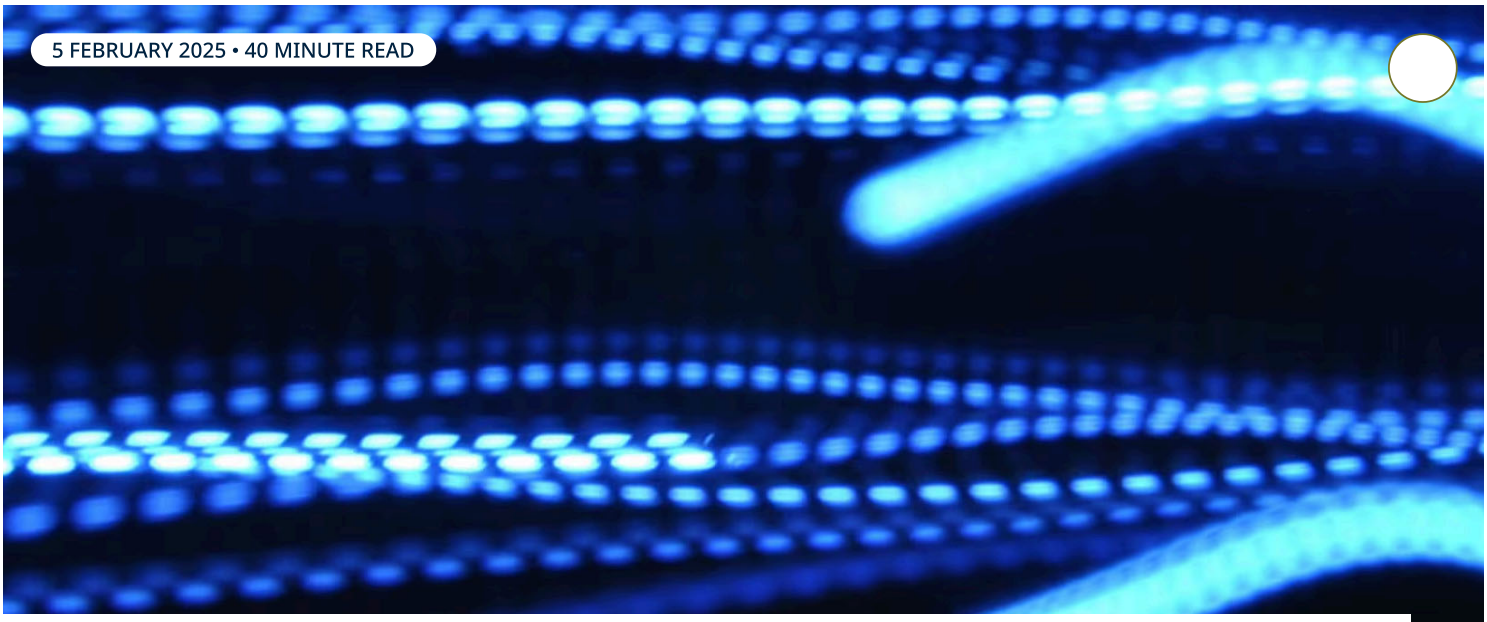
AI OUTLOOK SERIES



European Commission publishes guidelines on Prohibited AI Practices

Navigating the EU's AI Act: Key Guidelines on Prohibited AI Practices

Written by: Danny Tobey, Ashley Carr, Karley Buckley, Kaylee Hoffner, Kyle Kloeppel

On February 4, 2025, two days after the European Union Artificial Intelligence Act (EU AI Act)'s ban on AI Systems that leverage Prohibited Practices went into effect, the European Commission published Commission Guidelines on Prohibited Artificial Intelligence Practices in draft form (Draft Guidelines).

These Draft Guidelines provide additional clarification and context for the types of AI practices that are prohibited under the Act. They provide direction to 1) surveillance authorities in their enforcement efforts and 2) Deployers and Providers in their efforts to comply with the Act. While not binding or authoritative, the Draft Guidelines are intended to promote consistent application of the EU AI Act across the European Union (EU). The Draft Guidelines have been approved by the European Commission but have not been formally adopted at this time.

Under Article 5 of the EU AI Act, AI Systems that leverage Prohibited Practices are considered to pose an unacceptable risk to fundamental rights and EU values.

Enforcement of the EU AI Act is assigned to market surveillance authorities designated by the Member States and the European Data Protection Supervisor. There are heavy penalties for non-compliance with provisions dealing with Prohibited Practices, including fines of up to EUR35 million or 7 percent of global annual turnover of the preceding year (whichever is higher).

Key takeaways

- The Draft Guidelines aim to increase clarity and provide insight into the Commission's interpretation of Prohibited Practices under Article 5 of the Act.

- The Draft Guidelines are lengthy but are still in draft form and, even when finalized, will be non-binding. All guidance provided therein is subject to the formal requirements set forth in the Act.

- Though the Draft Guidelines are not comprehensive, they are a helpful step in assessing whether an AI System qualifies as prohibited under the Act.

Selected clarifications and examples from the Draft Guidelines

For each of the Prohibited Practices outlined in Article 5 of the Act, the Draft Guidelines provide an overview of the main components of the provision, practical examples, clarification of practices that are out of scope from the prohibitions, and measures that can be taken to avoid providing or using AI systems in ways that are likely to be prohibited. The Draft Guidelines also highlight where the prohibitions overlap or are related to other Union legal acts.

The below table is intended to highlight key clarifications and examples from the Draft Guidelines based on questions that we frequently receive from clients. The examples assume that all of the required elements of each prohibited category are otherwise met (except where noted). Additionally, it is important to note that the overarching exceptions set forth in Article 2 (eg, national security) apply to the prohibited practices and are relevant to the practical application of these categories.

Prohibited Practice	Key elements of the Prohibited Practice's definition under the AI Act *Placing on the market, putting into service, or use of:*	Key insights from the Draft Guidelines	Selected examples from the Draft Guidelines *(not exhaustive)*
Subliminal techniques, harmful manipulation and deception (Art. 5(1)(a))	An AI System that: • Deploys subliminal, purposefully manipulative, or deceptive techniques… • With the objective or effect of materially distorting behavior… • By appreciably impairing a person's ability to make an informed decision and causing them to make a decision they would not otherwise have made… • In a manner that is reasonably likely to cause them or another significant harm	• Human intent to deceive is NOT required; requires only that the AI System deploy the manipulative techniques [69, 76] • Visual labeling of "deep fakes" mitigates risk of deception and harmful distorting effects [71] • The Guidelines look to EU consumer protection law to interpret the concept of "material distortion of behavior" [80] • AI applications that "are not reasonably likely to cause significant harms	• An AI System that displays text in a video that is technically visible but flashes too quickly for the conscious mind to register, while still being capable of influencing attitudes or behaviors • An AI System that embeds images within other visual content that are not consciously perceived but that may still be processed by the brain and influence behavior

		are in principle outside the scope of the prohibitions" [134] • The Guidelines differentiate manipulation from "lawful persuasion" – *ie*, "presenting arguments or information in a way that appeals to reason and emotions, but explains the AI System's objectives and functioning" [128]	• An AI chatbot that impersonates a friend of a person or a relative with synthetic voice, causing scams and significant harms
Harmful exploitation of vulnerabilities (Art. 5(1)(b))	An AI System that: • Exploits the vulnerabilities of a person/group due to age, disability, or social or economic situation, and • Has the objective or effect of materially distorting the person's behavior, and • Does so in a manner that causes or is reasonably likely to cause significant harm	• "Vulnerabilities" is broadly defined to include "cognitive, emotional, physical, and other forms of susceptibility" [102] • Susceptibility must be the result of ("due to") the person belonging to one of the enumerated groups: age, disability, or socioeconomic situations [102]	• An AI System used to target older people with deceptive personalized offers or scams – *eg*, to influence them to buy "expensive medical treatments" or "deceptive investments schemes." • An AI system that targets young users and uses addictive reinforced schedules with the objective of keeping them dependent on the application
Social scoring (Art. 5(1)(c))	An AI System that: • Evaluates or classifies people based on social behavior or personality characteristics and • Leads to either detrimental or unfavorable treatment of certain persons or groups in social contexts	• Profiling of natural persons under EU data protection law, when conducted through AI systems, may be covered by this prohibition [154] • Need not be AI alone, but AI must play an "important role" in the social score [161]	• A tax authority uses an AI predictive tool (considering income as well as unrelated data, such as social habits) on all taxpayers' tax returns to select tax returns for closer inspection • An insurance company collects

	that is: ◦ Unrelated to the contexts in which the data was originally generated/collected or ◦ Disproportionate to their social behavior or its gravity	• Applies even if the social score "is produced by an organization(s) different from the one that uses the score" [162]	spending and other financial information from a bank (unrelated to the determination of eligibility for life insurance) and uses it to determine the price of the premium to be paid
Criminal offense risk assessment and prediction (Art. 5(1)(d))	An AI System that: • Assesses or predicts the risk of a person committing a crime… • Based solely on profiling the person or assessing personality traits or characteristics… • Subject to narrow exceptions for individuals already verifiably linked to criminal activity	• While focus is on law enforcement, the prohibition does apply to private actors, particularly where they are: ◦ Acting on behalf of law enforcement [207, 208] ◦ Assessing or predicting the risk of a person committing a crime for legal compliance purposes (*eg*, anti-money laundering) [209] • Generally excludes systems that profile entities or organizations rather than individual people [215]	• A law enforcement authority uses an AI system to predict criminal behavior for crimes such as terrorism solely based on individuals' age, nationality, address, type of a car, and marital status
Untargeted facial scraping to develop facial recognition databases (Art. 5(1)(e))	An AI System that: • Creates or expands facial recognition… • Through the untargeted scraping of facial images…	• Facial recognition database is one "capable of matching a human face from a digital image or video frame against a database of faces" [226]	• A facial recognition software company collects pictures of faces that have been scraped from social media, transforming facial features into mathematical representations, and

	• From the internet or CCTV footage	• Database need not be solely used for this purpose [226] • "Untargeted" means "without a specific focus on a given individual or group", regardless of respect of opt-out protocols [228] • Tools "instructed to collect images or video containing human faces only of specific individuals" are targeted and outside the scope of the prohibition [229] • Does not apply to untargeted scraping of other biometric data (*eg*, voices) [234]	hashing and indexing them for future comparison to determine whether new images match faces in the database
Emotion recognition in the workplace or education (Art. 5(1)(f))	AI System that: • Infers emotions of a person in the workplace or educational setting… • Except where used for certain medical or safety reasons	• Includes "both AI systems identifying or inferring emotions or intentions" [245] • Emotions and intentions defined broadly and should not be interpreted restrictively [247] • Inferences from written text would not be based on biometric data, but inference from key stroke or body postures or movements would be prohibited [251] • Workplace examples are focused on use of technology in reference to employees and workers (*eg*, rather	• An AI System that infers that an employee is unhappy, sad, or angry towards customers (*eg*, from body gestures, a frown, lack of a smile) **Example of safety/medical exception (*not* prohibited):** • An employer using AI-enabled devices or digital assistants at the workplace to measure anxiety based on measured stress levels when deploying dangerous machines or dealing with dangerous chemicals (due to exceptions for safety and medical reasons)

		than customers) [254]	
Biometric categorization to deduce or infer status in sensitive groups (Art. 5(1)(g))	An AI System that: • Uses an individual's biometric data to deduce their status as a member of sensitive groups (*eg*, race, political opinions, trade union membership) • Except that this prohibition excludes labeling or filtering of lawfully acquired biometric datasets and categorizing biometric data for law enforcement	• Categorization is "not about identifying an individual or verifying their identity, but about assigning an individual to a certain category" [276] • People must be "individually" categorized for prohibition to apply [282]	• A social media platform that uses AI to categorize persons according to assumed political or sexual orientation based on biometric data from photos they have uploaded to the platform, in order to send them targeted political messages • A biometric categorization system that claims to be capable of determining someone's race based on their voice
Real-time **remote biometric identification** (Art. 5(1)(h))	• Using real-time remote biometric identification systems in publicly accessible spaces for the purposes of law enforcement… • Unless strictly necessary for a narrowly defined set of objectives such as targeted search for specific victims of abduction or trafficking or prevention of substantial and imminent threat to lives or safety of people	• Only prohibited category that applies solely to "use" rather than placing on the market • Access control purposes (*eg*, face identification to enter a restricted area) are outside the prohibition [305]	• Using real-time facial recognition systems to monitor all persons on a public street for general security, crime prevention, and overcrowding concerns

DLA Piper is here to help

DLA Piper's team of lawyers and data scientists assist organizations in navigating the complex workings of their AI Systems to help ensure compliance with current and developing regulatory requirements. We continuously monitor updates and developments arising in AI and its impacts on industry across the world.

At the *Financial Times*'s 2024 North America Innovative Lawyer awards, DLA Piper was conferred the Innovation in New Services to Manage Risk award for its AI and Data Analytics practice.

For more information on AI and the emerging legal and regulatory standards, please visit DLA Piper's focus page on AI.

Gain insights and perspectives that will help shape your AI Strategy through our AI ChatRoom series.

For further information or if you have any questions, please contact any of the authors.

Related insights

Publication

SEC settlement highlights continued enforcement focus on AI washing

28 JANUARY 2025 • 5 MINUTE READ

Publication

White House AI Executive Order sets its sights on free-market innovation

27 JANUARY 2025 • 8 MINUTE READ

Related capabilities

Artificial Intelligence and Data Analytics

People