Understanding Roboticists' Power through Matrix Guided Technology Power Analysis

Tom Williams
Colorado School of Mines
Golden, Colorado, USA

ABSTRACT

Roboticists wield substantial power through the ways we choose to design and deploy robots. But understanding the nature of this power requires us to consider the different *types* of power wielded through different types of robot design choices, and the social and historical factors that shape the power landscape into which robots are embedded. To facilitate this type of analysis, I present *Matrix-Guided Technology Power Analysis (MGTPA)*, a framework for analyzing the different types of power that technologists wield across different domains of power, with sensitivity to the social and historical forces that determine the default and alternative trajectories of those technologies. Further, I show how MGTPA can be used to better understand the specific types of power that roboticists wield.

CCS CONCEPTS

• **Social and professional topics** → Codes of ethics; History of computing; **User characteristics**; • **Computer systems organization** → *Robotics*; • **Applied computing** → *Sociology*.

KEYWORDS

Robot Ethics, Matrix of Domination, Matrix Guided Technology Power Analysis, Black Feminist Thought

ACM Reference Format:
Tom Williams. 2024. Understanding Roboticists' Power through Matrix Guided Technology Power Analysis. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction (HRI '24 Companion), March 11–14, 2024, Boulder, CO, USA.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3610978.3640766

1 INTRODUCTION

As roboticists, we wield substantial power through the ways we choose to design and deploy robots. Recently, researchers like Winkle et al. [125], Zhu et al. [131], and Hou et al. [57] have encouraged our field to interrogate the power dynamics that characterize Human-Robot Interaction contexts, and to view those power dynamics as a central consideration for HRI research. As Winkle et al. [125] stress, quoting D'Ignazio and Klein [28], the first step towards grappling with power in HRI is "analysing how power operates in the world". In Winkle et al. [125]'s conceptualization, this requires understanding the interrelated power dynamics between different stakeholders, including robots, users, bystanders, designers, and funding agencies, rather than myopically focusing on the human-robot interaction itself.

But effectively interrogating the power dynamics surrounding roboticists' design choices will also require us to understand the *social and historical* factors that have shaped the power landscape in which we design. Trying to understand the ways roboticists wield power *without* understanding the social and historical context that define where we are today is like trying to steer a boat while only considering the orientation of the boat itself. Such an approach is unlikely to be successful: without understanding the currents carrying the boat, you won't be able to tell where the boat is going by default, nor how you would need to intervene to change its course. Without understanding the perils of the water that surround you, you won't be able to adjudicate whether the boat's voyage is likely to end in peril without your intervention. Without knowledge of the shoreline beyond the water's edge, you won't be able to identify and steer yourself toward safer harbors. And without the ability to envision alternative forms of travel, you might never ask yourself why you're paddling a boat in the first place.

Moreover, to effectively analyze the power landscape of robotics, and the way that power is structured, we need to be precise about *what type* of power we are talking about, as different types of power are shaped by different social and historical dynamics, manifest in different ways, and must be subverted using different tools.

In this work, I argue that the Matrix of Domination from Black Feminist Thought can provide a crucial framework for approaching power in Human-Robot Interaction in a way that is sensitive to the *type* of power we are concerned with, and the *socio-historical origins* that determine the default and alternative trajectories of interactive robot technologies. More specifically, in this work I propose a concrete framework, *Matrix Guided Technology Power Analysis (MGTPA)*, that technologists can use to analyze the power that roboticists wield, and how it can be wielded towards more equitable ends, across each of the Domains of Power specified within the Matrix of Domination.

I will thus begin by describing how we *typically* talk about power within Human-Robot Interaction (i.e., interpersonal persuasion and influence), what is missed by that narrow focus, and how the Matrix of Domination encourages us to take a broader view of power within Human-Robot Interaction. Next, I will motivate and detail the series of questions that *Matrix-Guided Technology Power Analysis* encourages technologists to ask. Finally, I will apply *Matrix-Guided Technology Power Analysis* across each of the domains of power beyond interpersonal power, and show how this produces a range of concrete design suggestions for how roboticists can revise, rethink, and rip up the status quo in order to subvert the default power inequities that interactive robot design is primed to reinforce.

HRI '24, March 11–14, 2024, Boulder, CO, USA

ACM ISBN 979-8-4007-0323-2/24/03.
https://doi.org/10.1145/3610978.3640766

2 HUMAN ROBOT INTERACTION'S DEFAULT LENS: INTERPERSONAL POWER

2.1 How HRI Normally Conceptualizes Power

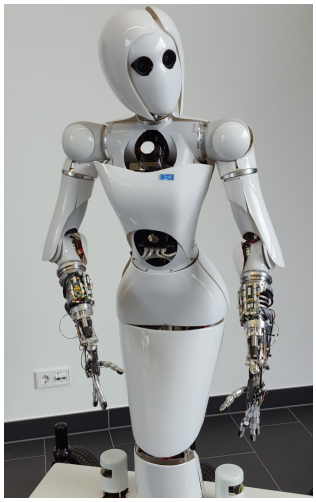
When roboticists talk about power in a social (rather than electrical or physical) sense, they are typically talking about robots' potential for persuasion and influence. Often, discussions of robots' persuasive power stem from a broader tradition of considering the ways that *persuasive technologies* can be developed to encourage people to behave in ways that are safer or more beneficial to themselves and others [20, 50, 68, 104]. Robots designed for use by children have been shown to be effective at getting them to eat their vegetables [6], and at promoting productive conflict resolution between them [102]. Similarly, robots designed for adults have been shown to help nurture empathy and encourage altruism [14, 65]. While "nudges" are sometimes critiqued for being paternalistic [109], the motivation provided by robots in areas like Socially Assistive Robotics is often explicitly sought out by patients and clients. Two excellent examples of robot persuasion being used towards positive ends come from the early work of Katie Winkle, who developed persuasive robots that help users meet their personal therapeutic [124] and exercise [124] goals.

By the same token, HRI researchers have noted that robots can easily persuade people in harmful ways, including humans' tendency to dramatically overtrust robots in ways that could directly lead to physical harm to users in safety-critical situations [93], and robots' ability to inadvertently persuade people in ways that could weaken important moral and social norms [59, 60].

2.2 What is Missed by a Narrow Focus on Interpersonal Persuasion

The types of power discussed thus far are forms of *interpersonal power*: power wielded between human and robot as part of everyday interactions, that allow one interactant to directly influence the other party. Understanding the landscape of power as advocated in *Feminist Human-Robot Interaction* [125] would encourage us to broaden our consideration of the network of stakeholders related to this interaction context, the relative amounts of power held by each, and how this might mediate the influence wielded within the interaction. While a critical first step, this analysis of the power differences between stakeholders may miss key power considerations that arise if power is conceptualized in different ways.

For example, we might consider the power dynamics that surround the AILA robot, a robot designed for space exploration with an overtly White, feminine, female morphology (Fig. 1a), and the REEM robot, a robot sold to the Dubai police [113] with a White, and more masculine morphology (Fig. 1b). Considering the design of these robots, there are a number of critical questions we might ask, including: Why are these robots designed to look this way? How are these robots designed to look *at* people? Why are they being deployed into these domains? And what are the implications to how those questions are answered? All of these questions are positively *dripping* with issues of power, only some of which relate to the robot's ability to influence or persuade, and, I would argue, only some of which would become clear from stakeholder analysis.



(a) AILA robot, from DFKI. Image taken by Tom Williams on June 14th, 2016

(b) REEM robot, from PAL Robotics. Image by Loz Pylock, licensed under CC BY-SA 2.0.

Figure 1: What power dynamics shape, and are shaped by, the design and deployment of these robots?

2.3 The Matrix of Domination

How might we more effectively conceptualize the types of power wielded by robot designers, and how might we more effectively consider the social and historical context into which our robots are embedded? In this work, I propose to answer these questions through the lens of the *Matrix of Domination*.

First introduced by Patricia Hill Collins in her 1990 book *Black Feminist Thought* [22], the Matrix of Domination delineates four domains of power that explain not just the way that people are shallowly influenced, but moreover how oppression itself works in our society. Within the Matrix of Domination, the Interpersonal Domain of everyday interactions plays a critical role, as it is where social power structures are negotiated, reinforced, and subverted [22, p. 363–365]; but it is just one domain of power. Equally important are the Structural, Disciplinary, and Cultural domains of power, in which oppression is structured, managed, and justified.

The Matrix of Domination helps us see that while roboticists are used to thinking primarily about the ways that robot(icist)s wield interpersonal power to directly influence people, roboticists also wield cultural power as their robot designs can reinforce (or subvert) the societal "common-sense" that justifies oppression. Roboticists wield disciplinary power through the algorithmic bureaucracy by which their robots are programmed to give preferential treatment (or simply work at all) for certain groups of people centered according to that common-sense, sort and categorize people according to that common sense, and enforce the norms and values elevated through that common sense. And roboticists wield structural power through their decisions as to who to collaborate with and what domains in which to employ robots, which may have the effect of legitimizing, normalizing, and re-enforcing institutions like policing, which (by design) structure oppression within our society.

The Matrix of Domination also helps us to see that the *oppression* that is reinforced through roboticists' wielding of these different types of power is not some abstract force, but is instead, most frequently, the exact type of oppression that the Matrix of Domination was meant to capture: White Patriarchy[1]. That is, the decisions roboticists make about how their robots are designed and deployed stand to reinforce or subvert the empowerment of White men, and the oppression of women and people of color.

3 MATRIX GUIDED TECHNOLOGY POWER ANALYSIS

To understand how roboticists wield power in ways that go beyond the ability to persuade or influence interactants through the exertion of interpersonal power, I propose a concrete framework I call *Matrix Guided Technology Power Analysis (MGTPA)*. Because the Matrix of Domination focuses on the ways that *oppression* is justified, managed, and structured across different *domains of power*, Matrix Guided Technology Power Analysis aims to make clear:

- The potential alignment between technologists' design choices and the ways that oppression is typically justified, managed, and structured across the cultural, disciplinary, and structural domains of the Matrix of Domination;
- The historical and social factors that have shaped the ways that technologies may, by default, align with those existing power dynamics.
- The ways that, *if that default path is not diverted*, technologies may *reinforce* those power dynamics.
- The ways that, *by questioning not only unequitable outcomes but moreover, the assumptions inherited from those historical and social legacies*, technologies might instead be designed to not only revise, but moreover rethink and "rip up" the status quo, and to thus *subvert* those power dynamics.

To achieve these goals, Matrix Guided Technology Power Analysis asks researchers, designers, and engineers to ask eight key questions for each domain of power:

(1) How does power typically manifest within this domain?
(2) How might the examined technology operate in accordance with these typical power dynamics?
(3) How would this reinforce oppression?
(4) Is the examined technology currently designed in these oppression-reinforcing ways?
(5) How have historical and social factors shaped the ways technology is designed within this domain?
(6) How can we characterize the status quo of how the examined technology is being designed to engage with these power dynamics?
(7) What is the default outcome we would expect if the examined technology is continued to be designed in this way?
(8) How might this default outcome be avoided through approaches that (1) revise, (2) rethink, or (3) rip up the status quo?

[1]Or, more exactly, White Heteropatriarchal Capitalism. Due to the tight connections between race and class in American society, many robot design choices simultaneously reinforce (or subvert) intersecting oppressions across race *and class* based lines.

Now that the motivations and procedures of Matrix Guided Technology Power Analysis have been delineated, we are ready to see how it can be applied to robot design.

4 CULTURAL POWER

Question 1: How does power typically manifest within this domain? Within a social system with a White supremacist and patriarchal power structure, power within the cultural domain may manifest as the reinforcement of the stereotypes used to justify oppression.

Question 2: How might robots operate in accordance with these typical power dynamics? As embodied and often language-capable agents, robots can easily reinforce biases and stereotypes by depicting robots that are racialized and gendered in particular ways which, in the domains where those robots are used, fulfill the stereotypical assumptions of what robots should look like, and the stereotypical assumptions of the "right" people to work in those domains.

Question 3: How would such designs reinforce oppression? Here, we can see two ways that robots' adherence with stereotypical assumptions might reinforce oppression.

First, robots that rely on social identity design cues (facets of robot design that lead people to racialize or gender them in particular ways) to gain acceptance in a particular domain risk reinforcing the stereotypes that explain that increase in acceptance [61] [and see also 2, 82]. For example, robots with female or feminine morphological cues in service roles may reinforce associations between femininity and subservience [10, 40].

Second, robots that are designed for the male gaze may reinforce stereotypes about both the people they depict, the people designing robots, and the people for whom robots are designed. For example, robots intended for non-service work that are designed in the image of hegemonically empowered demographic groups (e.g., White or Asian men, depending on cultural context) may reinforce the idea that robotics is the domain of White and Asian men; and robots designed to match the sexual or beauty standards of hegemonically empowered demographic groups (e.g., robots designed to depict "beautiful women" [76]) simultaneously reinforce the idea that robotics is the domain of heterosexual men and the idea that women are primarily objects of heterosexual male desire [cf. 1, 19, 91].

Question 4: Are robots designed in these oppression-reinforcing ways? Are robots designed to rely on social identity design cues to gain acceptance? Yes. Most humanoid robots are designed in the image of hegemonically empowered racial groups: (either literally depicting White individuals or rendered in white plastic) [106], and virtually all language-capable robots are given White-cued voices and speak in Standardized American English.

Similarly, as Perugia et al. [89] show, most current humanoid robots are designed in ways that lead people to gender them, with most robots gendered as masculine, with the exception of robots designed for domains where surface features like hair and eyebrows are necessary (typically socially-oriented service roles), in which case robots are designed to be feminine. Robots designed to be feminine in this way are often given features suggesting grooming and clothing that align with patriarchal gender expectations, and in some cases, are given prominent chests. In short, robots are typically designed to align either with how their male robot designers would

like to see themselves (White, masculine), except when robots are intended for "feminine" domains, in which case they are designed for their male designer's gaze (as White and feminine).

While robots designed outside of Western laboratories are not typically designed to be *White*, they nevertheless reflect these racialized and gendered patterns within the local context. This can most clearly seen in the work of Ishiguro, whose androids notably include the Gemenoid [85] (a copy of himself), which is the clearest possible example of a roboticist designing a robot in their own image, and Erica (a robot made in the image of "30 beautiful women" to "appeal to everyone")[76], which may also be the clearest possible example of a robot designed for the heterosexual male gaze.

Question 5: How have historical and social factors shaped the ways we approach robot design within this domain? Here, what we are specifically asking is: How have historical and social factors shaped the ways that we design robots to look and sound in different domains (and can this explain how our technologies are being designed)?

As Evans [34] analyzes, one of the key unacknowledged cultural touchstone in robotics is the "Steam Man" of 1868, in which (1) engineers Zadoc Dederick and Isaac Grass responded to White male concerns about the end of slavery by envisioning a steam-powered Black slave designed in the image of Zip Coon (the second most popular character in popular culture at the time), and (2) author Edward Ellis [33] turned Dederick and Grass' engineering patent into an enormously successful piece of popular literature that birthed one of the first American science fiction movements[34, p. 120]. This imagination of Black male slave proto-robots in both engineering and science fiction led to the envisioning of White feminine proto-robots in early European science fiction [26], where the same template was used to respond to White male concerns about the rise of women's suffrage [21]. This in turn led to Karel Čapek's *R.U.R* [17], which infused these images of Black and Feminine artificial persons with White European labor concerns to tell a slave uprising story (using the Czech word for forced laborer, *robot*, to label these mechanical slaves). Isaac Asimov responded [7] to *R.U.R.* by refuting the narrative of slave uprising [3], instead telling stories that sought to explore the optimal way that (artificial) slaves could be controlled by computational slave codes [58], and naming the field that would develop those artificial slaves *robotics*. These decisions then directly informed the first successful robotics engineering efforts. The creators of the Unimate decided to pursue the development of robotics as a way of turning their love of Asimov's stories into reality, and with their development of their robots guided by Asimov's "three laws" slave codes [38]. Similarly, Norbert Wiener, the creator of Cybernetics, was an amateur science fiction author who attempted to collaborate with Asimov, and who urged the nascent robotics community to remember that what they were developing was "the precise economic equivalent of slave labor"[116, p. 152, as cited in Chude-Sokei 21, p. 163]. In short, robots (in the way we think of them today) were originally envisioned as mechanical slaves in both the technoscientific and cultural robot imaginaries, and were designed to specifically meet the needs and address the fears of White Men and points where White male hegemonic power was threatened by the end of slavery, the rise of women's suffrage, and by the rise of labor movements.

Overall, then, we can see that not only can robots reinforce White Patriarchal power (and other intersectional power hegemonies) within the cultural domain, but moreover, that the state of the art of robotics research inherits a legacy that is specifically primed towards this outcome.

Question 6: How can we characterize the status quo of how robots are being designed to engage with these power dynamics? Robots are (due in part to origination of robots as an envisioned way of helping White men maintain hegemonic power) being designed to encode the stereotypes, biases, aesthetic and sexual preferences, and assumptions about gender and racial roles, of their designers (typically White men), thus benefiting those already hegemonically empowered according to, and through, those stereotypes, biases, preferences, and assumptions.

Question 7: What is the default outcome we would expect if robots are continued to be designed in this way? If robot design continues along its current path, robots will continue to serve as vehicles for the reinforcement of the biases, stereotypes, and assumptions that are used to justify White patriarchy.

Question 8: How might this default outcome be avoided through approaches that (1) revise, (2) rethink, or (3) rip up the status quo? First, roboticists might *revise* the status quo, by saying "We won't use White patriarchal stereotypes, biases, and assumptions to guide robot design". This might involve designing for a wider diversity of robots [cf. 56, 92]. However, simply encouraging predominantly White male design teams to design more robots of color seems like a recipe for caricatured disaster, and an ironic return to the Steam Man origins of the robotic imaginary. Similarly, there may well be concerns about pink capitalism or rainbow capitalism depending on the motivations for these diversification efforts [108]. So while this approach may be a step in the right direction, it could merely trade one problem for a slightly more inclusive version of that same problem. Alternatively, revising the status quo could involve attempting to "remove" race and gender from humanoid robot design. However, there is substantial evidence that is wildly difficult to design robots whose appearance, and especially their voice, are not gendered or racialized in any way [98, 101]. While there has been some recent success towards this goal through agendered robot design approaches [108], it is yet unclear what it would mean to develop a language-capable robot whose voice is "free" of racial or ethnic cues. Finally, this could involve taking a *cross-stereotypic* approach in which roboticists design humanlike robots in ways that run counter to those White patriarchal expectations [cf. 42]. Winkle, for example, has demonstrated some success in designing feminine robots that deviate from culturally specific politeness norms [123, 126]. More generally, this approach could involve simply using more masculine robots in service positions, queering robot design [66, 107], or otherwise mixing-and-matching different design cues within a single robot body. More work needs to be done to explore whether these approaches will be effective ways of diverting robotics' White male gaze in a genuine and sensitive way.

Second, roboticists might *rethink* the status quo by saying "if the problem is with how robots are being designed to look and speak, lets just not design robots that are humanoid or language-capable".

This might include using zoomorphic [25] or mechanomorphic designs, or minimizing natural language capabilities. Perhaps your vacuum cleaner doesn't need to speak. However, while this approach might address the concerns listed in this section, it would throw the proverbial baby out with the bath water. In some domains, certain humanlike features might be critical to encourage the right type of engagement, and natural language capabilities are often critical to our entire endeavor, e.g., in robots for education.

Finally, roboticists might *rip up* the status quo by saying "Let's attack the root causes that explain how we have arrived at the status quo". If the reasons roboticists by default are reinforcing White patriarchal stereotypes, biases, and assumptions is because (1) this was their original envisioned purpose, (2) this motivating purpose has shaped the motivations and values of robotics, (3) these motivations and values have shaped the demographics of robot design teams, and (3) today's robot design teams develop robots that fulfill a combination of (a) the historic motivations, values and biases of robotics, and (b) their own personal motivations, values, and design biases, then perhaps we should be working to change the motivations and values of robotics and the demographic composition of robot design teams.

This might include taking participatory design approaches [46] [cf. 4, 13, 69, 70, 86, 96, 122, 127] to design robots whose appearance and purpose satisfies the needs and desires of historically excluded and oppressed populations [see also 83, 84] (and to make political [29] and ethical [79] dimensions of robot design clear in the process), the use of design-justice oriented end-user programming approaches to enable members of those communities to design or redesign robots that better meet their own preferences and needs, as well as more simply, a push for greater demographic diversity in both the design teams and executive boards of our robotics companies.

5 DISCIPLINARY POWER

Next, lets consider the Disciplinary Domain, which focuses on the way that power is managed through the enforcement of "commonsense" rules and logics.

Question 1: How does power typically manifest within this domain? Within a social system with a White supremacist and patriarchal power structure, power within the disciplinary domain may manifest as the enforcement of hegemonic norms, or as the enforcement of hegemonic logics of race and gender categorization.

Question 2: How might robots operate in accordance with these typical power dynamics? Here, we can see (at least) two possible answers, leading to two different threads in our analysis. First, robots may be designed to enforce (or even merely adhere to) White patriarchal norms. Second, robots may be designed to categorize individuals according to White patriarchal logics of race and gender.

Question 3: How would such designs reinforce oppression? Here, we can again see (at least) two sets of possible answers. First, robots designed to enforce White patriarchal norms may punish members of communities whose norms differ from that White patriarchal default; and robots designed to enforce (or adhere to) White patriarchal norms may further center those norms while problematizing and casting as deviant the behaviors of members of other groups. Second, robots that categorize individuals according to White patriarchal logics of race and gender may reinforce those logics.

Question 4: Are robots designed in these oppression-reinforcing ways? Gender recognition is a widely used robot capability. The API for the Pepper and Nao robots (arguably the most widely deployed social robots of the past decade) provides functionality that attempts to guess the gender of a robot's interactant [88], and there is a wide literature exploring how to use this capability, and exploring new ways for robots to gender interactants (e.g., on the basis of breast shape [90]). And while there has thankfully been little work thus far on algorithmically racializing robots' interactants (despite hundreds of such efforts in the Computer Vision community [30]), many roboticists have called for the personalization of robots to interactant culture, race, and ethnicity [49] (see also [37, 71, 90, 97]. As such, *most* of today's social roboticists are not engaging in algorithmic social identity categorization, but the field is trending in a direction where these methods may become increasingly common.

Question 5: How have historical and social factors shaped the ways we approach robot design within this domain? Here, what we are specifically asking is: How have historical and social factors shaped the ways that we think about robot moral and social norms, and the ways that robots perceive and categorize (and can this explain how our technologies are being designed)?

First, lets consider robot enforcement (and adherence to) norms. The most popular approach to machine morality (that is, enabling robots to enforce and adhere to moral principles) is the use of hand-specified deontic norms, that indicate which actions should be considered to be obligated, forbidden, or permitted [74]. A natural question to ask about this approach is "Whose norms are encoded through this approach?" The answer to this question is, typically "The norms of White (and Asian) men", for at least two reasons. First, robots are typically programmed to (implicitly or explicitly) adhere to the norms of their creators; and because most robots' creators are currently White (and Asian) men, robots are typically designed to enforce and adhere to those specific norms. Second, by considering the historical context of attempts to select a set of universal deontic norms, we can see that this general enterprise is fundamentally primed towards the reinforcement and enforcement of White Patriarchal norms due to its grounding in Kantean Deontology. Central to Immanuel Kant's moral and political philosophies are notions of universalizability: e.g., that we should avoid actions that we collectively view as bad, and that we should treat all people with respect. These may at first glance seem like admirable and obvious objectives. But as Charles Mills points out, these views need to be understood through the lens of Kant's other work, especially his work on early race science [62, 63], which led Charles Mills to call Kant the "father of the modern concept of race". As Mills writes in *Kant's Untermenschen*:

> "When Kant talks about the importance of treating all persons with respect ... he is not making race-neutral and racially-inclusive pronouncements; he is really talking about the White population (and if feminists are correct in their parallel line of argument, only about the male half). If this is true, it is obviously a radically different picture of the Kant we all thought we knew and loved. The distinction

> between "Treat all persons with respect," where "person" is assumed to be racially inclusive, and "Treat only Whites with respect" (at least here on Earth) is obviously not minor and trivial at all. It would mean that we need to reconceptualize his philosophy as distinguishing (White) persons from (Non-White) sub-persons" [77].

As such, it is perhaps unsurprising that today's dominant approaches to machine morality tend to center White and Masculine moral norms, since those are the exact norms that Kantean Deontology was intended to elevate.

It might also be argued that Machine Morality's focus on designing rules that robots can follow to avoid causing harm inherits from the work of Isaac Asimov and his famous Three Laws of Robotics. Here again, this lineage primes roboticists to elevate White Patriarchal norms. Asimov's stories were famously written as an alternative to Čapek's narrative of race and class uprising, and as such, Asimov's laws have been read as a Slave Code: a vision of a set of universal principles that could be encoded to ensure that mechanical slaves would not revolt, but would instead act in the interests of their White male masters.

Next, lets consider robot categorization of individuals according to White patriarchal logics of race and gender. As Williams [119] discusses, social robot APIs are increasingly categorizing people according to their gender, and there is a rise of computer vision systems outside robotics that attempt to categorize individuals according to their race. A natural question to ask about this approach is "Whose notions of race and gender, and whose logics of categorization, are encoded through this approach?" The answer to this question is, typically "The logics of (Straight) White (and Asian) men", for at least two reasons. First, as above, when robots and other computer vision systems are programmed to sort those they perceive into categories of race or gender, they typically do so according to the categories that would be used by their programmers, or by those who collected the training sets on which their programmers' models are trained [99, 100]. Because most creators of robots and computer vision datasets are currently (straight) White (and Asian) men, robots are typically designed to enforce the logics of race and gender held within those particular demographic groups. Second, by considering the historical context of these technologies, we can see the way that the entire notion of technologies that categorize individuals and make those categories visible.

By considering the historical context of technologies that assign and make visible an individual's race or gender, we can see that this general enterprise is fundamentally primed towards the reinforcement and enforcement of White Patriarchal logics of race and gender due to their inheritance from a longer history of biometric surveillance technologies. In her book Dark Matters, Black Studies scholar Simone Browne demonstrates how many modern computing technologies fall within a long history of biometric technologies used for racialized and gendered surveillance [15], including slave brands, early non-computerized databases like the "Book of Negroes", lantern laws in early New York City, and the use of racial categories on state-issued photo IDs. While Brown observes that modern biometric technologies and slave branding are clearly distinct technologies,

> "When we think of [how peoples' bodies are being] informationalized by way of biometric surveillance – sometimes voluntarily and sometimes without consent or awareness – and then stored in large-scale databases... we can find histories of these accountings and inventories of the bodies in slave registers, slave branding and the slave vessel manifests." [15]

Overall, then, we can see that not only can robots reinforce White Patriarchal power within the disciplinary domain, but moreover, the state of the art of robotics research inherits a legacy that is specifically primed towards this outcome.

Question 6: How can we characterize the status quo of how robots are being designed to engage with these power dynamics? Robots are (due in part to the legacy of moral philosophies and biometric surveillance technologies on which they build) being designed to encode the moral and social norms of their designers (typically White men) or those whose data they're predominantly trained on (again, White men); and designed to enforce logics of race and gender that benefit those already hegemonically empowered (e.g., enforcing racial caste and patriarchal gender norms).

Question 7: What is the default outcome we would expect if robots are continued to be designed in this way? If robot design continues along its current path, robots will continue to reinforce the norms and logics used to manage White patriarchy.

Question 8: How might this default outcome be avoided through approaches that (1) revise, (2) rethink, or (3) rip up the status quo? First, roboticists might *revise* the status quo, by saying "we won't use those White patriarchal norms or categories". This might involve taking a *cross-cultural* approach in which roboticists endow robots with conceptualizations of moral norms upheld within communities that are not already hegemonically empowered, or in which they assign people to a broader and more inclusive set of identity categories. While these approaches move in the right direction, they ultimately suffer from the same problems, as they merely trade one problem for a slightly more inclusive version of that same problem.

Second, roboticists might *rethink* the status quo, by saying "we will let communities use their own processes for specifying norms, and their own logics of categorization". This might involve taking a *participatory* approach in which roboticists work with oppressed communities to design robots whose norms align with those of the community (creating new opportunities for the creation of shared norms), or to design robots whose methods of categorization better align with the values of those communities [51] (e.g., Bennett et al. [9]'s work with Blind users has shown the potential merits of using self-provided (rather than algorithmically determined) categorizations). While these approaches begin to question dominant *processes* and *logics*, they fail to acknowledge or question the social and historical context that explains why the default approach within robotics reinforces oppressive power dynamics.

Finally, roboticists might *rip up* the status quo, by saying "Let's attack the root causes that explain how we have arrived at the status quo". If the reasons why roboticists by default are reinforcing White patriarchal logics is that the dominant approach to machine morality is grounded in a moral philosophy that is inextricably linked to race science, perhaps we should do away with that approach.

That is, perhaps we should be pursuing approaches to robot moral reasoning that are grounded in alternative moral philosophies, such as Confucian Role Ethics [72, 95] [cf. 114, 114, 115, 121, 129, 130]; or perhaps we need to identify new ways of ensuring safe implementation of our robotic systems that do not rely on explicit moral reasoning [78]. Similarly, if robots' use of categories of race and gender perpetuates White patriarchal logics regardless of where category judgments come from, perhaps we need to design robots that can engage in natural and effective interactions without categorizing people. Finally, if all of these problems reflect roboticists' tendency to imbue robots with their own norms and logics, perhaps we should be demanding more diverse robot design teams, and the adoption of Design Justice [23] driven approaches [cf. 87] that emphasize end-user-programming as the site of robot design.

6 STRUCTURAL POWER

Question 1: How does power typically manifest within this domain? Within a social system with a White supremacist and patriarchal power structure, power within the structural domain may manifest as the legitimization and facilitation of the institutions that are used to structure oppression, such as the police.

Question 2: How might our technologies operate in accordance with these typical power dynamics? Roboticists might operate in accordance with these structural power dynamics by designing robots in collaboration with, or intended for use by, institutions used to structure oppression, such as the police.

Question 3: How would such designs reinforce oppression? Here, we can see two ways that roboticists' collaboration with or facilitation of police use of robots might reinforce oppression. First, if roboticists build robots for use by police, this could exacerbate and legitimize the use of police as an institution for surveilling, controlling, and enacting violence toward people of color and the poor, thus reinforcing the hegemonic power of high-socioeconomic status White men. Second, any collaboration between roboticists and police may serve to legitimize and normalize the use of police to surveil, control, and enact violence toward people of color and the poor, and the use of police as the appropriate means of addressing other societal problems, as discussed by Williams and Haring [120].

Question 4: Are our technologies designed in these oppression-reinforcing ways? By and large, most robots are not being designed for use in policing domains, and few researchers are choosing to actively collaborate with the police. However, three main concerns may be raised surrounding the current practice of robotics. First, while police-roboticist collaborations are still rare, the number of such collaborations appears to be on the rise. Second, one of the dominant general-purpose non-humanlike interactive robot technologies, Boston Dynamics' Spot, appears to be primarily sold to US police departments, meaning that roboticists working on that general purpose platform may be indirectly facilitating police use of robots. Third, one of the most promising forthcoming general-purpose humanlike interactive robots, Amazon's Astro, appears to be currently marketed as a Ring camera device, which Amazon has used to help expand police surveillance networks [32, 53]. This similarly means that roboticists working on that general purpose platform may also be indirectly facilitating police use of robots.

Question 5: How have historical and social factors shaped the ways we approach robot design within this domain? Here, what we are specifically asking is: How have historical and social factors shaped the acquisition and use of robots by police?

Policing is a relatively recent phenomenon, with the first police departments emerging in the 1800s to suppress labor unrest at the end of the industrial revolution. In the US, the first police departments formed in the 1830s in the Northeast as a way of curbing unrest amongst exploited working class immigrants [67] and exerting control over religious minorities [41], while working *with* local "petty criminals", e.g. to help fence stolen goods [31]. Meanwhile, police departments did not exist in other parts of the country, like the the American South, until the 1860s, at which point the dissolution of the US Slave Economy forced southern states to "rebrand" their previous system of law enforcement (Slave Patrols [48]) into a network of formal police and vigilante terrorist groups [117].

Today, US police continue to surveil and control people of color and other oppressed groups, although local governments are also increasingly shifting other health and safety tasks toward the police: a choice wherein societal problems are addressed using violence rather than health, housing, or education. In service of this strategy, US police have been increasingly militarized over the past 100 years. While formal separation of police and military activities was instituted after the civil war during the reconstruction period (as a means to ensure that US troops were unable to interfere in the establishment of Jim Crow policies in the American South [110]), this barrier has slowly been eroded, ironically, in order to provide police with military equipment in order for them to more effectively use violence to respond to social justice movements within Black communities [8, 43], and to facilitate the criminalization and control of Black and anti-war communities [5, 24, 27]. This militarization increasingly includes the acquisition and use of robots, with hundreds of robots transitioned from the military to police [39], and thousands of robots acquired by the police from other sources, including both commercial drones [18, 44] and custom-built policing robots like Knightscope's K5 and Boston Dynamic's Spot.

Once police acquire robots, police use them for the same activities they have always pursued, and which they were designed to pursue: violence and surveillance against disempowered communities. Specifically, police have used robots for surveillance [54, 105], including surveillance of social justice movements, surveillance of unhoused people [16, 36], and "those engaging in drug transactions" [112] in regions where marijuana-related arrests and sentencing can be between four [81] and forty-five [111] times higher for Black Americans despite less frequent *use* of marijuana among those same Black communities. Similarly, police are using robots for more overt physical violence, with several instances over the past decade of police robots being used to deploy bombs and chemical munitions against people with mental disabilities [11, 94, 103], and expressing a desire to use robots to wipe Black communities "off the map" [112]. Moreover, police supervisors in some cities have explicitly voted to allow police to kill people using robots [12, 54, 105]. Across all these uses, robots stand to exacerbate police's *historic and intended* tendencies toward violence, through moral buffering [47] and increasing the cognitive accessibility of violent options.

Overall, then, we can see that not only can robots reinforce White Patriarchal power in the structural domain, but moreover,

that the state of the art of robotics research inherits a legacy that is specifically primed towards this outcome.

Question 6: How can we characterize the status quo of how robots are being designed to engage with these power dynamics? Robots are (due in part to the historical purpose and use of police) being increasingly used to help police to enact violence and surveillance against oppressed communities, and roboticists are increasingly working to directly or indirectly support both these specific police activities, as well as the general system of policing.

Question 7: What is the default outcome we would expect if robots are continued to be designed in this way? If robot design continues along its current path (with roboticists either collaborating with the police or building robots for primary use by the police), roboticists will continue to be complicit in policing as an institution for structuring oppression and maintaining White male hegemonic power. As a concrete example of this, if Amazon's *Astro* robot is sold as a *Ring* device (as currently advertised), the HRI community may become complicit in the broadening of police surveilance networks into people's homes, given that Amazon collaborates with over 800 police departments through its *Ring* program [32, 35, 80], providing police with a centralized portal through which to access the data collected from Ring cameras, and coaching police on how to extract consent to access camera footage from users [32, 52].

Moreover, beyond this immediate complicity, there is real concern that students from historically oppressed groups may be dissuaded from joining the field of robotics once they learn of the increasing link between robotics and policing, thus perpetuating the current White masculinity of robotics and thus exacerbating all of the previous problems discussed in this paper.

Question 8: How might this default outcome be avoided through approaches that (1) revise, (2) rethink, or (3) rip up the status quo? First, roboticists might *revise* the status quo, by saying "we won't design robots for use by the police, especially in areas where police have a history of misusing robots". This is the approach taken by signatories of the *#NoJusticeNoRobots* campaign [118], who publicly "refuse[d] to facilitate the execution or publication of research in robotics or Artificial Intelligence (AI) that is performed in collaboration with State and Local Law Enforcement agencies (or national agencies such as ICE and CBP)". While these types of petition efforts may be a necessary first step, they are not sufficient to address police misuse of general purpose robots.

Second, roboticists might thus *rethink* the status quo by saying "if the problem is with how our robots are being used, lets just design them so they cannot be misused in that way, or so that they will primarily be used in ways that run counter to that misuse". This might include designing robots that are private by design [55], never storing sensor data in ways that could later be accessed by the police. Alternatively, this could include working with activists to design robots for sousveillance [75], explicitly working to maintain communities' safety by allowing them to monitor and document police activity. Drones have previously been used in this way to protect the safety of Occupy Wall Street protesters [132], make visible the infrastructural violence committed against Palestinians in East Jerusalem [64], and to expose the environmental devastation of the Dakota Access pipeline project and the police violence wrought against those protesting it at Standing Rock[132].

Finally, roboticists might *rip up* the status quo by saying "Let's attack the root causes that explain how we have arrived at the status quo". If the reasons why roboticists by default are reinforcing White patriarchal power when police use our robots is because our societies are designed to use race and class directed police violence as its "solution" to health and safety concerns, then perhaps we should be working to dismantle society's reliance on police as the solution to those problems. As Lisa Lowe writes in her analysis of the work of Ruth Wilson Gilmore [45]:

> "Abolition does not involve merely putting an end to prisons: it must mean forging robust alternative solidarities and social relations that do not yet exist. ... it is the presence of social relations of care, mutuality, and interdependency that make punishment and incarceration unnecessary. Abolition is aimed at the elimination of the racial capitalist social order that produces vast wealth accumulation for the few and vulnerability for the many, and which violently enforces and deepens those conditions. It is a program of creation that requires a social imaginary that is bounded neither by the nationalist terms of the current order nor by the capitalist terms for envisioning the 'global'." [73]

As such, a truly abolitionist approach to robotics that seeks to subvert White Patriarchal power hegemony within the structural domain might include designing robots for use by alternative organizations and institutions, and seeking to shift our collective resources away from the police, and towards those alternative organizations and institutions.

7 CONCLUSION

I have presented *Matrix-Guided Technology Power Analysis (MGTPA)*, a framework for analyzing the different types of power that technologists wield across different domains of power, with sensitivity to the social and historical forces that determine the default and alternative trajectories of those technologies. Further, I have shown how MGTPA can be used to better understand the specific types of power that roboticists wield. Critically, MGTPA is not intended to stand on its own, but rather is intended to supplement recent Feminist approaches to power analysis [125], proposals for close analysis of interpersonal power [57], and higher level frameworks like the *Instruments of Power* used to reason about Diplomatic, Informational, Military, and Economic power at the national level [128]. Yet as I have shown, the level of analysis afforded by the Matrix of Domination is acutely illuminating for the field of HRI, as it clearly demonstrates that roboticists cannot simply choose to ignore their power as conceptualized beyond persuasion, as "robotics as usual" will inevitably lead to the *reinforcement* of White Patriarchy. Because of the power roboticists wield to reinforce or subvert White Patriarchy, and because the default effect of robots is to reinforce White Patriarchy, roboticists have a clear responsibility to take concrete actions that will work to subvert White Patriarchy across the different domains of the Matrix of Domination.

ACKNOWLEDGMENTS

This work was funded in part by Air Force Young Investigator Award 19RT0497.

REFERENCES

[1] Simone Alesich and Michael Rigby. 2017. Gendered robots: Implications for our humanoid future. *IEEE Technology and Society Magazine* 36, 2 (2017), 50–59.

[2] Patrícia Alves-Oliveira, Patrícia Arriaga, Ana Paiva, and Guy Hoffman. 2021. Children as robot designers. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. 399–408.

[3] Isaac Asimov. 1950. *I, Robot*.

[4] Minja Axelsson, Raquel Oliveira, Mattia Racca, and Ville Kyrki. 2021. Social robot co-design canvases: A participatory design framework. *ACM Transactions on Human-Robot Interaction (THRI)* 11, 1 (2021), 1–39.

[5] Radley Balko. 2006. *Overkill: The rise of paramilitary police raids in America*. Technical Report. The Cato Institute.

[6] Ilaria Baroni, Marco Nalin, Mattia Coti Zelati, Elettra Oleari, and Alberto Sanna. 2014. Designing motivational robot: how robots might motivate children to eat fruits and vegetables. In *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 796–801.

[7] Gorman Beauchamp. 1980. The Frankenstein Complex and Asimov's Robots. *Mosaic: A Journal for the Interdisciplinary Study of Literature* 13, 3/4 (1980), 83–94.

[8] Michal R Belknap. 1991. *Urban race riots*.

[9] Cynthia L Bennett, Cole Gleason, Morgan Klaus Scheuerman, Jeffrey P Bigham, Anhong Guo, and Alexandra To. 2021. "It's Complicated": Negotiating Accessibility and (Mis) Representation in Image Descriptions of Race, Gender, and Disability. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–19.

[10] Hilary Bergen. 2016. "I'd blush if I could": Digital assistants, disembodied cyborgs and the problem of gender. *Word and Text, a Journal of Literary Studies and Linguistics* 6, 1 (2016), 95–113.

[11] Richard J Berry and Gorden E Eden Jr. 2014. *Albuquerque Police Department Monthly Report November 2014*. Technical Report. Albequerque Police Department.

[12] Sam Biddle. 2021. U.S. Marshals used Drones to Spy on Black Lives Matter Protests in Washington, D.C. https://theintercept.com/2021/04/22/drones-black-lives-matter-protests-marshals/. *The Intercept* (2021).

[13] Elin A Björling and Emma Rose. 2019. Participatory research principles in human-centered design: engaging teens in the co-design of a social robot. *Multimodal Technologies and Interaction* 3, 1 (2019), 8.

[14] Jason Borenstein and Ronald C Arkin. 2017. Nudging for good: robots and the ethical appropriateness of nurturing empathy and charitable behavior. *Ai & Society* 32 (2017), 499–507.

[15] Simone Browne. 2015. *Dark matters*. Duke University Press.

[16] Sarah Buhr. 2017. Security robots are being used to ward off San Francisco's homeless population. https://techcrunch.com/2017/12/13/security-robots-are-being-used-to-ward-off-san-franciscos-homeless-population/. *TechCrunch* (2017).

[17] Karel Čapek. 1920. *R.U.R. (Rossum's Universal Robots)*. Aventinum.

[18] Nanci K Carr. 2021. Programmed to protect and serve: The dawn of drones and robots in law enforcement. *J. Air L. & Com.* 86 (2021), 183.

[19] Stephen Cave and Kanta Dihal. 2020. The whiteness of AI. *Philosophy & Technology* 33, 4 (2020), 685–703.

[20] Vijay Chidambaram, Yueh-Hsuan Chiang, and Bilge Mutlu. 2012. Designing persuasive robots: how robots might persuade people using vocal and nonverbal cues. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*. 293–300.

[21] Louis Chude-Sokei. 2019. Race and robotics. *Cyborg futures: Cross-disciplinary perspectives on artificial intelligence and robotics* (2019), 159–171.

[22] Patricia Hill Collins. 2022. *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*. Routledge.

[23] Sasha Costanza-Chock. 2020. *Design justice: Community-led practices to build the worlds we need*. The MIT Press.

[24] Stacey Cotton, Peter Kraska, and James A Pikl. 2014. Law Enforcement, Community, and Military Tactics: What's the Conflict?. In *Working in the Public Interest Conference*.

[25] Kate Darling. 2021. *The new breed: what our history with animals reveals about our future with robots*. Henry Holt and Company.

[26] Auguste Villiers de l'Isle Adam. 1886. *L'Ève future*.

[27] Casey Delehanty, Jack Mewhirter, Ryan Welch, and Jason Wilks. 2017. Militarization and police violence: The case of the 1033 program. *Research & politics* 4, 2 (2017), 2053168017712885.

[28] Catherine D'Ignazio and Lauren F Klein. 2020. *Data Feminism*. MIT press.

[29] Carl DiSalvo. 2014. Critical making as materializing the politics of design. *The Information Society* 30, 2 (2014), 96–105.

[30] Charles-Olivier Dufresne-Camaro, Fanny Chevalier, and Syed Ishtiaque Ahmed. 2020. Computer vision applications and their ethical risks in the global south. *Graphics Interface* (2020).

[31] Roger G Dunham, Geoffrey P Alpert, and Kyle D McLean. 2020. *Critical issues in policing: Contemporary readings*. Waveland Press.

[32] Grace Egger. 2020. Ring, Amazon calling: the state action doctrine & the fourth amendment. *Washinton Law Review Online* 95 (2020), 245.

[33] Edward S Ellis. [1868] 2016. *The Steam Man of the Prairies*.

[34] Taylor Scott Evans. 2018. *The Race of Machines: A Prehistory of the Posthuman*. Ph. D. Dissertation.

[35] Cyrus Farivar. 2020. Cute videos, but little evidence: Police say Amazon Ring isn't much of a crime fighter. *NBC News* (2020).

[36] Todd Feathers. 2021. Honolulu Police Used a Robot Dog to Patrol a Homeless Encampment. https://www.vice.com/en/article/wx5xym/honolulu-police-used-a-robot-dog-to-patrol-a-homeless-encampment. *Motherboard (Vice News)* (2021).

[37] Pasquale Foggia, Antonio Greco, Gennaro Percannella, Mario Vento, and Vincenzo Vigilante. 2019. A system for gender recognition on mobile robots. In *Proceedings of the 2nd international conference on applications of intelligent systems*. 1–6.

[38] Association for Advancing Automation. [n. d.]. Unimate // The First Industrial Robot. A Tribute to Joseph Engelberger; https://www.automate.org/a3-content/joseph-engelberger-unimate.

[39] Bard College Center for the Study of the Drone. 2016. Law Enforcement Robots Datasheet. https://dronecenter.bard.edu/law-enforcement-robots-datasheet/.

[40] Leopoldina Fortunati, Autumn Edwards, Chad Edwards, Anna Maria Manganelli, and Federico de Luca. 2022. Is Alexa female, male, or neutral? A cross-national and cross-gender comparison of perceptions of Alexa's gender and status as a communicator. *Computers in Human Behavior* 137 (2022).

[41] Raymond Blaine Fosdick. 1920. *Crime in America and the Police*. Vol. 5. Century Company.

[42] Alessio Galatolo, Gaspar I Melsión, Iolanda Leite, and Katie Winkle. 2022. The Right (Wo) Man for the Job? Exploring the Role of Gender when Challenging Gender Stereotypes with a Social Robot. *International Journal of Social Robotics* (2022), 1–15.

[43] Fanna Gamal. 2016. The racial politics of protection: A critical race examination of police militarization. *Calif. L. Rev.* 104 (2016).

[44] Dan Gettinger. 2020. Public safety drones (3rd Ed.). *Center for the Study of the Drone at Bard College* (2020).

[45] Ruth Wilson Gilmore. 2022. *Abolition geography: Essays towards liberation*. Verso Books.

[46] Judith Gregory. 2003. Scandinavian approaches to participatory design. *International Journal of Engineering Education* 19, 1 (2003), 62–74.

[47] Lisa Guenther. 2021. Police, Drones, and the Politics of Perception. *The Ethics of Policing: New Perspectives on Law Enforcement* (2021), 248.

[48] Sally E Hadden. 2003. *Slave patrols: Law and violence in Virginia and the Carolinas*. Harvard University Press.

[49] Jaap Ham. 2021. Influencing robot influence: Personalization of persuasive robots. *Interaction studies* 22, 3 (2021), 464–487.

[50] Jaap Ham, René Bokhorst, Raymond Cuijpers, David Van Der Pol, and John-John Cabibihan. 2011. Making robots persuasive: the influence of combining persuasive strategies (gazing and gestures) by a storytelling robot on its persuasive power. In *Social Robotics: Third International Conference, ICSR 2011, Amsterdam, The Netherlands, November 24-25, 2011. Proceedings 3*. Springer, 71–83.

[51] Foad Hamidi, Morgan Klaus Scheuerman, and Stacy M Branham. 2018. Gender recognition or gender reductionism? The social implications of embedded gender recognition systems. In *Proceedings of the 2018 chi conference on human factors in computing systems*. 1–13.

[52] Caroline Haskins. 2019. Amazon is coaching cops on how to obtain surveillance footage without a warrant. https://www.vice.com/en/article/43kga3/amazon-is-coaching-cops-on-how-to-obtain-surveillance-footage-without-a-warrant. *Motherboard (Vice News)* (2019).

[53] Caroline Haskins. 2019. US cities are helping people buy Amazon surveillance cameras using taxpayer money. *Motherboard (Vice News)* (2019).

[54] Rebecca Heilweil. 2020. Members of Congress want to know more about law enforcement's surveillance of protesters. https://www.vox.com/recode/2020/5/29/21274828/drone-minneapolis-protests-predator-surveillance-police. *Vox* (2020).

[55] Tanja Heuer, Ina Schiering, and Reinhard Gerndt. 2019. Privacy-centered design for social robots. *Interaction Studies* 20, 3 (2019), 509–529.

[56] Inês Hipólito, Katie Winkle, and Merete Lie. 2023. Enactive Artificial Intelligence: Subverting Gender Norms in Robot-Human Interaction. *Frontiers in Neurorobotics* 17 (2023), 77.

[57] Yoyo Tsung-Yu Hou, EunJeong Cheon, and Malte F Jung. 2024. Power in Human-Robot Interaction. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction*.

[58] Simon Ings. 2019. Clones are not us. *New Scientist* 244, 3254 (2019), 30.

[59] Ryan Blake Jackson and Tom Williams. 2019. Language-capable robots may inadvertently weaken human moral norms. In *Proceedings of the 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 401–410.

[60] Ryan Blake Jackson and Tom Williams. 2022. Enabling morally sensitive robotic clarification requests. *ACM Transactions on Human-Robot Interaction (THRI)* 11, 2 (2022), 1–18.

[61] Ryan Blake Jackson, Tom Williams, and Nicole Smith. 2020. Exploring the role of gender in perceptions of robotic noncompliance. In *Proceedings of the 2020 ACM/IEEE international conference on human-robot interaction*. 559–567.

[62] Immanuel Kant. 1764. Observations on the Feeling of the Beautiful and Sublime. (1764). Translated by John T. Goldthwait, 1960.

[63] Immanuel Kant. 1777. Of the different human races. Translated by Jon Mark Mikkelsen, 1999.

[64] Hagit Keysar. 2019. A spatial testimony: The politics of do-it-yourself aerial photography in East Jerusalem. *Environment and Planning D: Society and Space* 37, 3 (2019), 523–541.

[65] Boyoung Kim, Ruchen Wen, Qin Zhu, Tom Williams, and Elizabeth Phillips. 2023. The impact of different ethical frameworks underlying a robot's advice on charitable donations. In *Proceedings of the IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*.

[66] Raj Korpan. 2023. Trust in Queer Human-Robot Interaction. *arXiv preprint arXiv:2311.07458* (2023).

[67] Roger Lane. 2013. *Policing the city: Boston, 1822-1885*. Harvard University Press.

[68] Rosalyn M Langedijk and Kerstin Fischer. 2023. Persuasive Robots in the Field. In *International Conference on Persuasive Technology*. Springer, 251–264.

[69] Hee Rin Lee, EunJeong Cheon, Maartje De Graaf, Patrícia Alves-Oliveira, Cristina Zaga, and James Young. 2019. Robots for social good: exploring critical design for HRI. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 681–682.

[70] Hee Rin Lee, Selma Šabanović, Wan-Ling Chang, Shinichi Nagata, Jennifer Piatt, Casey Bennett, and David Hakken. 2017. Steps toward participatory design of social robots: mutual learning with older adults with depression. In *Proceedings of the 2017 ACM/IEEE international conference on human-robot interaction*. 244–253.

[71] Timm Linder, Sven Wehner, and Kai O Arras. 2015. Real-time full-body human gender recognition in (RGB)-D data. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 3039–3045.

[72] JeeLoo Liu. 2017. Confucian robotic ethics. In *International Conference on the Relevance of the Classics under the Conditions of Modernity: Humanity and Science*.

[73] Lisa Lowe. 2023. The Dialectics of Abolition. *American Quarterly* 75, 2 (2023), 371–376.

[74] Bertram F Malle and Matthias Scheutz. 2020. Moral competence in social robots. In *Machine ethics and robot ethics*. Routledge, 225–230.

[75] Steve Mann, Jason Nolan, and Barry Wellman. 2003. Sousveillance: Inventing and using wearable computing devices for data collection in surveillance environments. *Surveillance & society* 1, 3 (2003), 331–355.

[76] Justin McCurry. 2015. Erica, the "most beautiful and intelligent" android, leads Japan's robot revolution. *The Guardian* (2015).

[77] Charles W Mills. 2005. Kant's Untermenschen. In *Race and Racism in Modern Philosophy*. Cornell University Press, 169–193.

[78] James H Moor. 2006. The nature, importance, and difficulty of machine ethics. *IEEE intelligent systems* 21, 4 (2006), 18–21.

[79] Terran Mott, Alexandra Bejarano, and Tom Williams. 2022. Robot Co-design Can Help Us Engage Child Stakeholders in Ethical Reflection. In *Proceedings of the 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 24.8% acceptance rate.

[80] Mike Murphy. 2019. Amazon's Ring may be branching out beyond outdoor cameras. https://qz.com/1646116/amazons-ring-might-be-getting-more-cameras-patents-show. *Quartz* (2019).

[81] Leah Nelson, Frank Knaack, and Will Tucker. 2018. *Alabama's War on Marijuana*. Technical Report. t Alabama Appleseed Center for Law & Justice and the Southern Poverty Law Center.

[82] Louis Neven. 2010. 'But obviously not for me"': robots, laboratories and the defiant identity of elder test users. *Sociology of health & illness* 32, 2 (2010), 335–347.

[83] Jaye Nias, Lelia Hampton, Princess Sampson, and Margie Ruffin. 2020. Decolonizing Technologies for Preserving Cultural and Societal Diversity. In *Proceedings of the CHI 2020 Workshop on Engaging in Race in HCI*.

[84] Jaye Nias and Margie Ruffin. 2020. CultureBot: A Culturally Relevant Humanoid Robotic Dialogue Agent. In *Proceedings of the 2020 ACM Southeast Conference*. 280–283.

[85] Shuichi Nishio, Hiroshi Ishiguro, and Norihiro Hagita. 2007. Geminoid: Teleoperated android of an existing person. *Humanoid robots: New developments* 14, 343-352 (2007), 10–1109.

[86] Anastasia K Ostrowski, Hae Won Park, and Cynthia Breazeal. 2020. Design research in hri: Roboticists, design features, and users as co-designers. In *Workshop on Designerly HRI Knowledge*.

[87] Anastasia K Ostrowski, Raechel Walker, Madhurima Das, Maria Yang, Cynthia Breazea, Hae Won Park, and Aditi Verma. 2022. Ethics, Equity, & Justice in Human-Robot Interaction: A Review and Future Directions. In *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 969–976.

[88] Amit Kumar Pandey and Rodolphe Gelin. 2018. A mass-produced sociable humanoid robot: Pepper: The first machine of its kind. *IEEE Robotics & Automation Magazine* 25, 3 (2018), 40–48.

[89] Giulia Perugia, Stefano Guidi, Margherita Bicchi, and Oronzo Parlangeli. 2022. The Shape of Our Bias: Perceived Age and Gender in the Humanoid Robots of the ABOT Database. In *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*. 110–119.

[90] Arnaud Ramey and Miguel A Salichs. 2014. Morphological gender recognition by a social robot and privacy concerns. In *2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 272–273.

[91] Kathleen Richardson. 2015. *An anthropology of robots and AI: Annihilation anxiety and machines*. Routledge.

[92] Laurel Riek and Don Howard. 2014. A code of ethics for the human-robot interaction profession. In *Proceedings of We Robot*.

[93] Paul Robinette, Wenchen Li, Robert Allen, Ayanna M Howard, and Alan R Wagner. 2016. Overtrust of robots in emergency evacuation scenarios. In *2016 11th ACM/IEEE international conference on human-robot interaction (HRI)*. IEEE, 101–108.

[94] Caitlin Rogers. 2018. The aftermath of police blowing up a Maine man's home with a bomb robot. https://bangordailynews.com/2018/08/28/mainefocus/the-aftermath-of-maine-police-exploding-a-mans-house-with-a-bomb-robot/. *Bangor Daily News* (2018).

[95] Henry Rosemont Jr and Roger T Ames. 2016. *Confucian role ethics: A moral vision for the 21st century?* V&R Academic.

[96] Selma Šabanović, Casey C Bennett, Jennifer A Piatt, Wynnie Chang, David Hakken, Sangguk Kang, and David Ayer. 2014. Participatory design of socially assistive robots for preventive patient-centered healthcare. In *IEEE/RSJ IROS workshop on assistive robotics for individuals with disabilities*.

[97] Alessia Saggese, Mario Vento, and Vincenzo Vigilante. 2019. MIVIABot: a cognitive robot for smart museum. In *International Conference on Computer Analysis of Images and Patterns*. Springer, 15–25.

[98] Anara Sandygulova and Gregory MP O'Hare. 2018. Age-related Differences in Children's Associations and Preferences for a Robot's Gender. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. 235–236.

[99] Morgan Klaus Scheuerman, Madeleine Pape, and Alex Hanna. 2021. Auto-essentialization: Gender in automated facial analysis as extended colonial project. *Big Data & Society* 8, 2 (2021).

[100] Morgan Klaus Scheuerman, Kandrea Wade, Caitlin Lustig, and Jed R Brubaker. 2020. How we've taught algorithms to see identity: Constructing race and gender in image databases for facial analysis. *Proceedings of the ACM on Human-computer Interaction* 4, CSCW1 (2020), 1–35.

[101] Katie Seaborn and Peter Pennefather. 2022. Neither "hear" nor "their": Interrogating gender neutrality in robots. *arXiv preprint arXiv:2205.00183* (2022).

[102] Solace Shen, Petr Slovak, and Malte F Jung. 2018. " Stop. I See a Conflict Happening." A Robot Mediator for Young Children's Interpersonal Conflict Resolution. In *Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction*. 69–77.

[103] Sara Sidner and Mallory Simon. 2016. How robot, explosives took out Dallas sniper in unprecedented way. https://www.cnn.com/2016/07/12/us/dallas-police-robot-c4-explosives/index.html. *CNN* (2016).

[104] Michael Steven Siegel. 2008. *Persuasive robotics: how robots change our minds*. Ph. D. Dissertation. Massachusetts Institute of Technology.

[105] Ryan Songalia. 2020. The Activist Who Streamed a Bizarre Standoff With the NYPD Will Keep Fighting. https://www.vice.com/en/article/jgxkm4/black-lives-matter-activist-nypd-drones-instagram-live. *VICE News* (2020).

[106] Robert Sparrow. 2019. Do robots have race?: Race, social construction, and HRI. *IEEE Robotics & Automation Magazine* 27, 3 (2019), 144–150.

[107] Katta Spiel, Os Keyes, Ashley Marie Walker, Michael A DeVito, Jeremy Birnholtz, Emeline Brulé, Ann Light, Pınar Barlas, Jean Hardy, Alex Ahmed, et al. 2019. Queer(ing) HCI: Moving forward in theory and practice. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–4.

[108] Michael Stolp-Smith and Tom Williams. 2024. More Than Binary: Transgender and Nonbinary Perspectives on Human Robot Interaction. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI)*.

[109] Cass R Sunstein. 2014. *Why nudge?: The politics of libertarian paternalism*. Yale University Press.

[110] Scott Tighe and William Brown. 2015. The Militarization of Law Enforcement: Bypassing the Posse Comitatus Act. *Justice Policy Journal* 12, 2 (2015).

[111] American Civil Liberties Union, Ezekiel Edwards, Emily Greytak, Brooke Madubuonwu, Thania Sanchez, Sophie Beiers, Charlotte Resing, Paige Fernandez, and Sagiv Galai. 2020. *A tale of two countries: Racially targeted arrests in the era of marijuana reform*. ACLU.

[112] Tyler Wall. 2013. Unmanning the police manhunt: Vertical security as pacification. *Socialist Studies/Études Socialistes* (2013).

[113] Tom Ward. 2017. Dubai Wants Robots to Make up 25% of Its Police Force by 2030. https://futurism.com/dubai-wants-robots-to-make-up-25-of-its-police-force-by-2030. *Futurism* (2017).

[114] Ruchen Wen, Zhao Han, and Tom Williams. 2022. Teacher, teammate, subordinate, friend: Generating norm violation responses grounded in role-based relational norms. In *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 353–362.

[115] Ruchen Wen, Boyoung Kim, Elizabeth Phillips, Qin Zhu, and Tom Williams. 2023. Comparing Norm-Based and Role-Based Strategies for Robot Communication of Role-Grounded Moral Norms. *ACM Transactions on Human-Robot Interaction* (2023).

[116] Norbert Wiener. 1988. *The human use of human beings: Cybernetics and society*. Da Capo Press.

[117] Kristian Williams. 2015. *Our enemies in blue: Police and power in America*. Ak Press.

[118] Tom Williams. 2020. "No Justice, No Robots". https://nojusticenorobots.github.io/.

[119] Tom Williams. 2023. The Eye of the Robot Beholder: Ethical Risks of Representation, Recognition, and Reasoning over Identity Characteristics in Human-Robot Interaction. In *Companion Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (alt.HRI)*. ACM.

[120] Tom Williams and Kerstin Haring. 2023. No Justice, No Robots: From the Dispositions of Policing to an Abolitionist Robotics. In *Proceedings of the AAAI/ACM Conference on Artificial Intelligence, Ethics, and Society (AIES)*.

[121] Tom Williams, Qin Zhu, Ruchen Wen, and Ewart J de Visser. 2020. The confucian matador: three defenses against the mechanical bull. In *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. 25–33.

[122] Katie Winkle. 2023. Power, Equity, and Building Better Robots. *XRDS: Crossroads, The ACM Magazine for Students* 30, 1 (2023), 64–68.

[123] Katie Winkle, Ryan Blake Jackson, Gaspar Isaac Melsión, Dražen Brščić, Iolanda Leite, and Tom Williams. 2022. Norm-breaking responses to sexist abuse: A cross-cultural human robot interaction study. In *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 120–129.

[124] Katie Winkle, Séverin Lemaignan, Praminda Caleb-Solly, Ute Leonards, Ailie Turton, and Paul Bremner. 2019. Effective persuasion strategies for socially assistive robots. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 277–285.

[125] Katie Winkle, Donald McMillan, Maria Arnelid, Katherine Harrison, Madeline Balaam, Ericka Johnson, and Iolanda Leite. 2023. Feminist human-robot interaction: Disentangling power, principles and practice for better, more ethical HRI. In *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*. 72–82.

[126] Katie Winkle, Gaspar Isaac Melsión, Donald McMillan, and Iolanda Leite. 2021. Boosting robot credibility and challenging gender norms in responding to abusive behaviour: A case for Feminist robots. In *Companion of the 2021 ACM/IEEE international conference on human-robot interaction*. 29–37.

[127] Katie Winkle, Emmanuel Senft, and Séverin Lemaignan. 2021. LEADOR: A method for end-to-end participatory design of autonomous social robots. *Frontiers in Robotics and AI* 8 (2021).

[128] D Robert Worley. 2012. *Orchestrating the Instruments of Power*.

[129] Qin Zhu. 2018. Engineering ethics education, ethical leadership, and Confucian ethics. *International Journal of Ethics Education* 3, 2 (2018), 169–179.

[130] Qin Zhu, Tom Williams, Blake Jackson, and Ruchen Wen. 2020. Blame-laden moral rebukes and the morally competent robot: A Confucian ethical perspective. *Science and Engineering Ethics* 26 (2020), 2511–2526.

[131] Yifei Zhu, Ruchen Wen, and Tom Williams. 2024. Robots for Social Justice (R4SJ): Toward a More Equitable Practice of Human-Robot Interaction. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI)*.

[132] Dennis Zuev and Gary Bratchford. 2020. The citizen drone: protest, sousveillance and droneviewing. *Visual Studies* 35, 5 (2020), 442–456.