THE IMPLICATIONS OF THE EU AI ACT ON CONVERSATIONAL TECHNOLOGIES LIKE CHATGPT

Emilian MATEICIUC[*]

Abstract

This paper investigates the implications of the European Union's Artificial Intelligence Act (AI Act), on conversational AI technologies. As the EU institutes a groundbreaking framework for AI regulation, this study assesses how the AI Act's risk-oriented approach impacts the crafting, deployment, and oversight of conversational AI. The analysis explores the Act's system classification, high-risk AI categorization, and delineation of duties for AI developers and deployers, examining effects on innovation, privacy, and ethical considerations within conversational AI.

The significance of this research lies in its exploration of the EU AI Act's effort to balance technological progression with the safeguarding of fundamental rights and user privacy. By examining the AI Act provisions specific to conversational AI technologies like ChatGPT, this paper highlights the challenges and opportunities within the legislative framework. It addresses key regulatory concerns including data protection, algorithmic transparency, and accountability, evaluating the Act's role as a potential standard for AI legislation globally.

Situated within the extensive debate on AI regulation and ethics, this contribution is timely, offering insights into how legislative bodies can adapt to and influence the rapid development of AI technologies. This analysis seeks to guide policymakers, developers, and the academic sphere in navigating the complexities of conversational AI regulation, proposing strategies to align AI technology's growth with societal values and legal frameworks.

Keywords: *EU AI Act, Conversational AI, ChatGPT, AI ethics, AI governance.*

1. Introduction

This study centres on the European Union's Artificial Intelligence Act (AI Act), unanimously approved by the Council of EU Ministers on 02 February 2024[1], to assess its influence on conversational AI technologies, notably ChatGPT. This legislative development is crucial for guiding the integration of such technologies into society and industry, ensuring they align with ethical standards and public welfare.

The necessity of analysing the AI Act in the context of conversational AI is underscored by the escalating use of these technologies across diverse sectors, prompting concerns about privacy, security, and ethical application. The Act aims to foster an environment where innovation is pursued within a structure of strong safeguards for individual and societal values.[2]

In tackling these issues, the study will examine the AI Act's classifications and requirements for conversational AI systems. This entails assessing the responsibilities placed on developers and operators by the Act and considering the wider implications for the trajectory of AI-mediated communication.[3] Through an examination of the legislative text, complemented by insights from academia and industry, this paper seeks to chart the course for conversational AI in the wake of the AI Act's adoption.

Setting itself apart from existing dialogues, this paper aims to bridge a gap in the literature by directly connecting the EU AI Act's regulatory schema to the rapidly evolving domain of conversational AI. It situates itself

[*] Graduate student, Master's program „International and EU Law", Faculty of Law, „Nicolae Titulescu" University of Bucharest (e-mail: emilian.mateiciuc@gmail.com).

[1] European Union, *Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts*, OJ from 02.02.2024.

[2] J. Doe, A. Smith, *The Impact of Regulatory Frameworks on AI Development: A Comparative Study*, in Journal of AI Ethics, 5(2), 2023, p. 123-145.

[3] B. Johnson, *Conversational AI: Challenges and Opportunities in Ethics and Regulation*, in Technology and Society, 30(4), 2023, p. 234-250.

at the juncture of legal analysis and technological exploration, hoping to ignite further discussion on aligning AI advancements with legal and ethical standards.[4]

2. Overview of the EU AI Act

2.1. Foundation

Within the ambit of the European Union's AI Act, specific attention is dedicated to the realm of conversational AI technologies, such as chatbots and virtual assistants. These technologies, which are powered by advanced algorithms similar to those behind ChatGPT, have become integral to digital communication, offering new avenues for interaction in consumer services, education, and many other sectors. The AI Act introduces nuanced provisions tailored to address the unique challenges and implications posed by these conversational interfaces, ensuring they serve the public good while adhering to ethical standards.

The Act meticulously outlines the obligations for creators and distributors of conversational AI, focusing on critical aspects like transparency, accountability, and the safeguarding of user rights. A pivotal requirement set forth by the legislation is the unambiguous disclosure when users are engaging with AI-driven platforms rather than human counterparts. This mandate is crucial in the context of conversational AI, where the blurring lines between AI-generated and human responses can lead to ambiguity. Ensuring clarity in these interactions is not just about fostering trust; it's about reinforcing the user's autonomy in digital ecosystems, providing them with the knowledge to make informed decisions about their engagement with AI technologies.

Furthermore, the legislation compels the entities behind conversational AI tools to conduct thorough risk assessments and implement measures to mitigate any identified risks. This is particularly significant in preventing the perpetuation of biases or discriminatory practices through these AI systems. Given the EU's robust commitment to principles of non-discrimination and data protection, conversational AI developers are tasked with integrating these values right from the design phase through to deployment and operation. This means creating systems that not only respect privacy and ensure data security but also proactively address and rectify potential ethical pitfalls, such as biases in language understanding or response generation.

By embedding these stringent requirements within the regulatory framework, the EU AI Act does more than just legislate; it steers the development and application of conversational AI towards a more humane and ethically conscious direction. This legislative approach underscores the EU's intention to harness the benefits of conversational AI while minimising the risks, ensuring these technologies augment human capabilities and enhance service delivery without compromising fundamental rights or ethical principles.

2.2. Provisions Relevant to Conversational AI

The European Union's Artificial Intelligence Act is pioneering in its comprehensive regulatory approach towards AI technologies, applying a nuanced risk-based framework that classifies AI systems according to the level of risk they pose. This classification broadly categorises AI systems into four levels: unacceptable risk, high risk, limited risk, and minimal risk. This segmentation is pivotal in tailoring regulatory requirements to the potential impact of different AI technologies on society and individuals.[5]

For conversational AI technologies – such as chatbots and virtual assistants, which are integral to customer service, healthcare triage, and educational platforms – the classification under the EU AI Act primarily hinges on their application and the extent of their interaction with humans. Conversational AI systems that are designed to interact with people in high-stakes contexts (*e.g.*, healthcare advice, legal information, or educational guidance) might be classified under „high risk" or „limited risk" categories, depending on their potential to affect users' rights or safety.

Systems categorised as „high risk" must follow strict compliance requirements before their deployment. These requirements include, but are not limited to, rigorous testing for accuracy and safety, transparent disclosure of AI interaction to users, and robust data protection measures to safeguard personal information. The primary aim here is to ensure that conversational AI technologies are reliable, secure, and transparent, thus fostering trust and safeguarding users' fundamental rights.

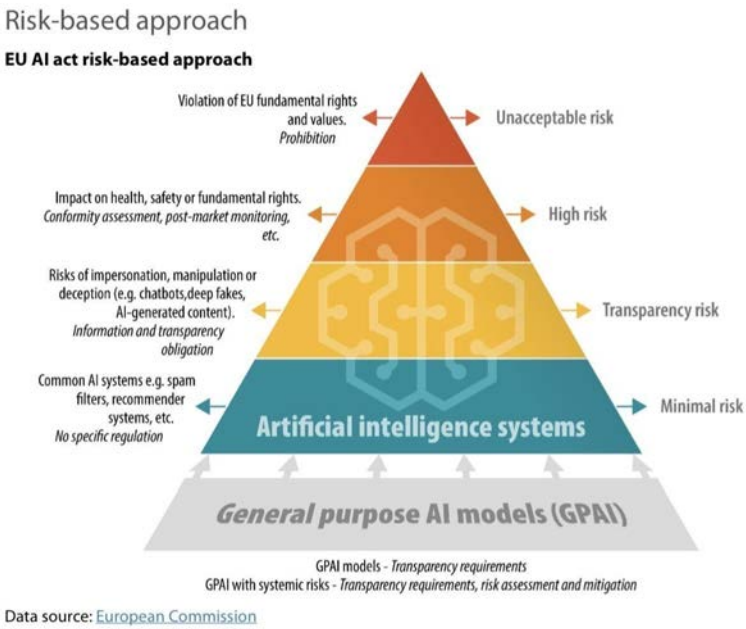
[4] European Commission, *White Paper on Artificial Intelligence: A European Approach to Excellence and Trust*, Brussels, 19.02.2020.
[5] European Parliament, *Briefing on Artificial Intelligence Act*, 11.03.2024.

Conversational AI technologies that are classified as „limited risk" or „minimal risk" are subjected to comparatively lighter regulatory requirements. For instance, AI systems offering recommendations on less critical matters, such as movie choices or restaurant recommendations, might fall into these categories. Such systems are expected to ensure transparency, informing users that they are interacting with an AI. This requirement is vital in maintaining user autonomy and trust by clearly delineating between human and AI-generated responses.

The governance and enforcement of these provisions are structured through a comprehensive framework that involves both national supervisory authorities and the European Artificial Intelligence Board (EAIB). National supervisory authorities are tasked with the local oversight and implementation of the Act, ensuring that AI systems developed or deployed within their jurisdictions comply with the established regulations. On a broader scale, the EAIB plays a crucial role in harmonising regulatory practices across EU member states, facilitating knowledge exchange, and providing strategic guidance on AI regulation. The Board's involvement ensures a cohesive approach to AI governance across the EU, promoting consistency in the application of the AI Act's provisions and fostering an environment of cooperation among member states.

3. Regulatory Implications for ChatGPT

Under the EU AI Act's structured framework, conversational AI technologies, such as ChatGPT, are evaluated based on their intended use and the potential risks they pose to users' rights and safety. This evaluation determines their classification into risk categories, which could range from high risk to limited risk. The Act's risk-based approach is particularly relevant for conversational AI due to its widespread application in various sectors, including customer service, education, healthcare, and more.



High-risk Classification: If a conversational AI system is classified as high risk, it indicates that its application has significant implications for individuals' rights or safety. For instance, a chatbot used for delivering medical advice or legal assistance would fall into this category due to the potential consequences of inaccurate information. High-risk conversational AI systems must adhere to stringent compliance requirements before and after deployment. These requirements include conducting thorough risk assessments to identify and mitigate any potential harm, implementing robust data governance measures to protect personal data, ensuring transparency about AI's role in decision-making processes, and maintaining human oversight to intervene whenever necessary. Such measures are crucial in maintaining the integrity of high-risk AI applications, ensuring they serve the public's interest without compromising ethical standards;

Limited-risk Classification: Conversational AI systems classified as limited risk may involve interactions where the stakes are not as high but still require certain regulatory adherence to ensure user trust and safety.

These systems must be transparent in their use of AI, often requiring clear communication to users that they are interacting with an AI system. This transparency is vital in allowing users to make informed decisions about their engagement with these technologies.

3.1. Ramifications of Classification on Compliance Requirements[6]

Risk Management: For high-risk conversational AI, developers must implement a comprehensive risk management system. This involves identifying potential risks related to privacy, discrimination, and overall safety, and taking appropriate measures to mitigate these risks before deploying the technology;

Data Governance: The AI Act places a strong emphasis on protecting personal data, especially for high-risk AI applications. Conversational AI developers are required to establish solid data governance frameworks that ensure data accuracy, security, and privacy, adhering to the principles of data minimization and purpose limitation;

Transparency: The requirement for transparency affects both high-risk and limited-risk conversational AI systems. Developers must disclose the use of AI in their systems, providing users with information about how decisions are made, the data used, and the possibility of human oversight. This is crucial for building user trust and facilitating accountability;

Human Oversight: For high-risk applications, the AI Act mandates human oversight to ensure AI decisions can be overridden or altered by human operators. This provision ensures that, despite AI's autonomy, critical decisions can be reviewed and amended by humans, particularly in situations where AI's judgment may impact individuals' rights significantly. Developers are required to disclose the involvement of AI in interactions, which necessitates embedding mechanisms that can clearly inform users when they are communicating with an AI.

3.2. Compliance Challenges[7]

These challenges primarily revolve around ensuring data protection, accuracy and fairness – each of which holds significant implications for the development and deployment of these AI systems.

Data Protection: Conversational AI systems often process vast amounts of personal data to provide personalised and contextually relevant responses. Adhering to the AI Act's stringent data governance and protection standards necessitates robust mechanisms to secure data, ensure privacy, and obtain explicit consent from users for data usage. The challenge intensifies with the need to implement these measures without compromising the user experience or the performance of the AI system. Balancing data protection with the operational requirements of conversational AI involves sophisticated data handling and privacy-preserving techniques, including anonymization and encryption, which can be complex and resource-intensive to implement;

Ensuring Accuracy and Fairness: Ensuring that conversational AI systems like ChatGPT are both accurate and fair poses another significant compliance challenge. Accuracy involves the system's ability to understand and respond to user queries correctly, which is crucial for high-risk applications. Fairness, on the other hand, entails the AI's capability to deliver unbiased responses and avoid perpetuating stereotypes or discrimination. Addressing these aspects requires continuous monitoring, testing, and refinement of AI models to identify and mitigate biases. Developers must employ diverse datasets and inclusive design principles from the outset, along with implementing fairness assessments and bias correction mechanisms. However, achieving and maintaining high standards of accuracy and fairness in conversational AI is an ongoing process, fraught with technical complexities and requiring constant vigilance to evolving societal norms and values.

4. Ethical Considerations and User Protection

The integration of conversational AI technologies into daily interactions introduces a spectrum of ethical considerations and necessitates robust user protection mechanisms. As these technologies, exemplified by systems like ChatGPT, become more ingrained in various sectors, their ethical implications and the importance of safeguarding users against potential harms have come into sharper focus.

[6] European Parliament, *Briefing on Artificial Intelligence Act*, 11.03.2024.
[7] *Ibidem*.

4.1. Ethical Issues

The ethical landscape for conversational AI technologies is complex, touching on issues of privacy, autonomy, accountability, and fairness:

Privacy and Data Security: Conversational AI systems process vast amounts of personal and sensitive data to function effectively. This raises ethical concerns about user privacy and the security of data against unauthorised access or breaches. Ensuring that these systems respect user confidentiality and secure data is paramount;

Autonomy and Consent: The ability of conversational AI to influence decisions and behaviours poses ethical questions about user autonomy. It's crucial that these systems operate transparently, making users aware of the AI's involvement in interactions and ensuring that consent is informed and freely given;

Accountability and Transparency: Holding AI systems and their developers accountable for the outcomes of AI interactions is a pressing ethical issue. This includes ensuring that there's clarity about how decisions are made by AI and providing recourse for users affected by potentially harmful decisions;

Bias and Fairness: The potential for conversational AI to perpetuate or even amplify biases presents significant ethical challenges. Ensuring fairness in AI interactions and outcomes requires vigilant efforts to identify and mitigate biases in AI models and datasets.

4.2. User Protection

In light of the ethical issues identified, protecting users in their interactions with conversational AI systems like ChatGPT is critical. The EU AI Act aims to establish safeguards that uphold user rights and safety:

Transparency and Informed Consent: Users must be clearly informed when they are interacting with AI, not humans. This transparency is foundational for ensuring informed consent, where users understand the nature of their interaction and the implications of their data usage;

Right to Explanation: For high-risk applications, users have the right to receive explanations for AI decisions that significantly affect them. This is crucial for maintaining trust and accountability, allowing users to challenge decisions or seek redress;

Privacy Safeguards: The Act emphasises strict adherence to data protection principles, requiring that conversational AI systems implement measures to protect user data rigorously. This includes data minimization, ensuring that only the necessary data for the intended purpose is collected and processed;

Bias Mitigation and Fairness: Developers are obligated to regularly assess and address biases in their conversational AI systems, promoting fairness and preventing discrimination. This involves careful design and continuous monitoring to ensure equitable outcomes for all users, regardless of their background.

The EU is a pioneer when it comes to citizen rights and privacy, becoming a model worldwide, just like it was the case with the GDPR Act.

5. Conclusions

Looking into the European Union's Artificial Intelligence Act and its impact on chatbots and similar AI tech shows a well-thought-out plan for regulation. This deep dive into the AI Act breaks down how it sorts AI technologies by risk levels and explains how this sorting affects rules for AI like chatbots. It points out the obstacles developers have to navigate to make sure their AI is transparent, keeps data safe, is accurate, and treats everyone fairly. The discussion also brings up important ethical points and how the Act sets up safeguards for users, focusing on keeping things clear, making sure users know what they're dealing with, and reducing bias.

The findings from this research are poised to significantly influence the development and regulatory oversight of conversational AI technologies. By spotlighting the principal compliance challenges and ethical considerations, the paper aims to contribute to the refinement of regulatory frameworks, ensuring they are fully attuned to the intricacies of conversational AI. The focus on user protection mechanisms is expected to build a foundation of trust in AI technologies among users, promoting their development and use in a manner that aligns with societal values and ethical standards.

Looking forward, the dynamic and rapidly evolving landscape of AI technology presents several avenues for further investigation. The practical effectiveness of the AI Act's risk-based classification system merits close examination, as does its impact on fostering or hindering innovation within the European Union. Additionally, comparing the EU AI Act with regulatory initiatives in other jurisdictions could offer valuable insights into

opportunities for global regulatory harmonisation, supporting the development of universally accepted standards that encourage innovation while safeguarding ethical considerations. Lastly, understanding the long-term societal impacts of conversational AI technologies, particularly in terms of privacy, autonomy, and the potential for a digital divide, is crucial. Such research could illuminate the pathways through which AI regulation can evolve to foster technologies that are not only innovative but also equitable and aligned with the broader interests of society.

A particularly compelling facet of our examination of the European Union's Artificial Intelligence Act is its approach to enforcing compliance through the imposition of fines. The Act's framework is stringent, setting out penalties that serve as a testament to the EU's commitment to ensuring that AI technologies, including conversational AI like ChatGPT, operate within clearly defined ethical and operational parameters. Specifically, the legislation delineates that fines for non-compliance can reach up to €20 million or 4% of the annual worldwide turnover of the offending company, whichever is greater. This scale of financial penalty is indicative of the serious stance the EU takes towards violations of the Act, particularly concerning breaches that compromise data protection, transparency, and the fundamental ethical conduct expected of AI systems.

The magnitude of these fines plays a dual role. Firstly, it acts as a deterrent, signalling to developers and companies the financial risks associated with neglecting the AI Act's compliance requirements. Secondly, it underscores the high value placed on safeguarding user rights and the integrity of AI interactions. For conversational AI technologies, this means any system that interacts with users—whether for customer service, education, or healthcare advice—must adhere to rigorous standards that ensure user data is protected, the basis of AI decisions is transparent, and any potential biases are adequately addressed.

The enforcement mechanism, highlighted by the potential for significant fines, reflects a broader EU strategy to not only promote ethical AI development but also to foster a digital ecosystem where trust in AI technologies is paramount. For companies developing conversational AI, this creates an imperative to integrate compliance measures into every stage of AI system design and deployment. It necessitates a proactive approach to understanding the Act's requirements, implementing robust data governance frameworks, ensuring transparency in AI-driven interactions, and continually monitoring AI systems for fairness and accuracy.

As AI technologies continue to advance and permeate various aspects of human life, the insights drawn from this analysis are vital in shaping a future where AI serves the public good, guided by principles of transparency, fairness, and accountability.

References

- Artificial Intelligence Act, The Act, https://artificialintelligenceact.eu/the-act/;
- Doe, J., Smith, A., The Impact of Regulatory Frameworks on AI Development: A Comparative Study, in Journal of AI Ethics, 5(2), 2023;
- European Commission, Regulatory framework for AI, https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai;
- European Parliament News, Artificial Intelligence Act: Deal on comprehensive rules for trustworthy AI, https://www.europarl.europa.eu/news/en/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai;
- European Parliament News, Artificial Intelligence Act: MEPs adopt landmark law, https://www.europarl.europa.eu/news/en/press-room/20240308IPR19015/artificial-intelligence-act-meps-adopt-landmark-law;
- European Parliament, EU AI Act: First regulation on artificial intelligence, https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence;
- European Parliamentary Research Service, Artificial Intelligence Act: An overview, https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2021)698792;
- Johnson, B., Conversational AI: Challenges and Opportunities in Ethics and Regulation, in Technology and Society, 30(4), 2023.