

Artificial Intelligence – governance

Risk expert

Ariel Conn

Founder and President,
Magnitude 10 Consulting

Governance of AI in the EU

The European Union has taken center stage in efforts to regulate artificial intelligence and its impact on the digital world. The EU is currently in various stages of development of three major Acts that will influence how the world uses AI: The Artificial Intelligence Act, The Digital Services Act and The Digital Markets Act. Though each of these are designed with EU citizens in mind, we can expect each will influence the global use of AI. As was the case with the General Data Protection Regulation (GDPR), the three Acts will require any business and/or website operating in the EU to adhere to the new regulations, even if it's based in another country. Also similar to the GDPR, many analysts predict that these Acts will become gold standards for regulating AI and digitalization. So what does each of these Acts do?

The Artificial Intelligence Act (AI Act)

The EU AI Act is a proposal by the European Commission, which will regulate AI use according to four risk-based categories.

The first, “unacceptable risk,” includes AI uses that violate fundamental rights. AI systems that fall into this category are prohibited outright and include AI programs that have “significant potential to manipulate persons through subliminal techniques,” especially those that target vulnerable groups. Also prohibited would be “AI-based social scoring for general purposes done by public authorities” and law-enforcement use of “‘real time’ remote biometric identification systems in publicly accessible spaces.”

The second category of AI risk is “high risk.” Whether or not an AI system is classified as high risk will depend on the intended purpose of the systems, and whether the system creates “a high risk to the health and safety or fundamental rights of natural persons.” The Act provides specific rules to increase transparency, traceability, and robustness for such systems.

The third and fourth categories are “low and minimal risk.” These include AI systems such as chatbots (low risk) and spam filters (minimal risk), with regulation



of low-risk systems focusing primarily on transparency for the user about their interaction with the system.

Penalties for noncompliance would be fines of up to €30 million (US\$31 million) or up to 6% of a company's total annual revenue. While the Act would create the strongest laws to prevent harm from AI, some fear it doesn't go far enough, especially with respect to restricting facial recognition in public, which could be used for mass surveillance.

As of this writing, the European Parliament and EU Member States were in the process of amending the document. A final version of the Act will likely take at least another year or longer to be completed and agreed upon, and after that, companies will have a couple of years to come into compliance.

The digital services package

In 2020, the European Commission proposed a digital services package which would include The Digital Services Act and The Digital Markets Act. Earlier this year, the European Commission, the European Parliament, and EU Member States announced their agreement on both Acts.

The Digital Services Act (DSA) is a comprehensive regulation that seeks to provide more safeguards for users and consumers of online marketplaces. The DSA requirements will be dependent on the size of the online platforms, with the most stringent regulations being applied to very large online platforms and very large online search engines, defined as having over 45 million users in the EU.

Among the new obligations outlined in the DSA are rules that will

- Require platforms to quickly remove illegal content;
- Provide alternatives to recommendation algorithms based on profiling, and increase transparency around such systems;
- Prohibit “dark pattern,” which can manipulate consumers and users into purchases or information sharing;
- Increase protections for minors, such as prohibiting targeted advertising to minors based on their personal data;
- Require regular analysis of algorithms to ensure risk reduction, and much more.

The Digital Markets Act (DMA) is an effort by the EU to make online marketplaces more open to competition. It specifically targets the largest platforms or “gatekeepers,” the big tech companies – especially browsers, messengers, and social media – worth billions and with more than 45 million users in the EU each month. One of the primary goals of the DMA is to prevent the largest companies from stifling competition from smaller companies or newcomers to the space. The DMA will also require the “gatekeepers” to enable interoperability, which would, for example, allow users of different messaging platforms to interact with each other.

As with the AI Act, both the DSA and the DMA come with steep fines if companies don't comply. Both digital acts are expected to be finalized within a matter of months (as of this writing), but companies will still have a couple of years to adapt before they'll face penalties.



Governance of AI outside the EU

In the last few years, some of the world's most prominent companies and organizations have developed ethical principles for AI, including groups like Google, SAP, the European Commission's High Level Expert Group on Artificial Intelligence, the Organization for Economic Co-operation and Development (OECD), IEEE's Ethically Aligned Design, the UK House of Lords, the US Department of Defense, UNESCO, NATO, and many more.

However, in all cases, these efforts have been little more than advisory, offering guidelines and suggestions rather than concrete laws and regulations. This situation has proven woefully insufficient as companies like Google, Facebook, and many others have garnered negative public attention for their struggles to address ethics and discrimination, even within their own organizations.

In the US, the White House has called for an AI Bill of Rights, the Algorithmic Accountability Act of 2022 was introduced to Congress in February, and other federal organizations are also looking into a handful of initiatives to address issues of bias and discrimination. These are all in early stages, and it's unclear what impact they'll have or if they'll pass.

Notwithstanding efforts in the EU, which won't go into effect for at least a couple more years, companies continue to be expected to develop AI for good with little real oversight or direction.

Autonomous weapons

Autonomous weapons systems (AWS) are generally considered to be weapons that could select and engage a target, without a person overseeing the decision-making process. AWS have triggered intense ethical and legal debates around the world, as people try to define the extent to which an algorithm can (or should) decide who lives, who dies, and how.

The International Committee of the Red Cross has recommended "that States adopt new legally binding rules," providing three specific suggestions for aspects of autonomy that should be ruled out or regulated. The IEEE Standards Association recently published a document outlining over 60 ethical and technical challenges associated with the development, use and governance of AWS. The International Committee of the Red Cross has recommended "that States adopt new legally binding rules," providing three specific suggestions for aspects of autonomy that should be ruled out or regulated. The IEEE Standards Association recently published a document outlining over 60 ethical and technical challenges associated with the development, use and governance of AWS.

However, though member states of the United Nations Convention on Conventional Weapons have considered this question for nearly a decade, they have yet to find consensus on the development and use of such weapons.

Autonomous weapons pose another threat too: if countries race to develop more powerful autonomous weapons, they could inadvertently find themselves in a race for advanced AI more generally. In such a situation, developers may cut corners or get sloppy in their efforts to be the first to create something new, and the resulting artificial intelligence systems are more likely to behave unpredictably or cause problems in some way.