Volume 7, No. 4 | July–August 2024

RAIL

The Journal of Robotics, Artificial Intelligence & Law

fastcase FULL COURT PRESS

RAIL
The Journal of Robotics, Artificial Intelligence & Law
Volume 7, No. 4 | July–August 2024

THE JOURNAL OF ROBOTICS, ARTIFICIAL INTELLIGENCE & LAW (ISSN 2575-5633 (print) /ISSN 2575-5617 (online) at $495.00 annually is published six times per year by Full Court Press, a Fastcase, Inc., imprint. For customer support, please contact Fastcase, Inc., 729 15th Street, NW, Suite 500, Washington, D.C. 20005, 202.999.4777 (phone), or email customer service at support@fastcase.com.

Publishing Staff
Publisher: Leanne Battle
Production Editor: Sharon D. Ray
Cover Art Design: Juan Bustamante

Cite this publication as:

The Journal of Robotics, Artificial Intelligence & Law (Fastcase)

A Full Court Press, Fastcase, Inc., Publication

Editorial Office

729 15th Street, NW, Suite 500, Washington, D.C. 20005
https://www.fastcase.com/

POSTMASTER: Send address changes to THE JOURNAL OF ROBOTICS, ARTIFICIAL INTELLIGENCE & LAW, 729 15th Street, NW, Suite 500, Washington, D.C. 20005.

Articles and Submissions

Direct editorial inquiries and send material for publication to:

Steven A. Meyerowitz, Editor-in-Chief, Meyerowitz Communications Inc., 26910 Grand Central Parkway, #18R, Floral Park, NY 11005, smeyerowitz@meyerowitzcommunications.com, 631.291.5541.

Material for publication is welcomed—articles, decisions, or other items of interest to attorneys and law firms, in-house counsel, corporate compliance officers, government agencies and their counsel, senior business executives, scientists, engineers, and anyone interested in the law governing artificial intelligence and robotics. This publication is designed to be accurate and authoritative, but neither the publisher nor the authors are rendering legal, accounting, or other professional services in this publication. If legal or other expert advice is desired, retain the services of an appropriate professional. The articles and columns reflect only the present considerations and views of the authors and do not necessarily reflect those of the firms or organizations with which they are affiliated, any of the former or present clients of the authors or their firms or organizations, or the editors or publisher.

QUESTIONS ABOUT THIS PUBLICATION?

For questions about the Editorial Content appearing in these volumes or reprint permission, please contact:

Leanne Battle, Publisher, Full Court Press at leanne.battle@vlex.com or at 202.999.4777

For questions or Sales and Customer Service:

Customer Service
Available 8 a.m.–8 p.m. Eastern Time
866.773.2782 (phone)
support@fastcase.com (email)

Sales
202.999.4777 (phone)
sales@fastcase.com (email)

ISSN 2575-5633 (print)
ISSN 2575-5617 (online)

Landmark Law on Artificial Intelligence Is Approved by the European Parliament

Charlotte E. Walker-Osborn, Christiane Stuetzle, Lokke Moerel, Marijn Storm, and Stephan Kreß*

In this article, the authors discuss in detail the EU's new Artificial Intelligence Act.

The highly anticipated EU Artificial Intelligence Act is finally here! With extraterritorial reach and wide-reaching ramifications for providers, deployers, and users of artificial intelligence (AI), the Artificial Intelligence Act was finally approved by the European Parliament (EP) on March 13, 2024. The text of the approved version is based on the political agreement that the EP reached with the Council of the European Union in December 2023. Members of the EP passed the law with 523 votes in favor, 46 against, and 49 abstentions. The Act aims to safeguard the use of AI systems within the European Union as well as prohibiting certain AI outright.

The AI Act applies to:

- Providers placing AI systems or models on the market in the European Union or putting into service AI systems or placing on the market general purpose AI (GPAI) models in the European Union, irrespective of whether those providers are located within or outside the European Union;
- Deployers of AI systems that have their place of establishment in or who are located within the European Union;
- Providers and deployers of AI systems that have their place of establishment or who are located in a third country in situations where the output produced by the AI system is used in the European Union;
- Importers and distributors of AI systems into or within the European Union;
- Product manufacturers who place an AI system on the market or put into service an AI system within the European

Robotics, Artificial Intelligence & Law / July–August 2024, Vol. 7, No. 4, pp. 259–274.

ISSN 2575-5633 (print) /ISSN 2575-5617 (online)

Union together with their product and under their own name or trademark;
- Authorized representatives of AI systems where such providers are not established in the European Union; and
- Affected persons or citizens located in the European Union.

Following a final linguistic check, the Act requires formal endorsement by the European Council, as such, it is expected to be finally adopted before the end of the EP's legislature in June 2024.

The AI Act will enter into force 20 days after its publication in the *Official Journal of the European Union*. It will be fully applicable 24 months after its entry into force. However, certain provisions will come into force and need to be complied with sooner.

Timeline and Transition Period

After the Council of the European Union formally endorses the AI Act, it will be published in the *Official Journal* and enter into force 20 days later. The AI Act provides various transition periods for specific requirements, including:

- *Prohibited AI practices*: The transition period for prohibited AI practices, including certain uses of GPAI systems, is six months;
- *High-risk AI practices*: The transition period for the requirements for high-risk AI systems is 24 months; and
- *General purpose AI*: The transition period for the requirements on GPAI models (as defined below) and GPAI models that pose systemic risk (GPAI-SR) is 12 months.

No fines will be imposed for any violation of the GPAI requirements for a further 12 months, creating a de facto additional grace period.

Preexisting AI Systems

After the transition periods have passed, the AI Act will also apply to AI systems that are already available on the EU market if, after this transition period, a substantial modification is made to the AI system. The AI Act will apply to preexisting GPAI models

after 36 months, regardless of whether they are subject to substantial modifications.

Scope

Definition of AI Systems

The AI Act applies to AI systems. An "AI system" is defined in the text of the Act as "a machine-based system designed to operate with varying levels of autonomy, that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments." The term aligns both with the updated Organisation for Economic Co-operation and Development definition of AI systems issued in 2023 as well as the definition set out by the Biden administration in its Executive Order 14110 on Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence published in October 2023.

Geographic Scope and Scope of AI Systems Caught by the AI Act

The AI Act has extraterritorial scope. This means that organizations outside the European Union will have to comply with the law in certain specified circumstances as well as those within the European Union. The Act applies to providers, deployers, and users of AI systems.

The AI Act applies to:

- Providers placing AI systems or models on the market in the European Union or putting into service AI systems or placing on the market GPAI models in the European Union, irrespective of whether those providers are located within or outside the European Union;
- Deployers of AI systems that have their place of establishment in or are located within the European Union;
- Providers and deployers of AI systems that have their place of establishment or are located in a third country in

situations where the output produced by the AI system is used in the European Union;
- Importers and distributors of AI systems into or within the European Union;
- Product manufacturers who place an AI system on the market or put it into service an AI system within the European Union together with their product and under their own name or trademark;
- Authorized representatives of AI systems where such providers are not established in the European Union; and
- Affected persons or citizens located in the European Union.

It is therefore extremely wide-reaching. It is noteworthy that the AI Act applies to the outputs of AI systems used within the European Union, even if the AI providers or deployers are themselves not located in the European Union.

Tiered Approach to Regulating AI Systems

The AI Act will prohibit certain AI systems in the European Union. It also sets out various categories or tiers of AI systems that each carry distinct levels of obligations as well as potential fines for noncompliance.

Prohibited AI Practices

Certain AI practices that are deemed to pose an unacceptable risk to individuals' rights will be banned. These prohibited AI practices include:

- Using AI to exploit the vulnerabilities of individuals;
- Using AI to manipulate individuals using subliminal techniques;
- Social scoring (with limited exceptions);
- Predicting the likelihood of an individual committing a criminal offense based solely on profiling of their personality traits and characteristics;
- The use of biometric identification systems in publicly accessible spaces for law enforcement (with limited exceptions);

- The use of emotion recognition within the workplace and educational institutions; and
- Untargeted scraping of facial images from the internet or closed-circuit television footage to create facial recognition databases.

The last of these prohibitions, in particular, may have wide-reaching impacts for existing trained models that have incorporated these practices already as well as for the necessary engineering approach going forward.

High-Risk AI Systems

The AI Act places several detailed obligations on what it categorizes as "high-risk AI." Examples of high-risk AI uses include use of AI systems in critical infrastructure, education and vocational training, employment, essential private and public services (such as health care and banking), certain systems in law enforcement, migration and border management, justice, and democratic processes (for example, influencing elections).

For high-risk AI systems, organizations must assess and reduce risks, maintain use logs, be transparent (see more on transparency below) and accurate, and ensure human oversight. Individual citizens will have a right to submit complaints to the relevant market surveillance authority and to receive explanations about decisions based on high-risk AI systems that affect their rights.

Provisions Relating Specifically to GPAI

The final text of the AI Act includes a new regime for providers of GPAI models. The AI Act defines a GPAI model as "an AI model, including where such an AI model is trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable of competently performing a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications, except AI models that are used for research, development or prototyping activities before they are released on the market."

GPAI models will be subject to its own risk-based, two-tier approach, with a set of requirements that apply to all GPAI models and more stringent requirements applicable only to GPAI-SR. Separate requirements apply to GPAI systems (i.e., AI systems based on GPAI models). GPAI systems can qualify as high-risk AI systems, if they can be used directly for at least one purpose that is classified as high risk.

The providers of all GPAI models must:

- Create technical documentation of the GPAI model and provide this to the AI Office or competent local supervisory authority upon request;
- Create documentation for third parties who use the GPAI model to create their own AI systems;
- Implement a policy to respect EU copyright law and specifically text and data mining (TDM) opt-outs (see application of copyright law to AI systems and rights to opt out below); and
- Make available a summary about the content used to train the GPAI model.

GPAI with Systemic Risk

Providers of GPAI-SR will be subject to additional requirements. GPAI models pose a systemic risk if, for example, they have high-impact capabilities in the sense that the cumulative amount of compute used for their training measured in FLOPS is greater than 10^25. This would be a lower threshold than the 10^26 FLOPS threshold for the reporting obligation under the U.S. Executive Order on AI.

In addition to the requirements outlined above, providers of GPAI-SR models must:

- Perform model evaluation to identify and mitigate systemic risk (e.g., negative effects on democratic process and dissemination of illegal, false, or discriminatory content), and continuously assess and mitigate such systemic risk;
- Assess and mitigate possible systemic risks at an EU level, which may stem from the availability and/or use of the GPAI model that poses systemic risk;

- Report serious incidents to the AI Office and the competent local supervisory authority; and
- Ensure an adequate level of cybersecurity for the GPAI model that poses systemic risk and physical infrastructure.

Regulation of GPAI Models

Providers of both GPAI and GPAI-SR models may rely on a code of practice to demonstrate compliance with the AI Act requirements, until a harmonized standard is published. Providers of GPAI should furthermore be able to demonstrate compliance using alternative adequate means, if codes of practice or harmonized standards are not available, or they choose not to rely on those.

The AI Office will facilitate the drawing up of a code of practice and will invite providers of GPAI models, competent national authorities, and other relevant stakeholders (e.g., academia, civil society organizations, industry groups) to participate. Providers of GPAI models may also draft their own code of practice. Completed codes of practice will have to be presented to the AI Office and AI Board for assessment, and to the European Commission (EC) for approval. The EC can decide to give any code of practice general validity within the European Union, such as allowing providers of GPAI models to rely on a code of practice prepared by another provider of a GPAI model.

If no code of practice has been completed and approved when the GPAI requirements become effective (i.e., 12 months after the GPAI chapter of the AI Act becomes effective, expected to be around the end of Q2 2025), the EC may adopt common rules for the implementation of the GPAI and GPAI-SR obligations by adopting an implementing act.

Deepfakes and Chatbots

Crucially, the AI Act specifically requires that (save for certain public interest exemptions) artificial or manipulated images, audio, or video content (deepfakes) need to be clearly labeled as such. This is particularly important in a year when so many elections are taking place given the potential influencing power of such deepfakes. Similarly, when AI is used to interact with individuals

(e.g., via a chatbot), it must be clear to the individual that they are communicating with an AI system.

Application of Copyright Law to AI Systems and Rights to Opt Out

The AI Act obliges GPAI providers to implement a policy to respect EU copyright law. Copyright law applies to the field of AI, both to use of copyrighted works for training purposes and the potential infringing outputs of the GPAI models.

The AI Act includes a training data transparency obligation, which initially related to copyrighted training data but covers all types of training data in the final version. Providers of GPAI models have to make publicly available a sufficiently detailed summary of the content used for training, which should be comprehensive enough to facilitate parties with legitimate interests, including copyright holders, to exercise and enforce their rights under EU law, but also take into account the need to protect trade secrets and confidential business information. The AI Office is due to provide a summary template that will give more insight here as to what will be expected.

For use of copyrighted works for training purposes, the AI Act explicitly mentions that GPAI providers must observe opt-outs made by rights holders under the TDM exception of Article 4(3) of Directive (EU) 2019/790.[1] This exception entails that where an opt-out has been effectively declared in a machine-readable form by an organization, the content may not be retrieved for AI training. This provision clarifies that the TDM exception also applies to the use for training GPAI, but it leaves a variety of practical uncertainties open (e.g., technical standards for opt-out, scraping of content from websites where the rights holder is unable to place an opt-out, declaring an opt-out after the AI has already been trained with the data, and evidentiary challenges when enforcing rights in training data). The final text does set an expectation for providers of GPAI to use of "state-of-the-art technologies" to respect opt-outs. It is noteworthy that a recital underlines that any provider placing a GPAI model on the EU market must be copyright compliant in the meaning of this provision, even indicating that AI training conducted outside the European Union must observe TDM opt-outs.

As to potential copyright issues relating to the output of the AI models, the AI Act does not provide clarifications as to the copyright position. It should be noted that there are already a number of litigations in play regarding this area both in Europe and beyond. Therefore, many follow-up questions remain outstanding, such as whether prompts likely to cause infringing outputs should be blocked from processing, how to reliably assess AI output under copyright law (e.g., as a parody or pastiche), the allocation of liability between provider and user, notice and takedown procedures, and so on.

The bottom line remains that the existing copyright frameworks within the European Union and the accompanying technical side do not yet have a tailor-made response to copyright issues related to training data and the impact on the usability of the respective AI system. Over time, courts or private actors may shape solutions both within the European Union and globally.

Limited Exemptions for Free and Open-Source AI Models

The AI Act contains exceptions for free and open-source AI models. The requirements of the AI Act do not apply to AI systems released under free and open-source licenses except:

- If they are placed on the market or put into service as high-risk AI systems,
- If they qualify as a prohibited AI practice, or
- If the transparency requirements for deepfakes or chatbots apply to the system (see above).

Governance and Enforcement

There are four key aspects of future governance under the AI Act:

- National competent authorities will supervise and enforce the AI Act's application and implementation regarding conformity of high-risk systems. Each Member State must establish or designate at least one notifying authority (responsible for conformity assessment bodies and

their monitoring) and one market surveillance authority (responsible for ex post monitoring).

- The EC and the AI Office: The EC established a dedicated AI Office in February 2024, which is tasked with central oversight and enforcement of the rules regarding GPAI, facilitating those requirements (e.g., by issuing template documents), and developing EU expertise and capabilities in the field of AI. In addition, the EC is responsible for the overall framework of the AI Act, including evaluating new AI technologies and updating the criteria for qualifying as GPAI models that pose systemic risk or high-risk AI over time.
- The Scientific Panel: A scientific panel of independent experts will be formed to assist the AI Office. The panel will have an advisory role, contributing to the development of evaluation methodologies for GPAI and advising on the selection and impact of GPAI. The panel will monitor safety issues in the market and will launch qualified risk alerts to the AI Office that may trigger investigations.
- The AI Board: The AI Board will consist of one representative from each EU Member State and will serve as a coordination platform and advisory body to the EC and is tasked with supporting a consistent application of the AI Act across EU Member States by developing standards and issuing guidance.

Fundamental Rights Impact Assessments Are Required, But Not Always

The AI Act requires a fundamental rights impact assessment to be conducted for high-risk AI systems, but only by public authorities, or by private actors when they use AI systems for credit scoring or for risk assessment and pricing in relation to life and health insurance. A fundamental rights impact assessment must include:

- A description of the deployer's processes in which the high-risk AI system will be used in line with its intended purpose;
- A description of the period of time and frequency in which each high-risk AI system is intended to be used;

- Understanding the categories of natural persons and groups likely to be affected by the use of the AI system in the specific context;
- Understanding the specific risks of harm likely to impact the identified categories of persons or a group of persons, taking into account the information given in the provider's instructions;
- A description of the implementation of human oversight measures, according to the instructions of use; and
- The measures to be taken if these risks materialize, including internal governance and complaint mechanisms.

Where the deployer is already required to carry out a data protection impact assessment under the EU General Data Protection Regulation (GDPR), the fundamental rights impact assessment must be conducted in conjunction with the data protection impact assessment.

Compliance with this obligation will be facilitated by the AI Office, which has been tasked with developing a template for the fundamental rights impact assessment.

Enforcement and Increased Penalties

The maximum penalties for noncompliance with the AI Act were increased in the political agreement on the EU AI Act reached in December 2023. There are a range of penalties and fines depending on the level of noncompliance. At their highest level, an organization can be fined an astounding €35 million or 7 percent of global annual turnover.

As with the GDPR, these levels of fines mean that organizations have a strong financial imperative to comply with the AI Act's provisions and with ethical and societal rationales.

Interaction with Data Protection Laws

Since the first draft of the AI Act, it was made clear that it would act as a "top up" of the GDPR in relation to personal data and that the GDPR remains applicable. The final text clarifies that both individuals and supervisory authorities keep all their rights

under data protection laws, such as the GDPR and the ePrivacy Directive, and that the AI Act does not affect the responsibilities of providers and deployers of AI as controllers or processors under the GDPR. The responsibilities under the GDPR are relevant because many of the risk-management obligations under the AI Act are similar to obligations that already exist under the GDPR. The AI Act, however, has a far broader scope and will also apply if the AI system:

- Is trained with regular data (i.e., not personal data subject to the GDPR),
- Is trained with personal data not subject to GDPR, or
- If the provider bringing the AI system on the EU market does not qualify as a "controller" under GDPR.

Many risk-management obligations under the AI Act cover similar topics as those under the GDPR. The obligations under the AI Act, however, are more detailed and wider in scope (applying to all data). By way of example:

- The AI Act's requirement to implement a risk-management system shows many similarities to the GDPR's requirement to conduct a data protection impact assessment (DPIA);
- The AI Act's requirement to implement data governance measures, including the use of appropriate data sets, is similar to the GDPR's requirement to ensure the fair and lawful use of personal information; and
- The AI Act's requirement to ensure appropriate human oversight, proportionate to the risk posed by the AI system, is similar to the GDPR's prohibition on automated decision making (with limited exceptions).

Controllers under the GDPR who currently train (or further train) AI systems with personal data or use AI systems processing personal data will therefore be able to leverage their GDPR compliance efforts toward complying with the AI Act's risk management obligations. Currently, it appears that the only specific requirement in the AI Act that fully overlaps with the GDPR is the right granted to individuals to an explanation for decisions based on high-risk systems that impact the rights of individuals (see below).

Limited Rights for Individuals

The initial draft of the AI Act did not bestow any rights directly on individuals. The final text changes this, by granting individuals the right to:

- Obtain an explanation of a decision made by a deployer of high-risk AI system on the basis of the output from such high-risk AI system, where the decision has legal effects or similarly significantly affects that person. The individual is in this case entitled to obtain a clear and meaningful explanation on the role of the AI system in the decision-making procedure and the main elements of the decision taken. The AI Act does not shed much light on the exact content of such explanation, but merely indicates that it should be clear and meaningful and should provide a basis for affected individuals to exercise their rights. This right is broader than the right to an explanation of an automated decision under the GDPR. Whereas the right under the GDPR is limited to decisions based solely on automated processing, the right under the AI Act extends to all decisions taken based on the output of a high-risk AI system (i.e., also including human decisions based on AI outputs); and
- Complain to a supervisory authority if the individual considers that there is an infringement of the AI Act.

The AI Act and National Security

The AI Act includes an exemption for AI systems that exclusively serve military, defense, or national security purposes.

The AI Act does "not apply to areas outside the scope of EU law" and in any event should not affect member states' competences in national security, "regardless of the type of entity entrusted by the Member States to carry out tasks in relation to those competences." The consolidated text clarifies that only if AI systems exclusively serve military, defense, or national security purposes, the AI Act does not apply. If an AI system is used for other purposes as well (e.g., civilian or humanitarian), or gets repurposed at a later stage,

providers, deployers, and other responsible persons or entities must ensure compliance with the regulation.

The exemption, however, remains unclear in the sense that the notion of "national security" is not clearly defined under EU law, and Member States apply different concepts and interpretations. To rely on the exemption for national security purposes other than military or defense, companies need to be mindful to ensure the respective purpose is indeed exclusively a "national security" use case in each Member State where an EU nexus exists. The recitals of the AI Act suggest that any use for "public security" would be distinct from a "national security" purpose, which appears inconsistent with the goal not to interfere with member state competences and to align the AI Act with other recent EU legislation like the Data Act, which exempts national security and public security purposes altogether.

The AI Act indeed foresees specific derogations with regard to public security. For example, it recognizes the interests of law enforcement agencies to quickly respond in duly justified situations of urgency for exceptional reasons of public security by using high-risk AI tools that have not passed the conformity assessment.

Looking at the broader implications of the AI Act in the area of national security, it may have an indirect impact on how AI tools will be viewed by member state authorities regulating their national security interests. For example, the AI Act may help by framing the undefined notion of AI included in the current EU framework regulation for screening of foreign direct investments. According to this framework, member states may take critical technologies, including artificial intelligence, into account when determining whether an investment is likely to affect security or public order.

Allocation of Responsibilities Across the AI Value Chain

Another key aspect that the AI Act includes is the allocation of compliance obligations along the AI value chain. All draft versions of the AI Act provided a mechanism where the obligations of the AI provider automatically transfer to certain deployers or to other operators. The final provides that any distributor, importer, deployer, or other third party shall be considered a provider of a high-risk AI system for the purposes of the AI Act, and shall be

subject to the respective provider obligations, in certain defined circumstances, namely if:

- They put their name or trademark on a high-risk AI system already placed on the market or put into service (the AI Act does, however, clarify that the parties may have contractual arrangements in place that allocate the obligations otherwise between them, which will likely only refer to their internal relationship and not affect their external relationships);
- They make a substantial modification to a high-risk AI system in a way that it remains a high-risk AI system; or
- They modify the intended purpose of an AI system, including a GPAI system, which has not been classified as high-risk and has already been placed on the market or put into service in such manner that the AI system becomes a high-risk AI system.

In these circumstances, the provider that initially placed the relevant AI system on the market or put it into service shall no longer be considered a provider of that specific AI system for purposes of the AI Act. In essence, all AI Act obligations in relation to the modified/rebranded AI system will switch to the new provider. This would apply even where, for example, a non-compliance of the AI system with the AI Act was already triggered by the original provider. Thus, the new provider may be responsible for compliance shortfalls of the original provider.

The original provider shall, however, closely cooperate and make available necessary information and provide reasonably expected technical access and other assistance required for new provider to fulfill its obligations under the AI Act.

The AI Act also retains the further EP proposal to obligate providers of high-risk AI systems and third parties that supply AI systems, tools, services, components, or process to such providers for integration into the high-risk AI system to specify details for their cooperation by written agreement. The terms of that agreement must allow the provider of the high-risk AI system to fully comply with its obligations under the AI Act.

In addition to the cooperation obligations, the final text stipulates specific technical documentation requirements for GPAI models to facilitate integration into downstream AI systems.

This area, as with any value chain proposition, needs careful forethought, both in the adoption and use of AI within one's own systems or for original providers allowing such use. Contractual provisions will be key here.

What's Next?

As mentioned above, the AI Act is expected to be formally endorsed by the Council before the end of the European Parliament's legislature in June 2024. The AI Act will then be subject to various transition periods (see Timeline and Transition Period above).

Notes

* The authors, attorneys with Morrison & Foerster LLP, may be contacted at cwalker-osborn@mofo.com, cstuetzle@mofo.com, lmoerel@mofo.com, mstorm@mofo.com, and skress@mofo.com, respectively.

1. https://eur-lex.europa.eu/eli/dir/2019/790/oj.