

Could artificial intelligence help us build a technological world that is more ethical?

December 13 2022, by Juan F. Samaniego



Artificial intelligence could serve to empower the user. Credit: Lucrezia Carnelos on Unsplash

A robot may not injure a human being or, through inaction, allow a human being to come to harm. A robot must obey the orders given it by human beings except where such orders would conflict with the First Law. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

The Three Laws of Robotics were set out by Isaac Asimov eighty years ago, long before artificial intelligence became a reality. But they perfectly illustrate how humans have dealt with the [ethical challenges](#) of technology: by protecting the users.

However, the ethical challenges facing humanity, whether they are related to technology or not, are not really a technological problem, but rather a social one. As such, technology in general, and artificial intelligence in particular, could be used to empower users and help us move toward a world that is more ethically desirable. In other words, we can rethink the way we design technology and artificial intelligence and draw on them to build a more ethical society.

This is the approach put forward by Joan Casas-Roma, a researcher at the SmartLearn group belonging to the Faculty of Computer Science, Multimedia and Telecommunications at the Universitat Oberta de Catalunya (UOC), in his open-access "Ethical Idealism, Technology and Practice: a Manifesto," published in *Philosophy & Technology*. In order to understand how to implement this [paradigm shift](#), we need to go back in time a little.

Artificial intelligence is objective, right?

When Asimov first set out his laws of robotics, the world was a very low-tech place compared to the present day. It was 1942 and Alan Turing had only just finished formalizing the algorithmic concepts that would be key to the development of modern computing decades later. There were no computers, no internet, let alone artificial intelligence or autonomous robots. But Asimov was already anticipating the fear that humans would succeed in making machines so intelligent that they would end up rebelling against their creators.

But later, in the early days of computing and data technologies in the

1960s, these issues were not among the key concerns of science. "There was a belief that, because the data were objective and scientific, the resulting information was going to be true and of high quality. It was derived from an algorithm in the same way that something is derived from a mathematical calculation. Artificial intelligence was objective and therefore helped us to eliminate human bias," explained Joan Casas-Roma.

But this was not the case. We came to realize that the data and the algorithms replicated the model or worldview of the person who was using the data or who had designed the system. In other words, the technology itself was not eliminating human biases, but rather transferring them to a new medium. "Over time, we have learned that artificial intelligence is not necessarily objective and, therefore, its decisions can be highly biased. The decisions perpetuated inequalities, rather than fixing them," he said.

So, we have ended up at the same point that was anticipated by the Laws of Robotics. Questions about ethics and artificial intelligence were brought to the table from a reactive and protective point of view. When we realized that artificial intelligence was neither fair nor objective, we decided to start acting to contain its [harmful effects](#). "The ethical question of artificial intelligence arose from the need to build a shield so that the undesirable effects of technology on users would not continue to be perpetuated. It was necessary to do so," said Casas-Roma.

As he explains in the manifesto, the fact of having to react in this way has meant that over the past few decades we have not explored another fundamental question in the relationship between technology and ethics: what ethically desirable consequences might a set of artificial intelligences with access to an unprecedented amount of data help us to achieve? In other words, how can technology help us move toward the construction of an ethically desirable future?

Toward an idealistic relationship between ethics and technology

One of the European Union's major mid-term goals is to move toward a more inclusive, more integrated and more cooperative society in which citizens have a greater understanding of global challenges. To achieve it, technology and artificial intelligence could be a major obstacle, but they could also be a great ally. "Depending on how people's interaction is designed with [artificial intelligence](#), a more cooperative society could be promoted," said Casas-Roma.

There has been an undeniable boom in online education in recent years. Digital learning tools have many benefits, but they can also contribute to a sense of isolation. "Technology could encourage a greater sense of cooperation and create a greater sense of community. For example, instead of having a system that only automatically corrects exercises, the system could also send a message to another classmate who has solved the problem to make it easier for students to help each other. It's just one idea to understand how technology can be designed to help us interact in a way that promotes community and cooperation," he said.

According to Casas-Roma, an ethical idealist perspective can rethink how technology and the way users use it can create new opportunities to achieve ethical benefits for the users themselves and society as a whole. This idealistic approach to the ethics of technology should have the following characteristics:

- **Expansive.** Technology and its uses should be designed in a way that enables its users to flourish and become more empowered.
- **Idealist.** The end goal that should always be kept in mind is how technology could make things better.
- **Enabling.** The possibilities created by technology must be

carefully understood and shaped to ensure that they enhance and support the ethical growth of users and societies.

- **Mutable.** The current state of affairs should not be taken for granted. The current social, political and economic landscape, as well as technology and the way it is used, could be reshaped to enable progress toward a different ideal state of affairs.
- **Principle-based.** The way technology is used should be seen as an opportunity to enable and promote behaviors, interactions and practices that are aligned with certain desired ethical principles.

"It's not so much a question of data or algorithms. It is a matter of rethinking how we interact and how we would like to interact, what we are enabling through a technology that imposes itself as a medium," concluded Joan Casas-Roma.

"This idea is not so much a proposal concerning the power of technology, but rather the way of thinking behind whoever designs the technology. It is a call for a paradigm shift, a change of mindset. The ethical effects of technology are not a technological problem, but rather a social problem. They pose the problem of how we interact with each other and with our surroundings through technology."

More information: Joan Casas-Roma, Ethical Idealism, Technology and Practice: a Manifesto, *Philosophy & Technology* (2022). [DOI: 10.1007/s13347-022-00575-7](https://doi.org/10.1007/s13347-022-00575-7)

Provided by Universitat Oberta de Catalunya

Citation: Could artificial intelligence help us build a technological world that is more ethical? (2022, December 13) retrieved 27 April 2024 from <https://techxplore.com/news/2022-12-artificial-intelligence-technological-world-ethical.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.