Dawn of the EU's AI Act: political agreement reached on world's first comprehensive horizontal AI regulation

Alert

14 December 2023

---

11 min read

Clara Hainsdorf |

Tim Hickman |

Dr. Sylvia Lorenz |

Jenna Rennie

White & Case Tech Newsflash

On Friday, December 8, 2023 – after months of intensive trilogue negotiations – the European Parliament and Council reached political agreement on the European Union's Artificial Intelligence Act ("EU AI Act"). Hailed by European Commission President Ursula von der Leyen as a "global first",[1] this "historic"[2] Act positions the EU as a frontrunner of AI regulation, being the "very first continent to set clear rules for the use of AI".[3] With this landmark piece of legislation, the EU seeks to create a far-reaching and comprehensive legal framework for the regulation of AI systems across the EU – with the aim of ensuring that AI systems are "safe"[4] and "respect fundamental rights and EU values",[5] while looking to encourage AI investment and innovation in Europe. Once the consolidated text is finalized in the coming weeks, the majority of the EU AI Act's provisions will apply two years after its entry into force.[6]

Overall Approach and Key Issues Debated

- The EU AI Act is intended to ensure the safety of AI systems on the EU market and provide legal certainty for investments and innovation in AI, while minimizing associated risks to consumers as well as compliance cost for providers. The EU AI Act prominently features a risk-based approach, defining four different risk classes, each of which covering different use cases of AI systems. While some AI systems are banned entirely, barring narrow exceptions, the EU AI Act imposes specific obligations on the providers and deployers of so-called high-risk AI systems, including testing, documentation, transparency, and notification duties.

- In the marathon trilogue negotiations between the EU Commission, Parliament and Council leading up to last week's political agreement, the list of prohibited and high-risk AI systems, including the classification of and exceptions for biometric identification systems, as well as the enforcement structure and mechanisms of the EU AI Act were amongst the most contentious issues. Furthermore, the regulation of so-called general purpose AI models, like foundation models and generative AI, which was first introduced in the EU Parliament's [negotiating position](#) from June 2023, was fiercely debated in the final stages of the trilogue and deemed particularly controversial due to fears that excessive regulation could hinder innovation and harm European companies.[7]

Scope of Application

- Because the final text of the EU AI Act has not yet been published, details of the recent political consensus remain unknown. This concerns, inter alia, the precise definition of "AI systems", which reportedly aligns with the approach proposed by the ("OECD")[8] and could therefore match the definition stipulated in the EU Parliament's negotiating position, which was narrower than the broad definition set out in the EU Commission's April 2021 [proposal](#). Additionally, details on the addressees of the EU AI Act have not been disclosed yet, though the regulation will likely apply mainly to providers and deployers of AI systems. However, it has been confirmed that the EU AI Act will not apply to AI systems "which are used exclusively for military or defence purposes" or to "AI systems used for the sole purpose of research and innovation".[9]

Risk-based Approach to AI Regulation

- At its core, the EU AI Act will adopt a risk-based approach, classifying AI systems into four different risk categories depending on their use cases: (1) unacceptable-risk, (2) high-risk, (3) limited-risk, (4) minimal/no-risk. The Act's focus will likely lie on unacceptable-risk and high-risk AI systems, with both risk classes having received much attention in the EU Parliament's and Council's amendments and during the trilogue negotiations.

- First, AI systems that create an **unacceptable risk**, contravening EU values and considered to be a clear threat to fundamental rights, will be **banned** in the EU. As per the political agreement, the EU AI Act will prohibit:

- "biometric categorisation systems that use sensitive characteristics (e.g. political, religious, philosophical beliefs, sexual orientation, race);

- untargeted scraping of facial images from the Internet or CCTV footage to create facial recognition databases;

- emotion recognition in the workplace and educational institutions;

- social scoring based on social behaviour or personal characteristics;

- AI systems that manipulate human behaviour to circumvent their free will;

- AI used to exploit the vulnerabilities of people (due to their age, disability, social or economic situation)"; and[10]

- "certain applications of predictive policing".[11]

- While **biometric identification systems** will be banned in principle, an agreement has been reached on narrow **exceptions** for the use of such systems in publicly accessible spaces **for law enforcement purposes**.[12] Accordingly, they may only be used after prior judicial authorization and for the prosecution of a strictly defined list of crimes.[13] Moreover, national data protection authorities will need to be notified when biometric identification systems are being used.[14] Post-remote biometric identification systems shall be deployed solely for the "targeted search of a person convicted or suspected of having committed a serious crime".[15] As for real-time biometric identification systems, they will need to comply with a set of strict conditions and their use will "be limited in time and location, for the purposes of:

- targeted searches of victims (abduction, trafficking, sexual exploitation);

- prevention of a specific and present terrorist threat; or

- the localisation or identification of a person suspected of having committed one of the specific crimes mentioned in the regulation (e.g. terrorism, trafficking, sexual exploitation, murder, kidnapping, rape, armed robbery, participation in a criminal organisation, environmental crime)".[16]

- Second, certain AI systems with "significant potential harm to health, safety, fundamental rights, environment, democracy and the rule of law" will be classified as **high-risk**, including:

- "certain critical infrastructures for instance in the fields of water, gas and electricity";

- "medical devices";

- "systems to determine access to educational institutions or for recruiting people";

- "certain systems used in the fields of law enforcement, border control, administration of justice and democratic processes"; and

- "biometric identification, categorisation and emotion recognition systems".[17]

- For AI Systems classified as high-risk AI, **comprehensive mandatory compliance obligations** with respect to, inter alia, risk mitigation, data

governance, detailed documentation, human oversight, transparency, robustness, accuracy, and cybersecurity will apply.[18] High-risk AI systems will also be subject to conformity assessments to evaluate their compliance with the Act, with an emergency procedure "allowing law enforcement agencies to deploy a high-risk AI tool that has not passed the conformity assessment procedure in case of urgency".[19] Furthermore, mandatory **fundamental rights impact assessments** and a right of citizens "to launch complaints about AI systems and receive explanations about decisions based on high-risk AI systems that impact their rights" are also part of the political agreement.[20] So-called regulatory sandboxes and real-world-testing will allow for the development and training of AI systems before they are placed on the market.[21]

- Third, AI systems classified as **limited-risk**, including chatbots, and certain emotion recognition and biometric categorization systems, as well as systems generating deepfakes, will be subject to **more minimal transparency obligations**.[22] The transparency requirements will include, inter alia, informing users that they are interacting with an AI system and marking synthetic audio, video, text and images content as artificially generated or manipulated for users and in a machine-readable format.[23]

- Finally, all other AI systems not falling under one of the three main risk classes, such as AI-enabled recommender systems or spam filters, are classified as **minimal/no-risk**. The EU AI Act allows the **free use** of minimal-risk AI systems, while voluntary codes of conduct are encouraged.[24]

Safeguards for General-Purpose AI Models

- After prolonged debates on the regulation of so-called "foundation models" in the EU AI Act, the EU Parliament and Council reportedly now reached a compromise in the form of an amended tiered approach. The compromise refers to the term of general-purpose AI ("**GPAI**") systems/models and distinguishes between obligations on two tiers: (1) a number of horizontal obligations that apply to all GPAI models, and (2) a set of additional obligations for GPAI models with systemic risk.

- With respect to the first tier, all GPAI model providers will have to adhere to transparency requirements by, inter alia, drawing up technical documentation.[25] They will also need to comply with EU copyright law and provide detailed summaries about the content used for training.[26] GPAI models of the lower tier will be exempt from the transparency requirements while they are in the R&D phase or if they are open source.[27]

- With respect to the **second tier**, GPAI models are set to be designated as encompassing **systemic risk** when meeting certain criteria.[28] GPAI models that have been classified as systemic risk will be subject to more stringent obligations, which include "conduct[ing] model evaluations, assess[ing] and mitigat[ing] systemic risks, conduct[ing] adversarial testing, report[ing] to

the Commission on serious incidents, ensur[ing] cybersecurity and report[ing] on their energy efficiency".[29] GPAI models with systemic risk may comply with the EU AI Act by adhering to codes of practice, at least until harmonized EU standards are published.[30] The codes of practices will be "developed by industry, the scientific community, civil society and other stakeholders together with the Commission".[31] Furthermore, a scientific panel of independent experts will issue alerts on systemic risks and support the classification and testing of the models.[32]

Enforcement Framework and Penalties

- It is anticipated that the EU AI Act will be primarily enforced through national competent market surveillance authorities in each Member State.[33] Additionally, a European AI Office – a new body within the EU Commission – will take up various administrative, standard setting and enforcement tasks, including with respect to the new rules on GPAI models, to ensure coordination at European levels.[34] The European AI Board, comprised of member states' representatives, will be kept as a coordination platform and to advice the Commission.[35]

- Fines for violations of the EU AI Act will depend on the type of AI system, size of company and severity of infringement and will range from:

    o 7.5 million euros or 1.5% of a company's total worldwide annual turnover (whichever is higher) for the supply of incorrect information;[36] to

    o 15 million euros or 3% of a company's total worldwide annual turnover (whichever is higher) for violations of the EU AI Act's obligations;[37] to

    o 35 million euros or 7% of a company's total worldwide annual turnover (whichever is higher) for violations of the banned AI applications.[38]

    o Notably, one outcome of the trilogue negotiations is that the EU AI Act will now provide for more proportionate caps on administrative fines for smaller companies and startups.[39] Furthermore, the EU AI Act will allow natural or legal persons to report instances of non-compliance to the relevant market surveillance authority.[40]

What's Next?

With political agreement reached, the EU AI Act will shortly be officially adopted by the EU Parliament and Council and published in the EU's Official Journal to enter into force. The majority of the Act's provisions will apply after a two-year grace period for compliance.[41] However, the regulation's prohibitions will already apply after six months

and the obligations for GPAI models will become effective after 12 months.[42] By joining to the [AI Pact](), which will be launched by the European Commission, AI developers can commit to implementing key provisions of the EU AI Act voluntarily prior to the respective deadlines.[43] During the grace period, much work will need to be done at both Member State and Union levels to establish effective oversight structures and publish guidance on the implementation of the EU AI Act.

With the imminent entry into force of the landmark EU AI Act, the EU seeks to position itself at the forefront of responsible AI development and to ensure that governance keeps pace with innovation in this rapidly evolving sector. Given the stated aim of the EU AI Act in ensuring that AI systems in the EU are "safe, transparent, traceable, non-discriminatory and environmentally friendly",[44] the efficacy of the EU AI Act will no doubt be compared to and measured against approaches adopted in other leading AI nations such as the UK and the US, and international efforts to set out guardrails for AI such as at the G7, G20, OECD, Council of Europe, and the UN.

1 X, Ursula von der Leyen (@vonderleyen), "The EU AI Act is a global first […]", 8 December 2023.

2 X, Thierry Breton (@ThierryBreton), "Historic! The EU becomes the very first continent to set clear rules for the use of AI […]", 8 December 2023.

3 ibid.

4 European Council, Press Release, "Artificial intelligence act: Council and Parliament strike a deal on the first rules for AI in the world", 9 December 2023 ("**European Council Press Release**").

5 ibid.

6 ibid.

7 Euractiv, "EU's AI Act negotiations hit the brakes over foundation models", updated 15 November 2023.

8 European Council Press Release, supra note 4.

9 ibid.

10 European Parliament, Press Release, "Artificial Intelligence Act: deal on comprehensive rules for trustworthy AI", 9 December 2023 ("**European Parliament Press Release**").

11 European Commission, Press Release, "Commission welcomes political agreement on Artificial Intelligence Act", 9 December 2023 ("**European Commission Press Release**").

12 European Parliament Press Release, supra note 10.

13 ibid.

14 Press Conference on the Political Agreement, "Part 7 - Q&A", 9 December 2023.

15 European Parliament Press Release, supra note 10.

16 ibid.

17 European Commission Press Release, supra note 11.

18 ibid.

19 European Council Press Release, supra note 4.

20 European Parliament Press Release, supra note 10.

21 ibid; European Commission Press Release, supra note 11.

22 European Council Press Release, supra note 4.

23 ibid; European Commission Press Release, supra note 11.

24 European Commission Press Release, supra note 11.

25 European Parliament Press Release, supra note 10.

26 ibid.

27 Press Conference on the Political Agreement, "Part 8 - Q&A", 9 December 2023.

28 Press Conference on the Political Agreement, "Part 6 - Q&A", 9 December 2023.

29 European Parliament Press Release, supra note 10.

30 ibid.

31 European Commission Press Release, supra note 11.

32 ibid.

33 ibid.

34 ibid.

35 European Council Press Release, supra note 4.

36 ibid.

37 ibid.

38 ibid.

39 ibid.

40 ibid.

41 European Commission Press Release, supra note 11.

42 ibid.

43 ibid.

44 European Parliament, "EU AI Act: first regulation on artificial intelligence", 14 June 2023.