Hey 👋

I'm Oliver Patel, author and creator of ***Enterprise AI Governance***.

This free newsletter delivers practical, actionable and timely insights for AI governance professionals.

My goal is simple: to empower you to understand, implement and master AI governance.

If you haven't already, sign up below and share it with your colleagues. Thank you!

For more frequent updates, be sure to follow me on LinkedIn.

This week's edition features:

✅ Top 5 practical tips for prohibited AI compliance
✅ What does the EU AI Act say on prohibited AI?
✅ The European Commission's new Guidelines on prohibited AI (140 pages distilled)

Top 5 practical tips for prohibited AI compliance

Many readers of Enterprise AI Governance are experienced lawyers, compliance executives, and governance leaders seeking practical advice and actionable insights.

Therefore, before providing an overview of the prohibited AI practices and the European Commission's new Guidelines on the topic, I will frontload this week's edition with my practical tips for meeting Article 5 of the AI Act head on.

1. **Deliver company-wide communications, awareness, and training campaigns on prohibited AI.** AI governance leaders need to constantly remind themselves that although AI is everywhere, the AI Act is still a relatively niche topic. Do not assume that everyone already knows about prohibited AI, just because it is top of mind for us. Indeed, the legal concept of prohibited AI practices is new for us all. Although there are many uses of AI that would of course be illegal, the AI Act is the first EU law which explicitly prohibits specific AI use cases and systems. Therefore, invest ample time in rolling out dedicated training and communications campaigns, so that your organisation is fully aware of these new provisions, their scope, and practical impact.
2. **Determine whether prohibited in the EU means prohibited worldwide. And be prepared to justify your decision.** A realistic operational scenario that could arise is a team or department in your organisation, based outside of the EU, wanting to use an AI system, which is, or could potentially be, prohibited under

the AI Act. The use of this AI system is confined to one specific non-EU location. No AI outputs will be used in the EU and no EU employees, customers and stakeholders will be involved in any way. It is feasible that this use case could be perfectly legal in the relevant jurisdiction. Furthermore, perhaps it is not even particularly problematic, from an ethical or cultural standpoint. This raises a key question for global organisations: should prohibited in the EU mean prohibited worldwide? And, if so, what is the justification for this? On the one hand, the EU has prohibited certain AI practices due to their morally egregious nature and/or the potential for harm. However, on the other hand, no jurisdiction has a monopoly on morality.

3. **Do not assume that there are no activities in your organisation that breach these provisions. Be wary of the unknown unknowns.** You may have a well-established and embedded AI governance process, which captures all new AI projects, initiatives, and vendor applications. It may be that nothing has ever come through which relates to, or closely resembles, any of the prohibited AI practices. Unfortunately, this does not necessarily mean that no such development, deployment, or use of prohibited AI is occurring. Assumptions are dangerous and you don’t know what you don’t know. That is why dedicated training, compliance, and audit exercises are crucial.

4. **Conduct dedicated compliance initiatives, such as assurance evaluations and audits.** In large organisations, where it is difficult to keep track of and monitor everything, there is a need for bespoke compliance initiatives aimed at deterring, preventing, identifying, and halting AI activities which risk falling into the prohibited category. These initiatives are in addition to your BAU AI governance risk assessment processes. Ad hoc mechanisms, such as senior leadership attestation, assurance monitoring and evaluation, and internal and external audits, should all form part of your arsenal. Such initiatives will focus minds, hold leaders to account, and steer behavioural and cultural change in the right direction.

5. **Pay close attention to tools which may have emotional recognition capabilities, as there will be plenty of grey areas for companies.** Every prohibited AI category has grey areas and exceptions, but perhaps none more so than emotional recognition. For most responsible organisations, it is reasonable to suggest that they are unlikely to develop, deploy, or use AI systems in most of the prohibited AI categories. Emotional recognition could be an outlier, as it is something that could be inadvertently or unintentionlly used, or turned on, in applications designed for recruitment, coaching, training, productivity, and employee and customer engagement or support. Furthermore, it may be that similar tools can lawfully be used to predict and monitor performance, sentiment, or behaviour, but close attention will need to be paid to ensure this does not cross over into emotional recognition. It is also not necessarily a capability which is

deemed socially, culturally, or morally unacceptable in every jurisdiction worldwide.

Bonus tip: without a robust AI governance framework and risk assessment process, which is based upon a solid policy foundation, sponsored by senior leadership, and embedded across the organisation, you will be much less likely to identify and prevent any prohibited AI practices. Always prioritise establishing the fundamental building blocks of AI governance, before turning your attention elsewhere.

What does the EU AI Act say on prohibited AI?

Article 5 of the EU AI Act may only take up three out of 144 pages, but it is probably the most consequential section. It outlines the AI practices that are now prohibited under EU law.

The provisions on prohibited AI practices became applicable on 2nd February 2025, exactly 6 months after the AI Act entered into force on 1st August 2024.

The prohibited AI practices are deemed to be particularly harmful, abusive, and contrary to EU values. Breaching these rules carries the largest potential enforcement fines of up to 7% of global annual turnover. This is a highly dissuasive figure which should focus minds in the boardroom.

Below, I will provide a digestible summary of the eight categories of prohibited AI practice.

For some of these, there are specific and limited exceptions, which must be carefully considered and reviewed on a case-by-case basis. For example, the practices of predictive policing and the use of emotion recognition systems are not outright prohibited in every scenario.

Also, some of these categories proved highly contentious during the AI Act's legislative process and trilogue negotiations. In particular, the question of whether law enforcement authorities should be permitted to use AI systems for real-time remote biometric identification in public was totemic. The European Parliament argued for an outright prohibition, whereas member states wanted their authorities to retain such capabilities. A compromise between these two positions was eventually struck.

Now the dust has settled, it is prohibited, under EU law, to sell, make available, or use AI systems for any of the practices listed below.

Causing harm by deploying subliminal or deceptive techniques

- AI systems which deploy subliminal, manipulative, or deceptive techniques in order to distort the behaviour of an individual or group, in a way which causes, or is likely to cause, significant harm.
- This can include persuading or manipulating people to engage in unwanted behaviours.
- This practice is prohibited even if there was no intention to cause harm.

Causing harm by exploiting vulnerabilities

- AI systems which exploit the vulnerabilities of an individual or group (e.g., age, disability, or socioeconomic status) to distort that individual or group's behaviour in a way which causes, or is likely to cause, significant harm.
- This practice is also prohibited even if there was no intention to cause harm.

Social credit scoring systems

- AI systems which evaluate and score people based on social behaviour and personal characteristics, with the score leading to detrimental or unfavourable treatment which is disproportionate, unjustified, and/or unrelated to the context of the original data.
- Such AI systems can lead to discrimination, marginalisation, and social exclusion.

Predictive policing

- AI systems which are used to assess or predict the risk of an individual committing a criminal offence, where the assessment is based solely on automated profiling of their traits and characteristics.
- It is unlawful to use any of the following traits as the sole basis for assessing and predicting whether someone will commit a crime: nationality, birth location, residence location, number of children, level of debt etc.
- However, predictive policing AI systems are classified as high-risk (i.e., not prohibited), when they are not based solely on the type of automated profiling described above.

Emotion recognition in the workplace and education

- AI systems which infer or detect emotions, based on biometric data (e.g., facial images, fingerprints, or physiological data) in the context of the workplace and educational institutions.
- This includes emotions like happiness, excitement, sadness, and anger. However, it does not include physical states like pain or fatigue.
- There are exemptions for emotional recognition systems used for medical or safety purposes. In such permissible scenarios, an emotion recognition system would be classified as high-risk.

Creating facial recognition databases via untargeted scraping

- AI systems which create or expand databases for facial recognition, via untargeted scraping of facial images from the internet or CCTV footage.

Biometric categorisation to infer specific protected characteristics

- AI systems which categorise individuals based on biometric data, to infer protected characteristics like race, political opinions, trade union membership, religious beliefs, sexual orientation, or sex life.
- If biometric categorisation systems are used to infer characteristics or traits which are not protected, then those AI systems are classified as high-risk.

Law enforcement use of real-time biometric identification in public

- Law enforcement use of ’real-time’ remote biometric identification systems in public (e.g., facial recognition used to identify and stop potential suspects in public) is prohibited, apart from in very specific and narrowly defined scenarios. This prohibition is intended to safeguard privacy and prevent discrimination.
- Acceptable scenarios for leveraging AI in this way include searching for victims of serious crime, preventing imminent threats to life (e.g., terrorist attacks), or locating suspects or perpetrators of serious crimes (e.g., murder).

- However, independent judicial or administrative authorisation must be granted, before such AI systems are used. The use must only occur for a limited time period, with safeguards.
- National regulators and data protection authorities must be notified each time an AI system is used in this way. The European Commission will publish an annual report tracking and documenting these use cases.
- The development and placing on the market of AI systems intended for this purpose is, however, not prohibited. But there are prohibitions on the use of such systems.

The European Commission's new Guidelines on prohibited AI

Last week, the European Commission published Draft Guidelines on prohibited AI practices. These shed light on how organisations can interpret, practically implement, and comply with these important provisions.

The guidelines are also designed to assist regulatory authorities reviewing such cases, although those authorities are not legally obliged to take these guidelines into account.

At 140 pages—which is not too dissimilar to the overall length of the EU AI Act—we get a glimpse of the task ahead. To be facetious, if the EU keeps up its rate of 140 pages of guidance for every 3 pages of legal text, then AI governance professionals may have over 6700 pages of overall text which they will need to understand, in order to fully grasp this new law.

On a more serious note, this highlights that over the coming months and years, there will be a huge amount of additional documentation and analysis, such as official guidelines, recommendations, opinions, standards, codes of practice, enforcement decisions, and court judgements, which we will all need to keep up with. Luckily, you've come to the right place 😉

Summary of key points

- Regulatory authorities should use these guidelines in relevant cases and investigations. However, despite the various examples provided, there will always need to be a detailed, case-by-case assessment of the AI system in question.

- Although much enforcement of Article 5 will be actioned by regulators at the member state level, they must keep the Commission and other regulators informed regarding any cases with cross-border impact. Furthermore, regulatory authorities should strive for harmonisation on prohibited AI enforcement, via collaboration at the European AI Board.

- The prohibitions apply to any AI system, whether it has a narrow ‘intended purpose’ which is prohibited’, or whether it is a general-purpose AI system used in a manner which is prohibited. Deployers must therefore ensure that they do not use or customise general-purpose AI systems in a prohibited way.

- There are other scenarios, under EU law, where the use of an AI system could be prohibited, even if this is not explicitly stipulated in the AI Act. For example, there are other conceivable uses of AI which could breach the EU’s Charter of Fundamental Rights.

Example of potential prohibited scenarios for each category

In the guidelines, the European Commission provided some illustrative examples of what could be prohibited, across each category. I have provided a selection below.

Causing harm by deploying subliminal or deceptive techniques

- An AI companionship app imitates how humans talk, act, and respond. It uses human-like traits and emotional signals to affect how users feel and think. This can make people emotionally attached to the app, leading to addiction-like behavior. In some cases, this might cause serious problems, like suicidal thoughts or even harm to others.

Causing harm by exploiting vulnerabilities

- An AI chatbot targets and radicalises socio-economically disadvantaged people, encouraging them to harm or injure other people, by tapping into their fears, vulnerabilities, and sense of social exclusion.

Social credit scoring systems

- A national labour agency uses AI to determine whether unemployed people should receive state employment benefits. As part of the scoring process, data is used and relied upon which is unrelated to the purpose of the evaluation, such as marital status, health problems, or addictions.

Predictive policing

- A law enforcement authority uses AI to predict and determine that an individual is more likely to commit a terrorism offence, based solely on personal characteristics like their age, nationality, address, type of car, and marital status.

Emotion recognition in the workplace and education

- Inferring emotions from written text (e.g., sentiment analysis) is not emotion recognition. However, inferring emotions from keystrokes, facial expressions, body postures, and movement is emotional recognition, as it is based on biometric data.

Creating facial recognition databases via untargeted scraping

- A facial recognition company collects photos of people's faces using an automated tool that searches social media platforms. It gathers these images along with other data, such as the image's URL, location data, and the person's name. The company then processes the photos to extract facial features and converts them into mathematical data for easy storage and comparison. When someone uploads a new photo to their facial recognition system, it is checked for any matches in the facial database.

Biometric categorisation to infer specific protected characteristics

- An AI system which categorises individuals' social media profiles based on their assumed sexual orientation, which is predicted by analysing biometric data from their photos, is prohibited.

ICYMI: get the full 17-page report which compares AI laws in the EU, China and U.S.A 👇