GREEN
EUROPEAN
JOURNAL

Europe’s AI (Balancing) Act

Article by Seden Anlar
April 11, 2024

After a years-long legislative journey, the European Parliament approved the first comprehensive regulation of artificial intelligence globally. Though an important step in regulating rapidly evolving technology, the AI Act makes major concessions to both corporate interests and law enforcement authorities.

Welcome to the future, where science fiction scenarios are no longer confined to the silver screen. Since ChatGPT’s public release in November 2022, artificial intelligence (AI) has infiltrated many aspects of our lives, including healthcare, education, and, inevitably, politics. Yet, this is just the opening act.

From students relying on AI to craft essays to doctors placing their trust in AI diagnoses, the influence of artificial intelligence is undeniable. In Europe, this trend is also visible in political communications, where AI-driven tools are now commonplace. Deepfake videos can blur the lines between reality and deception, potentially impacting public perception. In Belgium, for instance, the Flemish Christian Democrat party “resurrected” the late Prime Minister Jean-Luc Dehaene in a deepfake campaign video. Disturbingly, instruments of harassment like “deep nudes” and deepfake porn have emerged, highlighted by incidents involving figures like Italian Prime Minister Giorgia Meloni.

As the AI “boom” unfolded, Europe was grappling with a regulatory void, prompting member states to turn to the EU for comprehensive legislation. The outcome is the EU AI Act, passed by the European Parliament in March 2024 and hailed as the world’s first comprehensive, horizontal, and binding AI regulation.

The long road to regulation

In contrast to the rapid advancement of AI, the regulatory journey has beenanything but smooth sailing. Brussels first proposed AI regulations as early as 2019, aiming to lead global efforts in monitoring emerging technologies. This was a significant promise from European Commission President Ursula von der Leyen, who identifies as a “tech optimist”. These pledges materialised with the introduction of the AI Act in April 2021.

The early drafts of the law focused on AI systems performing specific functions, such as scanning resumes and job applications. However, the unexpected surge in general-purpose AI models, particularly OpenAI’s ChatGPT, caught EU policymakers off guard, which led to hurried adjustments and set off the development of an altered AI Act.

With artificial intelligence poised between immense benefits and significant risks, a key question that emerged and dominated the discourse on the AI Act debate was: Should regulation encourage innovation or focus on mitigating potential harms?

Lobbies were lobby-ing

From the start, the European Parliament leaned towards preemptive regulation. Its June 2023 position on the AI Act called for the tight regulation of foundation models like GPT in the ChatGPT, regardless of their assigned risk category or their purpose – a stance driven by concerns over the vast amount of training data required to build them, as well as their impact on privacy.

Recognising potential threats to their business models, tech giants have aimed to influence this process. While publicly advocating for AI regulation, privately they resisted any significant restriction of foundational models. Recent research by the Corporate Europe Observatory revealed that 66 per cent of AI-related meetings involving members of the European Parliament in 2023 were with corporate stakeholders, compared to 56 per cent between 2019 and 2022. Recognising the European Parliament as a challenging target, Big Tech swiftly shifted its focus to the European Commission and EU member states: 86 per cent of high-level Commission officials' AI meetings in 2023 were with industry representatives.

In late 2023, tech companies amped up both private and public pressures. Controversies peaked when OpenAI's CEO Sam Altman hinted at possibly withdrawing ChatGPT from Europe (a stance he later retracted), highlighting the delicate balance between regulatory expectations and industry interests.

Lobbying efforts weren't confined to offshore entities; European AI startups like Mistral AI and Aleph Alpha increased pressure on their national governments, particularly in France and Germany. Consequently, these EU member states advocated for a more innovation-friendly approach and a lighter touch, fearing excessive regulation might stifle European competitiveness and innovation. They even proposed major exemptions for foundation models from AI Act regulations.

Sovereignty lost in clouds?

France's concerns underscore a broader issue: the EU AI Act's strict regulations could further undermine digital sovereignty, pushing reliance towards non-European AI solutions. Ironically, despite these concerns, France continues to rely on international companies for data storage, highlighting the ongoing tension between autonomy and globalisation in the digital domain. Recently, Paris-based Mistral AI has announced that its AI model Large, a competitor to GPT-4, will be hosted on Microsoft's Azure cloud platform rather than on the French provider OVHcloud.

The human side

Beyond technical and economic considerations, the AI Act delves into significant ethical and human rights issues. In Europe and around the world, AI systems are increasingly deployed for harmful forms of state surveillance, including biometric identification, emotion recognition, and predictive policing. These technologies often disproportionately affect marginalised communities, providing those in control with unchecked authority, suppressing democratic liberties, facilitating widespread surveillance, undermining legal rights, and reinforcing existing oppressions and inequalities.

Recognising the threat such technology poses to democracy, even before the release of the original AI Act proposal, civil society organisations have consistently advocated for a human-centric approach to AI, aiming to place fundamental rights at the heart of the legislation and lead the battle against biometric mass surveillance. In November 2021, over 100 civil society organisations called for concrete changes to the proposed Act to prioritise fundamental rights.

Based on these concerns, in the initial drafts of the Act, the European Parliamentadvocated for stricter

biometric restrictions. However, France, aiming to utilise AI in the fight against crime and terrorism, lobbied aggressively, exerting significant pressure on the Parliament to soften the proposed measures.

As debates intensified and the process dragged on, civil society voiced concerns over further delays in passing the AI Act. A recent study from the European Council on Foreign Relations predicts a "major shift to the right, with populist parties gaining votes and seats" in June's EU elections. While member states under the current mandate have already shown resistance to restrictions and oversight of their use of AI in this process, such opposition is expected to intensify.

Pressured from different sides, the European Parliament approached the final trialogue negotiations with the member states and the Commission with a strong proposal. In the end, however, safeguards for human rights were significantly diluted or entirely omitted, with massive loopholes for public authorities and relatively weak regulation of the largest foundation models that dominate the digital sphere (and pose the greatest harm).

The Act unpacked

In its final form, the AI Act adopts a "risk-based approach" to products or services. It categorises AI systems according to their potential societal impact: the greater the risk, the stricter the regulations, with some applications being completely banned due to their dangers. The result is a complex bowl of rules – almost like a regulatory salad.

Unacceptable uses (kind of)

At its core, the EU AI Act aims to protect against AI uses that covertly manipulate decision-making or exploit vulnerabilities through deceptive techniques, such as subliminal messaging or emotion recognition within workplaces or schools. By banning these and other "unacceptable" uses, the Act sets a clear ethical boundary.

However, the Act is surrounded by controversies and marked by significant compromises. Despite its firm stance on certain practices, it introduces exemptions, particularly for law enforcement and migration authorities. Emotion recognition technologies, banned in general contexts, are permitted for migration control purposes. Civil society argues this could lead to surveillance abuses, particularly against marginalised communities and people on the move, echoing concerns of AI-facilitated racial profiling and unwarranted surveillance.

Emotion recognition technologies could lead to surveillance abuses, particularly against marginalised communities and people on the move.

Moreover, the Act prohibits real-time biometric facial recognition in public areas – also known as face scanning and facial recognition – yet carves out exceptions for law enforcement in scenarios involving severe crimes like terrorism or the search for missing persons, contingent upon judicial approval. This seemingly controlled use has been criticised heavily, especially considering that in 2021 alone, more than 6,000 suspects were targeted by European arrest warrants, indicating a potentially broad application. As Ella Jakubowska of European Digital Rights (EDRi), pointed out, this is especially worrying in light of the global trend of mislabeling human rights defenders, journalists, and even climate

activists as terrorists. Amnesty International claimed that loopholes in the Act greenlight “dystopian digital surveillance”.

Regulated surveillance

Beyond unacceptable risks, the AI Act classifies certain uses as high-risk, requiring compliance with European Commission guidelines and relevant standards. High-risk applications, such as those in medical devices and critical infrastructure like water or electricity, as well as in education and employment, must meet rigorous requirements, including using high-quality data and ensuring clarity for users.

While real-time biometric identification is prohibited except in certain circumstances, biometric identification through recorded video lands in the high-risk category – permitted under tight regulations but not completely banned. This distinction hinges on technical and procedural nuances: real-time scanning is seen as overly intrusive, whereas reviewing recorded footage is believed to allow for more control and oversight.

However, in the context of human rights, the impact of both surveillance types is equally alarming. The digital rights network EDRI highlighted that “the fear of being pervasively watched and tracked does not diminish if authorities or companies take longer to review footage. In fact, the threat to individual freedoms and democracy might be even more severe with ‘post’ processing” – with governments, police forces, companies, or malicious entities having access to individuals’ highly sensitive personal data for years to come.

Deep fakes, shallow laws

Another AI technology that escaped strict regulation is deepfakes – capable of creating fake images, videos, or audio of real people, places, or events. Despite being considered high-risk by many, the current AI Act categorises deepfakes as posing limited risk, only requiring them to be labelled as artificially manipulated. The legislation falls short of providing a comprehensive framework for holding the creators of deepfake technology accountable. Rather, it favours preventive strategies over punitive measures, suggesting developers might have to embed strong safeguards, such as advanced watermarking or detection algorithms against malicious use.

What about ChatGPT?

The AI Act targets general-purpose “models” – that is, the behind-the-scenes technology that powers AI tools like ChatGPT or Google’s Bard – rather than the consumer-facing applications. Developers of these models are required to maintain detailed technical documentation and assist companies or individuals deploying their tools in understanding their functionality and limitations. They must also summarise the copyrighted materials (e.g., texts, images) used in training the models, and collaborate with the European Commission and national enforcement authorities regarding compliance.

Certain general-purpose models are designated as posing a “systemic risk” due to their extensive influence and potential to trigger catastrophic events. Developers of these models must take extra steps to minimise those risks, set up safety measures, and report any incidents to the newly established “AI Office” of the Commission, which is tasked with overseeing compliance with the regulations.

Brussels effect?

The AI Act has been criticised not only for potential human rights infringements in Europe but also for its geographical limitations, as it neglects the global consequences of technology developed within the bloc. This oversight implies that AI systems created in the EU could be exported and potentially contribute to human rights abuses abroad. This situation reflects an underlying imbalance: the EU's detailed internal AI regulations starkly contrast with the absence of guidelines for the external application of these technologies. As things stand, there is a glaring disconnect between the EU's internal regulatory efforts and its international human rights obligations.

The AI regulations set in Brussels could end up setting the standard worldwide.

However, this concern may not be as significant as it seems. In practice, many companiesprefer to conform all of their products to the EU's strict standards rather than designing separate features just for the EU market. As a result, the AI regulations set in Brussels could end up setting the standard worldwide.

Rethinking AI governance

The legislative journey of the EU AI Act has exposed significant flaws in EU policymaking. The interests of national governments, law enforcement agencies, and Big Tech lobbies often managed to overshadow the public interest and human rights, pressuring the European Parliament to give up essential protections. This situation underscores a larger issue: as sociologist Evgeny Morozov puts it, AI legislation tends toward market-driven, profit-centric approaches that favour surveillance.

The numerous exemptions risk undermining the Act's objectives, particularly concerning transparency and oversight for high-risk AI systems used by law enforcement and migration authorities. The national security exemption stands out, allowing member states to bypass regulations for activities deemed related to national security. In particular, law enforcement and migration authorities, armed with these exemptions, may further entrench a culture of impunity, continuing the deployment of harmful AI against marginalised communities.

Rejecting this status quo is critical as it sustains existing political dynamics and exacerbates inequalities. The long-held narrative of "innovation versus regulation", promoted by tech giants to evade substantial accountability, reveals a concerning concentration of power. Tech platforms should be open, accessible, and governed democratically rather than dominated by a few entities. Moving from Big Tech to 'Big Democracy" involves reshaping our digital social contract to highlight social and environmental justice, rectify inequalities, and advocate for digital citizenship, data sovereignty, and privacy.

AI legislation tends toward market-driven, profit-centric approaches that favour surveillance.

Europe's tech policy should actively build a digital society that prioritises collective well-being over private interests. This includes demanding transparency in AI training data and digital content production to safeguard individual and author rights and foster an environment that encourages innovation and collective creativity.

Looking ahead, civil society coalition Access Now, which works towards these principles, suggests that immediate actions are necessary at both EU and national levels to document and mitigate AI-related harms, especially in areas like migration and policing, to protect against rights violations. It is time to reevaluate resource allocation towards technologies that support rights rather than violate them, and to promote a human-first approach in our engagement with AI.

What now?

As artificial intelligence rapidly evolves, the AI Act's implementation looms, with a complete regulatory framework expected by mid-2026. Yet, significant uncertainties persist, especially concerning the law's practical application and the finalisation of technical standards and guidelines

In terms of enforcement, immediate actions will address certain AI applications, particularly biometric identification, which will be subject to strict review and potential prohibition. Moreover, each member state will set up their own AI watchdog, where citizens can file complaints if they believe they have been the victims of rule violations. Meanwhile, Brussels' AI Office is going to be enforcing and supervising the law for general-purpose AI systems.

While the AI Act provides a measure of certainty for today, the dynamic AI landscape suggests that the regulatory journey, initiated in 2019 (a time when the capabilities of today's AI could hardly have been anticipated), is far from over. Given the rapid advancement of technology against the slower pace of the EU's legislation and implementation processes, we are confronted with a crucial question: Can the regulations established today remain relevant tomorrow, or will they serve as a flexible framework that can guide the EU through future technological developments? Will the periodic reviews set up to ensure the continued relevance of the AI Act be sufficient?

Ultimately, the success of the AI Act – and, by extension, the future of AI governance in Europe – will depend not only on the provision's ability to withstand the test of time, but also on the will of European policymakers and societies to direct the development of artificial intelligence towards the public good, ensuring that technology serves humanity – not the other way around.

Seden Anlar is a Brussels-based journalist, podcast host/producer, moderator, and political communications specialist. She writes and tells stories about intersectional climate and social justice-related issues. Previously, she has worked for political campaigning organisations such as Climate Action Network (CAN) Europe, political publications like the Green European Journal, and produced podcasts such as "Green Space" for the Green Party of England and Wales and "Changing the Table Podcast" for the Migration Policy Group.

Published April 11, 2024
Article in English
Published in the *Green European Journal*
Downloaded from https://www.greeneuropeanjournal.eu/europes-ai-balancing-act/