ICLE International Center for Law & Economics

What You Need to Know About the EU’s New AI Regulation
April 23, 2021
Jason Pilkington

The European Commission this week published its proposed Artificial Intelligence Regulation, setting out new rules for “artificial intelligence systems” used within the European Union. The regulation—the commission’s attempt to limit pernicious uses of AI without discouraging its adoption in beneficial cases—casts a wide net in defining AI to include essentially any software developed using machine learning. As a result, a host of software may fall under the regulation’s purview.

The regulation categorizes AIs by the kind and extent of risk they may pose to health, safety, and fundamental rights, with the overarching goal to:

- Prohibit “unacceptable risk” AIs outright;
- Place strict restrictions on “high-risk” AIs;
- Place minor restrictions on “limited-risk” AIs;
- Create voluntary “codes of conduct” for “minimal-risk” AIs;
- Establish a regulatory sandbox regime for AI systems;

- Set up a European Artificial Intelligence Board to oversee regulatory implementation; and
- Set fines for noncompliance at up to 30 million euros, or 6% of worldwide turnover, whichever is greater.

AIs That Are Prohibited Outright

The regulation prohibits AI that are used to exploit people's vulnerabilities or that use subliminal techniques to distort behavior in a way likely to cause physical or psychological harm. Also prohibited are AIs used by public authorities to give people a trustworthiness score, if that score would then be used to treat a person unfavorably in a separate context or in a way that is disproportionate. The regulation also bans the use of "real-time" remote biometric identification (such as facial-recognition technology) in public spaces by law enforcement, with exceptions for specific and limited uses, such as searching for a missing child.

The first prohibition raises some interesting questions. The regulation says that an "exploited vulnerability" must relate to age or disability. In its announcement, the commission says this is targeted toward AIs such as toys that might induce a child to engage in dangerous behavior.

The ban on AIs using "subliminal techniques" is more opaque. The regulation doesn't give a clear definition of what constitutes a "subliminal technique," other than that it must be something "beyond a person's consciousness." Would this include TikTok's algorithm, which imperceptibly adjusts the videos shown to the user to keep them engaged on the platform? The notion that this might cause harm is not fanciful, but it's unclear whether the provision would be interpreted to be that expansive, whatever the commission's intent might be. There is at least a risk that this provision would discourage innovative new uses of AI, causing businesses to err on the side of caution to avoid the huge penalties that breaking the rules would incur.

The prohibition on AIs used for social scoring is limited to public authorities. That leaves space for socially useful expansions of scoring systems, such as consumers using their Uber rating to show a record of previous good behavior to a potential Airbnb host. The ban is clearly oriented toward more expansive and dystopian uses of social credit systems, which some fear may be used to arbitrarily lock people out of society.

The ban on remote biometric identification AI is similarly limited to its use by law enforcement in public spaces. The limited exceptions (preventing an imminent terrorist attack, searching for a missing child, etc.) would be subject to judicial authorization except in cases of emergency, where ex-post authorization can be sought. The prohibition leaves room for private enterprises to innovate, but all non-prohibited uses of remote biometric identification would be subject to the requirements for high-risk AIs.

Restrictions on 'High-Risk' AIs

Some AI uses are not prohibited outright, but instead categorized as "high-risk" and subject to strict rules before they can be used or put to market. AI systems considered to be high-risk include those used for:

- Safety components for certain types of products;
- Remote biometric identification, except those uses that are banned outright;
- Safety components in the management and operation of critical infrastructure, such as gas and electricity networks;
- Dispatching emergency services;
- Educational admissions and assessments;
- Employment, workers management, and access to self-employment;
- Evaluating credit-worthiness;
- Assessing eligibility to receive social security benefits or services;
- A range of law-enforcement purposes (e.g., detecting deepfakes or predicting the occurrence of criminal offenses);
- Migration, asylum, and border-control management; and
- Administration of justice.

While the commission considers these AIs to be those most likely to cause individual or social harm, it may not have appropriately balanced those perceived harms with the onerous regulatory burdens placed upon their use.

As Mikołaj Barczentewicz at the Surrey Law and Technology Hub has pointed out, the regulation would discourage even simple uses of logic or machine-learning systems in such settings as education or workplaces. This would mean that any workplace that develops machine-learning tools to enhance productivity—through, for example, monitoring or task allocation—would be subject to stringent requirements. These include requirements to have risk-management systems in place, to use only "high quality" datasets, and to allow human oversight of the AI, as well as other requirements around transparency and documentation.

The obligations would apply to any companies or government agencies that develop an AI (or for whom an AI is developed) with a view toward marketing it or putting it into service under their own name. The obligations could even attach to distributors, importers, users, or other third parties if they make a "substantial modification" to the high-risk AI, market it under their own name, or change its intended purpose—all of which could potentially discourage adaptive use.

Without going into unnecessary detail regarding each requirement, some are likely to have competition- and innovation-distorting effects that are worth discussing.

The rule that data used to train, validate, or test a high-risk AI has to be high quality ("relevant, representative, and free of errors") assumes that perfect, error-free data sets exist, or can easily be detected. Not only is this not necessarily the case, but the requirement could impose an impossible standard on some activities. Given this high bar, high-risk AIs that use data of merely "good" quality could be precluded. It also would cut against the frontiers of research in artificial intelligence, where sometimes only small and lower-quality datasets are available to train AI. A predictable effect is that the rule would benefit large companies that are more likely to have access to large, high-quality datasets, while rules like the GDPR make it difficult for smaller companies to acquire that data.

High-risk AIs also must submit technical and user documentation that detail voluminous information about the AI system, including descriptions of the AI's elements, its development, monitoring, functioning, and control. These must demonstrate the AI complies with all the requirements for high-risk AIs, in addition to documenting its characteristics, capabilities, and limitations. The requirement to produce vast amounts of information represents another potentially significant compliance cost that will be particularly felt by startups and other small and medium-sized enterprises (SMEs). This could further discourage AI adoption within the EU, as European enterprises already consider liability for potential damages and regulatory obstacles as impediments to AI adoption.

The requirement that the AI be subject to human oversight entails that the AI can be overseen and understood by a human being and that the AI can never override a human user. While it may be important that an AI used in, say, the criminal justice system must be understood by humans, this requirement could inhibit sophisticated uses beyond the reasoning of a human brain, such as how to safely operate a national electricity grid. Providers of high-risk AI systems also must establish a post-market monitoring system to evaluate continuous compliance with the regulation, representing another potentially significant ongoing cost for the use of high-risk AIs.

The regulation also places certain restrictions on "limited-risk" AIs, notably deepfakes and chatbots. Such AIs must be labeled to make a user aware they are looking at or listening to manipulated images, video, or audio. AIs must also be labeled to ensure humans are aware when they are speaking to an artificial intelligence, where this is not already obvious.

Taken together, these regulatory burdens may be greater than the benefits they generate, and could chill innovation and competition. The impact on smaller EU firms, which already are likely to struggle to compete with the American and Chinese tech giants, could prompt them to move outside the European jurisdiction altogether.

Regulatory Support for Innovation and Competition

To reduce the costs of these rules, the regulation also includes a new regulatory "sandbox" scheme. The sandboxes would putatively offer environments to develop and test AIs under the supervision of competent authorities, although exposure to liability would remain for harms caused to third parties and AIs would still have to comply with the requirements of the regulation.

SMEs and startups would have priority access to the regulatory sandboxes, although they must meet the same eligibility conditions as larger competitors. There would also be awareness-raising activities to help SMEs and startups to understand the rules; a “support channel” for SMEs within the national regulator; and adjusted fees for SMEs and startups to establish that their AIs conform with requirements.

These measures are intended to prevent the sort of chilling effect that was seen as a result of the GDPR, which led to a 17% increase in market concentration after it was introduced. But it’s unclear that they would accomplish this goal. (Notably, the GDPR contained similar provisions offering awareness-raising activities and derogations from specific duties for SMEs.) Firms operating in the “sandboxes” would still be exposed to liability, and the only significant difference to market conditions appears to be the “supervision” of competent authorities. It remains to be seen how this arrangement would sufficiently promote innovation as to overcome the burdens placed on AI by the significant new regulatory and compliance costs.

Governance and Enforcement

Each EU member state would be expected to appoint a “national competent authority” to implement and apply the regulation, as well as bodies to ensure high-risk systems conform with rules that require third party-assessments, such as remote biometric identification AIs.

The regulation establishes the European Artificial Intelligence Board to act as the union-wide regulatory body for AI. The board would be responsible for sharing best practices with member states, harmonizing practices among them, and issuing opinions on matters related to implementation.

As mentioned earlier, maximum penalties for marketing or using a prohibited AI (as well as for failing to use high-quality datasets) would be a steep 30 million euros or 6% of worldwide turnover, whichever is greater. Breaking other requirements for high-risk AIs carries maximum penalties of 20 million euros or 4% of worldwide turnover, while maximums of 10 million euros or 2% of worldwide turnover would be imposed for supplying incorrect, incomplete, or misleading information to the nationally appointed regulator.

Is the Commission Overplaying its Hand?

While the regulation only restricts AIs seen as creating risk to society, it defines that risk so broadly and vaguely that benign applications of AI may be included in its scope, intentionally or unintentionally. Moreover, the commission also proposes voluntary codes of conduct that would apply similar requirements to “minimal” risk AIs. These codes—optional for now—may signal the commission’s intent eventually to further broaden the regulation’s scope and application.

The commission clearly hopes it can rely on the “Brussels Effect” to steer the rest of the world toward tighter AI regulation, but it is also possible that other countries will seek to attract AI startups and investment by introducing less stringent regimes.

For the EU itself, more regulation must be balanced against the need to foster AI innovation. Without European tech giants of its own, the commission must be careful not to stifle the SMEs that form the backbone of the European market, particularly if global competitors are able to innovate more freely in the American or Chinese markets. If the commission has got the balance wrong, it may find that AI development simply goes elsewhere, with the EU fighting the battle for the future of AI with one hand tied behind its back.

View Article