Guidelines on prohibited AI practices

Transformative Legal Experts

From the AI Act to the Guidelines on prohibited AI practices

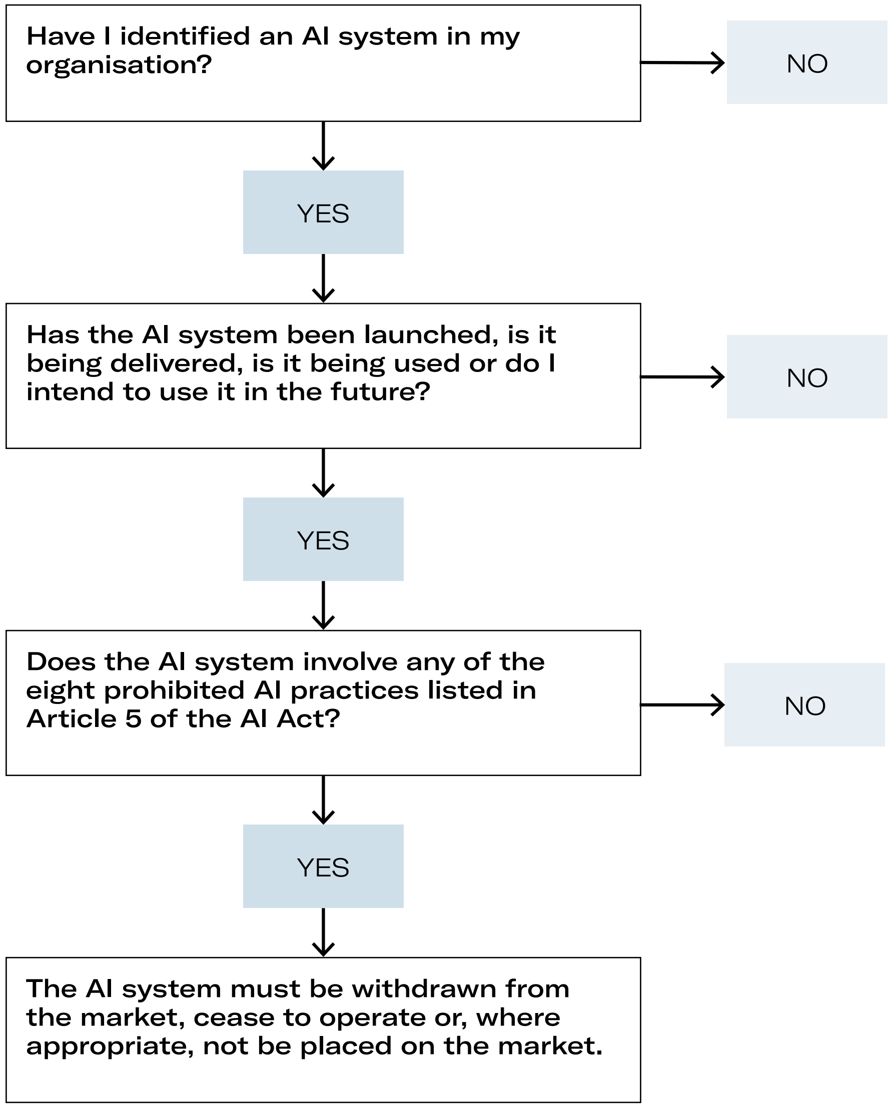
Following the entry into force of the Artificial Intelligence Act ("**AI Act**"), organisations have been required to comply with the rules on prohibited AI practices since 2 February this year.

The European Commission has now published guidelines on prohibited AI practices[1] ("**Guidelines**"). We will now look at some of the practical implications of the Guidelines for organisations.

What is the role of the Guidelines?

The Commission's Guidelines help to define and identify practices that are considered prohibited in the field of AI Systems[2], as they pose unacceptable risks[3] to the security and fundamental rights of citizens.

More specifically, Article 5 of the AI Act lists eight prohibited AI practices and the European Commission recently published its guidelines on these practices. It remains for organisations, working with IT and legal teams, to identify which AI systems[4] are prohibited and should therefore cease to operate, not be launched and/or be withdrawn from the market.



Have I identified an AI system in my organisation?

NO

YES

Has the AI system been launched, is it being delivered, is it being used or do I intend to use it in the future?

NO

YES

Does the AI system involve any of the eight prohibited AI practices listed in Article 5 of the AI Act?

NO

YES

The AI system must be withdrawn from the market, cease to operate or, where appropriate, not be placed on the market.

[1] Brussels, 04.02.2025, C(2025) 884 final, COMMUNICATION TO THE COMMISSION, Approval of the content of the draft Communication from the Commission - Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act).

[2] AI systems can be confused with other technologies such as traditional software, deterministic algorithms or even pre-programmed script-based virtual assistants or chatbots.

[3] The AI Act defines four levels of risk: minimal risk; limited risk with transparency obligations; high risk; unacceptable risk. AI systems are subject to different rules depending on the level of risk identified on a case-by-case basis. For AI practices with unacceptable risk, the rule is simple: prohibition. In other words, AI systems that pose an unacceptable risk to the safety and fundamental rights of citizens are prohibited from being placed on the market, put into service or used.

[4] Definition of what an AI system is: see Article 3(1) of the AI Act and the Guidelines on the definition of an artificial intelligence system.

Who should read the Guidelines?

The Guidelines are particularly useful for organisations that develop, supply and/or use AI systems, even if they were placed on the market before the date of application of the AI Act rules on prohibited AI practices.

In addition, although they are non-binding, they will of course be taken into account by courts and regulators.

How can I identify prohibited AI practices?

Having identified an AI system, organisations should establish whether the AI system involves practices that fall within the list of eight prohibited practices (see Article 5 of the AI Act).

To help organisations identify prohibited AI practices and act appropriately to comply with applicable laws, the Commission has provided some examples, either in terms of the techniques used in their development, or in terms of the effects and/or harm these practices may cause.

The Commission gives some specific examples of AI practices prohibited either because of the techniques used to develop them or because of the effects and/or harm they may cause.

```
def create_vpc(vpc_name):
    ec2 = boto3.client("ec2")
    response = ec2.create_vpc(CidrBlock="10.0.0.0/16")
    vpc_id = response["Vpc"]["VpcId"]
    ec2.create_tags(Resources=[vpc_id], Tags=[{"Key": "Name", "Value": vpc_name}])
    return vpc_id

def create_rds_instance(db_instance_identifier, vpc_id):
    rds = boto3.client("rds")
    response = rds.create_db_instance(
        DBInstanceIdentifier=db_instance_identifier,
        Engine="postgres",
        VpcSecurityGroupIds=[vpc_id],
        StorageType="gp2",
        StorageEncrypted=True,
        DBInstanceClass="db.t2.micro",
        MultiAZ=False,
        EngineVersion="11.5",
        MasterUsername="admin",
        MasterUserPassword="Password1",
    )
    return response["DBInstance"]["DBInstanceArn"]

def backup_rds_instance(db_instance_arn, backup_bucket_name):
    veeam = veeam_365_sdk.Client()
    veeam.backup_rds_instance(db_instance_arn, backup_bucket_name)

def connect_to_hpe_data_lake(vpc_id, db_instance_identifier):
```

Here are a few examples:

LIST OF PROHIBITIONS (SEE ARTICLE 5 OF THE AI ACT):	KEY WORDS, CONCEPTS AND COMPONENTS OF THE PROHIBITIONS:	AI SYSTEMS THAT USE THE FOLLOWING PRACTICES ARE PROHIBITED (SEE EXAMPLES IN THE GUIDELINES):	EXCEPT IN THE FOLLOWING CASES:
Manipulative techniques and deceptive techniques [5(1)(a) of the AI Act]	○ Subliminal techniques ○ Purposefully manipulative techniques ○ Deceptive techniques ○ Material distortion of behaviour	○ Use of subliminal techniques e.g. visual or auditory subliminal messages, sub-visual or sub-audible stimuli, distraction of attention, manipulation of time perception to create impatience and user dependency, brain spyware. ○ Use of purposefully manipulative techniques e.g. an AI system using sound or background images to induce mood swings, increase user anxiety and distress. ○ Use deceptive techniques, such as presenting false and/or misleading information e.g. AI chatbot impersonating a human, causing fraud and significant damage.	○ A system that exhibits manipulative behaviour in a purely incidental way, provided that appropriate preventive and mitigating measures have been taken in the event that harm is reasonably likely to occur e.g. a generative AI system that hallucinates[5] and therefore presents misleading information. ○ A system that uses lawful persuasion techniques and whose operational objectives are transparent and respect people's autonomy e.g. a system that analyses emotions to improve interactions with customers and provide them with support (outside the scope of the prohibition).
Harmful exploitation of vulnerabilities [5(1)(b) of the AI Act]	○ Exploitation ○ Vulnerabilities ○ Children ○ Older people ○ People with disabilities ○ Social or economic situation ○ Material distortion of behaviour ○ Reasonably likely to cause significant harm	○ Exploitation of the vulnerabilities of an individual or group of people with the objective, or the effect, of materially distorting the behaviour of that person or group, likely to cause significant harm e.g. an AI chatbot that targets disadvantaged groups by inciting them to commit acts of violence; or an AI system that exploits the cognitive vulnerabilities of the elderly by targeting them with more expensive medical treatments.	○ A system that uses persuasion but not manipulation e.g. AI systems that help children learn in school and games (outside the scope of the prohibition).

[5] Term used to describe a technical defect in a generative AI system.

LIST OF PROHIBITIONS (SEE ARTICLE 5 OF THE AI ACT):	KEY WORDS, CONCEPTS AND COMPONENTS OF THE PROHIBITIONS:	AI SYSTEMS THAT USE THE FOLLOWING PRACTICES ARE PROHIBITED (SEE EXAMPLES IN THE GUIDELINES):	EXCEPT IN THE FOLLOWING CASES:
Social scoring: evaluation or classification [5(1)(c) of the AI Act]	○ Social behaviour ○ Personal or personality characteristics ○ Definition of profiles ○ Detrimental or unfavourable treatment ○ Unjustified or disproportionate treatment	○ Developing and using a system to assess and classify individuals on the basis of behaviour and/or personal characteristics over time that may lead to unfavourable or harmful results e.g. an AI system used by the tax authority that uses tax return data to select specific individuals for audit; or a bank using an AI system to determine creditworthiness and decide whether or not a particular individual should receive a mortgage based on unrelated personal characteristics. ○ The development and use of a system that scores and classifies individuals on the basis of behaviour and/or personal characteristics over time may lead to unjustified or disproportionate results e.g., a government agency using an AI system to control fraud in the process of awarding scholarships by using marital status, parents' education level, as a discriminator of the level of fraud.	○ Social classification system with fair and proportionate treatment e.g. an AI system used by a company with a legitimate interest to detect financial fraud, where the assessment is based on relevant data such as transactional behaviour (outside the scope of the prohibition).
Assessing and predicting the risk of a criminal offence [5(1)(d) of the AI Act]	○ Risk assessment ○ Prediction of criminality ○ Definition of profiles ○ Personality traits and characteristics ○ Human assessment ○ Objective and verifiable facts directly linked to a criminal activity	○ Risk assessment to predict the likelihood of a person committing a criminal offence e.g. an algorithm that can predict crime based on nationality and ethnicity. ○ Assessment based solely on profiling and/ or assessment of personality traits without the inclusion of objective and verifiable data relating to criminal activity e.g. tools that predict crime based on personality traits such as age or marital status.	○ An AI system used to assist in the assessment of a person's involvement in criminal activity based on objective and verifiable facts directly related to the criminal activity e.g. using an AI system to profile and categorise reasonably suspicious dangerous behaviour in a crowd indicating that someone is preparing to commit a crime and is likely to do so (outside the scope of the prohibition).
Untargeted scraping of facial images from the Internet / CCTV to develop facial recognition databases [5(1)(e) of the AI Act]	○ Facial recognition ○ Untargeted scraping ○ Databases ○ Facial images ○ Internet ○ CCTV	○ Untargeted scraping of facial images from the Internet or CCTV footage to gather as much information as possible without focusing on a specific group or individual e.g. facial recognition software trained on social network images. ○ Creation or expansion of facial recognition databases capable of matching a human face in an image/video with a face in the database e.g. software that analyses security images to create databases without the consent of individuals.	○ Systems using untargeted collection of biometric data other than facial images e.g. voice samples (outside the scope of the ban).

LIST OF PROHIBITIONS (SEE ARTICLE 5 OF THE AI ACT):	KEY WORDS, CONCEPTS AND COMPONENTS OF THE PROHIBITIONS:	AI SYSTEMS THAT USE THE FOLLOWING PRACTICES ARE PROHIBITED (SEE EXAMPLES IN THE GUIDELINES):	EXCEPT IN THE FOLLOWING CASES:
Recognition of Emotions [5(1)(f) of the AI Act]	○ Inference of emotions ○ Biometric data ○ Workplace ○ Educational institutions ○ Medical reasons ○ Safety reasons	○ Detecting or inferring emotions and intentions in work environments or educational institutions based on biometric data e.g. AI system that monitors facial expressions and micro-expressions to assess the level of engagement and interest of students in a classroom.	○ The ability of a system to detect or infer emotions and intentions based on biometric data in the workplace or educational settings for medical or safety reasons where there is an explicit need under labour law e.g. an AI system used to measure stress levels of workers where stress is a risk to workers on dangerous machinery.
Biometric categorisation [5(1)(g) of the AI Act]	○ Categorisation ○ Biometric data ○ Classification ○ Deduction or inference ○ Sensitive characteristics	○ Using biometrics to infer sensitive characteristics to assign labels that could lead to discriminatory treatment e.g. an AI system that classifies individuals based on skin colour and associates these profiles with crime statistics. ○ Individual categorisation of a person based on their biometric data e.g. an AI system that tries to deduce a person's ethnicity from their religious orientation, voice, tattoos or facial features.	○ Whether they label or filter biometric data collected in accordance with applicable law, or whether they categorise for objective technical reasons to identify and prevent specific risks e.g. AI system that categorises patients from images according to skin colour to diagnose oncological problems.
'Real-time' remote biometric identification ("RBI") [5(1)(h) of the AI Act]	○ Identification ○ RBI system ○ Biometric data ○ In real time ○ Remote ○ Publicly accessible spaces ○ For law enforcement purposes	○ Real-time remote biometric identification (RBI) in public places for law enforcement purposes e.g. the police using real-time RBI systems to identify a shoplifter and compare their facial images with other criminal records held by the criminal investigation department..	○ RBI systems for specific purposes such as searching for missing persons or preventing terrorist attacks e.g. genetic monitoring to prevent imminent threats.

Why should organisations be aware of these Guidelines?

Although these Guidelines are not binding, it is important to bear them in mind. As a soft law instrument, they are an important interpretative tool for public and private entities, as well as for regulators and courts, in identifying what are considered prohibited AI practices.

Organisations should audit and, where necessary, review AI systems that have been placed on the market, are in service or are currently in use. They should also evaluate those systems that are being developed and/or deployed, or that they intend to develop in the future.

It is important to remember that failure to comply with the prohibitions set out in Article 5 and now specified in the Commission's Guidelines constitutes a serious infringement. Anyone who commits such a serious infringement will be subject to the maximum fine of up to €35 million or, if the offender is a company, up to 7% of its current worldwide turnover for the preceding financial year, whichever is the greater.

Organisations will be required to review the list of prohibited practices annually, which may be amended from year to year[6].

As a soft law instrument, they are an important interpretative tool for public and private entities, as well as for regulators and courts, in identifying what are considered prohibited AI practices.

[6] Article 112 of the AI Act.

About PLMJ

→ Who we are

> "Strengths include geographical presence, immediate response, deep knowledge and experience in market concepts and themes."

CLIENT REFERENCE FROM
THE LEGAL 500

About the Technology, Media and Telecommunications practice

→ What we do

KEY CONTACTS

Pedro
Lomba

Partner and head of the Technology, Media and Telecommunications practice.

pedro.lomba@plmj.pt

Benedita Cunha
Pinto

Associate in the Technology, Media and Telecommunications practice.

benedita.cunhapinto@plmj.pt

February 2025