Policy Brief

To: Thomas Schneider, Chair of the Committee on Artificial Intelligence of the Council of Europe, and the **European Commission**

Consultancy Group 25: Paul Arnold, Annabel Claire Duport, Pirun Chan, Veronika Zadernivska, Jiang Feng

Disclaimer: *Due to the timeframe of the course, the policy paper only reflects the political debate up to the 5th December. Therefore, the recent trialogue negotiations between the 6th and 9th of December are not part of the analysis.*

SHIELDING EU SECURITY: MITIGATING AI RISKS FOR STRATEGIC PROTECTION

WHERE TO BEGIN

The emergence of **Generative AI** signifies a ground-breaking **technological leap** but brings forth critical concerns around **privacy**, **security**, and **discrimination**. Evaluation of the EU's **AI Act (AIA)** reveals incomplete coverage, leaving **regulatory gaps** due to **legislative ambiguities** and conflicting interests. To address this, we propose immediate interventions: a **tiered risk-based model**, an independent **AI Office establishment**, and strengthened collaboration with **social media platforms**. These recommendations aim to fortify **regulatory frameworks**, fostering a balanced environment that guards against AI misuse while nurturing **innovation**.

WHAT WE ARE ALREADY FACING

The rise of **Generative AI** signifies a profound transformation in how daily tasks are executed, epitomized by the introduction of **ChatGPT** in **November 2022** by **OpenAI**.

However, the rapid expansion of Generative AI raises substantial concerns, including **harmful content**, **potential biases**, and **over-reliance on technology**.

Instances of inadvertent **copyright violations** or **academic plagiarism** by users of ChatGPT have been resolved through transparency. Yet, more severe cases involving **misinformation**, **misleading content distribution**, and even **criminal acts** using this technology highlight the inadequacy of classifying Generative AI as "limited risk."

These serious issues advocate for a broader conversation beyond mere transparency, emphasizing the insufficiency of the current classification in addressing the significant challenges posed by Generative AI.

CURRENT EU`S REACTION TO AI ISSUES

- EU's Act on AI**, introduced in 2021, categorized AI systems into four risk levels and lacked specificity on** Generative AI**.**
- June 2023 amendments **proposed stricter obligations for Generative AI:** transparency measures**,** data governance**, and** mandatory registration**.**
- Deadlock in trilogue negotiations **due to an alliance opposing strict rules for Generative AI, revealing** tensions between **economic interests** and **civil rights protection.**

FLAWS IN RESPONSES TO AI THREATS

- Alliance's proposal **delegates civil rights protection to tech companies, risking** democratic accountability **and prioritizing profit motives.**
- Lack of democratic enforcement **raises concerns of unnoticed** wrongdoing **and** misconduct**, challenging the EU's goal as a leading AI regulator.**
- EU Parliament's proposal **aims to prevent Generative AI misuse but faces challenges in** content labeling **and** stricter transparency rules **potentially burdening small enterprises.**
- Risk of AI innovation leaving the EU **for regions with lax regulations poses a threat, emphasizing the need for a delicate balance between** fostering innovation **and** protecting rights**.**

WHAT WE RECOMMEND

The outlined strategies propose comprehensive solutions to tackle the challenges posed by Generative AI:

1. **Implement a tiered Risk-Approach for Generative AI Models:** Advocating for explicit and strict obligations for Generative AI providers aligned with the EU Parliament's proposal. A tiered approach based on the size or power of the AI models should aim to prevent disproportionate restrictions on smaller AI developers, fostering innovation while ensuring compliance for established players.

2. **Establish an AI Office for Cooperation and Innovation:** Proposing the establishment of an independent AI Office to address unique challenges in sectors like healthcare, gaming, and the military. Emphasizing the importance of frequent regulatory reviews and collaboration with AI providers to create a dynamic and globally harmonized regulatory framework.

3. **Strengthen cooperation with businesses on Social Media Awareness:** Urging the extension of transparency guidelines to regulate the distribution of AI-generated content, irrespective of model size. Emphasizing collaboration between the EU and social media platforms for real-time fact-checking, mandatory labeling, and public awareness initiatives to combat identity theft, misinformation, and copyright infringement.