Hogan Lovells

Engage

Legal insight and analysis

EU Lawmaker proposes to regulate generative AI – among other significant changes to the forthcoming AI Act

15 May 2023

On Thursday, 11 May, the Committees of the European Parliament for the Internal Market and Civil Liberties voted for far-reaching amendments to the EU's proposed artificial intelligence regulation ("AI Act"). The Committees' proposals seek to further develop the framework for regulating the risks associated with AI that have dominated the political debate in this area for several years, while also addressing the emerging and important concerns associated with generative AI and other forms of general purpose AI ("foundation models").

The proposed amendments also expand the scope of obligations for so-called "users" of AI systems – and establish "trustworthiness requirements" that shall, if implemented, apply to any AI system – irrespective of whether it is considered "high-risk" or not. Whilst the EU lawmakers deserve respect for their comprehensive and courageous amendment proposals, some of them may simply not go in the right direction. Supporting the competitiveness of EU research institutions and business requires careful consideration – as the EU cannot afford to fall back further in the global concert of fast paced AI innovation.

Next steps

The revised draft AI Act as proposed by the relevant Committees, will now go to a plenary vote by the European Parliament which is scheduled to take place in June 2023 and is expected to be approved. Given the substance and scope of the proposed changes, intensive trilateral negotiations will then likely follow between the European Council, Commission and Parliament, who will need to reach agreement on the final version of the Act. It is possible that a successful resolution may be reached before the end of 2023 – so that the world's first comprehensive piece of AI specific regulation can enter into force in early 2024.

Key amendments

The Parliament's proposals introduce a number of amendments that are particularly controversial and that are likely to dominate the upcoming discussion, namely (1) new rules on foundation models and generative AI, (2) new trustworthiness requirements for all AI systems, (3) an expansion of obligations that apply to the users of AI systems, (4) changes to the scope of those AI systems that are labelled "high-risk" and (5) additions to the list of prohibited AI practices.

1. Foundation Models and Generative AI

The proposed amendments introduce new rules for foundation models – that is AI models trained on a very broad range of sources and large amounts of data - for a generality of different applications. Such foundation models typically serve as basis for a wide range of downstream tasks and can be made available for such specific, dependent applications through means of "open source" or application program interface (API). A particularly visible and indeed omnipresent form of foundation models are what is called "generative AI", where the model is designed to generate content of any kind like texts, code, images, animations, videos or music.

The proposed provisions impose specific obligations on the providers of such foundation models to:

- guarantee robust protection of fundamental rights, health and safety and the environment, democracy and rule of law;
- assess and mitigate risks associated with the model;
- introduce practical measures in the design and development of the model to ensure certain standards are met;
- register the model in the newly introduced EU database.

The proposed provisions impose additional and farther reaching transparency requirements for "generative" foundation models, such as

- disclosing that AI generated the content;
- designing the model to prevent it from generating illegal content;
- publishing summaries of copyrighted data used for training; and
- supporting innovation and protecting citizens' rights

Exemptions to those obligations and requirements for foundation models in general and generative AI in particular are made for research activities and AI components provided under open-source licenses.

2. Challenges for the proposed rules for Foundation Models and Generative AI

Regulating Foundation models in general and for generative AI in particular is dearly needed and makes much sense. Yet, such regulation is anything but easy and here are some of the challenges that the attempt of the EU legislator may encounter in the discussion that will now ensue:

- does it really make sense for a general purpose AI – that is open to just any form of application – to "*guarantee the strong protection*" of "*human rights, health, safety, democracy, the environment and the rule of law*"? Whilst all these aspects clearly deserve attention, their specific content is very broad and difficult to define. Accordingly, it may make more sense to stick with the previous approach taken in the AI Act, as originally proposed – and that is to address all those questions on the level of application, rather than the model from which an application was generated. On the foundational level, a more subtle approach may be useful to not suffocate their development.

- The request for risk assessment/mitigation strategies for foundation models may just not be workable in a realistic way. Again, it is the question whether a model that serves as a platform for downstream applications is the right place for any such obligations. This may be particularly relevant as many risks only show and emerge during training phases for AI systems – which typically happens at the application level.

- The burden imposed on foundation models by the proposed provision may disproportionally disadvantage small and medium sized AI businesses as only big corporations may have the resources to comply with these wide-sweeping duties.

- The request for entities using generative AI to publish summaries of the copyright protected content they used may simply not be feasible. Aside from the gigantic size of the task itself, the exemptions provided by EU copyright law for mining copyrighted content may stand in the way.

- The request to disclose that any specific content was generated by AI may also not work. First, users may simply not comply which leads to a kind of predictable level of non-compliance which is unhelpful for just any piece of legislation. Whilst it is a good aim to go for utmost transparency, more care needs to be applied to make the obligations workable.

- The request to design the model in a way to prevent illegal content is well intended. Yet the legislator will need to consider whether it makes more sense to align these obligations with the similar obligations that exist already in form of "content moderation obligations" under the EU Digital Services Act. The phenomenon is very similar – if not the same: a platform is provided that may be utilized by third party applications to misuse it. For those forms of misuse, a consistent approach may make sense – and may actually be needed given the overlapping and relatedness – and also with a view to the specific political and societal dangers that derive from it.

3. General AI Principles

Previous versions of the AI Act from the European Commission and Council have predominantly focused on introducing obligations in relation to 'high-risk' use cases of AI. However, the Parliament's amendments propose to significantly widen the scope of the regulation, by introducing a set of general principles for the development and use of AI. These principles are intended to apply to all AI systems, irrespective of the risk that they pose. They will require organizations to exercise best efforts in developing and using AI systems in accordance with the following requirements:

- Human agency and oversight
- Technical robustness and safety
- Privacy and data governance
- Transparency
- Diversity, non-discrimination and fairness
- Social and environmental wellbeing

4. Additional User Obligations

While providers of high-risk AI systems (i.e., developers) are subject to the primary obligations under the AI Act, the Parliament's amendments also propose to broaden the range of requirements that apply to organizations that deploy these systems. These organizations have been referred to as 'users' in past versions of the AI Act, but are now referred to as 'deployers'.

These additional requirements include, for example:

- Undertaking a detailed impact assessment, which takes into account the risks that the AI system poses to the fundamental rights of individuals alongside the relevant mitigations that will be implemented.
- Introduction of an AI governance system, which addresses procedures for compliant-handling and redress.
- Implementation of human oversight over the relevant AI system.
- Provision of algorithmic transparency information to end-users.

5. High-Risk AI Systems

The comprehensive regulation of the specific risks deriving from so called "high-risk AI systems" has been the main focus and objective of the AI Act. Whether or not a specific AI qualifies as "high-risk" depends on its specific sphere of application and each use-case is explicitly listed within the AI Act. AI systems used in fields such as medical devices, automotive vehicles, educational assessment, job recruitment, credit assessments, critical infrastructure and health insurance have already been identified as "high-risk" in previous drafts.

New to the list of these “high-risk AI systems” are those applications that aim to “influence voters in political campaigns.” This addition was clearly needed and makes sense under any perspective. Whilst AI (including all sorts of data analytical methods) will clearly be used in the context of elections, it worthwhile for these uses to be particularly closely scrutinized by means of regulation. At the moment, there seems to be hardly any field of AI application that is in more need for a useful regulation.

6. Prohibited AI Practices

Art. 5 of the previous draft AI Act provided (already before the latest amendment proposals for wide-ranging prohibitions) of many AI practices that were regarded to be overly intrusive, discriminatory or otherwise abusive. Those prohibitions included in particular any AI practice to (1) apply forms of “social scoring” (2) exploit personal vulnerabilities, (3) discriminate or otherwise unduly categorize people according to gender, race, age, etc. or (4) undertake real-time biometric identification in publicly accessible spaces.

These already existing prohibitions were the target of intense criticism by many human rights groups and subject to multiple efforts to stretch the legal protection against any form of intrusive AI practice. These efforts proved successful and the latest amendment proposal by the EU Parliament Committees are now significantly more restrictive. The list of additional or significantly extended prohibitions now also cover systems for:

- retrospective (non-real time) remote biometric identification - with the only exception of law enforcement for the prosecution of serious crimes and only after judicial authorization;
- predictive profiling for law enforcement purposes (based on objective cluster criteria such as location/movement, conduct, past criminal behavior, etc.);
- emotion recognition (ie analysis of human conduct such as facial expressions, body language, gestures, and voice tones to assess their emotional state) in law enforcement, border management, workplace, and educational institutions; and
- indiscriminate scraping of biometric data from social media or CCTV footage to create facial recognition databases.

It is to be expected that some of these amendment proposals will meet the fierce resistance of some of the EU member states – that will exercise their say through the EU Council in the upcoming trialogue negotiations. The questions that will arise are very delicate and difficult indeed: is it acceptable to use the power of AI to go through publicly available information to build new databases to help identifying prospective wrongdoers? Is a strong limitation of the retrospective analysis of public space footage sensible in the context of crime investigation? Surely, these and related questions will continue to stir intensive discussion – by legislature and in society at large.

We will continue to inform you about the progress of this and other AI related pieces of legislation.

Authored by Leopold von Gerlach and Dan Whitehead.

Contacts

Leopold von Gerlach
Partner
Hamburg
leopold.vongerlach@hoganlovells.com

Dan Whitehead
Counsel
London
dan.whitehead@hoganlovells.com