Kommentar

Thilo Hagendorff

Ethics of Machine Learning. A Critical Appraisal of the State of the Art

Abstract

When assessing AI ethics, one can oscillate between two momentums, between ethics as genuine effort to put principles into practice and ethics as mere marketing strategy, as "toothless" discourse. Between the two poles, a gradual transition exists. This transition occurs, among other reasons, since AI ethics research is conducted by public as well as private institutions. Especially with regard to the latter, AI ethics can be repurposed for marketing strategies that aim at signaling some kind of pseudo trustworthiness of AI products to the public. These two momentums are in permanent conflict with each other. The paper describes this tension and evaluates both sides of the conflict.

Bei der Beurteilung von KI-Ethik kann zwischen zwei Momenten oszilliert werden, nämlich zwischen Ethik als echtem Bemühen, Prinzipien in die Praxis umzusetzen, und Ethik als bloßer Marketingstrategie, als "zahnlosem" Diskurs. Zwischen diesen beiden Polen gibt es einen fließenden Übergang. Dieser Übergang findet unter anderem deshalb statt, weil die Forschung zur KI-Ethik sowohl von öffentlichen als auch privaten Institutionen betrieben wird. Vor allem in Bezug auf letztere kann KI-Ethik für Marketingstrategien umfunktioniert werden, die darauf abzielen, der Öffentlichkeit eine Art Pseudo-Vertrauenswürdigkeit von KI-Produkten zu signalisieren. Diese beiden Momente stehen in einem permanenten Konflikt zueinander. Der Beitrag beschreibt diese Spannung und bewertet beide Seiten des Konflikts.

Introduction

Not many other technology trends have caused such a widespread discourse on their ethical implications like artificial intelligence. Artificial intelligence is named among other large-scale, high-risk technologies like nuclear energy or genetic engineering.[1] Many of those risk discourses are far-fetched and have their origins in science fiction narratives that are appealing to mass media but have nothing to do

1 See Ulrich Beck: *World at Risk*, Cambridge 2008. Nick Bostrom: “Existential Risks. Analyzing Human Extinction Scenarios and Related Hazards,” *Journal of Evolution and Technology* 9/1

with actual technical possibilities and realistic trajectories of future technological improvements. Besides exaggerated proclamations on the future of artificial intelligence – which, by the way, historically repeat themselves during the hype phases of the technology[2] –, one can identify various down-to-earth, but nevertheless significant ethical issues and problems that are associated with machine learning applications. I do not want to use the term 'artificial intelligence' from now on, because what is actually meant when using the term are methods of machine learning, comprising statistical techniques like regression methods, Bayes algorithms, support vector machines, decision trees, as well as artificial neural networks of all kinds. All those methods have their limits,[3] but are nevertheless quite powerful and can have far-reaching ethical implications when implemented in certain societal contexts, especially in the case of algorithmic decision making in high stakes areas like the police, healthcare, legal or education system. To cope with those challenges, a somewhat virulent discourse on machine learning ethics either aims at actually throwing a spanner in the works and stopping dangerous developments and, in addition, at promoting value alignment and beneficial applications, or it merely serves the purpose of giving compliance signals to a worried public in order to stifle critique. When assessing machine learning ethics, one can always oscillate between those two momentums, between ethics as genuine effort to put principles into practice[4] and ethics as marketing strategy[5]. In the following, those two momentums shall be described and assessed in more detail.

Putting principles into practice

The modern machine learning ethics discourse, a subfield in the wider strand of technology ethics, started off several years ago with the composition of specialized codes of ethics. Those ethics guidelines found their inspiration in existing deliberations on robot, data, and algorithm ethics. Eventually, machine learning ethics stand in

---

(2009), pp. 1–37. Nick Bostrom: *Superintelligence. Paths, Dangers, Strategies*, Oxford: Oxford UP 2014. Steve Omohundro: "Autonomous technology and the greater human good," *Journal of Experimental & Theoretical Artificial Intelligence* 26/3 (2014), pp. 303–315.

2 See Mikel Olazaran: "A Sociological Study of the Official History of the Perceptrons Controversy," *Social Studies of Science* 26/3 (1996), pp. 611–659.

3 See Thilo Hagendorff and Katharina Wezel: "15 challenges for AI: or what AI (currently) can't do," *AI & SOCIETY – Journal of Knowledge, Culture and Communication* 35 (2020), pp. 355–365.

4 See Jessica Morley, Luciano Floridi, Libby Kinsey and Anat Elhalal: "From What to How. An Overview of AI Ethics Tools, Methods and Research to Translate Principles into Practices," *arXiv* (2019), pp. 1–21, https://arxiv.org/pdf/1905.06876.pdf (checked 5/1/2021).

5 See Rodrigo Ochigame: "The Invention of 'Ethical AI.' How Big Tech Manipulates Academia to Avoid Regulation," *The Intercept* (2019), https://theintercept.com/2019/12/20/mit-ethical-ai-artificial-intelligence/ (checked 1/7/2020).

many regards in a tradition of evolving buzzwords – ranging from knowledge discovery in databases, data mining, big data, data science to predictive analysis – which build the center and the identity of academic and public discourses. With the groundwork that has been done in those fields of technology ethics, machine learning ethics could draw on well-considered ethical principles while adding only a few specific ones like explainability, accountability, or debiasing tenets to adapt to new technological capabilities that are specific to artificial neural nets and other machine learning techniques. Hence, the state of the art comprises reflections on how classical ethical principles can be implemented in decision routines of autonomous machines[6] over meta-studies about machine learning ethics[7] or the empirical analysis on how trolley problems are solved[8] to reflections on specific problems[9] and comprehensive ethics guidelines[10]. Nearly one hundred of those guidelines are currently available, stemming from government agencies, research institutes, or companies.

While in many regards, stark differences between the guidelines exist, there are common denominators, too. When comparing the guidelines with each other, one can see that at least six principles occur in nearly all of them. Those principles comprise the demand for privacy protection, fairness, accountability, transparency, safety, and the promotion of the common good via machine learning applications. There

---

6 See Michael Anderson and Susan Leigh Anderson: "Towards Ensuring Ethical Behavior from Autonomous Systems. A Case-Supported Principle-Based Paradigm," *Artificial Intelligence and Ethics: Papers from the 2015 AAAI Workshop* (2015), pp. 1–10. Amitai Etzioni and Oren Etzioni: "Incorporating Ethics into Artificial Intelligence," *Journal of Ethics* 21/4 (2017), pp. 403–418. Han Yu, Zhiqi Shen, Chunyan Miao et al.: "Building Ethics into Artificial Intelligence" *arXiv* (2018), pp. 1–8, https://arxiv.org/abs/1812.02953 (checked 05/01/2021).

7 See Ville Vakkuri and Pekka Abrahamsson: "The Key Concepts of Ethics of Artificial Intelligence," *Proceedings of the 2018 IEEE International Conference on Engineering, Technology and Innovation* (2018), pp. 1–6. Marcelo Prates, Pedro Avelar and Luis C. Lamb: "On Quantifying and Understanding the Role of Ethics in AI Research. A Historical Account of Flagship Conferences and Journals," *arXiv* (2018), pp. 1–13, https://arxiv.org/abs/1809.08328 (checked 05/01/2021). Paula Boddington: *Towards a Code of Ethics for Artificial Intelligence*, Cham: Springer 2017. Daniel Greene, Anna Lauren Hoffman, Luke Stark: "Better, Nicer, Clearer, Fairer: A Critical Assessment of the Movement for Ethical Artificial Intelligence and Machine Learning." *Proceedings of the Hawaii International Conference on System Sciences* (2019), pp. 2122–2131, https://hdl.handle.net/10125/59651 (checked 05/01/2021). Judy Goldsmith and Emanuelle Burton: "Why Teaching Ethics to AI Practitioners Is Important," *The AAAI-17 Workshop on AI, Ethics, and Society* (2017), pp. 110–114.

8 See Edmond Awad, Sohan Dsouza, Richard Kim et al.: "The Moral Machine experiment," *Nature* 563/7729 (2018), pp. 59–64. DOI: 10.1038/s41586–018–0637–6.

9 See Peter Eckersley: "Impossibility and Uncertainty Theorems in AI Value Alignment (or why your AGI should not have a utility function)," *arXiv* (2018), pp. 1–13, https://arxiv.org/abs/1901.00064 (checked 05/01/2021).

10 See Anna Jobin, Marcello Ienca and Effy Vayena: "The global landscape of AI ethics guidelines," *Nature Machine Intelligence* 1/9 (2019), pp. 389–399. Jessica Fjeld, Nele Achten, Hannah Hilligoss, et al.: "Principled Artificial Intelligence. Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI," *Berkman Klein Center Research Publication* 1 (2020), pp. 1–39, http://dx.doi.org/10.2139/ssrn.3518482 (checked 05/01/2021). Thilo Hagendorff: "The Ethics of AI Ethics. An Evaluation of Guidelines," *Minds and Mac*hines 30 (2020), pp. 99–120.

seems to be a tacit consensus that the ethically sound usage of machine learning has to fulfill those six principles. However, one can question whether those tenets are the very relevant. When considering issues like dual-use problems, future of employment, the industry's diversity crisis, the decline of social cohesion through machine learning based information filters, and the like, it gets really hard to gauge which ones are more or less important. While the latter issues are mentioned only very seldomly in the guidelines, the former six principles are an integral part of the codes of ethics. A potential explanation for this lies in the observation that issues that can be operationalized mathematically, thus issues for which technical fixes can or have already been developed are mentioned more frequently than issues that require genuinely social solutions. Enormous technical efforts are undertaken to meet ethical targets in the fields of accountability and explainable AI,[11] fairness and discrimination aware machine learning[12] as well as privacy[13]. Many of those endeavors are unified under the FAT ML or XAI community.[14] Solutions for privacy, accountability, fairness, explainability, safety, or robustness issues can be implemented in terms of technical measures, in opposition to solutions for issues like the diversity crisis,[15] the use of autonomous weapon systems,[16] labor displacement,[17] missing legal norms,[18] and the like, which require political efforts and societal change. However, albeit the existence of various technical problem-solving approaches, the existing ethics guidelines almost never contain any technical instructions whatsoever.

Another striking aspect is the fact that the current machine learning ethics discourse has, without explicitly reflecting on it, a sole focus on one particular ethical approach, namely deontology. Deontology is, besides utilitarian approaches, indeed the most prominent ethical framework. Nevertheless, ethics obviously has a broa-

---

11 See Brent Mittelstadt, Chris Russell and Sandra Wachter: "Explaining Explanations in AI," *Proceedings of the Conference on Fairness, Accountability, and Transparency* (2019), pp. 279–288.

12 See Joy Buolamwini and Timnit Gebru: "Gender Shades. Intersectional Accuracy Disparities in Commercial Gender Classification," *Proceedings of Machine Learning Research* 81 (2018), pp. 1–15.

13 See Benjamin Baron and Mirco Musolesi: "Interpretable Machine Learning for Privacy-Preserving Pervasive Systems," *arXiv* (2017), pp. 1–10, https://arxiv.org/abs/1710.08464 (checked 05/01/2021).

14 See Michael Veale and Reuben Binns: "Fairer machine learning in the real world. Mitigating discrimination without collecting sensitive data," *Big Data & Society* 4/2 (2017), pp. 1–17. Andrew D. Selbst and Solon Barocas: "The Intuitive Appeal of Explainable Machines." *Fordham Law Review* 87 (2018), pp. 1085–1139.

15 See Tom Simonite: "AI is the Future – But where are the Women?," *Wired* (2018), https://www.wired.com/story/artificial-intelligence-researchers-gender-imbalance/ (checked 05/01/2021).

16 See Amanda Sharkey: "Autonomous weapons systems, killer robots and human dignity," *Ethics and Information Technology* 21/2 (2019), pp. 75–87.

17 See Erik Brynjolfsson and Tom Mitchell: "What can machine learning do? Workforce implications," *Science* 358/6370 (2017), pp. 1530–1534.

18 See Ryan Calo: "Artificial Intelligence Policy. A Primer and Roadmap," *SSRN Journal* (2017), pp. 1–28, http://dx.doi.org/10.2139/ssrn.3015350 (checked 05/01/2021).

der 'toolbox' than only the two aforementioned theories. However, this circumstance is completely ignored in machine learning ethics. Virtue ethical approaches – to name just one example of an ethical theory, nearly nobody takes notice of – does not focus on situation-independent, universal principles, rules, or imperatives but on character dispositions, personality traits, moral intuitions, or 'technomoral virtues' such as honesty, justice, courage, empathy, civility, and the like.[19] Moreover, virtue ethics can bridge over to considerations in the field of organization ethics, where individual character traits like cognitive moral development, idealism, or job satisfaction are equally taken into account alongside organizational environment characteristics like an egoistic or altruistic work climate, decent error cultures, organizational visions, and long-term targets, power hierarchies, employee participation, mechanisms for code enforcement, and many more.[20] Changing organizational structures in a way that the likelihood for the occurrence of unethical intentions and behavior is minimized is the ideal way for ensuring socially acceptable machine learning applications. In order to truly implement ethical decision making into organizations that are researching, developing, and distributing machine learning solutions, the existence of codes of ethics is only a minor factor that determines whether the likelihood of ethical behavior is strengthened or weakened.

Investigating the transition from abstract ethical principles to practice, the transition "from what to how" as "the second phase of AI ethics"[21] is the objective of the latest papers in machine learning ethics.[22] Researchers stress that principles alone cannot guarantee desired outcomes. A new focus must be put on individual and organizational practices. For that purpose, the development of machine learning applications is divided into different stages where at each step specific requirements of particular ethical frameworks have to be accomplished. For that purpose, more technologically specific guidance and explanations on how to fulfill abstract principles are gathered. Morley et al. distinguish between the business and use-case development, the design phase, the training, and test data procurement, the building of an artificial intelligent application, the testing and deployment of it, as well as the monitoring of its performance.[23] During each of this phases, tools and methods are to be identified that ensure high-level ethical values like beneficence, non-maleficence,

19 See Shannon Vallor: *Technology and the Virtues. A Philosophical Guide to a Future Worth Wanting,* New York: Oxford UP 2016.
20 See Jennifer J. Kish-Gephart, David A. Harrison and Linda Klebe Treviño: "Bad apples, bad cases, and bad barrels. Meta-analytic evidence about sources of unethical decisions at work," *The Journal of Applied Psychology* 95/1 (2010), pp. 1–31.
21 See Morley et al.: "From What to How," *arXiv* (2019), pp. 1–21.
22 See Brent Mittelstadt: "Principles alone cannot guarantee ethical AI," *Nature Machine Intelligence* 1/11 (2019), pp. 501–507.
23 See Morley et al.: "From What to How," *arXiv* (2019), pp. 1–21.

autonomy, justice, or explainability.[24] Although the push for a transition from principles to practices has to be appreciated, one can criticize that this emphasized step to applied machine learning ethics does not really change the established ethical approach besides the fact that more fine-grained principles are introduced. For instance, instead of simply stressing the importance of privacy, principles are introduced on how to fulfill requirements of the privacy-by-design,[25] privacy audit[26] or privacy impact assessment[27] framework. But basically, one does not move away from deontology or a principled approach. The principles are just refined in the sense of a conversion from high-level to micro ethics, from ethics to technology ethics, to machine ethics, to computer ethics, to information ethics, to data ethics etc.

Shortcomings of machine learning ethics

In a recent article entitled "The Invention of 'Ethical AI'", the author argues that tech companies manipulate academia by funding countless initiatives on AI ethics and other soft-law governance measures in order to avoid truly binding legal regulations that could hurt the pursuit of monetary profits.[28] Nowadays, the discourse on 'ethical AI', 'value alignment', 'beneficial AI', and the like is not just successfully established in academia, but also in the mainstream press. The prevalence of those terminologies may suggest to legislators that legally enforceable restrictions are not necessary, due to the self-governance of the industry. Despite this critique, it is striking how many codes of machine learning ethics have been published by IT companies, from Google to OpenAI, Microsoft, DeepMind, IBM etc. In comparison to ethics guidelines from scientific or governance institutions, industry guidelines comprise on average considerably less principles.[29] Also, by simply counting the number of words, one can see that industry guidelines are shorter than guidelines

24 Luciano Floridi, Josh Cowls, Monica Beltrametti et al.: "AI4People – An Ethical Framework for a Good AI Society. Opportunities, Risks, Principles, and Recommendations," *Minds and Machines* 28/4 (2018), pp. 689–707.

25 See Ann Cavoukian, Scott Taylor and Martin E. Abrams: "Privacy by Design: essential for organizational accountability and strong business practices." *Identity in the Information Society* 3/2 (2010), pp. 405–413.

26 See Eleni-Laskarina Makri and Costas Lambrinoudakis: "Privacy Principles: Towards a Common Privacy Audit Methodology," *Trust, Privacy and Security in Digital Business*, eds. Simone Fischer-Hübner, Costas Lambrinoudakis and Javier López, Cham: Springer 2015, pp. 219–234.

27 See Marie Caroline Oetzel and Sarah Spiekermann: "A systematic methodology for privacy impact assessments: a design science approach," *European Journal of Information Systems* 23/2 (2014), pp. 126–150.

28 Rodrigo Ochigame: "The Invention of 'Ethical AI.' How Big Tech Manipulates Academia to Avoid Regulation," *The Intercept* (2019), https://theintercept.com/2019/12/20/mit-ethical-ai-artificial-intelligence/ (checked 1/7/2020).

29 See Hagendorff: "The Ethics of AI Ethics," pp. 99–120.

from non-economic institutions. Moreover, a close link between business and science is revealed not just by the fact that all major machine learning conferences are sponsored by industry partners but also by the raising number of corporate-affiliated machine learning papers.[30]

Another conflict for ethical machine learning arises in view of frequent statements on "world leadership in AI"[31] where the USA, China, and Europe compete for the lead in research on and application of increasingly capable machine learning systems. This "AI race,"[32] with its side effects of reckless competitive thinking, poses a major threat to ethical considerations that are always under suspicion to hamper the needed technological progress for 'AI superiority'. As long as machine learning research, and development is not seen as a cooperative effort but as a fierce competition, ethics will be a tough act to follow. The race for the 'best' AI reduces the likelihood that technical precaution measures will be entrenched, it reduces the likelihood that truly benevolent machine learning systems are developed, it reduces dialogue between research groups, it intensifies in- and out-group-thinking, and the like. In short, the 'AI race,' regardless whether it is a mere narrative or harsh reality, hampers efforts to create ethically sound machine learning applications.

Apart from business imperatives and a fierce 'AI race' rhetoric, principled machine learning ethics can be called into question by scrutinizing its tangible ramifications. Here, the decisive criterion is whether ethical guidelines bring about change in individual decision-making. In this regard, a pretty sobering insight came from a recent study by McNamara et al.[33] The researchers critically reviewed the idea that ethical guidelines serve as a basis for ethical decision-making for software-engineers. Their main finding was that the effectiveness of ethical codes is almost zero and that they do not change the behavior of tech professionals. In the survey, 63 software engineering students and 105 professional software developers were scrutinized. They were presented with eleven software-related ethical decision scenarios, testing whether the influence of the ethics guideline of the Association for Compu-

---

30 See Yoav Shoham, Raymond Perrault, Eric Brynjolfsson et al.: *The AI Index 2018 Annual Report*, Stanford University 2018, pp. 1–94.

31 abacusnews.com: "China Internet Report 2018," *South China Morning Post* (2018) https://www.abacusnews.com/china-internet-report/china-internet-2018.pdf (checked on 7/13/2018).

32 Steven Cave and Seán S. ÓhÉigeartaigh: "An AI Race for Strategic Advantage: Rhetoric and Risks," *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society* (2018), pp. 36–40.

33 See Andrew McNamara, Justin Smith and Emerson Murphy-Hill: "Does ACM's code of ethics change ethical decision making in software development?," *Proceedings of the 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering,* eds. Gary T. Leavens, Alessandro Garcia, Corina S. Păsăreanu, New York: ACM Press 2018, pp. 1–7.

ting Machinery (ACM)[34] in fact influences ethical decision-making in six vignettes, ranging from responsibility to report, user data collection, intellectual property, code quality, honesty to customer to time and personnel management. The results are disillusioning. Across individuals who did and did not see the ACM ethics guidelines, no statistically significant difference in the responses for any vignette were found.[35] In view of that, it becomes clear that machine learning ethics lacks enforcement mechanisms that reach beyond a voluntary and non-binding self-commitment of organizations and research communities.

Conclusion

Machine learning ethics faces many challenges. Ethics is misused for marketing or whitewashing purposes, it is instrumentalized in order to prevent the legislation of binding norms, it stifles critical public discourse, it is hard to measure its concrete ramifications, practitioners see it as a mere surplus or 'add-on' to technical concerns. Notwithstanding these challenges, machine learning ethics must not be discounted. In various areas of research and development on machine learning technologies, ethically motivated efforts are undertaken to meet the goals that are outlined in many of the important AI codes.[36] This holds especially true with regard to fields where technical fixes exist or can be found. Those fields comprise efforts to solve accountability as well as explainability problems,[37] to protect informational privacy,[38] to avoid algorithmic discrimination,[39] or to ensure machine learning safety[40]. However, ethical issues that cannot be solved via technical fixes are put in second place. Among those issues are, for instance, the diversity crisis in the machine learning

34 See Don Gotterbarn, Bo Brinkman, Catherina Flick, et al.: "ACM Code of Ethics and Professional Conduct. Affirming our obligation to use our skills to benefit society," https://www.acm.org/binaries/content/assets/about/acm-code-of-ethics-booklet.pdf (checked on 2/1/2019).

35 See McNamara et al.: "Does ACM's code of ethics change ethical decision making?", p. 4.

36 See Ala-Pietlä Pekka, Wilhelm Bauer, Urs Bergmann et al.: "The European Commission's High-Level Expert Group on Artificial Intelligence. Ethics Guidelines for Trustworthy AI. Working Document for stakeholders' consultation." European Commission: Brussels 2018, pp. 1–37. Floridi et al.: "AI4People – An Ethical Framework for a Good AI Society," pp. 689–707. Partnership on AI: "About Us," 2018, https://www.partnershiponai.org/about/ (checked on 1/25/2019).

37 See Anton Vedder and Laurens Naudts: "Accountability for the use of algorithms in a big data environment," *International Review of Law, Computers & Technology* 31/2 (2017), pp. 206–224. Mittelstadt et al.: "Explaining Explanations in AI," pp. 279–288.

38 See Baron and Musolesi: "Interpretable Machine Learning," pp. 1–10.

39 Thilo Hagendorff: „Maschinelles Lernen und Diskriminierung. Probleme und Lösungsansätze," *Österreichische Zeitschrift für Soziologie* 44/1 (2019), pp. 53–66.

40 See Dario Amodei, Chris Olah, Jacob Steinhardt et al.: "Concrete Problems in AI Safety," *arXiv* (2019), pp. 1–29, https://arxiv.org/pdf/1606.06565.pdf (checked on 05/01/2021).

field,[41] the difficult question about the future of work under the terms of machine learning driven technologies of automation,[42] the problem of diminishing social cohesion and solidarity due to machine learning based information filters on social media platforms,[43] hidden ecological as well as economic costs of machine learning system,[44] and many more. When speaking about putting principles to practice, the mentioned fields are but a few of many where improvements are urgent. In this context, to make progress one should not rely solely on 'checkbox' guidelines for machine learning ethics. Rather, a transition is required from a purely deontologically oriented ethical approach to a broader approach that is based on virtues and personality dispositions, knowledge expansions, lifelong education, responsible autonomy, freedom of action, etc. Hence, it is not just necessary to close the gap between ethics and technical discourses, to build bridges between abstract values and technical implementations, but also to recognize the importance of 'soft' skills, technomoral virtues, empathy, a change of organizational structures, and the courage to genuinely act in accordance to moral values in a world that is driven by economic and race imperatives. Only then can machine learning ethics be truly and sustainably put into practice.

Acknowledgements

This research was supported by the Cluster of Excellence "Machine Learning – New Perspectives for Science" funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – Reference Number EXC 2064/1 – Project ID 390727645.

41 See Simonite: "AI is the Future – But where are the Women?," *Wired* (2018), https://www.wired.com/story/artificial-intelligence-researchers-gender-imbalance/ (checked 05/01/2021).
42 See Brynjolfsson and Mitchell: "What can machine learning do?," pp. 1530–1534.
43 See Matevž Kunaver and Tomaž Požrl: "Diversity in recommender systems – A survey," *Knowledge-Based Systems* 123 (2017), pp. 154–162.
44 See Emma Strubell, Ananya Ganesh and Andrew McCallum: "Energy and Policy Considerations for Deep Learning in NLP," *arXiv* (2019), pp. 1–6, https://arxiv.org/abs/1906.02243 (checked 05/01/2021). Lilly Irani: "The Hidden Faces of Automation," *XRDS* 23/2 (2016), pp. 34–37.