

A way to govern ethical use of artificial intelligence without hindering advancement

February 17 2023, by Rae Lynn Mitchell



Credit: Pixabay/CC0 Public Domain

Texas A&M University School of Public Health researchers are developing a new governance model for ethical guidance and enforcement in the rapidly advancing field of artificial intelligence (AI).

Known as Copyleft AI with Trusted Enforcement, or CAITE, the researchers believe this model will guard against the potential harms of AI without hindering technological advancements.

Cason Schmit, JD, assistant professor at the School of Public Health and director of the Program in Health Law and Policy, Megan Doerr of Sage Bionetworks and Jennifer Wager, JD, of Penn State discuss their new [model](#) in a recent article in the journal *Science*.

Artificial intelligence (AI) promises to revolutionize nearly every aspect of our daily lives. However, misuse of AI-based tools could cause harm, especially to communities already facing inequity and discrimination. This potential for harm calls for ethical guidance through regulation and policy. But the rapid advancement of AI and the often-inflexible nature of government regulation have made creating such ethical guidance challenging.

Schmit, Doerr and Wager developed the CAITE model to face these challenges. CAITE combines aspects of copyleft licensing and the patent-troll model, two methods of managing [intellectual property rights](#) that can be considered at odds with one another.

Copyleft licensing allows sharing of intellectual property under conditions like attributing the original creator or noncommercial use, and derived works must use the same license terms as the original. Creative Commons licenses are a type of copyleft licensing. However, copyleft licensing schemes usually have little enforcement power.

The other side of the CAITE model uses the patent troll approach, which uses enforcement rights to ensure compliance. A patent troll is an organization that owns intellectual property rights and sues others in court as a way to generate revenue rather than creating or licensing technology.

The CAITE model is built on an ethical use license. This license would restrict certain unethical AI uses and require users to abide by a code of conduct. Importantly, it would use a copyleft approach to ensure that developers who create derivative models and data must also use the same license terms as the parent works. The license would assign the enforcement rights of the license to a designated third-party known as a CAITE host. In this way, the enforcement rights for all these ethical use licenses would pool in a single organization, empowering the CAITE host as a quasi-government regulator of AI.

"This approach combines the best of two worlds: a model that is as fast and flexible as industry, but with enforcement teeth and power of a traditional government regulator," Schmit said.

The authors note that using a nongovernment party designated by the AI developer community could allow for greater flexibility in enforcement and trust in oversight. CAITE hosts can set consequences for unethical actions such as financial penalties or reporting instances of consumer protection law violations.

At the same time, the CAITE approach allows for leniency policies that can promote self-reporting and gives flexibility that typical government enforcement schemes often lack. For example, incentives for AI users to report biases that they discover in their AI models could enable the CAITE host to warn other AI users who are relying on those potentially dangerous AI models.

Schmit and colleagues point out that the CAITE approach, while flexible, will require participation of a large portion of the AI community. Additionally, pilot implementation of ethical policies built using the CAITE approach will require further research and funding. Implementing this model will also rely on AI community members from many different disciplines to develop its features and overcome

challenges that arise.

Although it will require significant community buy-in, and possibly government incentives, Schmit and colleagues state that industry will likely prefer the more flexible CAITE framework to the stringent and slow-to-adapt regulations that governments could eventually impose.

"Efforts to promote ethical and trustworthy AI must go beyond what is legally mandated as the baseline for acceptable conduct," Wagner said. "We can and should strive to do better than what is minimally acceptable."

Once implemented, CAITE will guard against the potential harms of AI without hindering technological advances. The researchers say that as AI continues to expand into more of our daily lives, the value of a responsive ethical framework will become crucial.

More information: C. D. Schmit et al, Leveraging IP for AI governance, *Science* (2023). [DOI: 10.1126/science.add2202](https://doi.org/10.1126/science.add2202)

Provided by Texas A&M University

Citation: A way to govern ethical use of artificial intelligence without hindering advancement (2023, February 17) retrieved 27 April 2024 from <https://techxplore.com/news/2023-02-ethical-artificial-intelligence-hindering-advancement.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.