EUROPEAN LAW BLOG
NEWS AND COMMENTS ON EU LAW

ELB Blogpost 32/2023, 24 July 2023
Tags: Artificial intelligence, digital governance, generative artificial intelligence, foundation models, law and technology
Topics: Artificial Intelligence, Data protection and digital governance

The EU AI Act at a crossroads: generative AI as a challenge for regulation

By Christian Djeffal

Advances in artificial intelligence (AI) have once again surprised people. New approaches to training very large context-aware systems have enabled generative AI systems (GAI), especially large language models (LLMs), which can produce content that is, in many cases, indistinguishable from the products of the human mind. ChatGPT, an LLM, has proven to be one of the fastest-growing consumer applications, and such popularity raises the question of how to govern and regulate AI even more pressing. Discussions in the AI community have been at a fever pitch. Two letters with humongous support from the AI community tried to stir debates. One from the Institute for the Future of Life called for a 6-month moratorium on experiments to figure out how to deal with systems adequately. Another letter stressed the risk of extinction to get policymakers to act. These debates on GAI and large AI systems have landed them in the middle of deliberations over a proposed AI Act. While there is a great deal of overlap between the European Commission's (Commission) original draft, the position of the Council of the European Union (Council), and the European Parliament's (Parliament) position, they diverge on how to deal with the models that are under discussion.

The European Commission's draft for an AI Act (DAIA) did not explicitly address GAI. However, with the rapid adoption of ChatGPT by consumers, it became a focus in later iterations of the legislative process. The mode of deliberation of the AI Act is the informal so-called fast-track legislative procedure, which consists of trilogues, i.e. informal consultations between the Commission, the Council, and the Parliament. It is a most exciting time to compare the approaches to learn about different ways EU law can and should tackle GAI. Therefore, this post will discuss the regulatory approaches and the importance of foresight and knowledge aspects of the proposed regulation.

Regulatory approach of the European Commission

The Commission's initial proposal and the Council's and EP's proposed amendments differ substantially in their treatment of GAI. The Commission's draft does not explicitly address this category of AI. The main approach to regulation comprises three structural dimensions: first the classification of AI according to risk levels, second according to the specific role of actors and third according to sectorial differences.. The risk levels have so far been at the centre of attention. There is generally a distinction between AI systems which cause:

- **Unacceptable risks,** leading to prohibitions of the use of such AI systems according to Title II (Article 5 DAIA)
- **High risks,** leading to regulations including a conformity assessment according to Title III, Chapters I&II (Arts. 6ff DAIA)
- **Limited risks,** requiring transparency according to Title IV
- **Low risks,** where establishing voluntary Codes of Conduct is to be encouraged according to Title IX

The central part of the proposed regulation addresses high-risk AI systems. High risk AI systems according to Article 6, are either part of a safety component of a product, or is a product itself, or specific applications in a sector designated as high risk according to Annex III. In the original Annex III, the following areas are designated as high risk: biometric identification and categorisation of natural persons; management and operation of critical infrastructure; education and vocational training; employment, workers management and access to self-employment; access to and enjoyment of essential private services and public services and benefits; law enforcement; migration, asylum and border control management; and administration of justice and democratic processes. In essence, the high-risk status is classified according to sector.

A sectoral approach is seen more broadly in the proposed Act. The other relations regarding sectors are:

- The applicability of DAIA in Article 2 is excluded for certain instruments;
- The specific content of the obligation can also be adapted to sectorial needs through further definition, for example, according to Article 40ff. DAIA;
- The definition of risks depends on certain sectors according to Article 6 (2) and Annex III DAIA
- The exact obligations in the conformity assessment also depend on sectors as is evidenced in Article 43 DAIA

- Aspects of enforcement can vary in relation to sectors according to Article 63ff. DAIA

Furthermore, obligations applying to high-risk systems depend on the actors they address. The main categories of actors are:

- Providers
- Product manufacturers
- distributors,
- importers, users
- or any other third-party

Chapter 3 of Title II describes the obligations that are tailored to the abilities and responsibilities of these actors. Therefore, in summary, the obligations which apply to specific AI systems and the nature of such obligations will depend on risk levels, the role of actors, and sectors. The Commission has created a complex but nuanced system of regulating AI in different circumstances. The general approach chosen by the Commission is apt to accommodate GAI, including LLMs. Its adaptability introduces a level of flexibility necessary to cover very different application areas, such as healthcare, education, and legal technology, by diverse actors. However, it also offers a typical frame of reference and a general framework. A purely sectoral regulation without a general framework would run the risk of multiplying the requirements for GAI in a way that they become too burdensome or even impossible to comply with. Therefore, the general approach of the Commission does the trick. Unsurprisingly, the Parliament and the Council did not radically alter but cautiously modify the draft in this respect.

Amendments of the Council: actors providing for general purpose AI

The Council has addressed the rise of GAI, such as large language models, by introducing a new category of AI systems, namely general purpose AI systems in its position. Article 3(1b) DAIA (Council position) defines such systems as:

> "intended by the provider to perform generally applicable functions such as image and speech recognition, audio and video generation, pattern detection, question answering, translation and others; a general-purpose AI system may be used in a plurality of contexts and be integrated in a plurality of other AI systems".

The problem identified by the Council relates to the general purpose nature of many AI systems, and the concerns over how to deal with AI systems which are open for various purposes. How can a technology that can simultaneously enable care robots and autonomous lethal weapons systems be regulated? The Council answers this and similar questions by providing that the rules on high-risk systems apply to general purpose AI that can be used in such contexts unless such uses are explicitly excluded (Article 4b, Council position). It also gives the Commission the authority to "specify and adapt" the requirements, which extends the Commission's competencies to alter the requirements for high-risk systems.

Furthermore, providers of general purpose AI are required to conduct conformity assessments (section 3, Council position). Most importantly, however, Article 4b is introduced which modifies the obligations for providers of general purpose AI systems, thereby creating a new role for such actors by limiting what such providers are obliged to do. While other providers of high-risk AI systems are generally subject to the obligations contained in Articles 16-25, providers of general purpose AI would only be bound by a lesser set of obligations:

- providing their name and trademark (Article 16 aa));
- conducting a conformity assessment (Article 16 e));
- registration (Article 16f));
- corrective actions (Article 16 g));
- CE marking (Article 16i));
- demonstrate conformity (Article 16 j));
- appointing an authorised representative (Article 25);
- EU declaration of conformity (Article 48);
- post market monitoring (Article 61); and
- sharing information with incoming competitors (Article 4b(5)).

The Council, therefore, adds a new role to the Commission's list of actors by limiting the obligations of providers of general purpose AI systems, and creating an additional obligation to share knowledge directly with competitors. In essence, 'by changing the Commission's approach to adopting a variation for certain systems, it suggests a way to adapt the AI Act's requirements for certain high-risk systems.

Amendments of the Parliament: a new risk class for foundation models

The Parliament attempts to address the issue of GAI by focusing on foundation models, effectively introducing a separate risk category. It defines foundation models as "an AI model trained on a wide range of data at scale, [which] is designed for the generality of output and can be adapted to a wide range of specific tasks" (Article 3(1c), Parliament position). The emphasis of this definition is less on the generality of the system and more on potential uses and the that such models can be further adapted to specific tasks. This emphasis becomes even more apparent when looking at Recital 60g: "There is considerable uncertainty as to how foundation models will evolve, both in terms of the typology of models and in terms of self-governance." The Parliament focuses more on the models themselves, providing several criteria for such models in Article 28b (1), which can be summarised as follows:

(a) Obligation to establish risk governance;

(b) Obligation to establish data governance;

(c) Requirements to have appropriate levels of performance, predictability, interpretability, corrigibility, safety and cybersecurity;

(d) Obligation to reduce energy use, resource use and waste, as well as to increase energy efficiency, and the overall efficiency of the system

(e) draw up extensive technical documentation and intelligible instructions for use;

(f) establish a quality management system; and

(g) register that foundation model in the EU database;

Article 4a(2) clarifies that the Parliament intends to limit the obligations of the operators of foundation models to specific obligations under Article 28b, and that they fall outside the framework for high-risk systems. Article 28b introduces the separate set of obligations for foundation models and providers thereof, and is summarized above. This distinct set of obligations effectively adds foundation models as a separate layer to the Commission's risk typology and reduces the responsibility of actors in that relation. It is a lighter version of for the obligations for high-risk AI systems.

Within the obligations created for providers of foundation models, the Parliament also adds a dedicated section on GAI (Article 28(b)(4), Parliament position), which reads

"*Providers of foundation models used in AI systems specifically intended to generate, with varying levels of autonomy, content such as complex text, images, audio, or video ("generative AI") and providers who specialise a foundation model into a generative AI system, shall in addition*

a) comply with the transparency obligations outlined in Article 52 (1),

b) train, and where applicable, design and develop the foundation model in such a way as to ensure adequate safeguards against the generation of content in breach of Union law *in line with the generally- acknowledged state of the art, and without prejudice to fundamental rights, including the freedom of expression,*

c) without prejudice to national or Union legislation on copyright, document and make publicly available a sufficiently detailed summary of the use of training data protected under copyright law."

This section is critical as it defines GAI and highlights transparency, the human rights implications of content generation, and intellectual property as essential additions to be addressed, alongside the more general obligations of GAI as foundation models.

Balancing innovation and responsibility

So far, the main challenge of analysing regulatory instruments on AI has been to weigh their negative impact on innovation against their positive impact on the responsible use of AI. The Council and Parliament positions have not radically changed the general framework established by the Commission. However, both proposed amendments did reduce the burden on providers of GAI. At first sight, the Parliament's amendments appear more innovation-friendly, as the catalogue of obligations in Article 28b for foundation models is a reduced version of the requirements for high-risk AI systems. However, easing the regulatory burden comes at the cost of further complicating the risk-based approach. The very general, short, and open description of the requirements for foundation models needs to be revised. the wording could lead to difficulties of understanding and interpreting the exact obligations. It could also happen that extensive obligations are introduced through tertiary legislation.

The advantage of the Council's approach lies in its stance towards responsibility. It modifies the obligations of providers of general purpose systems according to their potential to influence the systems in question. Finding tailored solutions for individual organisations and updating their processes is preferable to categorising systems in the dynamic field of AI as a set of emerging technologies. The Council's approach fits better with the attempt to provide general rules that cover many forms and configurations of AI technologies.

A general risk assessment obligation and the respective knowledge governance

One aspect that neither the Council nor the Commission addresses is the problem of how to deal with the open-endedness of general purpose technologies such as GAI. It is hard to foresee future uses that innovations will make possible. Nevertheless, the social impact of technologies very often comes from uses that were not anticipated when the technology was invented. In the early days of the Internet, little was known about the positive and negative impacts of e-commerce and social media. While regulation will certainly not be able to completely solve the problem of foresight, there are two general directions in which the draft could improve.

The first is to try to understand the implications of technologies at a very early stage. For general-purpose technologies such as GAI, it would be beneficial to more precisely define risk management processes for general-purpose systems. Assessments and mitigation measures should focus on the potential future impacts of the technologies and allow for continuous adaptation. Furthermore, such processes need to involve stakeholders. Such participatory mechanisms are already prescribed by Article 35(7) of the General Data Protection Regulation and Article 35 and Recital 90 of the Digital Services Act. It takes many perspectives and views to understand the impact of technology, so more is needed than for providers of general-purpose systems to reflect all by themselves.

The second important addition would involve knowledge management and the sharing of knowledge. There are already good ideas in the proposal and the amendments, such as the Council's proposal that providers of general-purpose AI systems should share knowledge on compliance with competitors entering the market in Article 4b(4). In its amendments to Annex VII, the Parliament also proposed an obligation to report information on foundation models to a database on high-risk AI systems. What still needs to be added is the mediation of knowledge between providers of AI and those who provide specific services based on this AI. Simple and effective obligations and infrastructures for vertical and horizontal knowledge sharing on risks and mitigation strategies would provide access to knowledge and reduce ignorance.

What is next?

Companies and innovators involved in the development of AI are actively seeking regulation. They are calling for a framework within which they can develop and use this powerful technology responsibly and minimize risks of unforeseen and potentially harmful consequences. This palpable demand creates a powerful and unique opportunity for EU lawmakers. They stand on the precipice of a new frontier, with the chance to create rules that have the potential to resonate globally. However, these rules should not spread because markets demand their adoption, and addressees bow to economic pressure. Instead, they should spread their influence because they come from a position of nuanced understanding of the profound implications of large AI systems, including GAI. By creating sensible and forward-looking regulation, EU lawmakers can redefine the narrative around AI. They can provide the blueprint for harnessing its potential while protecting humanity from its risks, fostering a future where AI technology is used responsibly and, in some cases, in support of the objectives laid out in the EU Treaties and the Charter.