Cultivating Ethical Engineers in the Age of AI and Robotics: An Educational Cultures Perspective

Stephen C. Rea
Engineering, Design & Society
Colorado School of Mines
Golden, Colorado, USA
srea@mines.edu

Qin Zhu
Humanities, Arts & Social Sciences
Colorado School of Mines
Golden, Colorado, USA
qzhu@mines.edu

Dean Nieusma
Engineering, Design & Society
Colorado School of Mines
Golden, Colorado, USA
nieusma@mines.edu

Kylee Shiekh
Applied Mathematics & Statistics
Colorado School of Mines
Golden, Colorado, USA
knshiekh@mines.edu

Tom Williams
Computer Science
Colorado School of Mines
Golden, Colorado, USA
twilliams@mines.edu

***Abstract*—This paper considers the cultivation of ethical identities among future engineers and computer scientists, particularly those whose professional practice will extensively intersect with emerging technologies enabled by artificial intelligence (AI). Many current engineering and computer science students will go on to participate in the development and refinement of AI, machine learning, robotics, and related technologies, thereby helping to shape the future directions of these applications. Researchers have demonstrated the actual and potential deleterious effects that these technologies can have on individuals and communities. Together, these trends present a timely opportunity to steer AI and robotic design in directions that confront, or at least do not extend, patterns of discrimination, marginalization, and exclusion. Examining ethics interventions in AI and robotics education may yield insights into challenges and opportunities for cultivating ethical engineers. We present our ongoing research on engineering ethics education, examine how our work is situated with respect to current AI and robotics applications, and discuss a curricular module in "Robot Ethics" that was designed to achieve interdisciplinary learning objectives. Finally, we offer recommendations for more effective engineering ethics education, with a specific focus on emerging technologies.**

Keywords—ethics, education research, AI, robotics, emerging technologies

I. Introduction

This paper considers the cultivation of ethical identities among future engineers and computer scientists, particularly those whose professional practice will extensively intersect with emerging technologies enabled by artificial intelligence (AI). Many current engineering and computer science students will go on to participate in the development and refinement of AI, machine learning, robotics, and related technologies, thereby helping to shape the future directions of these applications. AI is a wide-ranging class of tools that can be used for data analysis and integration in support of decision-making in real-world applications across a variety of sectors, including finance, national security, health care, criminal justice, transportation, and risk mitigation [1]. Increasingly, these applications are integrated [2] [1] with legacy technological systems and infrastructures in the interest of improved efficiency, accuracy, and convenience. These integrations are occurring in both the private sector (e.g., customer service chatbots, algorithmic content recommenders, smart home devices, autonomous vehicle fleets) and the public sector (e.g., automated decisionmakers [ADMs] in entitlement distribution, predictive policing, government procurement services, and a range of military applications). Many such applications are generating substantial economic and social benefits, with one economist predicting that AI-related developments will result in an increased GDP of $3.7 trillion for North America by 2030 [1].

The proliferation of AI and robotics integrations and the associated growth in supporting educational programming provide a ripe opportunity to critically explore—and potentially inflect—the broad social impacts of these emerging technologies. According to the Bureau of Labor Statistics, employment in computer and information research, which includes AI and machine learning, is projected to grow 22 percent by 2030, much faster than the average for all occupations [3]. The U.S. federal government supports the rise in AI research and is projected to invest more than $6 billion in AI-related research-and-development projects in 2021. Further,

This work was supported by U.S. National Science Foundation grant 1909847 and extends prior research supported by grant 1635887.

[1] Alexandra Mateescu and M. C. Elish argue that it is more accurate to understand AI systems as being *integrated* rather than "deployed" or "applied" in order to emphasize how they are fundamentally sociotechnical and as such are inseparable from broader social and cultural processes.

978-0-7381-1423-1/21/$31.00 ©2021 IEEE

the U.S. Congress recently passed an omnibus appropriations bill to provide additional funding for military-related AI initiatives [4]. At the same time, researchers have demonstrated the actual and potential deleterious effects that these technologies have on individuals and communities with respect to fairness, accountability, and transparency as well as power, justice, and equity. Together, these trends present a timely opportunity to steer AI and robotic integrations in directions that confront, or at least do not extend, patterns of discrimination, marginalization, and exclusion. For example, scholars have called recently for the design of robots with an explicitly feminist, pro-social justice stance [5][6]; nascent efforts have attempted to design robots for and by communities of color [7] and related calls have advocated decolonizing AI and leveraging AI for decolonialist projects [8][9][10]. Systematically redirecting AI and robotic systems toward equity, however, begins with how we educate the engineers and computer scientists who will contribute to the development and integration of AI systems and robots of the future.

The next section provides additional background on the ethical challenges of emerging technologies and our team's broader research project on the cultural dimensions of engineering ethics education. This is followed by a brief overview of some of the ethical questions and implications specifically relevant to AI and robotics integrations. The fourth section presents a case study of an interdisciplinary computer-science-and-ethics course that was designed to integrate ethical sensitivity and reflexivity into the design of robots and human-robot interactions. Finally, the conclusion offers possible educational interventions for more effectively preparing future engineers to address ethical conundrums, particularly as they arise in AI and robotics but also across the spectrum of emerging technologies.

II. Background

Calls to address the ethical dimensions of various emerging technologies have grown alongside the development of these technologies and the growth of supporting educational programming [11]. For AI and robotics programs, these calls include a focus on "ethical" or "responsible" technology design as part of engineering and computer science curricula [12][13].[2] While there have been considerable efforts to incorporate ethics into engineering education over the past two decades, the successes and failures of such efforts are difficult to assess [14]. Engineering students often encounter ethics as standalone academic content via dedicated courses or modules. By separating and compartmentalizing lessons about engineering's ethical and social impacts from other coursework, educators (as agents) and curricula (as structures) communicate implicitly to students that such considerations are ancillary to "real" engineering: worth discussing, perhaps, but not fundamental to engineering or engineers' professional identities.

Another risk of ethics instruction in engineering is haphazard, patchwork, or opportunistic but superficial inclusion. For engineering programs, ABET accreditation requires learning outcomes including recognizing ethical and professional responsibilities and consideration of social and environmental impacts of engineering solutions [15]. This often results in ethics "add ins" in courses not otherwise attentive to engineering's social implications or context. Such check-box approaches to ethics instruction can communicate that having any kind of ethical discussion is better than nothing, regardless of whether such discussion is effectively structured or facilitated. Education researchers often describe such implicit messaging as part of a field's "hidden curriculum," in this case reflecting engineering's embedded values that segregate technical knowledge and practice from their social contexts and demote ethics to generalized professional competency rather than core "engineering" knowledge [16].

Our analysis of the ethical challenges of AI, robotics, and related emerging technologies draws on distinct, complementary perspectives among our team members: an anthropologist focused on digital culture, an engineering ethics scholar focused on emerging technologies, an STS scholar focused on institutional innovations in engineering education, an undergraduate student majoring in applied math and statistics and interested in educational research, and a computer scientist focused on natural language-based human-robot interaction. This group convened around a shared interest in engineering ethics education and attentiveness to the cultural dimensions of engineering education in particular. Our approach investigates the intersection of formal, planned educational activities (e.g., courses, curricula, and student and professional codes of ethics) and the full range of informal, unplanned yet widely shared experiences that shape learning outcomes (e.g., as conveyed by shared mindsets, implicit value judgments, and distributed social interactions).

This work extends earlier research by Dean Nieusma at his prior institution, which entailed extensive interviews with 18 undergraduate engineering students at a technological university in the northeastern U.S. [17]. Students at various levels (first-year to senior) and from a variety of engineering and science disciplines, including computer science, participated. Interviews followed a semi-structured protocol that was designed to elicit progressively detailed portrayals of students' understandings of and engagement with questions of ethics in both personal and educational contexts, with the latter including both formal and informal settings. Interviews were fully transcribed, and transcripts were coded independently using emergent themes by three of the authors. These proto-themes were compared and iterated to work progressively toward themes that both categorized and connected across the entire set of data [18].

Findings from these interviews encapsulate the tensions noted above around the lack of integration among technical and social/ethical dimensions of students' educational experience. The present analysis also draws on distinct prior research by other team members. Stephen C. Rea previously synthesized the literature on the social impacts of AI and machine learning to

[2] For the purposes of this paper, we will use the term "ethics" to encompass both ethics and responsibility frameworks, ignoring the sometimes-nuanced distinction: the former tends to emphasize the integration of well-defined, predetermined ethical values and principles into engineering design and education, whereas the latter focuses on analyzing, anticipating, and designing moral impacts of technologies on human perceptions and behaviors. We will also focus our attention on engineering education while recognizing its considerable but not complete overlap with computer science education in most programs in the U.S.

provide recommendations to the consumer financial services industry [19]. Tom Williams and Qin Zhu developed and co-taught a course entitled "Robot Ethics" at the Colorado School of Mines, which aims to cultivate students' moral sensitivity toward the ethical issues arising from the design and deployment of robotic systems [20]. We elaborate these contributions in the sections that follow.

While student interviewees in Nieusma's prior research project all agreed that engineers must consider the personal, professional, and social dimensions of their work—and how it intersects with ethical implications—most of them did not see ethics as central to, or even a notable aspect of, their formal education. As one student put it, "[Professors] want to teach us pure technical skills and [don't] really want us to focus on anything else because they're trying to get engineering through [to] us, not ethics. They think of us as engineering students, not as students trying to learn ethics at the same time." Another student concurred: "I feel like if they had more classes, they'd be able to go into [ethics], but at the same time they don't really have time for more classes. … If you make more [ethics classes mandatory], then you're cutting into the ability to actually learn what your major is." Sentiments like these reflect a belief common among engineering students and professors alike that ethics education is a nice supplement in principle, but that it should not displace any of the "real" engineering curriculum.

Students also conveyed how they perceive their relationship to ethics, both in the present and in the future. For example, in response to an interviewer's question about who defines problems for engineers, one student noted, "I think no one really defines the problem; I think it's just society. Engineers look at the problem and they try to look at it in a certain way so they can find a solution to it, but [they are] alienated from the system." This student's comments arguably represent a dominant mindset in engineering [21], in which "engineering" is a purely analytic activity distinct and set apart from critical reflection on the social world in which that activity manifests.

In this way, students developed an "invisibility" perception of ethics, where ethics was considered as either irrelevant or relevant but absent from their engineering courses. Students commented that "[they] don't really do anything in classes that would require a lot of ethical decisions"; "[ethics] was mentioned in philosophy classes but not really so much in pure engineering [classes]"; and "[ethics] has never been explicitly talked about in [our disciplinary engineering classes]." Engineering course instructors may thus be missing opportunities to engage students in surfacing ethical considerations because they seem to be irrelevant to purely technical problems, while then relying on extra efforts to produce educational opportunities that integrate ethics back into technical engineering problems.

While the social and ethical implications of technology are pertinent across engineering disciplines, they are notably germane to AI and robotics, because these emerging technologies have a confounding combination of immediate and far-reaching, nuanced and profound, actual and potential social impacts. Computer science—the field most closely involved in AI and robotics research—has grappled for decades with questions of ethical and responsible technology. It appears that there is growing recognition that future engineers whose work will be fundamentally mediated by AI and robotics must have grounding in ethics education in order to design and build new technologies in a responsible way. Examining ethics interventions in AI and robotics curricula may yield insights into challenges and opportunities for engineering ethics education that are more broadly applicable and oriented to the future.

III. Ethical and Social Impacts of AI and Robotics

Our interview data suggest that engineering students understand well that their work will impact society; in fact, this is a motivating factor for many students to pursue engineering degrees and careers in the first place. Many of our interviewees described being drawn to their fields of study by a desire to fix problems, to "do good" in the world [22]. At the same time, as one student put it, "I'm not quite sure we're all entirely ready enough to get into the consequences of our actions." Regarding AI and robotics specifically, belief in the (presumed) ethical neutrality of these technologies is prevalent. Among their greatest value propositions is that AI systems will remove human biases from classification and decision-making and that autonomous robotic systems will not make the kinds of errors—some deadly—that humans make. This faith in the ethical neutrality of AI and robotic systems becomes an obstacle to recognizing their fundamentally *sociotechnical* characteristics, that is, the inherent and usually complex intertwining of a given artifact's social and technical facets. We hope to show that, by exposing engineering students to these technologies' social entanglements, they can take a critical first step in developing their ability to devise ethical and responsible tools in the future.[3]

Creating ethical technologies requires recognizing the central role played by context. AI systems and the machine learning (ML) algorithms that power them do not operate in an ethical vacuum, free of human relationships and interests. No matter the specifics of their design or purpose, algorithms and the systems they support are always already embedded in social contexts that affect and are affected by them. Science and technology studies scholars call this the "mutual shaping" of technology and society [25]. Biases can be encoded in the datasets on which ML algorithms are trained, arising from poor sampling strategies, incomplete or erroneous information, or the social inequalities that already exist in the world. For example, the legacy of racism embedded in the United States' social hierarchies has ripple effects that can be observed in data about criminal justice sentencing [26], employment and hiring patterns [27], housing [28], and credit discrimination [29], to name but a few instances. And since ML algorithms and AI systems cannot

[3] Though beyond the purview of this paper, it bears mentioning that many critics reject the very possibility of "ethical" AI and/or robots, in large part because the social contexts of their integration are perceived to be unethical. For more detailed discussions in this vein, see [23] [24].

build themselves, the humans who construct them may, however unintentionally, introduce their own biases when deciding on a model's goals, selecting features, identifying which attributes are relevant, and developing classifiers.[4]

As law and technology experts Julia Powles and Helen Nissenbaum note, "Bias is a social problem, and seeking to solve it within the logic of automation is always going to be inadequate" [31]. They assert that focusing solely on building more effective, fairer AI is ultimately a distraction from addressing the foundations of structural inequality. This perspective is also echoed by scholars such as Bennett and Keyes [32] and Le Bui and Noble [33], who have argued for moving beyond the "fairness"-centered second wave of AI ethics research and towards a third wave grounded in social justice and similar approaches. Others commentators take this criticism further, emphasizing how AI integrations not only are incapable of fixing fundamentally unethical institutional processes, but also can exacerbate the structural inequalities that those processes uphold. Some also argue that bias and discrimination are built into the very business models that drive AI development [34]. Artist and scholar, Mimi Onuoha, has coined the term "algorithmic violence" to refer to the "violence that an algorithm or automated decision-making system inflicts by preventing people from meeting their basic needs." She elaborates: "[Forms of algorithmic violence] not only affect the ways and degrees to which people are able to live their everyday lives, but in the words of Mary K. Anglin, they 'impose categories of difference that legitimate hierarchy and inequality'" [35][see also 36].

The ethical and social impacts of robotics are arguably even more evident than for AI, as many contemporary robotics systems and devices are designed for contexts of intensive human-robot interaction (e.g., elder care, physical therapy and rehabilitation, and advanced manufacturing). And while there is an assumption of ontological separation between robotics and the social world (as with AI), the contexts that robots are integrated within are also shaped by patterns of inequality, marginalization, and discrimination—patterns they may ultimately reinforce. For example, Knightscope was inspired to develop its "fully autonomous security robots" [37] after the 2012 Sandy Hook massacre and 2013 Boston Marathon bombing "to reduce both crime and the economic burden it places on every facet of society … [by] integrat[ing] directly into existing security infrastructures, allowing the public and private sectors to proactively build stronger communities while empowering them to be safer" [38]. However, the use of its products in contexts like shopping malls complements existing surveillance practices with demonstrable patterns of race- and class-based discrimination [39].

Another example is that of lethal autonomous weapons systems, such as military drones, which are sometimes colloquially called "killer robots" [40]. The development and use of these AI-enabled weapon systems is sometimes predicated on the idea of waging "ethical war" [41] by eliminating human biases and errors in target discrimination and making more proportional military actions possible. However, just-war ethicists argue, "the idea that moral judgment in the complex on-the-ground ever-evolving circumstances in war is simply an enduring illusion, with AI being its most technologically advanced iteration to date" [42].

Robotic systems also unsettle where, when, and how accountability is determined if something should go wrong with their operation, as well as which avenues of recourse are available to injured parties. M. C. Elish analyzes where legal responsibility lies when autonomous vehicles are involved in accidents, proposing the concept of a "moral crumple zone": "how responsibility for an action may be misattributed to a human actor who had limited control over the behavior of an automated or autonomous system" [43]. Elish argues, "the moral crumple zone protects the integrity of the technological system, at the expense of the nearest human operator" [43]. The removal of human decision-makers, or their marginalization within human-in-the-loop systems, not only challenges legal precedents for assigning blame, but also raises important ethical questions regarding human rights and dignity among others [44].

IV. Case Study: Robot Ethics

Two general educational design approaches have been deployed in teaching ethics of AI and robotics: (1) standalone courses, typically offered as either computer science or philosophy courses and (2) ethics-focused modules in AI and robotics courses whose primary focus is typically technical [45]. From the perspective of moral psychology, there is a gap between ethical reasoning (i.e., knowing what is good versus bad and why) and ethical action (i.e., committing and acting to do good) [46]. Effective ethics education requires students to relate their moral learning experiences to their own everyday personal and professional experiences (that is, how they actually engage in their social worlds) [47]. Efforts to teach engineering ethics that focus mainly on *ethical reasoning* precariously assume that good ethical reasoning skills will lead to ethical actions and outcomes in real-world contexts. Teaching students various ethical reasoning tools (e.g., codes of ethics and ethical theories and frameworks) and asking them to apply these tools to hypothetical cases or future career practices can, ironically, risk undermining ethical action. This "applied ethics" approach can cause a disconnection between students' education experience and the realities of ethical decision making in contemporary engineering practice [48].

Ethical issues arising in real-world engineering practice are often messier than hypothetical cases, without clear hints about which ethical tools are ready to apply in particular situations. Instead, what is needed is high-level ethical sensitivity and the ability to act spontaneously in response to incomplete information and ambiguous signals. Such skills are typically cultivated by deep reflective engagement with everyday practical experience more than by memorizing theoretical ethics frameworks. In other words, ethics skills tuned to real-world action arise from effortful and meaningful engagements that "moralize" everyday decision making and the navigation of authentic social contexts. For engineers to be truly ethically competent, they need to develop a perceptual sensitivity that

[4] On how the composition of AI research and design teams affects outcomes, see [30].

allows them to see ethical issues as they naturally emerge from seemingly "ethically neutral" engineering practice. In this way, students (and instructors!) can learn to recast "purely technical" decisions as socially entangled and therefore moral in character. Arguably, perceptual sensitivity to the many ethical dimensions of engineering practice is much more important than moral reasoning skills applied to solve hypothetical ethical problems in inauthentically primed contexts.

These insights into ethics education apply equally to computer-science students in the Robot Ethics course referenced above. Tom Williams and Qin Zhu designed a curricular module to achieve interdisciplinary learning objectives, spanning technical dimensions, human-robot interaction, and normative theory of technology. Concerning the latter, students investigated how robots shape human behaviors due to their perception as moral and social agents, which can then carry over into human-human relationships. As described in [20], the Robot Ethics instructional team evaluated the efficacy of the module through a randomized controlled experiment performed as part of the Spring 2019 course offering. This instance of the course was designed to enable students to apply moral learning lessons to their everyday, situated experiences. It then required them to reflect on the relevance of the former to the latter. The course also asked students to empathize with potential users and their needs and to project the associated moral lessons to an immediate context of professional application, reflecting on the (often powerful) role of technical expertise in shaping the experience of users, and hence thereby shaping "society," albeit in a limited way.

In implementing the new module, an in-class lecture and student-run experiment investigated how the designer's choice of a wake word (e.g., "Alexa," "Okay, Google," and "Hey, Siri") affected user politeness, as directed both at the robot and at other humans. Undergraduate students served either as experimenters or as "wizards": Experimenters learned to guide participants through consent procedures, to brief participants on their experimental tasks, and to debrief participants on experimental motivations upon the study's completion, whereas wizards learned how to teleoperate the robot. Graduate students enrolled in the course served as "confederates"—actors who carried out participant requests. The experiment used a SoftBank Pepper robot, with participants interacting with the robot and a human confederate in a restaurant scenario. Depending on the experimental condition, participants were required to interact with the robot using either a traditional wake word (e.g., "Hey Pepper") or a "polite" wake word (e.g., "Excuse me, Pepper"). To investigate the normative influence of this design choice, the experiment collected linguistic statistics surrounding participants' use of politeness cues in their language towards both the robot and the human confederate throughout the length of the experimental interaction.

While the empirical results of the experiment indicated no learning gains resulting from the new module compared to a traditional, lecture-based curriculum, an intriguing finding arose: Students who explored the normative influence of technology through participation in the new module experiment showed higher retention of related content than those who acted merely as confederates, whose associated lessons were merely lecture-based. We do not assert that these learning gains were sufficient by themselves to justify the additional overhead effort imposed by the new curriculum; however, Williams and Zhu plan to revisit incorporation of the module after addressing some of the process-based lessons learned through the research. Another option is to explore incorporating experiment design and piloting [49], a typical requirement of many human-robot interaction courses. Because these experiments are only piloted rather than being run with real participants, they have little associated cost. And because they are designed by the students themselves, they lead to deeper engagement with the research questions under investigation, as suggested by our experimental ethics module. This combination of benefits is promising, so Williams and Zhu hope to try a pilot-oriented research approach in future offerings of AI and Robot Ethics courses.

V. Conclusion: Ethics Education as Ethical Inquiry

As part of its Global Initiative on Ethics of Autonomous and Intelligent Systems (A/IS), the IEEE has recognized the crucial need for guidelines to ensure that the creation of such systems—including AI and robotics—is aligned with ethical principles. Alongside industry standards, public policies, and societal outreach efforts, the IEEE recommends that "education and training on ethical considerations relating to A/IS … should be developed and presented to students of engineering, A/IS, computer sciences, and other relevant fields. These courses … should be embedded from the beginning to allow for absorption of the underlying ethical considerations as well as allowing for critical thinking to come to fruition once the students graduate" [50]. While we concur with this recommendation, as students and instructors of ethics teaching and learning, we are left unsatisfied. For us, such calls for emerging technology ethics education imply that the educational techniques needed to achieve these objectives are self-evident, that taking the appropriate course of action is a matter of will, not one of inquiry, experimentation, and exploration.

This same shortcoming—presuming that the appropriate course of action is already known and that we merely require the will to follow through—is evident too in much of engineering ethics education. To achieve the goals set out here, ethics instruction must extend beyond encyclopedic knowledge of ethical frameworks and their associated definitions. Rather, students will benefit more from engaging with ethical *inquiry*, that is, recognizing how ethical dilemmas emerge in practice and what they entail for the work that students may find themselves doing in the future. Inquiry is especially important for grappling with emerging technologies like AI and robotics, whose ethical implications are exceedingly complex and whose directions of application uncertain. Educators can leverage students' desire to "do good" by helping them foster their own ethics identities.

Our approach, and our recommendation, is to follow a student- and inquiry-centric approach to ethics education. This approach requires critically examining the role that implicit messaging has on students' perceptions of the relationship between ethics and engineering. This implicit messaging includes both that which occurs in the educational moment (e.g., what is communicated by instructors, teaching assistants, syllabi, and assignments) and across the entire curriculum and discipline. Such implicit messaging can work against students' capacity to recognize that developing their engineering competency entails not only technical knowledge but also cultivating ethical awareness and commitment. The Robot Ethics course accomplished this goal by including students in research experimentation into technology's normative influence on human behavior, thereby inviting students to critically examine the moral nature of specific technical decisions in robot design and, accordingly, the inseparability of these technical and moral dimensions.

Student-centric ethics education in engineering and computer science also requires attention to the myth of technology neutrality and the invisibility of context in much of students' problem solving in engineering coursework. Engineering curricula should emphasize a sociotechnical perspective on technology integration that treats ethics as an inherent and essential component of advancing emerging technologies. Students must confront the assumption that technological solutions exist for every problem and understand that, sometimes, the most ethical course of action is to *not* integrate a specific tool or technological system. Robot Ethics did this by integrating ethics-of-robot-interaction research as component of course requirements, thereby presenting a relational view of ethics that surfaced a normative theory of technology and attention to context.

Acknowledgments

The authors thank the anonymous undergraduate students who participated in the interview research as well as the participants in Robot Ethics during the Spring 2019 semester. This work was supported by National Science Foundation grant 1909847 and extends prior research supported by grant 1635887.

References

[1] D. M. West and J. R. Allen, "How artificial intelligence is transforming the world," Brookings Institution, Washington, D.C., 2018.

[2] A. Mateescu and M. C. Elish, "AI in context: the labor of integrating new technologies," Data & Society, New York, NY, 2019.

[3] Bureau of Labor Statistics, "Occupational outlook handbook: computer and information research scientists," US Department of Labor, 08-Sep-2021. [Online]. Available: https://www.bls.gov/ooh/computer-and-information-technology/computer-and-information-research-scientists.htm#tab-6.

[4] J. Harper, "Federal AI spending to top $6 billion," National Defense, 10-Feb-2021. [Online]. Available: https://www.nationaldefensemagazine.org/articles/2021/2/10/federal-ai-spending-to-top-$6-billion/.

[5] K. Winkle, G. I. Melsión, D. McMillan and I. Leite, "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, pp. 29-37, 2021.

[6] K. Winkle, R. B. Jackson, A. Bejarano and T. Williams, "On the flexibility of robot social identity performance: benefits, ethical risks and open research questions for HRI," in HRI Workshop on Robo-Identity, 2021.

[7] J. Nias, L. Hampton, P. Sampson and M. Ruffin, "Decolonizing technologies for preserving cultural and societal diversity," in CHI 2020 Extended Abstracts, 2020.

[8] S. Mohamed, M-T. Png and W. Isaac, "Decolonial AI: decolonial theory as sociotechnical foresight in artificial intelligence," Philosophy & Technology, vol. 33, no. 4, pp. 659-684, 2020.

[9] R. Eglash, A. Bennett, M. Lachney and W. Babbitt, "Race-positive design: a generative approach to decolonizing computing," in CHI 2020 Extended Abstracts, 2020.

[10] M. Ali, "Towards a decolonial computing," in Ambiguous Technologies: Philosophical Issues, Practical Solutions, Human Nature. International Society of Ethics and Information Technology, 2020, pp. 28–35.

[11] J. Borenstein and A. Howard, "Emerging challenges in AI and the need for AI ethics education," AI and Ethics, vol. 1, pp. 61-65, 2021.

[12] B. J. Grosz et al., "Embedded ethiCS: integrating ethics across CS education," Communications of ACM, vol. 62, no. 8, pp. 54-61, 2019.

[13] A. van Wynsberghe, "A method for integrating ethics into the design of robots," Industrial Robot, vol. 40, no. 5, pp. 433-440, 2013.

[14] D. Bairaktarova and A. Woodcock, "Engineering student's ethical awareness and behavior: a new motivational model," Science and Engineering Ethics, vol. 23, no. 4, pp. 1129–1157, 2016.

[15] ABET. (n.d.) Criteria for accrediting engineering programs, 2021 – 2022. ABET, Baltimore, MD. [Online]. Available: https://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-engineering-programs-2021-2022/.

[16] S. C. Rea, K. Shiekh, Q. Zhu and D. Nieusma, "The hidden curriculum and the professional formation of responsible engineers: a review of relevant literature in ASEE conference proceedings," in ASEE Virtual Annual Conference Content Access, Virtual, 2021.

[17] D. Nieusma and M. Cieminski, "Ethics education as enculturation: student learning of personal, social, and professional responsibility," in ASEE Annual Conference & Exposition, 2018.

[18] J. A. Maxwell, Qualitative Research Design: An Interactive Approach, 3rd ed. Thousand Oaks, CA: SAGE Publications.

[19] S. C. Rea, "A survey of fair and responsible machine learning and artificial intelligence: implications of consumer financial services," SSRN, January 2020. [Online]. Available: https://ssrn.com/abstract=3527034/.

[20] T. Williams, Q. Zhu and D. Grollman, "An experimental ethics approach to robot ethics education," in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, no. 9, 2020.

[21] D. Riley, "Mindsets in engineering," in Engineering and Social Justice. Williston, VT: Morgan & Claypool, 2008, pp. 33-45.

[22] M. S. Kleine and J. C. Lucena, "The world of 'engineering for good': towards a mapping of research, teaching, and practice of engineers doing good," in Middle Atlantic ASEE Section Spring 2021 Conference, Virtual, 2021.

[23] V. Dignum, Responsible Artificial Intelligence. Cham: Springer Nature, 2019.

[24] L. Rainie, J. Anderson and E. A. Vogels, "Experts doubt ethical AI design will be broadly adopted as the norm within the next decade," Pew Research Center, 16-Jun-2021. [Online]. Available: https://www.pewresearch.org/internet/2021/06/16/experts-doubt-ethical-ai-design-will-be-broadly-adopted-as-the-norm-within-the-next-decade/.

[25] D. MacKenzie and J. Wajcman, Eds., The Social Shaping of Technology. Buckingham, UK: Open University Press, 1999.

[26] J. Angwin, J. Larson, S. Mattu and L. Kirchner, "Machine bias," ProPublica, 23-May-2016. [Online]. Available: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing/.

[27] L. Yarger, F. C. Payton and B. Neupane, "Algorithmic equity in the hiring of underrepresented IT job candidates," Online Information Review, vol. 44, no. 2, pp. 383-395, 2019.

[28] R. Rothstein, The Color of Law: A Forgotten History of How Our Government Segregated America. New York: Liverlight, 2017.

[29] L. Rice and D. Swesnik, "Discriminatory effects of credit scoring on communities of color," Suffolk UL Rev., vol. 46, no. 3, pp. 935-966, 2013.

[30] J. Snow, "'We're in a diversity crisis': cofounder of Black in AI on what's poisoning algorithms in our lives," MIT Technology Review, 14-Feb-2018. [Online]. Available: https://www.technologyreview.com/s/610192/were-in-a-diversity-crisis-black-in-ais-founder-on-whats-poisoning-the-algorithms-in-our/.

[31] J. Powles and H. Nissenbaum, “The seductive diversion of ‘solving’ bias in artificial intelligence,” Medium.com, 07-Dec-2018. [Online]. Available: https://medium.com/s/story/the-seductive-diversion-of-solving-bias-in-artificial-intelligence-890df5e5ef53/.

[32] C. L. Bennett and O. Keyes, “What is the point of fairness? Disability, AI and the complexity of justice,” in ACM SIGACCESS Accessibility and Computing, no. 125, 2020.

[33] M. Le Bui and S. U. Noble, “We’re missing a moral framework of justice in artificial intelligence,” in The Oxford Handbook of Ethics of AI. New York: Oxford University Press, 2020, p. 163-179.

[34] A. M. Hamilton, “Silicon Valley pretends that algorithmic bias is accidental. It’s not,” Slate, 07-Jul-2021. [Online]. Available: https://slate.com/technology/2021/07/silicon-valley-algorithmic-bias-structural-racism.html/.

[35] M. Onouha, “Notes on algorithmic violence,” GitHub.com, 08-Feb-2018. [Online]. Available: https://github.com/MimiOnuoha/On-Algorithmic-Violence/.

[36] R. Benjamin, Race After Technology. Boston: Polity, 2019.

[37] Knightscope, n.d. The K5 ASR. Knightscope, Mountain View, CA [Online]. Available: https://www.knightscope.com/k5/.

[38] Knightscope, n.d. About us. Knightscope, Mountain View, CA [Online]. Available: https://www.knightscope.com/about-us/.

[39] C. J. Bates, “Regulating physical places with digital code,” Platypus: The CASTAC Blog, 31-May-2018. [Online]. Available: http://blog.castac.org/2018/05/physical-places-digital-code/.

[40] A. Krishnan, Killer Robots: Legality and Ethicality of Autonomous Weapons. New York: Routledge, 2009.

[41] E. Blue, M. Levine and D. Nieusma, Engineering and War: Militarism, Institutions, Ethics, Alternatives. Morgan & Claypool, 2014.

[42] J. Emery, “Killer robots: algorithmic warfare and techno-ethics,” Platypus: The CASTAC Blog, 07-Mar-2018. [Online]. Available: http://blog.castac.org/2018/03/killer-robots/.

[43] M. C. Elish, “Moral crumple zones: cautionary tales in human-robot interaction,” Engaging Science, Technology, and Society, vol. 5, pp. 40-60, 2019.

[44] R. Brownsword, “From Erewhon to AlphaGo: for the sake of human dignity, should we destroy the machines?” Law, Innovation and Technology, vol. 9, no. 1, pp. 117-153, 2017.

[45] E. Burton et al., “Ethical considerations in artificial intelligence courses,” AI Magazine, vol. 38, no. 2, pp. 22–34, 2017.

[46] J. Rest, D. Narvaez, M. Bebeau and S. Thoma, “A neo-Kohlbergian approach: the DIT and schema theory,” Educational Psychology Review, vol. 11, no. 4, pp. 291–324, 1999.

[47] M. W. Martin, Meaningful Work: Rethinking Professional Ethics. New York: Oxford University Press, 2000.

[48] R. McGinn, The Ethical Engineer: Contemporary Concepts & Cases. Princeton: Princeton University Press, 2018.

[49] J. S. Coker, “Student-designed experiments: a pedagogical design for introductory science labs,” Journal of College Science Teaching, vol. 46, no. 5, pp. 14-19, 2017.

[50] The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, 1st ed. IEEE, 2019.