Journal of Behavioral and Applied Management
2024, Vol. 24, No. 2, 131–134
https://jbam.scholasticahq.com/

The Ethical Considerations of Artificial Intelligence (AI)

Scott A. Jeffrey
Monmouth University

This paper examines the ethical implications of AI in 5 specific areas: Education, idea ownership, deep fakes, bias, and autonomous vehicles. The author examines these areas using the ethical traditions of deontology, consequentialism, and care ethics.

Artificial Intelligence (AI) has grown in application and importance recently. Some may immediately think about the movie "Terminator" in which Skynet AI became self-aware in 1997, beginning the war between computers and humans. While most do not believe that the rise of AI should cause that level of concern, the ethical implications of AI have not been thoroughly explored. This article attempts to address that gap and also direct the reader to other work in this area.

This article will focus on ethical considerations of AI in five areas. First, the manuscript will address academic dishonesty concerns that arise from generative AI; with a focus on ChatGPT and its use in post-secondary academics. Closely related to this is the area of creativity and ownership; for example ownership of ideas which are generated with the help of AI. The manuscriptI will then explore the concept of deep fakes, and how AI can help someone create a misleading video of another person.

The next topic will be bias and discrimination in AI. It is reasonably well known that facial recognition software does a very poor job of identifying non-white individuals (Marks, 2021). Beyond this problem, the use of AI in hiring can also reinforce biases in hiring engaged in by a firm. This problem tends to be exacerbated based on how the technology is "trained" and whose preferences are programmed into the software. This issue carries over into the final area that will be discussed, that of self-driving cars. Fully autonomous cars need to make decisions regarding what to do and the programmer will encroach in that decision making.

This manuscript will look at the ethical implications of AI through three "lenses". The first is Deontology, ethics based on duty and rules (Louden, 1996; Tarsney, 2018). The second will be consequentialism, looking at the net benefit to all stakeholders of decisions (Bourcier, 2020). Finally, we will look at Care Ethics. Also called feminist ethics, it is where a decision maker must demonstrate the most care for stakeholders (Edwards, 2009; Noddings, 2012). It is based loosely on virtue ethics, where caring would represent the idea of Aristotle's golden mean between caring too much and not caring all (Edwards, 2009; Golden Mean | Definition, Aristotle, Maimonides, Buddhism, Confucianism, & Facts | Britannica, n.d.).

Generative AI in Education

Many students are using applications such as ChatGPT to assist in writing college assignments (Li et al., n.d.). Many professors consider this to be a case of academic dishonesty as professors tend to be looking for original work generated by the student themself, usually called authentic work. The natural question that arises is whether using AI in this way is ethical. If you ask ten people, you will get ten answers to this question, and hopefully most will say "it depends".

As discussed above, most professors desire authentic work from their students, so if students are using generative AI to write their entire paper, most faculty would consider that unethical. If students were doing research to better understand the topic, most would not have any ethical problems with that use. To allow for this type of use, even the APA and MLA style guides provide a way to cite chatGPT ("How Do I Cite Generative AI in MLA Style?," 2023; How to Cite ChatGPT, n.d.) .

In addition to written assignments, generative AI can also help with quantitative assignments such as those found in accounting, finance, or economics. The ethics associated with this are harder to articulate, particularly if AI is used for homework assignments. The authenticity criteria could also work here. If generative AI such as chatGPT were used to complete an assignment, it would not be authentic work created by the student and by the definition above it would be unethical.

One silver lining is the issue of so-called hallucinations where AI may incorrectly generate an answer to a question. Since large language models (LLM) are trained on the internet, they may come across factually incorrect information which would then be delivered to the user. If the user is not careful, incorrect information may end up in the final product, tipping the recipient off that it was not authentic work.

Looking at this through the lens of deontology, it is pretty simple to say there should be a "rule" against cheating. Students have a duty to submit work that is their own, so having a large section of work created by AI would be unethical. Professors and other teachers have a duty to transfer knowledge to students so allowing students to cheat would violate that obligation. From a con-

Article citation: Jeffrey, S. A. (2024). The ethical considerations of artificial intelligence (AI). *Journal of Behavioral and Applied Management, 24*(2), 131–134.

sequentialist perspective, using generative AI to create an entire assignment would also be viewed as unethical. While the saving of time would be of benefit, the cost to potential learning from the assignment would outweigh any of these benefits.

The ethics of care is more complicated. The core idea in the ethics of care is that any action must stem from care. It seems relatively clear that a student using generative AI is acting out of care for themselves, at least their short-term selves. As previously discussed the student's learning may be negatively affected. This would violate the central tenet of care ethics in that this would harm the student. It is also important to look at this from the perspective of the instructor. Hopefully, instructors at all levels care about their students learning the material being delivered. If the use of generative AI negatively impacts learning, then a student using it to generate a final piece of work would not show care for the instructor and would therefore be unethical.

Creativity and Ownership

Closely related to the concept of academic integrity is ownership of material created using generative AI (Morgan, 2024). The US Patent office will not patent AI generated material because it views it as non-human created. Since AI "learns" by searching existing work, the newly generated item requires limited human work. Beyond this, there are few if any, legal requirements, however the ethical implications are clear. The important issue to consider is how much of the product is human created and how much is created by AI. Since patent protection brings financial gain, then that gain should not go to a person who received most of the design help from generative AI.

From a duty based perspective, it seems clear that this would be an unethical act based upon a simple rule that only one's own work should generate benefits. A consequentialist lens would calculate the net benefit of the action. If the new product or service is of value to society then that would go to the benefits. Since AI is machine based, one could say that no person is hurt by utilizing AI in generating this idea, therefore consequentialism would probably view this action as ethical.

Looking at this through the lens of care ethics is as in the other cases is more complicated as it is not immediately clear who should be cared about. If it is other possible inventors, then care ethics would suggest that using AI to design a product would harm those other possible inventors, making this unethical. If the general population was considered to be the relevant stakeholder, then using AI would be considered ethical if it brought the product or service to market sooner as this would demonstrate care for the people who might use the product or service.

Deep Fakes

As discovered by Taylor Swift and her fans, nude "fakes" (manipulated images) have become a serious issue for celebrities. Prior to the availability of AI tools, most fakes were limited to Photoshop or similar software. Because of this, making a good fake took a great deal of skill. Traditionally, these fakes were also static. Now with AI, anyone can make a very accurate looking video fake of anyone. To qualify as a deep fake, there is also a need for false audio. Beyond problems with celebrities, it can be a serious problem for politicians. If someone takes the time, they can make it seem like a politician is saying something that they did not say (Bond, 2024; Thompson & Maheshwari, 2023). Clearly this would be unethical as well as potentially dangerous. In fact, deep fakes have been used to spread misinformation (Dan et al., 2021; Lim et al., 2024).

From a deontological perspective, this would seem unethical as something fake would be considered dishonest (i.e. a lie). Using the consequentialist lens, a deep fake of a celebrity would be very damaging to that individual. Most likely more harm than any prurient gains delivered to viewers, leading to an overall negative outcome. Such an outcome would suggest any deep fake would be unethical. Looking at help with spreading untruths would suggest that the overwhelming damage done to the public would overwhelm any potential gain. This calculus would make the act unethical from the teleological perspective.

From a care ethics standpoint, the key stakeholder in celebrity fakes would be the celebrity him or herself. Since a nude fake would be damaging to someone's reputation, it would be easily viewed as unethical. In the case of a political deep fake, it would be clearly harmful to anyone who believed it, making it clearly unethical. It would likely also damage the reputation of any politician who had their identity and statements forged. Therefore care ethics would judge deep fakes as unethical.

Bias and Discrimination

Facial recognition has a problem with recognizing faces of non-white individuals (Chapman & Brustein, 2018). This could lead to more frequent arrests of innocent people of color (Marks, 2021). This is becoming more of a problem as more police departments and other law enforcement use AI to determine who to arrest. In addition, it has been shown that AI algorithms are less likely to approve a loan to individuals of color. AI systems have also been shown to display gender biases such as approving lower credit limits for women (Blascak & Tranfaglia, 2021).

From a purely deontological viewpoint both of these issues would violate issues of fairness, one of the central tenets of rule based ethics; therefore this would be considered unethical. Using a consequentialist lens, this would also be considered unethical. The benefits would accrue to those who use AI to be faster, but there would be overwhelming damage to those on the receiving end of these outcomes. Therefore on a net basis, there would be an overall negative assessment making this practice unethical.

Care ethics would draw a similar assessment of the ethics of using this AI. Caring for those who receive these

financial assessments would require an unbiased assessment of their credit worthiness. Since AI learns from past decisions about credit, use of AI in this context would be unethical from a care ethics viewpoint.

Self-Driving Cars (Autonomous Vehicles)

There have been many news stories about self-driving cars causing traffic fatalities and who can and should be blamed (Copp et al., 2023). Some ethicists have suggested that simply following the laws would maintain the ethics of self-driving cars (Designing Ethical Self-Driving Cars, 2023). On the surface, this looks like a classical trolley problem, where a person must decide whether to redirect a trolley that would kill five innocent people. However, if the trolley is redirected, it will kill one innocent person (Woollard, 2023). From a consequentialist viewpoint, this would be the ethical thing to do, as one life lost is better than five. A deontological perspective would likely disagree as most people would have a rule about actively causing a death. Woolard (2003) attempts to delineate this further by creating a distinction between allowing something to happen vs. performing an act. She claims that once agency is removed, the assessment is clearly different; that allowing something to happen is different than causing the act.

Following laws might not feel satisfying to many people who evaluate self-driving cars, but others may wonder whether stage four of Kohlberg's stages would be sufficient for all people (Kohlberg, 1973). Stage four of Kohlberg's stages is Law and Order morality, this implies that essentially following the law will ensure that actions are ethical. Looking at autonomous vehicles through the lens of care ethics, Care ethics would allow breaking the law if necessary to avoid harm to individuals. Care ethics would require social contract thinking (Kohlberg stage 5). This means of evaluation would allow for the breaking of traffic laws to avoid harm. As an example,x it would be permissible to cross a solid white line to avoid hitting a pedestrian provided there were no cars across the line.

The general problem is going to arise when the vehicle is programmed. The ethical views of those programming the vehicle might slip in. If the vehicles are programmed to be utilitarian, deontological views of the programmer might cause a change in the programming. This is hypothetical and could be solved by controls prior to shipping, but it still must be considered as an ethical problem.

Conclusion

Artificial Intelligence will begin to affect our lives in many ways and we should look carefully at all ethical issues before we accept it. AI will change the job skills required by employees, change the nature of decision making, and affect education at all levels.

This manuscript has looked at five potential uses of AI, taking a 10,000 foot view of some of the ethical concerns of AI in those contexts. The paper was selective, in that it only looked at five areas: Generative AI in education, creativity and ownership, deep fakes, bias and discrimination, and self-driving cars. It is obvious that AI will impact many more areas, with each requiring a deeper look at the issues involved.

References

Blascak, N., & Tranfaglia, A. (2021). *Decomposing gender differences in bankcard credit limits* (Working Paper (Federal Reserve Bank of Philadelphia) 21–35; Working Paper (Federal Reserve Bank of Philadelphia), pp. 21–35). Federal Reserve Bank of Philadelphia. https://doi.org/10.21799/frbp.wp.2021.35

Bond, S. (2024, February 8). *AI fakes raise election risks as lawmakers and tech companies scramble to catch up*. NPR. https://www.npr.org/2024/02/08/1229641751/ai-deepfakes-election-risks-lawmakers-tech-companies-artificial-intelligence

Bourcier, B. (2020). Jeremy Bentham's Principle of Utility and Taste: An Alternative Approach to Aesthetics in Two Stages. In A. Julius, M. Quinn, & P. Schofield (Eds.), *Bentham and the Arts* (1–xii, 299, pp. 227–243). University College London (UCL).

Chapman, L., & Brustein, J. (2018). *A.I. has a race problem*. Bloomberg Businessweek, 4575, 20–22.

Copp, C. J., Cabell, J. J., & Kemmelmeier, M. (2023). Plenty of blame to go around: Attributions of responsibility in a fatal autonomous vehicle accident. *Current Psychology: Research and Reviews, 42*(8), 6752–6767. https://doi.org/10.1007/s12144-021-01956-5

Dan, V., Paris, B., Donovan, J., Hameleers, M., Roozenbeek, J., van der Linden, S., & von Sikorski, C. (2021). Visual mis- and eisinformation, social media, and democracy. *Journalism & Mass Communication Quarterly, 98*(3), 641–664. https://doi.org/10.1177/10776990211035395

Designing Ethical Self-Driving Cars. (2023, January 23). https://hai.stanford.edu/news/designing-ethical-self-driving-cars

Edwards, S. D. (2009). Three versions of an ethics of care. *Nursing Philosophy, 10*(4), 231–240. https://doi.org/10.1111/j.1466-769X.2009.00415.x

Golden mean | Definition, Aristotle, Maimonides, Buddhism, Confucianism, & Facts Britannica. (n.d.). https://www.britannica.com/topic/golden-mean

How do I cite generative AI in MLA style? (2023, March 17). MLA Style Center. https://style.mla.org/citing-generative-ai/

How to cite ChatGPT. (n.d.). Https://Apastyle.Apa.Org. Retrieved June 26, 2024, from https://apastyle.apa.org/blog/how-to-cite-chatgpt

Kohlberg, L. (1973). Stages and Aging in Moral Development—Some Speculations. *The Gerontologist, 13*(4), 497–502. https://doi.org/10.1093/geront/13.4.497

Li, M., Enkhtur, A., Cheng, F., & Yamamoto, B. A. (n.d.). *Ethical Implications of ChatGPT in Higher Education: A Scoping Review*.

Lim, X.-J., Quach, S., Thaichon, P., Cheah, J.-H., & Ting, H. (2024). Fact or fake: Information, misinformation and disinformation via social media. *Journal of Strategic Marketing, 32*(5), 659–664. https://doi.org/10.1080/0965254X.2024.2306558

Louden, R. B. (1996). Toward a genealogy of `deontology'. *Journal of the History of Philosophy, 34*(4), 571. https://doi.org/10.1353/hph.1996.0070

Marks, P. (2021). Can the biases in facial recognition be fixed; Also, should they? *Communications of the ACM, 64*(3), 20–22. https://doi.org/10.1145/3446877

Morgan, A. (2024). Who Will Own AI-Built Designs? Generative AI could transform product design, but raises questions about creative ownership. Printed Circuit Design & Fab: *Circuits Assembly, 41*(1), 59–61.

Noddings, N. (2012). The Language of Care Ethics. *Knowledge Quest, 40*(5), 52–56.

Tarsney, C. (2018). Moral Uncertainty for Deontologists. *Ethical Theory & Moral Practice, 21*(3), 505–520. https://doi.org/10.1007/s10677-018-9924-4

Thompson, S. A., & Maheshwari, S. (2023, October 12). '*A.I. Obama' and Fake Newscasters: How A.I. Audio Is Swarming TikTok.* The New York Times. https://www.nytimes.com/2023/10/12/technology/tiktok-ai-generated-voices-disinformation.html

Woollard, F. (2023). The new trolley problem: Driverless cars and deontological distinctions. *Journal of Applied Philosophy, 40*(1), 49–64. https://doi.org/10.1111/japp.12610

Scott A. Jeffrey (sjeffrey@monmouth.edu)