GIBSON DUNN

The EU Agrees on a Path Forward for the AI Act

Client Alert | December 14, 2023

The agreement marks a watershed moment for AI regulation and is the most significant endorsement of so-called "comprehensive" AI regulation from a major political actor on the world stage. **I. Introduction** After almost 6 months of negotiations, in the late night and early morning hours of December 8-9, 2023, the European Commission, the Council and the Parliament reached political agreement on the provisional rules that will comprise the European Union's Artificial Intelligence Act (the "**AI Act**").[1] This agreement marks a watershed moment for AI regulation, and is the most significant endorsement of so-called "comprehensive" AI regulation from a major political actor on the world stage. The provisional agreement on the AI Act would:

- **Establish a broad and extraterritorial scope of application**,
- **Prohibit certain uses of AI entirely,** and
- Define a broad range of other uses as **"high-risk" and subject to stringent requirements**.

A number of procedural steps remain before the AI Act is finalized (as summarized in more detail in Section III); however, the staggered – and relatively rapid – planned enforcement of certain provisions bears note. Provisions related to prohibited AI systems are set to become enforceable six months after the Act is finalized; and provisions related to so-called General Purpose AI ("**GPAI**") become enforceable 12 months after this date. The rest of the AI Act is expected to become enforceable in 2026. The "long arm" of the AI Act will impact a broad range of business – including, but not limited to, those that intend to provide or deploy AI systems within the EU.[2] The distinct posture of the AI Act – based in part on fundamental and human rights jurisprudence – requires companies to think differently when preparing compliance strategies:

- Proactive engagement with **novel regulatory measures**, such as a fundamental rights impact assessment for certain AI systems, and
- Reimagining and documenting strategic decision-making related to **internal governance and compliance** in the face of unpredictable and uncertain go-forward risks.

This alert builds on our previous alerts which analyzed the European Commission's 2021 proposal on the AI Act, the European Council's common position (December 2022) and European Parliament's negotiating position (June 2023). This alert takes a closer look at some of the key areas of debate in the trilogue procedure and the political agreement that was reached. Negotiations on the final text of the AI Act remain underway, and the final wording of the provisional agreement is not yet public. The analysis below is based on public reports and our understanding of the substance of the agreement, but this may change based on a variety of factors. Keen followers of artificial intelligence would no doubt agree a lack of forward-looking certainty is the one guarantee in this area, and that applies equally to the AI Act's journey from initial gestation to "political agreement." **II. From negotiation to agreement: Outcome of the trilogue procedure**

a) Scope of Application

Related People

Vivek Mohan

Robert Spano

Kai Gesing

Joel Harrison

Christian Riis-Madsen

Stéphane Frank

Frances Waldmann

Christoph Jacob

Jonas L. Jousma

Yannick Oberacker

Ciara O'Gara

Tine M. Rasmussen

The AI Act will provide for an extra-territorial scope of application that includes obligations for providers and deployers of AI systems. This has significant consequences for businesses not established in the EU, if they place an AI system on the market or put it into service in the EU. Furthermore, the regulations apply when "outputs" produced by an AI system are (or are intended to be) used in the EU. A key issue yet to be determined relates to the scope of the "output" provisions and the extent to which the AI Act regulates activities across the AI supply chain. One example could be AI-assisted R&D that is outsourced outside Europe such that only the final product (e.g., products designed using AI that do not themselves fall within the definition of an AI system and are not, technically, "outputs" of an AI system) are marketed in the EU.

b) Definition of AI

The definition of AI, a key threshold for applicability of the AI Act, has been subject to significant change throughout the legislative process. As noted in our previous alert, what started out as an overly broad definition has been reportedly narrowed over time with the goal of ensuring that traditional computational processes and software are not inadvertently captured. While the final text of the AI Act has not yet been released, the language reportedly aligns with the definition of AI recently adopted by the OECD[3] and reflects the need to preserve the ability to adjust the definition as necessary to account for future developments in the fast-moving AI landscape. While the definition in the final text does not reflect the OECD's definition verbatim, it reflects its main elements, i.e., objective-based output generation that is able to influence its environment with varying degrees of autonomy.[4]

c) Prohibited AI systems

The AI Act classifies AI systems by risk level (unacceptable, high, limited, and minimal or no risk). AI systems that carry "unacceptable risk" are *per se* prohibited. In other words, the Act adopts bright line rules as to some uses of AI systems, based on fundamental rights concerns, differently from some other flagship regulations in the EU which usually allow for exceptions based on general rules. This framework will require particular diligence on behalf of businesses when classifying their AI systems for the purposes of the Act's risk-based approach. While the European Commission and Council advocated for a narrower list of prohibited AI systems, the Parliament's mandate included a longer list of prohibited AI systems, banning certain use cases entirely. The agreed AI Act prohibits, *inter alia*:[5]

- Manipulation of human behavior to circumvent end-users' free will;
- Social scoring;
- Certain applications of predictive policing;
- Emotion recognition systems used in the workplace; and
- Real time remote biometric identification for law enforcement purposes in publicly accessible spaces (with narrow exceptions).

The European Council resisted the full ban on real-time remote biometric identification in publicly accessible spaces proposed by the European Parliament. Instead, real-time remote biometric identification for law enforcement purposes is prohibited unless it fits within narrow exceptions, such as for uses tied to the prevention of terrorist attacks or to locate victims or suspects in connection with serious offences, such as terrorism. Under the agreed version of the AI Act, other forms of biometric identification that fall outside the scope of the prohibition (i.e., *ex-post* biometric identification) are considered "High-Risk" (for which, see below).

d) High-risk AI systems

High-risk AI systems, while permitted, are subject to the most stringent obligations. AI

systems may fall into this category if they pose a "significant risk" to an individual's health, safety, or fundamental rights, and are used, or intended to be used, for *inter alia* education, employment, critical infrastructure, public services, law enforcement, border control, and administration of justice. One new obligation for companies imposed by the AI Act will be to prepare fundamental rights impact assessments ('**FRIAs'**). Determining when and how to conduct an FRIA will raise many strategic and compliance-focused questions, taking account of the nature and scope of the AI system in question. Businesses will need to take steps to engage with this obligation as soon as possible. In response to concerns that the high-risk category may be over-inclusive, the agreed version of the AI Act adds the concept of a filter system, which provides a series of exemptions that would allow providers of AI systems to avoid the high-risk category based on self-assessments. The European Commission is expected to develop guidelines on the application of these filters. Currently, it has been reported that the filters exempt AI systems that are (i) intended to perform a narrow procedural task; (ii) intended to review or improve the result of a previously completed human activity; (iii) purely intended to detect decision-making patterns or deviations from prior decision-making patterns to flag potential inconsistencies; or (iv) used to perform preparatory tasks for an assessment relevant to critical use cases. It is not unlikely that the operation of this filter system may cause some problems in practice and lack of legal certainty as to the scope of the exemptions. Providers that make an assessment that their systems fall outside of the high-risk category may be obliged to provide, upon request, the results of their prior assessment as to classification to the respective national market surveillance authorities. The agreed AI Act also contemplates allowing these market surveillance authorities to carry out evaluations of AI systems where they have sufficient reason to believe they should be considered high-risk, and to impose fines where they have sufficient evidence that the AI system provider misclassified their system to avoid a high-risk classification.[6]

e) General Purpose AI (Foundation Models)

As explained in our previous alert, the Parliament's negotiating position introduced a regime for regulating GPAI models – or foundation models – consisting of models that "*are trained on broad data at scale, are designed for generality of output, and can be adapted to a wide range of distinctive tasks*".[7] Parliament also introduced certain separate testing and transparency requirements, with most of the obligations falling on any deployer that substantially modifies a GPAI system for a specific use case. The regulation of GPAI models was heavily debated during the trilogue procedure. The final text features a tiered approach (initially proposed by the European Commission) which places more onerous obligations on GPAI models that could pose "systemic" risks. The AI Act contains a presumption categorizing models as carrying a "systemic risk" where the model was trained using computing power greater than 10^25 floating point operations (FLOPs), indicating their capabilities and amount of underlying data. The European Commission has commented that "*these new obligations will be operationalized through codes of practices developed by industry, the scientific community, civil society and other stakeholders together with the Commission*."[8] The tiered approach represents a compromise between the stark opposition to *any* regulation of foundation models in the AI Act by some Member States, including France, Germany and Italy, and the European Parliament's preferred approach of establishing horizontal obligations which would apply equally to all foundation models.[9] While certain obligations will apply to GPAI models, e.g., transparency obligations relating to the training data as well as copyright safeguards or making AI-generated content recognizable, providers of foundation models that meet the "systemic risks" threshold will need to notify the Commission of their status and comply with the relevant obligations in the AI Act (similar to the notification mechanism in the EU Digital Services Act). The regime places onerous obligations on providers of foundation models that are similar to those that apply to high-risk AI systems, e.g., requirements to assess and mitigate the risks their models entail, comply with certain design, information and environmental requirements and register such models in an EU database.

f) Enforcement

The AI Act envisions a strict enforcement regime set to be overseen by national authorities designated by each EU Member State to supervise compliance within its territory as well as a centralized European Artificial Intelligence Office tasked with coordinating enforcement efforts. Notably, the AI Act does not contemplate a de-centralized system of enforcement or a "one-stop shop" as under the GDPR. However, the competent national authorities will be gathered in the European Artificial Intelligence Board to ensure consistent application of the law. The maximum fines for non-compliance with the AI Act can reach up to EUR 35 million or, if the offender is a company, up to 7% of its total worldwide annual turnover for the preceding financial year, whichever is higher.[10] Certain proportionate caps on fines will apply for small and medium-sized enterprises (SMEs) and start-ups. **III. Next Steps** The AI Act is an extensive and complicated piece of legislation, and its impact will be far-reaching. After the conclusion of the trilogue process, the AI Act is now subject to formal adoption by the European Parliament and the Council. Once adopted, the AI Act will be published in the Official Journal and will enter into force 20 days following publication. However, the AI Act will not become fully enforceable until two years *after* its entry into force—likely in 2026—with some exceptions for specific provisions such as prohibited AI systems and AI systems classified as GPAI, which will be applicable after 6 and 12 months, respectively. Now that political agreement has been reached, companies should be proactively engaging with the provisions of the AI Act and making preparations to ensure compliance by the applicable deadlines. _____ [1] Council of the EU, *Artificial intelligence act: Council and Parliament strike a deal on the first rules for AI in the world*, press release of 9 December 2023, available here. [2] The scope of some of the broadest jurisdictional hooks, including governing companies that are responsible for generating output from AI tools that have effect in the Union, remains to be seen. [3] See, Luca Bertuzzi, Euractiv, *OECD updates definition of Artificial Intelligence 'to inform EU's AI Act'*, Article of 9 November 2023, available here. [4] See, Luca Bertuzzi, Euractiv, *AI Act: EU policymakers nail down rules on AI models, butt heads on law enforcement,* Article of 9 December 2023, available here. [5] European Commission, *Commission welcomes political agreement on Artificial Intelligence Act*, Article of 9 December 2023, available here. [6] See, Luca Bertuzzi, Euractiv, *AI Act: Leading MEPs revise high-risk classification, ignoring negative legal opinion*, Article of 10 November 2023, available here. [7] European Parliament's compromise proposal, Art. 3(1c), available here. [8] European Commission, *Commission welcomes political agreement on Artificial Intelligence Act*, Article of 9 December 2023, available here. [9] See, Luca Bertuzzi, Euractiv, *EU's AI Act negotiations hit the brakes over foundation models*, Article of 10 November 2023, available here. [10] European Commission, *Commission welcomes political agreement on Artificial Intelligence Act*, Article of 9 December 2023, available here.

The following Gibson Dunn attorneys assisted in preparing this update: Vivek Mohan, Robert Spano, Kai Gesing, Joel Harrison, Christian Riis-Madsen, Nicholas Banasevic, Stéphane Frank, Frances Waldmann, Leon Freyermuth, Christoph Jacob, Jonas Jousma, Yannick Oberacker, Ciara O'Gara, Tine Rasmussen, and Hayley Smith.

Gibson, Dunn & Crutcher's lawyers are available to assist in addressing any questions you may have regarding these issues. Please contact the Gibson Dunn lawyer with whom you usually work, any of the leaders and members of the firm's Artificial Intelligence or Privacy, Cybersecurity & Data Innovation practice groups, or the following authors: Stéphane Frank – Brussels (+32 2 554 72 07, sfrank@gibsondunn.com) Kai Gesing – Munich (+49 89 189 33 180, kgesing@gibsondunn.com) Joel Harrison – London (+44 20 7071 4289, jharrison@gibsondunn.com) Vivek Mohan – Palo Alto (+1 650.849.5345, vmohan@gibsondunn.com) Christian Riis-Madsen – Brussels (+32 2 554 72 05, criis@gibsondunn.com) Robert Spano – London/Paris (+44 20 7071 4902, rspano@gibsondunn.com) Frances A. Waldmann – Los Angeles (+1 213.229.7914, fwaldmann@gibsondunn.com) **Artificial Intelligence:** Cassandra L. Gaedt-Sheckter – Co-Chair, Palo Alto (+1 650.849.5203, cgaedt-sheckter@gibsondunn.com) Vivek Mohan – Co-Chair, Palo Alto (+1 650.849.5345, vmohan@gibsondunn.com) Robert Spano – Co-Chair, London/Paris (+44 20 7071 4902, rspano@gibsondunn.com) Eric D. Vandevelde – Co-Chair, Los Angeles (+1 213.229.7186, evandevelde@gibsondunn.com) **Privacy,**

Cybersecurity and Data Innovation: Ahmed Baladi – Co-Chair, Paris (+33 (0) 1 56 43 13 00, abaladi@gibsondunn.com) S. Ashlie Beringer – Co-Chair, Palo Alto (+1 650.849.5327, aberinger@gibsondunn.com) Jane C. Horvath – Co-Chair, Washington, D.C. (+1 202.955.8505, jhorvath@gibsondunn.com) Alexander H. Southwell – Co-Chair, New York (+1 212.351.3981, asouthwell@gibsondunn.com)

Related Capabilities

Artificial Intelligence

Privacy, Cybersecurity, and Data Innovation