European Commission
Rue de la Loi 200
1000 Brussels
Belgium

25 June 2025

—Addressed to: President of the European Commission, Ursula von der Leyen—

Ensuring GPAI Rules Serve the Interests of European Businesses and Citizens

As 2 August 2025 nears, we–AI researchers and representatives from civil society, industry and academia–call on EU leaders to resist pressure from those attacking the rules on general-purpose AI (GPAI). The EU can show it is possible to offer industry innovation-friendly tools like the Code of Practice, without compromising on health, safety & fundamental rights. Robust GPAI rules will help ensure the EU AI Act effectively delivers on its objectives of promoting the uptake of human-centric and trustworthy AI, while mitigating systemic risks from GPAI. This is also a step towards technology sovereignty.

With this letter, we also reiterate calls from the European Parliament[1] to ensure the AI Office possesses adequate centralised expertise, capacity and mandate to act quickly in light of AI technological advancements. Without this, the EU will struggle to uphold the rule of law and protect its values, as AI is integrated throughout our society.

The GPAI obligations are critical to set the bar for responsible GPAI model development and deployment. Systemic risks from GPAI models continue to advance, with recent releases from OpenAI[2], Anthropic[3] and Google[4] showcasing new dangerous capabilities related to cyber, biological, radiological and nuclear threats. Reflecting warnings from the International AI Safety Report 2025[5], the pace of such developments will likely increase over the next few years[6]. At the same time, major GPAI model providers have drastically scaled back transparency and the rigour of safety-testing around model releases[7,8,9].

To facilitate fulfilling these obligations, the Code has been carefully crafted over 9 months by independent experts, involving over 1000 stakeholders, with most commitments applying to 5-15 large companies. Applying proportionate compliance burden, protecting downstream businesses, and overlapping with risk management practices already carried out by most large GPAI providers[10,11], its commitments set a crucial minimum bar for GPAI risk management.

Establishing sound and streamlined risk management, the Code also helps on other aims of the

---

[1] Euractiv: Getting serious about AI rules: Lack of enforcement capacity puts EU at risk
[2] OpenAI's Deep Research System Card
[3] Anthropic's Claude 4 Model Card
[4] Google's Gemini 2.5 Pro Preview System Card
[5] International AI Safety Report 2025
[6] Google's AlphaEvolve: A Gemini-powered coding agent for designing advanced algorithms
[7] FT: OpenAI slashes AI model safety testing time
[8] Fortune: Google's latest AI model is missing a key safety report in apparent violation of promises made to the U.S. government and at international summits
[9] Fortune: OpenAI updated its safety framework—but no longer sees mass manipulation and disinformation as a critical risk
[10] Existing Industry Practice for the EU AI Act's General-Purpose AI Code of Practice Safety and Security Measures
[11] SaferAI: G7 Hiroshima AI Process Code of Conduct and EU AI Act GPAI - Commonality Analysis

GPAI rules: reducing liability risks for EU companies, providing legal certainty, and most importantly fostering AI uptake in the EU. As suggested by A16z investor Anjney Midha, the main obstacle in Europe is not regulation but rather insufficient adoption[12]. Clear and robust rules provide the assurance and trust European businesses need, echoed by a Boston Consulting Group and MIT Sloan survey in which 73% of experts—including representatives from the Schwarz Group, H&M and EnBW—said GPAI providers can and must be held accountable[13]. In addition, a weak GPAI regime shifts the burden of AI Act compliance to startups and SMEs. Mitigating risks at the upstream model level once, where providers have more compliance resources, is far more efficient than doing so potentially thousands of times downstream.

As an important pillar of the European approach to ensuring democratic oversight over the most advanced AI, we urge you to further consider the following three elements as critical components of a future-proof governance regime for GPAI with systemic risk in Europe:

1. **Mandatory third-party testing in the Code of Practice.** GPAI models with systemic risk must undergo external assessment by qualified external assessors with sufficient time and access. This helps (i) prevent internal pressure to downplay risks in favour of rapid deployment, (ii) ensure mitigations are truly effective – especially as many current safeguards can be bypassed, and (iii) stimulate a third-party assurance ecosystem, providing trust to businesses and citizens that will in turn foster AI adoption and diffusion across sectors.
2. **Robust and multi-stakeholder review mechanisms.** Without a dynamic review mechanism that can respond to emerging risks and evolving state-of-the-art practices, the Code risks obsolescence before implementation. This mechanism must enable swift adaptation to new risk categories, emerging safety practices, and evolving technical standards. Moreover, and in line with the recommendations issued by the Chairs and Vice Chairs in draft three of the Code[14], an emergency update mechanism must be ensured to prevent imminent threats of large-scale irreversible harms or to mitigate their negative effects.
3. **Meaningful enforcement powers for the AI Office.** Without adequate capacity and expertise, the EU AI Office will not be able to keep pace with technological developments. The AI Safety unit (DG CNECT A3) should be scaled up to 100 staff, with the full implementation team expanding to 200. This is in line with the DSA team in terms of quantity. Such quantity must be matched by quality: world-leading AI safety and security experts should be targeted by the European AI Office and for other governance bodies like the Scientific Panel.

European citizens and businesses deserve AI that is developed and deployed with their safety, rights, and interests as primary considerations. The GPAI regime must establish this principle in practice. We call upon you to resist pressures that would compromise this essential regulatory framework, and to finalise a Code that reflects the EU's commitment to a positive AI future.

---

[12] A16z's Anjney Midha for Sifted

[13] MIT Sloan and BCG: How to Hold General-Purpose AI Producers Accountable

[14] Appendix 2: Code of Practice Draft Three

Sincerely,

- Daron Acemoglu - Nobel laureate, Economics 2024; Institute Professor, MIT
- Geoffrey Hinton - Nobel laureate, Physics 2024
- Stuart Russell - Distinguished Professor of Computer Science, University of California, Berkeley
- Prof. dr. Philipp Hacker - Europa-Universität Viadrina Frankfurt
- Prof. dr. Martin Vetterli, Professor and President Emeritus - Ecole Polytechnique Fédérale de Lausanne (EPFL)
- Prof. dr. Carina Prunkl - Utrecht University
- Prof. dr. Vincent Corruble - Sorbonne Université
- Prof. dr. Raja Chatila - Sorbonne Université
- Prof. dr. D.K. (Daniel) Mügge - University of Amsterdam
- Prof. dr. Teresa Scantamburlo - Ca' Foscari University of Venice
- Prof. dr. Federico L.G. Faroldi - University of Pavia
- Prof. dr. Lode Lauwaert - KU Leuven
- Prof. dr. Florian Tramèr – ETH Zurich
- Lennig Pedron – CEO Trust Valley initiative, program Director Ecole Polytechnique Fédérale de Lausanne (EPFL) Innovation Park
- Catelijne Muller, LL.M - President and Co-founder ALLAI
- Dr. Nathalie Smuha - Adjunct Professor of Law, NYU; Assistant Professor, KU Leuven
- Dr. Sören Mindermann - Scientific Lead, International AI Safety Report
- Dr. Sebastian Hallensleben
- Dr. Andrejs Vasiljevs - Co-Founder, Tilde, Latvia
- Dr. Joachim Bühler - CEO, TÜV-Verband
- Dr. Florian Mai - University of Bonn
- Dr. Leonard Dung - Ruhr-Universität Bochum
- Dr. Rania Wazir - Co-founder & CTO, leiwand.ai
- Dr. Nada Madkour - Non-Resident Research Fellow, Center for Long-Term Cybersecurity
- Dr. Deepika Raman - Non-Resident Research Fellow, Center for Long-Term Cybersecurity
- Adam Leon Smith - Project Leader, CEN-CENELEC JTC 21
- Demetrius Floudas - University of Cambridge; and Immanuel Kant Baltic Federal University
- Martin Hullin - Director, Digitalization and the Common Good, Bertelsmann Stiftung
- Robin Berjon - Technologist, Supramundane Agency
- David Evan Harris - University of California, Berkeley
- Jessica Newman - Founding Co-director, University of California Berkeley AI Policy Hub
- Krystal Jackson - Non-Resident Research Fellow, Center for Long-Term Cybersecurity
- Evan R. Murphy - Director, AI Governance & Safety Canada
- Lisa Soder - interface - Tech analysis and policy ideas for Europe

- TIC Council
- AppliedAI Institute for Europe

- FARI - AI for the Common Good Institute, Brussels
- Apart Research
- International Association for Safe and Ethical AI
- Center for Human-Compatible AI
- Pour Demain
- The Future Society
- Future of Life Institute
- Ada Lovelace Institute
- Avaaz
- AI Standards Lab
- Holtman Systems Research
- SaferAI
- Observatorio de Riesgos Catastróficos Globales
- ALLAI
- Centre pour la Sécurité de l'IA
- OAISIS/Third Opinion - Supporting AI Whistleblowers