Kluwer Competition Law Blog

A Deeper Look into the EU AI Act Trilogues: Fundamental Rights Impact Assessments, Generative AI and a European AI Office

Heidi Waem, Muhammed Demircan (DLA Piper) · Monday, November 13th, 2023

The World's first hard-law horizontal legislation on artificial intelligence is currently nearing political agreement between the European Union's (EU) three legislative branches, the European Parliament (**EP**), the Council of the European Union (**the Council**) and the European Commission (**EC**). Inter-institutional negotiations (or with their technical name, **trilogues**) continue, and there is an incredible amount of media coverage and buzz around the EU Artificial Intelligence Act (**EU AI Act**).

In this blog post, we aim to go beyond the surface. We will analyse a number of contentious matters from the EU AI Act that are currently being discussed in the hallways of the EU buildings. In particular, we will discuss i) Fundamental Rights Impact Assessments, ii) the rules over foundation models and generative AI that are being discussed, and iii) the enforcement and role of the European AI Office.

The EU AI Act, as a horizontally applicable legislation, is relevant for many subjects, ranging from the overall competitiveness of the EU to the protection of fundamental rights of ordinary citizens. Please note that the discussions are currently very fluid and subject to changes in the legislative process of the EU AI Act.

When do we expect the EU AI Act to be adopted?

The trilogues started on 14 June 2023 and are expected to last until 13 December 2023. There are already a lot of approved batches regarding the drafting of the EU AI Act. The general goal of the EU is to adopt the EU AI Act by the end of 2023. Whether this deadline will be met, will fully depend on how quick the possible agreement between the EP and the Council will occur on the contentious matters.

However, we might get to see the final text within the first quarter of 2024 as there are many more steps to be taken after a political agreement is reached between the EP and the Council, such as last votes and official translations to all of the EU's official languages. At this point, the end of 2023 seems to be an optimistic date.

Additionally, a grace period (a waiting period for legislation to come into force in full) is expected to apply to the EU AI Act from the date of publication in the Official Journal of the European Union. The duration of the grace period may vary between one to two years. The EP wants to limit this to one year due to the concerns that many high-risk AI systems will be introduced in the EU before the application date of the EU AI Act. However, the private sector and the Council are in favour of a two-year period to ramp up compliance efforts during this period as a longer grace period applies for compliance projects for providers and deployers of AI systems.

Fundamental Rights Impact Assessments: One of the heaviest responsibilities of deployers?

On 11 May 2023, the EP finalised its version of the EU AI Act to enter the trilogues with the Council and the European Commission (**EP Mandate**). The EP Mandate introduced a novel assessment for deployers of high-risk AI systems that adds on top of the impact assessments foreseen in the General Data Protection Regulation (**GDPR**) and the Digital Services Act.

The EP believes that the deployers, which are the stakeholders that will implement high-risk AI systems in real-life situations, have the best knowledge and ability to evaluate the impact of their high-risk AI systems on fundamental rights. Conducting Fundamental Rights Impact Assessments (**FRIA**) stands out as one of the most significant and relatively burdensome obligations that is imposed on the deployers of high-risk AI systems in the EP Mandate.

FRIAs surely have a legitimate goal. FRIAs are proposed by the EP to mitigate possible harms of AI in relation to individuals, beyond technical certification and risk mitigation methods. However, considering that many organisations may lack the necessary capabilities to first, fully understand the risks of a high-risk AI system and second, to assess their risks in terms of different fundamental rights, FRIAs have been a hot topic during the trilogues.

According to Article 29a of the EP Mandate of the EU AI Act, FRIAs shall be conducted **before** a high-risk AI system is put into use and shall incorporate the following matters:

"(a) a clear outline of the intended *purpose for which the system will be used;*

(b) a clear outline of the intended geographic and temporal scope of the system's use;

(c) categories of natural persons and groups likely to be affected by the use of the system;

(d) verification that the use of the system is compliant with relevant Union and national law on fundamental rights;

(e) the reasonably foreseeable impact on fundamental rights of putting the high-risk AI system into use;

(f) specific risks of harm likely to impact marginalised persons or vulnerable groups;

(g) the reasonably foreseeable adverse impact of the use of the system on the environment;

(h) a detailed plan as to how the harms and the negative impact on fundamental rights identified will be mitigated.

(j) the governance system the deployer will put in place, including human oversight, complaint-handling and redress."

Additionally, deployers are required to **notify** the national supervisory authority that will be tasked with the supervision and enforcement of the EU AI Act, defined by the Member States in the course of conducting an FRIA, and the views of the representatives of the relevant stakeholders such as consumer protection bodies shall be taken into account. Various organisations such as public authorities and as defined by the Digital Markets Act are expressly required to publish their FRIAs.

The EP, on the other hand, proposed an interesting link between the FRIAs and data protection impact assessments (**DPIA**). Article 29(6) requires that when a data controller is required to conduct a DPIA under the GDPR, a summary of this DPIA shall be published, "*having regard to the specific use and the specific context in which the AI system is intended to operate*".

In another proposed paragraph by the EP (Article 29a, para 6), there is also a similar interlink between the DPIAs and the FRIAs:

"Where the deployer is already required to carry out a data protection impact assessment under Article 35 of GDPR or Article 27 of EUI GDPR, the fundamental rights impact assessment referred to in paragraph 1 shall be conducted in conjunction with the data protection impact assessment. The data protection impact assessment shall be published as an addendum."

This reads as if the publication of the FRIAs would also depend on the regulation under the GDPR requiring a DPIA, aside from the status of the deployer, when the deployer is not a gatekeeper or a public authority. These – **not so clear** – links between the GDPR and the EU AI Act are yet to be refined in the definitive version of the EU AI Act.

We can observe the EP's efforts to underline the importance of safeguarding individuals' fundamental rights. However, many critics as well as the Council have stated the extra burden presented by FRIAs could scare companies away from the use of AI systems in the European market, which could stifle **competition and innovation**.

Although a final agreement has not yet been reached on the subject of FRIAs, it is now assured that they will make it to the final text. The Council, in response to the EP version of the FRIA-related provisions, has proposed three options:

- **The first option** is built upon the EP's version, but the FRIAs are restricted to new elements that were not previously dealt with for other chapters of the EU AI Act (which means not all fundamental rights would be under their scope), enabling users to cooperate with the AI providers for information that is missing from their instructions, and obliging the deployers to notify the national supervisory authority.
- **The second option** is wider, demanding deployers to submit information to the market authority via a template, such as an online form.
- **The third option** would streamline the FRIAs and merge them with the other requirements for high-risk AI systems, rather than via a distinct article as the EP recommended.

Yet all three options limit the obligation to conduct an FRIA **only to the public authorities** "on the ground that private companies will have to comply with similar obligations under the upcoming Due Diligence Directive". It is expected that the EP will strongly oppose this.

The EU has built significant know-how on impact assessments thanks to the DPIAs, which can be leveraged in the scope of the FRIAs, also in combination with the know-how of other international bodies such as the United Nations and the Council of Europe.

So, one way or another, FRIAs will be a tool that will make it to the final version of the EU AI Act. The hardships of concretely assessing the effects of complex AI systems on many fundamental rights with different elements at their core will be a challenging and long journey for negotiators to overcome. Further guidance from regulators and European-level bodies will be needed. However, in the end, FRIAs are not a rubber-stamping exercise, and it is in the best interest of every stakeholder to turn them into an effective tool both for boosting safe competition and for protecting fundamental rights.

Generative AI, Foundation Models, High-Impact Foundation Models: What will the EU agree upon in the end?

Since the meteoric entry of general-purpose AI systems into our lives, EU legislators have struggled to create a legislative framework that would meaningfully regulate such systems. Therefore, it is not a surprise that the initial version of the EU AI Act proposed in 2021, did not contain any specific obligations for these systems.

In order to close this gap, the EP has inserted massive obligations for so-called foundation models. Article 3 of the EP Mandate defines foundation models as "a*n AI system model that is trained on broad data at scale, is designed for generality of output, and can be adapted to a wide range of distinctive tasks*". **The EP** has also proposed **a lengthy list of obligations** for the providers of foundation models in Article 28(b). As the newly inserted Recital 79 of the EP Mandate reiterates, foundation models are perceived to be the founding blocks of many more specific AI systems, and they can notably be used in conjunction with high-risk AI systems.

Therefore, Article 28(b) of the EP Mandate requires providers of these systems to adhere to appropriate data governance design methods to mitigate the foreseeable risks to health, safety and fundamental rights. The obligations are set out in six different paragraphs, and they even include environmental protections such as *"making use of applicable standards to reduce energy use, resource use and waste, as well as to increase energy efficiency, and the overall efficiency of the system, without prejudice to relevant existing Union and national law"*. Another attention-grabbing provision requires providers to register the foundation model to an EU database (which would possibly cover all extensive technical, legal and risk mitigation documents). Easier said than done, these obligations were the first of their kind and drew a lot of pushback from AI developers and even within the EU circles.

On the other side of the coin, the Council merely made a single reference to generative AI systems by inserting it into the definition clause. Therefore, it was obvious that generative AI was going to be a key matter in the negotiation meetings in the trilogues.

While the trilogues were ongoing, President Biden issued an "Executive Order Safe, Secure, and Trustworthy Artificial Intelligence" (**Biden EO**). It is expected in Brussels that the Biden EO may affect the final text of the EU AI Act, to achieve a greater level of coherence with the USA on how the approach towards regulating AI systems. Although different to the approach proposed by the EP, the Biden EO also pays special attention to "*dual-use foundation models*" that "*exhibit, or*

could be easily modified to exhibit, high levels of performance at tasks that pose a serious risk to security, national economic security, national public health or safety, or any combination of those matters" including cyber, chemical/biological/radiological/nuclear weapons, and deception/manipulation risks. The Biden EO introduces new measures for AI foundation model developers to evaluate their models for potential risks and vulnerabilities. These measures are termed "red-teaming" and they are especially important for foundation models that have dual-use applications. The developers have to report the results of their "red-teaming" to the government and also show how they have addressed the issues posing the most salient risks and vulnerabilities of those systems.

The latest proposal brought by the Council during the trilogues that is on the table in Brussels is between the obligations imposed upon "*foundation models*", "*high-impact foundation models*" and "*general purpose AI systems*". Although not definitive, the possible definitions and obligations of this novel tiered approach are as such:

- **Foundation Models:** Transparency obligations, up-to-date technical documentation explaining the capacities and limitations of foundation models, compliance with the EU law (e.g., copyright laws, consent of individuals if their personal data is embedded in training data, opt-out for content creators, summary of how the training of the model was conducted).
- **High-Impact Foundation Models:** These systems are defined as the more capable systems trained with enormous amounts of data. However, it is yet to be seen how the criteria for differentiating between standard foundation models will be put on paper. If agreed, the European Commission will define and designate these models within 18 months from the EU AI Act's entry into force, in consultation with the new institution that the EU AI Act will incorporate, the European AI Office. The high-impact foundation models will have to undergo adversarial vetting, audit, incident reporting and risk assessments, among other obligations.
- **General Purpose AI Systems:** General purpose AI systems are systems that may be based on an AI model and they can include additional components such as traditional software and user interfaces to serve a variety of purposes, both for direct use as well as for their integration in other AI systems. If agreed as such, the providers of general-purpose AI systems have certain obligations when they license their systems to downstream operators that might use them for high-risk applications. These obligations include specifying the high-risk uses that are allowed or prohibited, providing technical documentation and relevant information for compliance, and taking measures to detect and prevent misuse.

Lastly, the European AI Office: Lessons from competition and data protection enforcement in the EU

Thanks to the competences of the European Union in the founding treaties, competition law is mainly enforced centrally by the European Commission if certain criteria and thresholds are met, despite the decentralisation model introduced via Regulation 1/2003. Yet, another example of an EU regulatory framework, data protection law is nearly fully enforced by national data protection authorities (apart from certain binding decision mechanisms at the EU level). Although it might be claimed that both enforcement models have their pros and cons, it is undeniable that the more complex a regulatory matter becomes, the harder it is for the Member States to implement and enforce them. Therefore, enforcement of – an overly complex, technical and first-of-its-kind law – the EU AI Act has also been a heated debate in the trilogues.

In fact, none of the institutions sought to make the EU AI Act an EU-enforced legislation. Yet, the EP pushed for a common approach by inserting a new body with new vested powers: the European AI Office. While the initial proposal created a "European AI Board" with a consultative nature, chaired by the European Commission, the EP has proposed to establish a European AI Office as a separate legal entity. Besides giving it consultative and cooperative powers, the EP wants to empower the European AI Office with the monitoring of foundation models in particular. The Council, on the other hand, has proposed a "with one representative from each of the Member States, and the European Data Protection Supervisor as an observer, in its mandated text of the trilogues. While the Council's version of the "European AI Board" carries the same name as the initial proposal of the EU AI Act, all the substantive provisions were rewritten, in a way to further empower the Member States.

In short, although the main enforcement powers will lie with the (it is not yet clear whether these authorities will be a single authority or a group of authorities in a Member State among them, one will take the supervisory role), **the EP still desires to create a strong EU presence in the governance of the EU AI Act** with the involvement of various EU representatives regarding cyber security, fundamental rights, aside from the Member States' representatives.

The Council, on the other hand, has limited the role of its version of the European AI Board to a mere consultation venue between the Member States.

The result here is also yet to be seen. However, it is clear that the EU legislators want to draw lessons from the enforcement model of both EU competition law and EU data protection law. This is clearly reflected in the statement of Drago? Tudorache, a Romanian politician and the co-rapporteur of the EU AI Act, who said that "no Member State alone would be able to properly handle big companies as these are very powerful actors".

Our view is that a high level of coherence between the Member States and a more centralised approach in comparison to the GDPR enforcement would only benefit the EU, as this would ensure a higher level of legal certainty for undertakings. Additionally, fractioning enforcement even further regarding the same legislation can scare away investors and developers of this crucial technology.

The EU AI Act is not a stand-alone piece, and the EU has been in an intense regulatory campaign on technology. Therefore, in order to mitigate the effects of the increasing number of legislative instruments, a united voice in enforcement will make the lives of every stakeholder easier, including the regulators.

Conclusion: History in the making

Although many chapters of the EU AI Act are not definitive yet, one thing is clear: history will record these discussions and future generations will read articles on how the first law on AI performed.

_____

To make sure you do not miss out on regular updates from the Kluwer Competition Law Blog, please subscribe here.
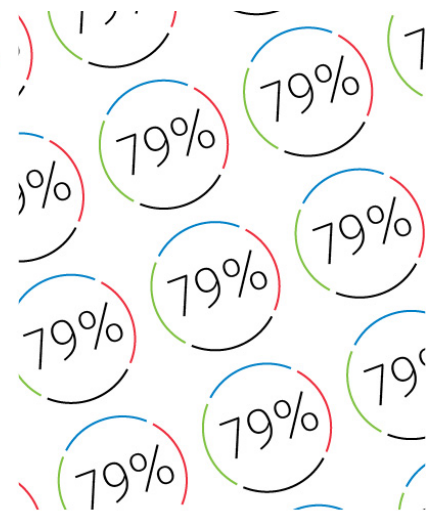
Kluwer Competition Law

The **2022 Future Ready Lawyer survey** showed that 79% of lawyers are coping with increased volume & complexity of information. Kluwer Competition Law enables you to make more informed decisions, more quickly from every preferred location. Are you, as a competition lawyer, ready for the future?

Learn how **Kluwer Competition Law** can support you.

This entry was posted on Monday, November 13th, 2023 at 9:00 am and is filed under Artificial Intelligence, Europe, European Commission, European Union, Regulation
You can follow any responses to this entry through the Comments (RSS) feed. You can leave a response, or trackback from your own site.