February 5, 2024

If you have any questions regarding the matters discussed in this memorandum, please contact the attorneys listed on the last page or call your regular Skadden contact.

This memorandum is provided by Skadden, Arps, Slate, Meagher & Flom LLP and its affiliates for educational and informational purposes only and is not intended and should not be construed as legal advice. This memorandum is considered advertising under applicable state laws.

One Manhattan West
New York, NY 10001
212.735.3000

Latest Text of EU AI Act Proposes Expanding Obligations for High-Risk and General AI Systems and Banning a Third Category

Executive Summary

- On 21 January 2024, a near complete draft version of the proposed text for the EU AI Act was unofficially shared with the public by a European media publication, after which a senior advisor in the European Parliament shared an updated draft of the legislation. This draft gives the most current indication as to what will, and what will not, be in the final act (the Act).
- On 2 February 2024, sources indicated that representatives from each of the EU's member states had approved the proposed text or a similar text. The approved text is expected to be presented to the EU Parliament for final approval in the coming months, and become law in spring 2024.
- The EU's "AI Pact" gives AI providers the opportunity to implement terms of the Act on a voluntary basis from the date the Act is adopted (which could be a matter of months after the EC publishes the official regulation). There may be significant market pressure on AI providers to join the pact as early adopters, which may not give the providers much time to prepare their systems and internal processes to be compliant.
- Although the text approved on 2 February 2024 (and therefore the final law's text) may differ from the currently available proposed text, this article provides an overview of key points companies should be aware of now to guide them in preparation for the impending regulation.
- On 24 January 2024, the European Commission announced that it will establish a new EU AI Office to (along with competent national authorities) monitor the effective implementation of and compliance with the Act and to receive certain notices from AI providers and deployers.

New Requirements Would Apply to Specialized AI Categories

Assuming the most recently proposed text is substantively adopted, the regulation would subject specialized categories of AI to new obligations/restrictions.

- **Some AI systems — those designated as Unacceptable Risk AI Systems (URAIs) — will be banned.** URAIs will be banned from the EU market. Except in limited, pre-authorised situations, this includes:
  - Social credit scoring.
  - Emotion recognition systems in work and educational contexts.

- AI that exploits people's vulnerabilities (*e.g.*, disability, age, gender).
- Systems that manipulate behaviour.
- Biometric categorisation using sensitive characteristics.
- Predictive policing.
- "Real-time" biometric information identification.

– **High Risk AI Systems (HRAIs) must follow additional requirements.**

- Assessing whether an AI system is an HRAI appears likely to be a complex assessment. Under the current draft, the determination depends on whether the AI system:
  - Is used as a component in, or constitutes, products covered in Annex II of the Act (which includes a range of products from medical devices to cableway installation); or
  - Performs functions listed under the use cases set out in Annex III of the Act, which include, among other functions:
    - Use for biometric purposes that are not otherwise considered a URAI system.
    - Critical infrastructure work.
    - Educational or vocational training.
    - Employment or self-employment purposes.
    - In relation to accessing essential public or private services.
    - Law enforcement.
    - Migration and asylum management.
    - In relation to judicial or democratic purposes.
- Especially in the context of the current status of AI systems and their use in the market, there are potentially wide exceptions to the use cases above. If the requirements for any of these exceptions can be met, the AI system would not be deemed an HRAI and would therefore not be subject to the relevant obligations. These conditions include that the AI system is used (i) for narrow procedural tasks, (ii) to improve the result of previously undertaken human activity or (iii) in a way that does not otherwise replace human review.
- Providers of HRAIs will likely be subject to specific requirements, including:
  - Ensuring the quality and accuracy of any training data and outputs.
  - Registering with and reporting to the new EU AI Office.
  - Undertaking impact assessments of fundamental rights and data protection, as well as implementing life cycle risk management systems.
  - Ensuring that downstream deployers are given adequate detail on the limitations and compliance requirements of the system, and that such deployers can ensure human oversight of the outputs of the HRAI.

Rules Would Apply to General Purpose AI Systems (GPAIs)

GPAIs are currently defined as AI systems based on models that display "significant generality" and are systems capable of performing "a wide range of distinct tasks," regardless of how they are placed in the market (*e.g.*, for download or through remote cloud access, although there are exceptions to certain transparency obligations for free and open-source models). This definition is broad enough to cover most existing foundation models.

The definition of a "Provider" of these systems is likely to include any party that puts its name or trademark on them or who makes a substantial modification to them — such that white labelling a third-party system would impose liability for the underlying system on both the third-party developer of the AI system and the party that commercializes the system using its trademark.

Providers of GPAIs placed onto the market would be under specific obligations, including to:

- Publish "sufficiently detailed" summaries of the training data.
- Implement a policy to respect the Copyright Directive, in particular by using "state of the art technologies" to identify and respect copyright holders' appropriate express reservation of rights to opt out of the use of their works for text and data mining (including for training GPAIs), "regardless of the jurisdiction in which the copyright-relevant acts underpinning the training" of the GPAI take place.
- Take appropriate steps to ensure content generated by GPAI is identified as such, including through watermarking, metadata identification and/or cryptographic methods.
- Upon request, make the GPAI available to the European Commission for evaluation and implement mitigation measures where a reasonably foreseeable risk is found that the GPAI can widely propagate negative effects on public health, safety, public security, fundamental rights or society as a whole.

Additional Rules Would Apply to GPAIs With Systemic Risks (GPAISRs)

GPAIs would be presumed to be GPAISRs if they have capabilities that match or exceed the capabilities recorded in the most advanced GPAIs based on certain indicators and technical benchmarks (including the amount of computing used to train the GPAI) as updated periodically.

GPAISR providers would need to notify the European Commission within two weeks of becoming aware that their GPAISR meets the requirements for such designation. Additionally, the European Commission has the power to independently identify and designate GPAISRs. The European Commission will maintain and publish an up-to-date list of all GPAISRs.

Providers of GPAISRs would be under specific obligations in addition to those applicable to GPAIs, including to:

- Perform and document model evaluations before placing GPAISRs on the market, including adversarial testing, alignment and fine-tuning of models.
- Monitor for and mitigate systemic risks through accountability policies, governance processes and post-market monitoring.
- Track and report serious incidents and possible corrective measures to the European Commission and competent national authorities.
- Ensure adequate cybersecurity (including against model theft and circumvention of safety measures) for the model and physical infrastructure.

Proposed Penalties

The proposed text shows that penalties have changed from those considered in previous proposals. The most recent include:

Breach	Maximum Penalty
Breaching the prohibitions on URAIs	Greater of €35 million and 7% of annual worldwide turnover
Noncompliance with any other obligations	Greater of €15 million and 3% of annual worldwide turnover
Supplying incorrect, incomplete or misleading information to regulators	Greater of €7.5 million and 1% of annual worldwide turnover

Potential Next Steps

AI system developers may want to consider the commercial implications of joining (or not joining) the AI Pact, including balancing the costs of voluntary compliance against the potential commercial and reputational pressures for AI systems to be seen as compliant.

If joining the AI Pact could be commercially beneficial or necessary, companies should consider whether signing up to do so is a viable option given the Act's text as it currently stands. What work would need to be undertaken to achieve compliance? System developers may wish to consider budgets and plans for this work now, noting the that final text of the Act may differ from the currently available text. To do this:

- System developers can assess existing AI systems and categorise them as URAIs, HRAIs, GPAIs and/or GPAISRs (noting that a GPAI or GPAISR may also be, depending on its use, an HRAI).
- If any HRAIs are identified, system developers can prepare for compliance with the mostly output- and notification-related obligations of the Act, including through complying with any third-party standards already in existence (*e.g.*, ISO/IEC 42001:2023) and updating internal policies, procedures and training programmes.
- Companies with HRAIs may want to consider if steps can be taken to allow an exception to this categorisation to apply, so as not to be categorised as an HRAI (noting that the requirements for the exceptions may change before the Act is finalized, and the system may be separately categorised as a GPAI in any event).
- If any GPAIs are identified, system developers may want to consider the input-related, technical processes that will need to be developed to ensure compliance, including to respect copyright holders' opt-out rights and publicise training datasets, as well as the output-related requirement to ensure adequate watermarking or metadata tagging of generated content.
- If any GPAISRs are identified, in addition to the above considerations applicable to GPAIs, system developers may want to consider implementing the processes for model evaluation (including red-teaming and alignment) , particularly as such models are likely to be dual-use foundation models subject to similar requirements in the United States under President Biden's Executive Order 14110 of October 30, 2023.
- Finally, any company using third-party AI tools may want to approach its vendors to request further due diligence information to allow the company to assess the likely category of the AI system (*i.e.*, URAI, HRAI, GPAI and/or GPAISR) and determine whether the vendor intends to sign up to the AI Pact.

Latest Text of EU AI Act Proposes Expanding Obligations for High-Risk and General AI Systems and Banning a Third Category

Contacts

David A. Simon
Partner / Washington, D.C.
202.371.7120
david.simon@skadden.com

Pramode Chiruvolu
Counsel / Palo Alto
650.470.4569
pramode.chiruvolu@skadden.com

Nicola Kerr-Shaw
Counsel / London
44.20.7519.7101
nicola.kerr-shaw@skadden.com

Eve-Christie Vermynck
Counsel / London
44.20.7519.7097
eve-christie.vermynck@skadden.com

Susanne Werry
Counsel / Frankfurt
49.69.74220.133
susanne.werry@skadden.com

Edmund Berney
Associate / London
44.20.7519.7186
edmund.berney@skadden.com