



## **EU On AI Ethics: Must ‘Enhance Positive Social Change’**

EU Technocrats define ethics for AI: “Environmental and societal well-being — AI systems should be sustainable and “enhance positive social change.” This is the coveted ‘Science of Social Engineering’ that harkens back to the original Technocracy movement in the 1930s. □ TN Editor

The European Union today published a set of guidelines on how companies and governments should develop ethical applications of artificial intelligence.

These rules aren’t like Isaac Asimov’s “Three Laws of Robotics.” They don’t offer a snappy, moral framework that will help us control murderous robots. Instead, they address the murky and diffuse problems that will affect society as we integrate AI into sectors like health care, education, and consumer technology.

So, for example, if an AI system diagnoses you with cancer sometime in the future, the EU's guidelines would want to make sure that a number of things take place: that the software wasn't biased by your race or gender, that it didn't override the objections of a human doctor, and that it gave the patient the option to have their diagnosis explained to them.

So, yes, these guidelines are about stopping AI from running amuck, but on the level of admin and bureaucracy, not Asimov-style murder mysteries.

To help with this goal, the EU convened a group of 52 experts who came up with seven requirements they think future AI systems should meet. They are as follows:

- **Human agency and oversight** — AI should not trample on human autonomy. People should not be manipulated or coerced by AI systems, and humans should be able to intervene or oversee every decision that the software makes.
- **Technical robustness and safety** — AI should be secure and accurate. It shouldn't be easily compromised by external attacks (such as adversarial examples), and it should be reasonably reliable.
- **Privacy and data governance** — Personal data collected by AI systems should be secure and private. It shouldn't be accessible to just anyone, and it shouldn't be easily stolen.
- **Transparency** — Data and algorithms used to create an AI system should be accessible, and the decisions made by the software should be "understood and traced by human beings." In other words, operators should be able to explain the decisions their AI systems make.
- **Diversity, non-discrimination, and fairness** — Services provided by AI should be available to all, regardless of age, gender, race, or other characteristics. Similarly, systems should not be biased along these lines.
- **Environmental and societal well-being** — **AI systems should be sustainable (i.e., they should be ecologically responsible) and "enhance positive social change"**
- **Accountability** — AI systems should be auditable and covered

by existing protections for corporate whistleblowers. Negative impacts of systems should be acknowledged and reported in advance.

You'll notice that some of these requirements are pretty abstract and would be hard to assess in an objective sense. (Definitions of "positive social change," for example, vary hugely from person to person and country to country.) But others are more straightforward and could be tested via government oversight. Sharing the data used to train government AI systems, for example, could be a good way to fight against biased algorithms.

These guidelines aren't legally binding, but they could shape any future legislation drafted by the European Union. The EU has repeatedly said it wants to be a leader in ethical AI, and it has shown with GDPR that it's willing to create far-reaching laws that protect digital rights.

But this role has been partly forced on the EU by circumstance. It can't compete with America and China — the world's leaders in AI — when it comes to investment and cutting-edge research, so it's chosen ethics as its best bet to shape the technology's future.

[Read full story here...](#)