More Than "If Time Allows": The Role of Ethics in AI Education

Natalie Garrett
Information Science
University of Colorado
Boulder, CO USA
natalie.garrett@colorado.edu

Nathan Beard
Information Studies
University of Maryland
College Park, MD USA
nbeard@umd.edu

Casey Fiesler
Information Science
University of Colorado
Boulder, CO USA
casey.fiesler@colorado.edu

ABSTRACT

Even as public pressure mounts for technology companies to consider societal impacts of products, industries and governments in the AI race are demanding technical talent. To meet this demand, universities clamor to add technical artificial intelligence (AI) and machine learning (ML) courses into computing curriculum—but how are societal and ethical considerations part of this landscape? We explore two pathways for ethics content in AI education: (1) standalone AI ethics courses, and (2) integrating ethics into technical AI courses. For both pathways, we ask: What is being taught? As we train computer scientists who will build and deploy AI tools, how are we training them to consider the consequences of their work? In this exploratory work, we qualitatively analyzed 31 standalone AI ethics classes from 22 U.S. universities and 20 AI/ML technical courses from 12 U.S. universities to understand which ethics-related topics instructors include in courses. We identify and categorize topics in AI ethics education, share notable practices, and note omissions. Our analysis will help AI educators identify what topics should be taught and create scaffolding for developing future AI ethics education.

CCS CONCEPTS

• Social and professional topics→ Computing education.

KEYWORDS

artificial intelligence; ethics education; curriculum

1 INTRODUCTION

Artificial intelligence (AI) is largely considered to be the next wave in computing. As industry looks to AI as a solution to various problems [14], the number of jobs in this sector is growing, and universities are clamoring to create curriculum that can meet the demand for computer scientists with AI expertise [35]. At the same time, the potential for negative impacts of AI continues to grow in rate and scale. The impact that AI is having on individuals and society as a whole is on the forefront of public discourse—but as universities respond to the need for technical talent in AI, how are they responding to the corresponding need to train technologists to consider ethical implications?

The mounting public pressure for companies to consider the societal impacts of their products is evident in the current discourse in news articles we see daily about ethics controversies [12], congressional hearings where tech leaders are questioned about their power [42], and policy initiatives that limit the scope of technologies like facial recognition [9]. The public is questioning the idea that technology is the solution to all of our problems [33] and recognizing the negative social and political impacts of algorithmic systems and machine learning (ML) [36]. The public—including journalists, activists, researchers, regulators, and professional organizations like AAAI and ACM—are calling for more thoughtful, ethical technologies. In a speech at the Electronic Frontier Foundation, danah boyd termed the public pressure as the "Great Reckoning" [5]. Meanwhile, tech leaders are questioned about their ethical practices in public hearings [42, 51], and the public is becoming more aware of the absence of critical ethical thinking by leadership in Silicon Valley. While tech companies may be slow to take action or critically consider the ramifications of what they create, some universities are responding by adding tech ethics courses to computing curriculum, particularly in increasingly important domains like AI, ML, and data science [11, 44].

However, if AI education is in the infancy stage of development, then AI ethics education is barely an embryo. While accredited computer science departments in the U.S. are required by the Accreditation Board for Engineering and Technology (ABET) to produce students with "an understanding of professional, ethical, legal, security and social issues and responsibilities," practices vary amongst universities [40]. As noted in recent work about the current state of general CS ethics education, as demand grows—particularly for integration of ethics into existing courses [15, 40]—an increasing number of instructors will be looking for guidance on what they should be teaching [11]. This guidance feels particularly critical for AI as possibly both the fastest growing technical area and the area with the most potential for pressing ethical concerns. As conversations about responsibility for the social and ethical implications of AI

AIES '20, February 7–8, 2020, New York, NY, USA

ACM 978-1-4503-7110-0/20/02 $15.00
https://doi.org/10.1145/3375627.3375868

continue, it is important to identify what kind of training our future technologists are receiving.

An understanding of how universities are training students to be the next AI professionals therefore is not only valuable to universities but also to industry and government. Over the past few years, companies have been attempting to take the lead through self-regulation and ethics mission statements, but they often differ in terms of topics and fail to consider some of the most critical ethical impacts [13, 51]. However, as the field evolves, it is fortunate that we can learn from current practices; in addition to offering a large number of general tech ethics classes [11], many universities have AI-specific ethics courses or (to a lesser extent) technical AI courses that include ethics content [40]. Therefore, to explore one component of the current state of AI ethics education, we examine a sample of these classes for patterns in what topics are currently being taught.

In this paper, we describe topics in a subset of two types of AI courses: standalone AI ethics courses, where the main learning objectives are related to ethics, and technical AI/ML courses, where the main learning objectives are technical in nature, but that also include some ethics content. Our goal is to spark conversation in the AI community about the ethics content we should teach in computing and beyond. Our exploratory, qualitative analysis of a total of 51 courses sheds light on topics covered, as well as common case studies, controversies, and readings used to discuss real world topics. Through this analysis, we are able to describe current trends in teaching AI ethics, to highlight not only what topics are commonly taught but also identify potential omissions. Ultimately, we are interested in what instructors consider to be key topics in AI ethics to inform future work and provide guidance for other universities and instructors entering this space.

2 BACKGROUND & RELATED WORK

The field of computer ethics is nearly as old as computing itself [47], though public and scholarly attention to this topic has experienced a resurgence, particularly with respect to emerging technologies like AI and ML [40, 44]. Conversations about ethical guidelines for AI are gaining traction as many organizations publish guidelines to signal their public commitment to ethical values or, some argue, to stave off potential government regulation [52]. However, though guidelines are a great start, they often do not lead to tangible implementation [31]. Stark and Hoffman [48] note that ethical codes should be seen only as a starting point to create the world we would like to live in; as Mittelstadt [31] puts it, ethics is a process and not a solution. Once guidelines are created, the real work begins.

The definition of "ethics" is also an ongoing conversation. A number of related concepts such as justice, power, responsibility, and values are important aspects of decisions about how to avoid harmful practices with technology, and how to consider its goals, politics, and consequences [43]. Prior work has considered subsets of, essentially, what we talk about when we talk about ethics in the context of AI—for example, in codes of ethics [48], ethics principles [13], and ethics guidelines [17]. Our work expands upon this space by considering what is being taught in AI ethics.

Strategies to accomplish the goals of ethics education are still up for debate. In the U.S., CS programs are required to include ethics in their curriculum for accreditation, yet universities and professors are left to determine how to implement it. The benefits of ethics education are clear; discipline-specific ethics education has proven to support not only moral development, but also helps convince students that ethics is a part of their profession and not just a public relations add-on [26]. Additionally, within CS, there have been calls for greater integration of ethics across the curriculum and into technical classes for decades [28]. However, though we are seeing some movement in this direction, it is still not common practice [11, 40].

Advocates for AI ethics education also argue that having students learn ethics as part of their technical curriculum is ideal because it takes ethics out of isolation and formalizes it [10]. Of course, once AI ethics is part of the curriculum, there is also the question of how to teach it. Some argue for project-based learning to help students conceptualize real-world societal impact of AI [1] or using science fiction to speculate about future technologies [6]. Though the analysis in this paper does not address pedagogy, the question of what topics are taught is still a step towards guidance for instructors. By benchmarking the content covered in current AI ethics courses, we can begin to identify patterns as well as omissions as society grapples with the consequences of AI.

3 METHODS

In educational research, the method of analyzing syllabi is common for assessing curricular requirements [8]. It has been used to provide insight into strategies for teaching computer science [3, 50], and evaluating a syllabus is considered an effective way to determine knowledge units in a course [50]. Previous work has used this method to determine the content and goals of general tech ethics classes [11]. Here, we focus explicitly on AI ethics and incorporate both standalone AI ethics classes and technical AI classes that include an explicit ethics component. To guide our analysis, our primary research question was: What ethics-related topics (including real-world examples and current events) are covered in AI ethics education, in both classes devoted to the topic and as part of technical classes? With this goal in mind, we first compiled a dataset of 51 course descriptions and syllabi, and then conducted a qualitative content analysis of course topics and readings.

3.1 Dataset

The dataset for this study was compiled from syllabi collected as part of two previous studies on computing ethics education; a subset of the authors were involved in both previous analyses. One study's dataset was collected from an uncurated, publicly available set of "tech ethics" syllabi crowdsourced from instructors; it contained 115 standalone tech ethics classes, a subset of which focused on AI or ML [11]. The other study involved a systematic collection of all AI, ML, and data science

courses from the top 20 U.S. university CS programs as identified by U.S. News and World Report [40]. Out of 186 courses analyzed, 36 mentioned an ethics-related topic (e.g., privacy, fairness, bias) in a course description or syllabus; these included both ethics-specific courses and technical courses that included some ethics content [40].

For the current study, we retrieved AI and ML specific courses from both of these datasets. After removing duplicates, the result was a total of 51 courses: 31 standalone ethics classes from 22 U.S. universities and 20 technical courses from 12 U.S. universities. This new analysis builds on this prior work and provides new insights by (1) deep diving into AI specifically, which allows for a higher level of granularity for AI topics than the analysis in Fiesler et al. [11], as well as additional analysis of readings; (2) analyzes specific topics in technical AI classes (which was not part of the analysis in Saltz et al. [40]); and (3) synthesizes insights across these two previously unrelated datasets in order to consider AI ethics in two different pedagogical contexts.

It is important to note that we are not making any claims about the representativeness of this dataset. Both of the original studies note limitations—particularly, that the standalone classes likely oversample from instructors who are active on social media and comfortable sharing their course materials [11], and the technical classes are limited to well-known U.S. CS programs [40]. Additionally, analysis of some of the technical courses is based on course descriptions rather than full syllabi. Even syllabi that might include detailed learning objectives, content, and weekly reading lists do not always contain fine-grained information about course content or insight into pedagogical aspects that are unique to the in-person experience and instructor-student relationship. A syllabus or course description that does not include ethics-related topics also does not necessarily mean that ethics is not part of the class—but it does indicate whether these topics are *explicit* parts of the course. Our goal with this analysis is not to make claims about the overall state of AI ethics education, but instead, to provide a snapshot of this space and an exploratory look at patterns and practices in this set of examples.

3.2 Data Analysis

After compiling this set of syllabi, we mined each for topics as listed in a schedule, as presented in a reading list, or as listed in a course description. Not all syllabi included all components, and some also included lecture slides or notes. To conduct a content analysis of this dataset, two researchers first conducted open coding and then worked together to create a codebook with thematic categories that could be applied to the entire dataset [27]. During the analysis, we discussed disagreements and edge cases and came to a consensus; given that many of the topics overlap, we used our best judgment to stay true to what we interpreted as the instructor's intent. When available, we used readings, lecture slides, and notes in the syllabus to clarify topics. We used affinity diagramming to group related topics into high-level categories before conducting formal coding of topics for each class [19]. For syllabi that included news articles, we used the codebook to categorize these as well. Though this analysis in which we categorized themes was qualitative, following formal coding we calculated relative frequencies that we include in our findings. During our analysis we recognized that the technical courses were different enough in scope from the AI ethics courses that they should be analyzed separately, which we did using the same methods but coming to a smaller set of high-level categories.

4 FINDINGS

4.1 AI Ethics Courses

Our affinity diagramming of topics in standalone AI ethics courses resulted in eight high level categories (listed in order from most frequent to least frequent): *bias, automation and robots, law & policy, consequences of algorithms, philosophy/morality, privacy, future of AI,* and *history of AI.* These categories are based on topics as described by the instructors in course syllabi, which means that they reflect an organizational scheme as well as content. For example, facial recognition might be discussed in a class session devoted to privacy, or bias, or law. These categories should therefore not be interpreted as having hard lines between them, but instead as a means of describing the overall space. Though for transparency we provide quantification of our topic mapping, we stress that because of these overlaps and fuzzy boundaries, the percentages should not be interpreted as a strong signal of relative importance. There were also a number of topics that did not fall under any of these categories; for the purposes of this description we excluded additional categories that were found in less than five syllabi.

Bias was the most represented topic, found in 87% of the syllabi, and in this category, we include concepts like discrimination and fairness. Some syllabi used a generic description of "bias" while others broke down the topic in a more granular way to include, e.g., gender bias, discrimination, and inequality. Exploring lists of readings for this topic, we observed that the COMPAS recidivism algorithm [25], the use of facial analysis to predict sexuality [34], and the case of Google mislabeling an African American woman as a gorilla [16] were all current events commonly used to demonstrate this concept.

The topic of *automation and robots* appeared in 71% of the syllabi. Instruction in this topic includes the various societal consequences of automation in, for example, work, labor, and the economy. Additionally, many courses covered the implications of technologies like autonomous weapons [54] and decisions self-driving cars must make [32]. It is worth noting that robots and the future of work were often discussed in tandem. Two examples are how self-driving trucks will change the trucking industry [39] and whether people will still find meaning in a future where humans do not have to work [18]. Some classes focused on robots specifically and included debates about robot rights [20] and how humans and robots can both learn to work together [24]. Readings also point to systemic problems like income inequality that will lead certain groups to benefit from AI developments while others bear the brunt of a changing economy as their jobs disappear [38].

The general topic of *law & policy* covers a wide range of topics, ranging from federal and state policies to general governance and questions of who should regulate AI. The inclusion of law (which,

of course, is not the same as ethics) is unsurprising, considering that many ethics-related courses cover both ethics and policy [11]. 55% of syllabi covered this topic, which included current events like law enforcement applications like predictive policing and explanations of current regulations like GDPR.

While only two courses mentioned traditional ethical theories (such as utilitarianism or deontology) in their syllabi, 45% of courses included topics in *Philosophy/Morality* generally. Morality, human responsibility, and even the meaning of life all appeared in these courses. Additionally, courses cover the potential existential threat of AI, how humans will decide to incorporate morals into AI systems, and how to maintain human dignity as AI becomes a larger part of our daily lives.

The combination of filter bubbles, recommender systems, propaganda, and the targeted nature of online advertising all demonstrate *consequences of algorithms*, which was covered in 45% of the syllabi. Beyond issues of bias (which we categorized into its own topic), society-level consequences like the impact of algorithms on democracy, civil rights, and confirmation bias are highlighted here. Current events showcase how predictive algorithms are helping social workers decide whether to involve child protective services [22] or algorithms that are used to predict who might be a victim of a crime [2]. Another common example is how Facebook and other online platforms limit what we see through algorithms that attempt to deliver the results we expect, limiting the variety of information people see [49].

Privacy was an explicit topic in 32% of courses, though we note that discussions of privacy also appeared in conjunction with other topics, like bias, regulation, and automation. Two courses had at least one lecture on differential privacy, and it should be noted that. Readings included articles about the "creepiness" of Facebook's "people you may know" feature, which has been known to recommend inappropriate connections—like between a mental health practitioner's patients [21]—or stories about Cambridge Analytica and the unexpected sharing of Facebook profiles [7]. Similarly, there are articles about database marketing [29] and how companies aggregate information about employees through surveillance and analytics [37].

The topic we labeled *Future of AI* was included in 26% of courses, and 19% of courses included the *History of AI*. Though there is of course overlap with other topics (particularly for the future of AI), we included these categories in part because references to "future" and "history" were so common in the context of topic descriptions, often without additional detail. The *history of AI* includes the history of computing generally as well as the rise of AI, often focusing on theory and the philosophical concepts around intelligence. *Future of AI* topics included projecting into future technologies, the possibility of superintelligence, the future of human jobs, and the potential of singularity. For example, one of the hot debates covered is between tech giants Elon Musk of Tesla and SpaceX and Mark Zuckerberg of Facebook: Musk believes that there should be proactive regulations for AI while Zuckerberg considers Musk's fear harmful [4]. This debate illustrates the hope and skepticism that exists in parallel when considering future AI technologies.

4.2 Technical AI/ML Courses

The systematic syllabi review described by Saltz et al. [40] revealed that out of nearly 200 AI/ML/data science courses, only 12% of technical courses also included some mention of an ethics-related topic in the syllabus or course description (and the methods note that they "erred on the side of inclusivity" in making this judgment). Our description below of 20 of these classes is therefore based on the somewhat rare inclusion of ethics as an explicit component technical AI classes.

Since these courses are technical in nature and do not focus on ethics, we would not expect the breadth of topics seen in the standalone AI ethics courses; therefore we analyzed these separately. The majority of these courses had limited information about topics (they typically did not include or did not list readings, for example); thus our descriptions are also limited. However, we are able to describe the most common topics that appeared. There were also some additional topics that only appeared only once in our dataset, including harassment, sustainability, and the ethics of data collection.

By far, the most common topics in Technical AI/ML courses were *bias, fairness,* and *privacy*. Bias and fairness as topics often went hand-in-hand, though bias was the most common. These topics are sometimes grouped together, and in the majority of the classes in which they appear, the context is *avoiding* bias, *promoting* fairness, and *protecting* privacy. In the context of machine learning, these are often treated as mathematical concepts [41], which may be what we are seeing here. Three classes specifically covered differential privacy, a machine learning framework that is used to mitigate the risk of exposing sensitive training data. However, it is important to note here that the majority of the syllabi listed topics but did not provide an explanation or further context.

There were also a number of vague descriptions that pointed to *general ethics* topics, which we applied as a category to a subset of these classes. "Ethics" as a field and concept is open to interpretation [43], and as shown in a recent analysis of general tech ethics syllabi [11], "ethics" as a topic on a syllabus could cover topics ranging from privacy and bias, to specific ethical theories, to more practical applications like professional ethics. While we cannot comment on the nature of what an instructor is actually teaching when they add the word "ethics" to a syllabus, we believe it is important to note the presence of ethics in a course.

While we had limited access to detailed syllabi or lecture materials, we did observe an interesting trend in timing. With the exception of one course, which mentions ethics on the first day, the majority of courses covered ethics-related topics within the last two classes. In fact, in one course, "ethics" was not assigned to a particular class, but only noted as a discussion topic "if time allows." However, we note again that these 20 classes were outliers for including ethics at all [40], and it is commendable that ethics is a part of these technical courses given the time constraints computing departments are under [45]. We believe that including ethics as an explicit part of the course is an important signal to students that ethics is important—and indeed, part of the technical practice of *doing* AI.

5 DISCUSSION

As AI classes become more common in university computing curricula, our analysis provides a window into how ethics might be included. In addition to providing general guidance for instructors about what topics could be taught in these classes, our findings suggest more concrete insights into the current state of AI ethics. However, it is also important to note that, beyond our analysis, AI ethics instruction may be taking place elsewhere; for example, in some CS programs, required general ethics courses also cover AI [11].

While our dataset does not allow for interpretations of pedagogy or what might be the most effective teaching method, we did note some common and notable practices. For example, for standalone AI ethics courses where reading lists were made available, the majority of the courses included news articles as reading assignments. This suggests incorporating current events as a common way to make the consequences of AI personal and real. Many of the articles were critical of Facebook, which is a platform that students might use and therefore understand the direct consequences of the company's sometimes questionable practices. Topics included fake news, privacy concerns about Facebooks "people you may know" feature, and how user news feeds are manipulated and may be shifting our perceptions of the world. Other commonly covered controversies concerned the most consequential real-world applications of AI, such as in policing and health.

Another potential way to make these issues more tangible is to include them as part of the technical practice of building AI. For example, in a subset of technical courses, we observed that learning around technical skills focused on societal considerations or what some instructors called "technology for social good." A potential barrier to including ethical or social implications as part of a technical class is that there is simply too much material to cover; one solution is to embed "technology for social good" into the kinds of assignments or learning objectives that already exist. In this context, students can learn that the technical skills needed to create AI applications could be applied to solve societal problems—while necessarily considering the implications of those solutions as part of the process.

Though we saw a variety of topics in this sample of syllabi, and our dataset cannot capture every topic currently taught in the context of AI ethics, it is worth noting a few omissions or topics that were rarely seen. To consider topics that might be absent here but appear in other parts of AI ethics discourse, we compared our dataset to a recent taxonomy of AI ethics guidelines [17]. A number of aspects of AI ethics appear in these guidelines that are absent or rare in our data. For example, only one course in our dataset mentions accessibility. Given the many barriers and biases that people with disabilities face, and the emergence of AI accessibility in the research literature [23], we would encourage more attention to this topic in the context of AI. Additionally, though gender in the context of bias was a common topic, we did not see diversity in the AI workforce called out specifically in these syllabi; however, we see this as important given the public discourse about the gender imbalance in tech companies and the lack of progress toward fixing this imbalance [53]. Finally, we note that the topic of sustainability is missing from both our dataset and the AI ethics guidelines. The environmental impact of computing is a significant and under-discussed topic within computing in general [30]; considering the computing power necessary for future AI applications, the environmental impacts of AI are likely to be of increasing importance as we move forward.

Finally, our analysis indicates that the most common ethics-related topics for technical courses are privacy and bias (sometimes as technical constructs), but there is no reason that a greater range of topics identified in the standalone classes couldn't be integrated into technical AI classes. As noted by Selbst et al. [41], even researchers and practitioners in the field of fairness in machine learning tend to abstract away the social context in which these systems are deployed, focusing instead on the model, the inputs, and the outputs. However, understanding social context is critical, particularly given the critical contexts in which AI is deployed—e.g., the justice system, military, health, all contexts that appeared frequently in the standalone AI ethics courses we analyzed.

As universities train the next generation of technologists who will build the systems that might cause the controversies of the future, it is critical that we train them consider the potential ramifications. When those building technologies understand that thinking through ethical and social consequences is as an essential part of technical practice, they will inevitably critically consider the technologies they are building. There is already evidence that programmers do feel some responsibility for the code they write [46], and universities can play a part in preparing engineers for the responsibility they will bear when the public holds them accountable. Though ethics is rarely explicitly called out as part of technical AI courses [40], we are seeing movement in the right direction and hope that our research can encourage and guide the evolution of AI education.

6 CONCLUSION

Now is the time to identify interventions to help those building AI systems consider consequences, and we hope that this exploratory work into current trends in AI ethics education will provide encouragement and inspiration towards that goal.

Ideally, future research will align the topics that are being taught in universities with those that are at the forefront of practical applications of AI. Any steps toward implementation will also need to consider organizational or practical barriers to incorporating ethics into AI education. As AI develops, we predict that engineers will be held more accountable for the technology they produce. We hope that this descriptive analysis is the first step in broader research in AI ethics education that leads to more accountability and responsibility in computing. In order for this to happen, the role of ethics in AI education has to go beyond "if time allows."

AKCNOWLEDGEMENTS

This work was partially funded through the Responsible Computer Science Challenge from Mozilla, Omidyar Network, Schmidt Futures, and Craig Newmark Philanthropies. A huge thanks to Jasmine Bilir, Faisal Lalani, and Rami Mouro for their research assistance on this project.

REFERENCES

[1] Safinah Arshad Ali, Blakeley H. Payne, Randi Williams, Hae Won Park, and Cynthia Breazeal. (2019) Constructionism, Ethics, and Creativity: Developing Primary and Middle School Artificial Intelligence Education. Presented in *the International Workshop on Education in Artificiation Intelligence K-12 (EDUAI '19)*. Palo Alto, CA, USA. Proceedings of IJCAI 2019.

[2] Anna Maria Barry-Jester, Ben Casselman, and Dana Goldstein. 2018. Should Prison Sentences Be Based On Crimes That Haven't Been Committed Yet? *FiveThirtyEight*. (2018) https://fivethirtyeight.com/features/prison-reform-risk-assessment/

[3] Brett A. Becker and Thomas Fitzpatrick. 2019. What Do CS1 Syllabi Reveal About Our Expectations of Introductory Programming Students?. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education (SIGCSE '19)*. ACM, New York, NY, USA, 1011-1017.

[4] Ian Bogost. 2017. Elon Musk and Mark Zuckerberg Debate Artificial Intelligence. *The Atlantic*. (2017) https://perma.cc/RK6C-S7Z2.

[5] boyd, danah. 2019. Facing the Great Reckoning Head-On. *Apophenia*. (2019) https://www.zephoria.org/thoughts/archives/2019/09/15/facing-the-great-reckoning-head-on.html.

[6] Burton, Emanuelle, Judy Goldsmith and Nicholas Mattei. 2015. Teaching AI ethics using science fiction. *AAAI Workshop - Technical Report* (2015), 33–37.

[7] Carole Cadwalladr and Emma Graham-Harrison. 2018. Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The Guardian*. (2018) https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election

[8] Felicia Chong. 2016. The Pedagogy of Usability: An Analysis of Technical Communication Textbooks, Anthologies, and Course Syllabi and Descriptions. Technical Communication Quarterly. 25, 1 (Jan. 2016), 12–28.

[9] Kate Conger, Richard Fausset and Serge F. Kovaleski. 2019. San Francisco Bans Facial Recognition Technology. *The New York Times*. (2019) https://www.nytimes.com/2019/05/14/us/facial-recognition-ban-san-francisco.html

[10] Eric Eaton, Sven Koenig, Claudia Schulz, Francesco Maurelli, John Lee, Joshua Eckroth, Mark Crowley, Richard G. Freedman, Rogelio E. Cardona-Rivera, Tiago Machado, and Tom Williams. 2018. Blue sky ideas in artificial intelligence education from the EAAI 2017 new and future AI educator program. *AI Matters*. 3, 4 (Feb. 2018), 23–31.

[11] Casey Fiesler, Natalie Garrett, and Nathan Beard. 2020. What Do We Teach When We Teach Tech Ethics?: A Syllabi Analysis. SIGCSE *Conference on Computer Science Education*. (2020).

[12] Casey Fiesler, and Blake Hallinan. 2018. "We Are the Product": Public Reactions to Online Data Sharing and Privacy Controversies in the Media. *CHI '18 Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (2018), 1–13.

[13] Daniel Greene, Anna Lauren Hoffmann, and Luke Stark. 2019. Better, Nicer, Clearer, Fairer: A Critical Assessment of the Movement for Ethical Artificial Intelligence and Machine Learning. *Proceedings of the 52nd Hawaii International Conference on System Sciences* (2019), 10.

[14] Larry Greenemeier. 2018. Can AI Really Solve Facebook's Problems? *Scientific American*. (2018) https://www.scientificamerican.com/article/can-ai-really-solve-facebooks-problems1/

[15] Barbara J. Grosz, David Gray Grant, Kate Vredenburgh, Jeff Behrends, Lily Hu, Alison Simmons, and Jim Waldo. 2019. Embedded EthiCS: Integrating Ethics Broadly Across Computer Science Education. *Commun. ACM* 62, 8 (2019), 54–61.

[16] Loren Grushe. 2015. Google engineer apologizes after Photos app tags two black people as gorillas. *The Verge*. (2015) https://www.theverge.com/2015/7/1/8880363/google-apologizes-photos-app-tags-two-black-people-gorillas..

[17] Thilo Hagendorff. 2019. The Ethics of AI Ethics -- An Evaluation of Guidelines. (Feb. 2019). https://arxiv.org/abs/1903.03425

[18] Yuval Noah Harari. 2017. The Meaning of Life in a World Without Work. *The Guardian*. (2017) https://www.theguardian.com/technology/2017/may/08/virtual-reality-religion-robots-sapiens-book

[19] Ali Haskins Lisle, Coleman Merenda, and Joseph Gabbard. 2019. Using affinity diagramming to generate a codebook: a case study on young military veterans and community reintegration. Qualitative Research (2019), 1–18.

[20] Nathan Heller. 2016. If animals have rights, should robots? *The New Yorker*. (2016), 1–14. https://www.newyorker.com/magazine/2016/11/28/if-animals-have-rights-should-robots.

[21] Kashmir Hill. 2017. Facebook Figured Out My Family Secrets, And It Won't Tell Me How. *Gizmodo*. (2017). https://perma.cc/NKG3-L8QL.

[22] Dan Hurley. 2018. Can an algorithm tell when kids are in danger? *The New York Times Magazine*. (2018) https://www.nytimes.com/2018/01/02/magazine/can-an-algorithm-tell-when-kids-are-in-danger.html

[23] Ben Hutchinson, Vinodkumar Prabhakaran, Emily Denton, Kellie Webster, Yu Zhong, and Stephen Denuyl. 2019. Unintended Machine Learning Biases as Social Barriers for Persons with Disabilities. *ACCESS*. Issue 125, (2019).

[24] Heather Knight. 2014. How Humans Respond to Robots : Building Public Policy through Good Design. https://www.brookings.edu/research/how-humans-respond-to-robots-building-public-policy-through-good-design/

[25] Jeff Larson, Surya Mattu, Lauren Kirchner and Julia Angwin. 2016. How We Analyzed the COMPAS Recidivism Algorithm. *ProPublica*. (2016) https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm

[26] Kristie J Loescher, Regina W Hughes, Frank Cavico, Jim Mirabella, and Pedro F Pellet. 2005. The Impact of an "Ethics Across the Curriculum" Initiative on the Cognitive Moral Development of Business School Undergraduates. *Teaching Ethics* (2005), 31–72.

[27] Kathleen M. MacQueen, Eleanor McLellan, Kelly Kay, and Bobby Milstein. 1998. Codebook Development for Team-Based Qualitative Analysis. Field Methods 10, 2 (1998), 31–36.

[28] C. Dianne Martin, Chuck Huff, Donald Gotterbarn, and Keith Miller. 1996. Implementing a Tenth Strand in the CS Curriculum. Commun. ACM 39, 12 (1996), 75–84.

[29] Alice Marwick. 2014. How your data are being deeply mined. *New York Review of Books*. 61, 1 (2014).

[30] Nick Merrill.. 2017. Better not to know? The SHA1 collision & the limits of polemic computation. *LIMITS 2017 - Proceedings of the 2017 Workshop on Computing Within Limits* (Jun. 2017), 37–42.

[31] Brent Mittelstadt. 2019. Principles Alone Cannot Guarantee Ethical AI. *Nature Machine Intelligence*. 1, (2019), 501–507.

[32] Moral Machine: 2017. *http://moralmachine.mit.edu/*. Accessed: 2019-11-04.

[33] Evgeny Morozov. 2013. *To Save Everything, Click Here: The Folly of Technological Solutionism*. Allen Lane.

[34] Heather Murphy. 2017. Why Stanford Researchers Tried to Create a 'Gaydar' Machine. *The New York Times*. (2017) https://www.nytimes.com/2017/10/09/science/stanford-sexual-orientation-study.html

[35] Dom Nicastro. 2019. How Are Universities Responding to the Tech Skills Gap? *CMS Wire2*. (2019). https://www.cmswire.com/digital-workplace/how-are-universities-responding-to-the-tech-skills-gap/

[36] Safiya Noble. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press.

[37] Don Peck. 2013. They're Watching You at Work. *The Atlantic*. (2013), 72–84. https://www.theatlantic.com/magazine/archive/2013/12/theyre-watching-you-at-work/354681/

[38] David Rotman. 2017. "The Relentless Pace of Automation". *MIT Technology Review*. (2017) https://www.technologyreview.com/s/603465/the-relentless-pace-of-automation/.

[39] Dominic Rushe. 2017. End of the road: will automation put an end to the American trucker? *The Guardian*. (2017) https://www.theguardian.com/technology/2017/oct/10/american-trucker-automation-jobs.

[40] Jeffrey Saltz, Michael Skirpan, Casey Fiesler, Micha Gorelick, Tom Yeh, Robert Heckman, Neil Dewar, and Nathan Beard. 2019. Integrating Ethics within Machine-learning Courses. *ACM Transactions on Computing Education* 19, 4 (2019), 1–26..

[41] Andrew D Selbst, danah boyd, Sorelle A Friedler, Suresh Venkatasubramanian, and Janet Vertesi. 2019. Fairness and abstraction in sociotechnical systems. *Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency* (2019), 59–68.

[42] Aarti Shahani. 2019. Congress' Power Struggle With Big Tech Will Be On Display At Hearings . *NPR*. (2019) https://www.npr.org/2019/07/16/742109310/congress-power-struggle-

with-big-tech-will-be-on-display-at-hearings

[43] Katie Shilton. 2018. Engaging Values Despite Neutrality: Challenges and Approaches to Values Reflection during the Design of Internet Infrastructure. Science, Technology, & Human Values. 43, 2 (2018), 247–269.

[44] Natasha Singer, N. 2018. Tech's ethical "dark side": Harvard, Stanford, and others want to address it. *The New York Times*. (2018) https://www.nytimes.com/2018/02/12/business/computer-science-ethics-courses.htm.

[45] Carol Spradling, Leen-Kiat Soh, and Charles Ansorge. 2008. Ethics Training and Decision-making: Do Computer Science Programs Need Help? *Proceedings of the 39th SIGCSE Technical Symposium on Computer Science Education* (New York, NY, USA, 2008), 153–157.

[46] Stack Overflow Developer Survey 2018: 2018. *https://insights.stackoverflow.com/survey/2018#ethics*. Accessed: 2019-11-04.

[47] Bernd Carsten Stahl, Job Timmermans, and Brent Daniel Mittelstadt. 2016. The Ethics of Computing: A Survey of the Computing-Oriented Literature. *ACM Comput. Surv.* 48, 4 (Feb. 2016), 55:1--55:38.

[48] Luke Stark and Anna Lauren Hoffmann. 2019. Data Is the New What? Popular Metaphors & Professional Ethics in Emerging Data Culture. *Journal of Cultural Analytics* 127, 127 (2019), 1–22.

[49] Zeynep Tufekci. 2014. Facebook and Engineering the Public. *The Message*. (2014) https://medium.com/message/engineering-the-public-289c91390225

[50] Manas Tungare, Xiaoyan Yu, William Cameron, Guofang Teng, A P Manuel, Lillian Cassel,Weiguo Fan, and Edward A Fox. 2007. Towards a Syllabus Repository for Computer Science Courses Categories and Subject Descriptors. In *Proceedings of the ACM SIGCSE Technical Symposium on Computer Science Education*. Covington, KY, 55–59.

[51] Siva Vaidhyanathan. 2019. Mark Zuckerberg Needs to Shut Up. *Wired*. (2019) https://www.wired.com/story/mark-zuckerberg-needs-to-shut-up/

[52] Ben Wagner. 2019. Ethics as an Escape from Regulation. In *Being Profiled*. E. Bayamlioğlu et al., eds. 84–89.

[53] Alison Wynn. 2019. Why Tech's Approach to Fixing Its Gender Inequality Isn't Working. *Harvard Business Review*. (2019) https://hbr.org/2019/10/why-techs-approach-to-fixing-its-gender-inequality-isnt-working

[54] The Case against Killer Robots: 2012. *Human Rights Watch*. https://www.hrw.org/report/2012/11/19/losing-humanity/case-against-killer-robots.