

FEBRUARY 28, 2025

European Commission Provides Guidance on Scope of AI Systems Under the EU AI Act

As part of a flurry of guidance relating to the EU AI Act (the ‘Act’), the European Commission (the ‘Commission’) released guidance on 6 February 2025 on the scope of ‘AI Systems’ as required under Article 96(1)(f) of the Act (the ‘Guidance’)¹. This quickly follows the Commission’s non-binding guidance on the prohibitions contained in Art. 5 of the Act². Although also not binding, the current guidance is also likely to be persuasive to competent authorities and courts when interpreting the Act. This alert summarises the Guidance for business developing or using AI tools and identifies key practical takeaways.

Practical Takeaways

1. **No exhaustive list of AI Systems.** The Commission repeats in multiple places that it is not possible to provide an exhaustive list of AI Systems covered by the Act. The Guidance aims to provide a step-based formulation to assist in determining what is a relevant AI System, but there is need for flexibility to cover rapid technological developments.
2. **Ability to infer is critical.** Many of the conditions for an AI System are stated to be fairly binary or easy to assess, with the most detailed consideration being a system’s ability to infer (and to a lesser degree its level of autonomy in doing so). The Guidance makes clear that this capability is important to differentiate AI Systems that are subject to need regulation from simple systems that perform basic optimisation or other functions.

Summary of Guidance

Article 3(1) of the Act defines ‘AI System’ as:

a machine-based system that is designed to operate with varying levels of autonomy and 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 Guidance breaks down this definition into the following seven elements, which are dealt with in turn: “(1) a machine-based system; (2) that is designed to operate with varying levels of autonomy; (3) that may exhibit adaptiveness after deployment; (4)

¹ Available here: <https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-ai-system-definition-facilitate-first-ai-acts-rules-application>.

² See here for Paul, Weiss’s commentary: <https://www.paulweiss.com/practices/litigation/artificial-intelligence/publications/european-commission-publishes-guidance-on-prohibited-ai-practices-under-the-eu-ai-act?id=56629>

and that, for explicit or implicit objectives; (5) infers, from the input it receives, how to generate outputs (6) such as predictions, content, recommendations, or decisions (7) that can influence physical or virtual environments.”

The Guidance emphasizes from the start that not each element need be continuously present throughout an AI System’s lifecycle (i.e., may only appear in one of the development or deployment stages), provided that they all appear at some point in the lifecycle.

1. **Machine-Based System.** The Guidance notes that all AI Systems must inherently run on machines because machines are required to enable AI to function. The term ‘machine’ is broad and covers both **hardware** (i.e., physical elements such as, processing units, memory, storage devices, networking units and input/output interfaces) and **software** (e.g., code, instructions, programs, operating systems and applications that handle how hardware performs)³. Reflecting on its introductory comment that AI Systems will evolve over time, the Guidance hypothesises that advanced quantum computing systems would still be categorised as machines for the purposes of the definition, as would a biological or organic system, provided they provide some computational capacity.
2. **Varying Degrees of Autonomy.** Recital 12 of the Act clarifies that the concept of ‘varying degrees of autonomy’ means that AI Systems should operate with *“some degree of independence of actions from human involvement and of capabilities to operate without human intervention”*. As such, the Guidance notes that the core concept in considering autonomy for AI Systems is the human-machine interaction and whether there is a “reasonable” degree of independence. Any system designed to operate solely with full manual human involvement and intervention is excluded, whether such involvement/intervention is direct or indirect (e.g., through systems-based controls that allow human delegation, or only require supervision by humans, of system operations). Conversely, a system that requires manual inputs but generates an output by itself will have the necessary degree of independence. The Guidance adds that where less human intervention is required, the risks are higher and so more built-in human oversight may be required as part of risk mitigation.
3. **Adaptiveness.** Again, Recital 12 of the Act provides clarification that ‘adaptiveness’ refers to self-learning capabilities that allow changes in system behaviour during use to produce different results based on the same inputs. However, the Guidance is clear that the use of the phrase “*may*” in the Act means that this is not a prerequisite for an AI System and a system can still be considered an AI System without this capability.
4. **AI System Objectives.** The Act notes that AI Systems must operate to one or more explicit or implicit objectives. The Guidance clarifies that **explicit objectives** refers to clearly stated goals that are directly encoded into the system (such as to optimise a cost function), whereas **implicit objectives** refers to goals that may be deduced from the systems behaviour or its underlying assumptions and arise from either the training or the deployment stage where the system interacts with the environment. The Guidance, by reference to Recital 12 and Article 3(12), also notes that an ‘objective’ is different to an ‘intended purpose’. The latter is the external context in which the system is intended to be deployed by the provider (i.e., the use case to assist a specific behaviour) as opposed to the objectives inherent to the system’s analysis of its input data. The intended purpose relies on both the inherent objective plus other factors such as the integration of the system into other behaviours.
5. **Inferring how to Generate Outputs.** The Guidance critically examines this criterion, stating that this inference ability is a “*key, indispensable*” factor in determining an AI System (to differentiate from “*simpler traditional systems or programming approaches*”). Recital 12 clarifies that this capability is to “*derive models or algorithms, or both, from inputs or data*”⁴. However, the Guidance notes that that ‘inference’ should not be narrowly interpreted as only the ability of a system to derivate outputs from given outputs, but also relates to the ‘building’ phase of AI development through the **AI techniques**

³ This aligns with the view put forward in a recent UK Court of Appeal judgment which held that both software and hardware can constitute compute programmes (*Emotional Perception v Comptroller* 18 July CA-2024-000036).

⁴ The Guidance notes that this aligns with the concept of ‘inference’ in the ISO/IEC 22989 standard, which defined inference as “*reasoning by which conclusions are derived from known premises*”.

incorporated into a system to enable inference (which then generate outputs, as discussed below). These AI techniques include ‘**machine learning**’ approaches and ‘**logic- and knowledge-based**’ approaches.

Machine learning This encompasses approaches that learn from data on how to achieve certain objective. It includes: (i) **supervised learning**, where the system learns from human-labelled data to pair input data with a correct output (e.g., an AI-enabled email spam detection system, image classification systems); (ii) **unsupervised learning**, where the system learns from unlabelled data and instead uses techniques such as clustering, anomaly detection and association rule learning to find patterns, structures and relationships in data without any explicit guidance (e.g., an AI System for drug discovery); (iii) **self-supervised learning**, as a form of unsupervised learning where the systems use unlabelled data and techniques such as contrastive learning to create its own labels or objectives (e.g., image recognition systems, predicative text models); (iv) **reinforcement learning**, where systems learn from data they have collected before through a reward function, learning from trail-and-error experience rather than labelled data or patterns; and (v) **deep learning**, where layered architectures (such as neural networks) automatically learn features from large data volumes to create high accuracy outputs.

Logic- and knowledge-based approaches This encompasses approaches that “*infer[s] from encoded knowledge or symbolic representations of the task to be solved*” (Recital 12). These systems use deductive or inductive engines, or techniques such as sorting, searching, matching and chaining, to apply logic or reasoning to new situations, learning from rules, facts and relationships encoded by humans, as opposed to learning from the underlying data (whether assisted by humans or not). Examples include classical language processing models regarding grammar or semantics rules to identify and then extract the meaning of a text, and medical diagnosis systems that draw conclusions based on encoded symptoms.

The Guidance notes that certain simpler systems (including those with only a limited capacity to infer) will not be ‘AI Systems’ including: (i) **systems aimed at improving mathematical optimisation** which may have the capacity to infer to improve efficiency but only at a level of basic data processing, such as physics-based simulations that are fed into established models (e.g., measure atmospheric processes to enable faster and more computationally efficient forecasts) and telecommunications systems that optimise bandwidth allocation based on predictions of resource requirements; (ii) **basic data processing systems** which follow predefined, explicit instructions or operations and do not undertake learning, reasoning or modelling during their lifecycle (e.g., database management systems and standard spreadsheet software); (iii) **systems based on classical heuristics** to more efficiently find approximate solutions where exact ones are impractical to find, which may employ pattern recognition but not via data-driven learning and do not show adaptability (e.g., a chess program, where the model cannot be adapted for anything other than chess once its inputs are defined); and (iv) **simple prediction systems** that use basic statistical learning rules (e.g., future stock price predictors which use historic average prices to make baseline predictions).

6. **Outputs That Are Predictions, Content, Recommendations or Decisions.** The Guidance follows the Act to split outputs in four categories which the Guidance notes, although not specific to AI Systems, are generated with more complexity in AI Systems.

Prediction An estimate as to an unknown value derived from known values. The Guidance notes that, although predictions are present in many types of non-AI based software, the differentiating factor is the complexity of the predictions an AI Systems can make (the real-time predictions in driverless cars).

Content New materials generated by AI Systems, such as text, images, videos or music. As the Guidance notes, although technically, the content may be seen as a type of prediction or decision, the Act treats them as a separate category of output (which is relevant for the rules around generative AI specifically, and other regulations such as the EU’s Digital Services Act).

Recommendation Specific suggestions for specific actions, products or services based on preferences, behaviours or other data inputs.

Decision Conclusions or choices made by a system, including recommendations which are then automatically applied without any human oversight.

7. **Influence over Environment.** The Guidance notes that this requires an ‘active’ element to a system which impacts the environment of deployment, including physical objects (such as robot arms) or virtual environments (such as digital spaces, data flows and software ecosystems). Interestingly, the Guidance does not interpret the phrase “can” in the same way as “may” is interpreted for adaptiveness, so it appears that the Commission’s view is that an influence on a physical or virtual environments is a required feature for an AI System.

Conclusion

In its press release, the Commission notes that the Guidance is designed to evolve over time and will be updated as necessary – specifically referencing new use cases. For now, it appears that the Commission is interpreting the concept of ‘AI System’ broadly, with the most complex consideration being whether the system is able to make inferences.

Interestingly, the concluding remarks note that, whether or not a system is determined to be an AI System under the Act, “the vast majority of systems ... will not be subject to any regulatory requirements” (e.g., in relation to prohibited practices, high-risk obligations, etc.). The Guidance also seeks to remove basic machine learning and optimisation tools from the Act’s scope and emphasises that high-risk AI systems already on the market prior to August 2026 will not be caught (as per Article 111(2) of the Act). When read in combination with the recent guidance on prohibited practices, the Commission seems to focus on the impacts of relevant systems, rather than their fundamental programming or technology, to assess the scope of compliance obligations and necessary degree of regulation. It will be interesting to review the Commission’s guidance on high-risk activities when released, to understand whether this direction of travel continues. More practically, five days after the Guidance was published, the Commission announced a €200 billion initiative to mobilise investment into AI (including specifically a €20 billion fund for AI gigafactories)⁵; the EU may be trying to tread a fine line between not stifling innovation while appeasing those member states that demand detailed regulatory oversight, especially given the regulatory divergence with other core jurisdictions⁶.

* * *

This memorandum is not intended to provide legal advice, and no legal or business decision should be based on its content. Questions concerning issues addressed in this memorandum should be directed to:

Jonathan H. Ashtor
+1-212-373-3823
jashor@paulweiss.com

John P. Carlin
+1-202-223-7372
jcarlin@paulweiss.com

Katherine B. Forrest
+1-212-373-3195
kforrest@paulweiss.com

Anna R. Gressel
+1-212-373-3388
agressel@paulweiss.com

Nicole Kar
+44-20-7601-8657
nkar@paulweiss.com

Henrik Morch
+32-2-884-0802
hmorch@paulweiss.com

John Patten
+44-20-7367-1684
jpatten@paulweiss.com

Audrey M. Paquet
+1-212-373-2397
apaquet@paulweiss.com

Alex Zapalowski
+44-20-7367-1697
azapalowski@paulweiss.com

Associates Edmund Berney, Scott Caravello, Ali Fazeli-Nia and Jason Gerson contributed to this Client Alert.

⁵ https://ec.europa.eu/commission/presscorner/detail/en/ip_25_467_

⁶ See, for example, the US and UK not signing the international agreement put forward at the Artificial Intelligence Action Summit in Paris on 10 – 11 February 2024: https://www.bbc.co.uk/news/articles/c8edn0n58gwo_