featureless humanoid forms that are supposed to represent the subject that is liberated or imperiled by automation's latest breakthroughs: the human *as such*. But the El Paso terrorist understood that this subject's generality is simply the ordinariness of white masculinity as a hegemonic norm and implied addressee of utopian and dystopian techno-speculations. The white supremacist's anger registered what Neda Atanasoski and Kalindi Vora's *Surrogate Humanity: Race, Robots, and the Politics of Technological Futures* describes as "white loss" (2), a racial narrative of technological freedom turned on its head so that the subject becomes object and a particular type of humanity imagines itself robbed of the right to be "great again" at the expense of dehumanized

others—others for whom technologies can substitute because they are already regarded as mere things of utility. If liberal political commentators have tended to use automation as a foil to Donald Trump's border policies—they claim that robots, not immigrants, are to blame for lost jobs—Atanasoski and Vora's analysis of the Latinx science fiction short "M.A.M.O.N." (2016), which depicts Trump as a giant robot guarding the border, suggests that the border *is* an automation policy. The whiteness of automation can thus be understood as the sociotechnical dimension of what Aziz Rana has called the "two faces of American freedom," or the entanglement of settler colonial and republican notions of self-rule with claims that slaves, Native Americans, Asians, and Latinxs are outsiders whose exclusion and internal domination are necessary for settler liberty.[4]

Atanasoski and Vora's major intervention in the automation debate is their argument that automation imaginaries are shaped by liberal humanism and the racial hierarchies embedded in it. The labor of dehumanized others whose "bodies are meant solely for work" (33) is the condition of possibility of the (white, male) liberal subject's freedom and capacity to "feel human." Treating automation as "technoliberalism," the authors claim that robots and other automated systems function in racial capitalism as technical surrogates for the slaves, servants, immigrants, housewives, and colonized others who have long been tasked with ameliorating the lives of more fully human subjects. While technoliberalism promises that automation will create postracial, postlabor freedom for all, Atanasoski and Vora claim that there is no liberal humanism without surrogacy: "*the liberal subject is an effect of the surrogate relation*" (5). The authors provide a wealth of diverse evidence for their view, but perhaps none is starker than Rastus the Mechanical Negro, a novelty robot created by Westinghouse Electric Corporation in 1930. Dressed in overalls, the dark-skinned robot spoke and bowed and wore an apple on its head that exploded when activated by a beam of light shot from an arrow. While the apple and arrow were ostensibly a whimsical reference to the William Tell legend, in which Tell rescues his son from cruel punishment, Westinghouse made a subtle change: Rastus responded to the exploding apple not with the gratitude of Tell's son or with anger but "with an exclamation of dismay."[5] Thus Rastus not only exhibits the directness with which racialized unfreedoms are built into robots but also shows how the wish for liberation from degrading labor is expressed as the desire for a new, more docile slave whom the "human" can playfully torment without the threat of retribution. Conversely, the reaction to a more recent video of the DARPA-funded Atlas robot showcases the racialized anxieties surrounding "killer robots." Released by Boston Dynamics,

the robot's developer, the video shows white employees hitting the humanoid robot with hockey sticks in order to demonstrate its ability to pick itself up. Atanasoski and Vora observe that the scene "startlingly conjures . . . the act of a machine standing up to a human master" (146). When YouTube users commented that the abused robot would soon start killing "humans," they inadvertently captured how the technoliberal dream of a new and improved slavery is simply the obverse side of the master's dread of slave revolt.

One strength of *Surrogate Humanity* is the range of technological discourses, objects, and processes in which the authors elucidate the logics of technoliberalism. Atanasoski and Vora demonstrate that the collaborative robot Baxter is a surrogate for Chinese labor that liberates the US worker from a monotony that is presumed to be proper over "there," in the imperial field of outsourced labor, but not "here." But since technical surrogacy remains an incomplete project, technoliberalism helps the liberal subject "feel human" not only by replacing but also by *concealing* the nonhuman other's still laboring body. For example, Alfred Club (since renamed Hello Alfred), an app-mediated hospitality service, offers emancipation from gendered and racialized social reproduction by enabling users to summon butler-like workers ("Alfreds") to do their laundry and perform other domestic labors without ever coming into contact with the user. "The innovation is the interface (the user interacts with a platform rather than a person), and thus the platform enables the fantasy that technology is performing the labor, though in fact it is being done by human beings" (93). In their chapter on artificial affect, Atanasoski and Vora argue that the designer of Kismet—a social robot with large, expressive eyes—was influenced by Charles Darwin's ideas about the unrestrained and primitive display of universal emotions in "savage" races. Designed to be fully emotionally transparent to human users, Kismet is a kind of "savage," a racialized exteriority by which the human subject becomes aware of its refined capacity for an interiority that the nonhuman other lacks. Chapter 5 presents drone warfare as a fantasy of "unmanned" colonial power, or colonialism without colonizers, while chapter 6 contends that human rights campaigns against killer robots affirm the liberal subject's capacity to feel empathy over the machine's amoral decisions over life and death—an ultimately pro-war position that presumes the colonial morality of killing less than human life. In all these cases, technoliberalism aims not to overcome the domination at the heart of liberal humanism but to perfect it by technical means. The way out of technoliberalism lies not in its expansion and greater diversity, which the authors critique in their epilogue on sex robots and pseudo-feminist AI, but through a rigorously antiracist and decolonial

feminism that "politically seeks to disrupt the categories of use, property, and self-possession rather than redress through inclusion" (196).

Sharing core themes, theoretical frameworks, and objects of analysis, *Surrogate Humanity* and Jennifer Rhee's *The Robotic Imaginary: The Human and the Price of Dehumanized Labor* make for good companion books. Since Rhee's text is more of a literary and filmic study, reading it together with *Surrogate Humanity* is like gaining a new perspective on a gemstone by turning it and watching light reflect from a different angle. Rhee also approaches robots and automated technologies through the figure of the human and its constitutive exclusions. Anthropomorphism is the core of the robotic imaginary, which Rhee defines as the "shifting inscriptions of humanness and dehumanizing erasures evoked by robots" (5). Marvin Minsky, the so-called father of AI, defined AI as "the science of making machines do things that would require intelligence if done by men" (10). As Minsky's gendered language indicates, the central metaphor of robotics and AI—their constant comparability to human skills—embeds norms of humanness and nonhumanness in technologies on the assumption that the human is already familiar and fully legible to particular subject positions. Drawing on the critique of liberal humanism in the work of Lisa Lowe and Denise Ferreira da Silva, who are also prominently cited in *Surrogate Humanity*, Rhee claims that the hegemonic norm in robotics is the liberal subject whose freedom requires the unfreedom and dehumanization of others, particularly as regards their labor. But Rhee is ultimately more invested in salvaging the human than Atanasoski and Vora are. Against the transparency of the liberal subject, Rhee affirms the opacity of the human. Codas to each chapter valorize robotic literature and art that, on the one hand, reckon with the histories of enslavement and colonialism that constitute the liberal subject, and on the other, represent "an understanding of the human through unrecognizability, difference, and unfamiliarity, rather than recognition, knowability, and givenness" (5).

The Robotic Imaginary contains four studies of the dialectic of anthropomorphism and dehumanization in care labor, domestic labor, emotional labor, and drone labor. In chapter 1, Rhee begins with an often-overlooked moment in Alan Turing's "Computing Machinery and Intelligence," the 1950 essay that introduces the Turing test for machine intelligence. Turing suggests that creating AI can be modeled on the education of a human child. Turing's observation implies that machine intelligence approaches the human when it is an outcome of gendered care labor and acquires the capacity to replicate this care in the service of others. The chapter explores how Turing's robotic

imaginary shapes actual and fictional AIs that are gendered female: Joseph Weizenbaum's AI therapist ELIZA (1966), Helen in Richard Powers's novel *Galatea 2.2* (1995), and Samantha in Spike Jonze's film *Her* (2013). In all three cases, Rhee shows the continuity between the devaluation of women's care labor and the ways in which automated systems act as disembodied, caring women (an argument that also explains why the digital assistants Siri, Alexa, and Cortana are female). The following chapter on domestic labor ingeniously connects the Cold War metaphors of containment and the "closed world," the computer science paradigm of symbolic AI, and Ira Levin's novel *The Stepford Wives* (1972). While symbolic AI reduces human intelligence to stereotypical representational models of the world that privilege the perspectives of highly educated, middle-class, white male computer scientists, *The Stepford Wives* is about men who kill their wives and replace them with domestic robots, thereby reconstructing and purifying symbolic AI's closed world of gender normativity. Reading Alex Garland's film *Ex Machina* (2014) as a reimagining of *The Stepford Wives*, Rhee notes that the film reiterates the novel's figuration of emancipation from domesticity in terms of elite white feminism. Ava, the white female robot in *Ex Machina*, escapes her closed world and enters human society by compelling Kyoko, a mute Asian female robot, to sacrifice her life in an attack on their male slave master. The scene crystallizes the robotic imaginary's racial and class hierarchies: "white, middle-and upper-class women's freedom and happiness is achieved by the exclusion, if not exploitation, of the very women of color and white working-class women who make possible their liberation from a closed world" (90). What is conspicuously absent in these cultural artifacts is shame, which chapter 4's reading of Philip K. Dick's novels *We Can Build You* (1972) and *Do Androids Dream of Electric Sheep?* (1968) presents as an ethical critique of demands to perform falsely universalized and standardized human emotions.

Rhee's final chapter on drone labor, the book's political and ethical centerpiece, exemplifies the rebel humanism that propels *The Robotic Imaginary*. Although military drone labor is feminized—pilots are sometimes portrayed as mere video gamers who lack the soldier's masculine strength and courage— drone labor also requires the repression of care. Pilots routinely describe their dehumanized targets as insects. Materializing and enacting decisions that certain lives are not worth mourning, as Judith Butler puts it, drones draw a bright line between the humanity that is meant to benefit from the automation of war and others who stand in, or adjacent to, the drone's crosshairs. Yet Rhee is skeptical of the attempt to humanize the victims of drone strikes through empathy. James Bridle's *Drone Shadow* (2012–17) series, which consists of large

outlines of military drones in public spaces, is meant to shock the Western subject into projecting itself into the position of the victim in the periphery. But this empathetic subject remains the liberal human who falsely imagines that dehumanizing violence happens only outside the imperial center and thus not in the very social hierarchies that generate the subject's feelings of relative safety. Rhee instead celebrates Bridle's *Dronestagram* (2012–15) and Teju Cole's *Seven Short Stories about Drones* (2013). *Dronestagram* consists of images taken from the perspective of the drone itself, and since "drone vision . . . is incompatible with the human" (161), the images absolutely block identification and empathy. Cole's short stories also circumvent empathy by inviting the reader's *mourning* for a victim who is no less human because illegible and unfamiliar to liberal subjectivity. Echoing Butler's claim that "the human comes into being, again and again, as that which we have yet to know," Rhee finally calls for "reconfigur[ing] the robotic imaginary to embrace difference and the unknown, thus refusing sameness, resemblance, or familiarity to the Western Subject as the defining characteristic of the human" (176).

Liberal humanism has its own mechanism of transcendence, namely, technologies that replace human bias with the machine's alleged objectivity and neutrality. The broad influence of this techno-deterministic common sense—what Meredith Broussard calls "technochauvinism"—makes it difficult to recognize and evaluate many of automation's racializing processes. Moreover, these processes are not only ideologically but also materially concealed in the so-called black box of technical design and implementation. Or more precisely: the *anti*-black box. The latter is one of Ruha Benjamin's synonyms for the New Jim Code, an elaboration of Michelle Alexander's concept of the New Jim Crow. While mass incarceration perpetuates Jim Crow within an apparently color-blind post–civil rights legal system, the New Jim Code works through "*new technologies that reflect and reproduce existing inequities but that are promoted and perceived as more objective or progressive than the discriminatory systems of a previous era*" (5–6). In *Race after Technology: Abolitionist Tools for the New Jim Code*, Benjamin claims that race itself is a technology, a tool "designed to stratify and sanctify social injustice as part of the architecture of everyday life" (17). By codifying race's technologies in objects and sociotechnical environments, and by presenting these technologies as progressive tools for overcoming inequality, the New Jim Code automates the production and reproduction of racial hierarchies, allowing them to operate more quickly, smoothly, and beyond the reach of contestation. And as Virginia Eubanks documents in *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor* (2017), when private corporations develop New Jim Code

systems for public institutions such as welfare agencies, the corporations effectively set racialized policy without public oversight and regulation—a form of control without the hassles of hegemony.[6]

But how does race "get into" robots? *Race after Technology* raises and answers fundamental questions like this one in accessible and convincing ways, making it ideal for the undergraduate classroom and activist circles. When engineers train machines on data that contain bias (input), the newly generated patterns that guide the system's automated decisions about new data (output) reflect and extend the initial bias, especially if engineers understand social progress and scientific objectivity as color blindness toward the unequal society that produced the data. "To the extent that machine learning relies on large, 'naturally occurring' datasets that are rife with racial (and economic and gendered) biases, the raw data that robots are using to learn and make decisions about the world reflect deeply ingrained cultural prejudices and structural hierarchies" (58). Benjamin's first chapter showcases a beauty contest in which machine learning technology judged user-submitted photographs. Treating beauty as an objective and universal quality that technology can assess more accurately than biased humans, the Beauty AI algorithms mimicked their training data and selected mainly white winners. Or consider PredPol, the predictive policing algorithm used by the Los Angeles Police Department. In chapter 2, Benjamin displays the PredPol algorithm: surely there can be no racism in math? But PredPol uses data generated by long-standing hypersurveillance of racialized neighborhoods and sends police to hot spots where they have a mandate to racially profile and discover all manner of "crimes" before they occur, effectively turning PredPol into a self-fulfilling prophecy. Thus, when an automated system's designers not only repudiate racism but even claim to be making postracial technology, we must evaluate what the technology *does*, how it racially filters and organizes social life, not designers' intentions. The glitch might be a useful tool for this task. If Google Maps tells you to turn on "Malcolm Ten Boulevard," the misrecognition of a major black political figure is "not an aberration but a form of evidence" (80) about the narrow perspectives of database engineers.

Sharing Atanasoski and Vora's and Rhee's skepticism toward expanding liberal humanism to include its others, Benjamin elucidates a double bind facing many racialized subjects. Systems from automatic water faucets to webcams have trouble recognizing dark skin. But while black people are relatively invisible to such technologies, they are also *too* visible to police. If facial recognition technology is recalibrated to recognize dark skin more efficiently, the danger is smoother insertion into racial technology and greater reification of racial difference. Benjamin points to one possible future of the New Jim Code: the

Zimbabwean government has agreed to make a database of Zimbabwean faces available to a Chinese startup that is building facial recognition technology for law enforcement. The database is supposed to improve the technology by helping it learn the differences between Asian and black faces, as if these were natural, not social and political, phenomena. For the New Jim Code not only is color-blind but also recodes racial technology as recognition and celebration of difference, especially if this helps racial capitalism construct and sell to niche markets. By analyzing family names, Diversity, Inc. serves companies that want to tailor their marketing to niche markets but are not permitted to collect racial data about customers. In the case of names like "Johnson" that are racially ambiguous, Diversity, Inc. correlates names with zip codes that reflect housing segregation: "racialized zip codes are the output of Jim Crow policies and the input of New Jim Code practices" (147). Inclusion is ultimately just another promise that technology is magic and that inequality could be overcome if we just had better tools. "The road to inequity," Benjamin writes, "is paved with technical fixes" (7).

If we are in Antonio Gramsci's time of monsters, in the interregnum between a decaying world and a new one that cannot be born, automation is a silver bullet. It promises to eliminate the age's monstrous symptoms and deliver a more efficient capitalist technocracy in which racism and other trenchant social inequalities have been engineered out of existence with neutral and inclusive technology. To be sure, automation is moving forward at a much slower pace than the loudest enthusiasts and doomsayers claim. There is much "fauxtomation" in automation.[7] Nonetheless, automation is a definite posthegemonic project, a way of black-boxing social control and bypassing the processes through which hegemonic groups have usually represented and won the consent of a social formation. If hegemony is a representational cultural politics, a project to "make things mean,"[8] then automation is posthegemonic insofar as technologists indirectly govern society by stripping technology of all meanings other than neutrality, efficiency, and convenience. This posthegemony is a technosocial assemblage that is under construction and gathering ever more allies, as the Latourians would say. *Surrogate Humanity*, *The Robotic Imaginary*, and *Race after Technology* are major contributions to the social and cultural study of robots in our third era of automation discourse because they demonstrate that automation is ultimately about the degrees of automaticity with which power can be produced, reproduced, and contested. The sociotechnical futures of work and play, policing and warfare, shopping and sex are not simply matters of "innovation." Since technologies inscribe and materialize the hierarchies of the US racial order, the stakes of automation are the stakes of power's possible futures.

Notes

1. Amy Sue Bix, *Inventing Ourselves Out of Jobs? America's Debate over Technological Unemployment, 1929–1981* (Baltimore: Johns Hopkins University Press, 2000); David F. Noble, *Forces of Production: A Social History of Industrial Automation* (New York: Knopf, 1984); Shoshana Zuboff, *In the Age of the Smart Machine: The Future of Work and Power* (New York: Basic Books, 1988); Ruth Schwartz Cowan, *More Work for Mother: The Ironies of Household Technology from the Open Hearth to the Microwave* (New York: Basic Books, 1983); Venus Green, *Race on the Line: Gender, Labor, and Technology in the Bell System, 1880–1980* (Durham, NC: Duke University Press, 2001).
2. Abigail W. Geiger, "How Americans See Automation and the Workplace in 7 Charts," Pew Research Center, April 8, 2019, www.pewresearch.org/fact-tank/2019/04/08/how-americans-see-automation-and-the-workplace-in-7-charts/.
3. Erik Brynjolfsson and Andrew McAfee, *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies* (Cambridge, MA: MIT Press, 2014); Martin Ford, *Rise of the Robots: Technology and the Threat of a Jobless Future* (New York: Basic Books, 2015).
4. Aziz Rana, *The Two Faces of American Freedom* (Cambridge, MA: Harvard University Press, 2010).
5. Catherine A. Stewart, *Long Past Slavery: Representing Race in the Federal Writers' Project* (Chapel Hill: University of North Carolina Press, 2016), 13.
6. Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor* (New York: St. Martin's, 2017).
7. Astra Taylor, "The Automation Charade," *Logic*, August 1, 2018, logicmag.io/failure/the-automation-charade/.
8. Stuart Hall, "The Work of Representation," in *Representation: Cultural Representations and Signifying Practices*, ed. Stuart Hall (London: Sage, 1997), 24.