Feedback on the Scientific Panel of Independent Experts Implementing Regulation

15th November 2024

Policy Team
policy@futureoflife.org

Feedback on the Scientific Panel of Independent Experts Implementing Regulation

Organization

Future of Life Institute

Point of Contact

policy@futureoflife.org

About the Organization

The Future of Life Institute (FLI) is an independent nonprofit organization with the goal of reducing large-scale risks and steering transformative technologies to benefit humanity, with a particular focus on artificial intelligence (AI). Since its founding, FLI has taken a leading role in advancing key disciplines such as AI governance, AI safety, and trustworthy and responsible AI, and is widely considered to be among the first civil society actors focused on these issues. FLI was responsible for convening the first major conference on AI safety in Puerto Rico in 2015, and for publishing the Asilomar AI principles, one of the earliest and most influential frameworks for the governance of artificial intelligence, in 2017. FLI is the UN Secretary General's designated civil society organization for recommendations on the governance of AI and has played a central role in deliberations regarding the EU AI Act's treatment of risks from AI. FLI has also worked actively within the United States on legislation and executive directives concerning AI. Members of our team have contributed extensive feedback to the development of the NIST AI Risk Management Framework, testified at Senate AI Insight Forums, briefed the House AI Task-force, participated in the UK AI Safety Summit, and connected leading experts in the policy and technical domains to policymakers across the US government.

The Future of Life Institute (FLI) welcomes the opportunity to provide feedback on the draft implementing regulation establishing a scientific panel of independent experts in artificial intelligence. While we commend the Commission's efforts to establish this crucial oversight mechanism, we identify several areas where the regulation could be strengthened to ensure the panel's effectiveness and enable comprehensive coverage of systemic risks stemming from general-purpose AI (GPAI) models and systems.

Article 3: Selection criteria and composition of the scientific panel

> ### ARTICLE 3(3)(A)
>
> "Experts shall be selected with regard to the need to ensure:
>
> (a) multidisciplinary, adequate, and up-to-date scientific or technical expertise in the field of artificial intelligence. Such expertise shall include, but is not limited to, adequate understanding or experience working with (i) machine learning architectures, training methods, and optimisation techniques; (ii) AI model and system evaluation, testing, and benchmarking; (iii) AI safety engineering, robustness assessments and systemic risk assessments; (iv) cybersecurity and privacy-preserving techniques; (v) expertise pertaining to applied sectors, fundamental rights and equality, AI ethics, social sciences, and other relevant fields. The composition shall ensure proportionate representation across these different fields of expertise, with particular attention to maintaining a sufficient depth of technical expertise to evaluate advanced general purpose AI models and their capabilities."

Specifying technical domains shows intent of purpose for a technical audience and helps attract the best talent. It ensures the panel can effectively assess current and emerging systemic risks stemming from GPAI models and systems. The composition of expertise should target those domains presenting the greatest systemic risks, while ensuring sufficient breadth and coverage of the diverse range of possible systemic risks. This should account for the cross-cutting nature of GPAI, including the interaction and possible amplification of systemic risks therein.

> ### ARTICLE 3(4)
>
> "The selection of experts shall ensure that at least one ~~and no more than three~~ nationals from each Member State of the Union and each member of the European Free Trade Association which is a member of the European Economic Area are appointed as experts to the scientific panel, provided that there are applicants from that country who satisfy the criteria stipulated in the call. ~~Nationals from Member States of the Union and members of the European Free Trade Association which are members of the European Economic Area shall constitute at least four-fifths of the experts on the scientific panel."~~

It is important to ensure that there is at least one national from each EU and EFTA country at a minimum to allow for fair and balanced European representation, and to attempt to avoid the possibility that certain Member States would be overly or disproportionately represented. This leaves 29 spaces in the scientific panel that can be allocated to qualified experts from the EU and EFTA – or from third countries. Given that advanced expertise on AI risks is generally scarce globally, and AI risks are borderless, this allows the panel to reflect the broader international AI community, while still guaranteeing that experts from the EU and EFTA have the casting vote in a simple majority (as recommended below) when deciding to issue a qualified alert, when recommending a Chair and Vice-Chair, or when adopting rules of procedure.

Article 4: Term of office

> **ARTICLE 4(3)**
>
> "Where an expert is dismissed during his or her term of office, a replacement for that expert shall be appointed by the Commission following a call for expression of interest ~~for the remainder of the term from the reserve list~~."

A reserve list for replacements would limit access to the most up-to-date scientific or technical expertise, as required by Article 68(2) of the AI Act. A fresh call for expression of interest provides opportunities for the involvement of other experts, who may have been unavailable during a previous call due to other commitments, or who may have gained relevant or further experience since the earlier selection process.

Article 7: Performance of tasks and preparation of documents

> **ARTICLE 7(2)**
>
> "For recommendations, opinions, or qualified alerts issued pursuant to Articles 68(3) and 90 of Regulation (EU) 2024/1689, the Secretariat, in consultation with the Chair of the scientific panel, may appoint a rapporteur and one or more contributors to prepare such recommendation, opinion or qualified alert. Where a rapporteur is appointed, the Secretariat shall ensure that contributors have complementary expertise to the rapporteur to support the preparation of recommendations, opinions or qualified alerts. When preparing recommendations, opinions or qualified alerts, the Scientific Panel shall also consult civil society, academia, industry representatives and other relevant parties."

Explicit reference to broad consultations is crucial to ensure balanced representation of societal interests, as qualified alerts guide the Commission's decision on whether models pose systemic risks, requiring increased oversight. Gathering diverse stakeholder input not only improves the quality and legitimacy of the panel's outputs, but also helps to identify complex interactions between GPAI models and systemic risks, which may be missed in more narrow consultations. The complementarity of diverse external input to the scientific panel's work also applies to its work internally, which should seek to leverage synergies between different domain experts to address the spectrum of systemic risks when preparing a recommendation, opinion or qualified alert.

> **ARTICLE 7(5)**
>
> "The list of attendees, the agenda, the minutes, and the conclusions of such hearings shall be made publicly available on a dedicated Commission website without undue delay and at the latest within two weeks."

This aligns with the Digital Services Act and its establishment of the European Board for Digital Services, whose procedures require publishing agendas, minutes, and attendance lists.

Article 10: Independence, impartiality and objectivity

> **ARTICLE 10(3)**
>
> "They shall make a declaration of interest indicating any interest which may compromise or may reasonably be perceived to compromise their independence, impartiality and objectivity, including any relevant circumstances relating to their close family members. All declarations of interest shall be made public on a dedicated Commission website without undue delay and at the latest within two weeks."

Declarations of interest must be publicly available, as per Article 68(4) of the AI Act. This ensures accountability, enables external verification of independence, and builds trust in the scientific panel's impartiality.

Article 12: Transparency

> **ARTICLE 12(1)**
>
> "The Secretariat shall make available to the public on a dedicated Commission website, without undue delay:
>
> - recommendations and opinions prepared by the scientific panel;
> - qualified alerts prepared by the scientific panel, excluding information that would identify a specific general purpose AI model or provider of a general purpose AI model;
> - the priorities and focus areas, including capabilities, systemic risks, and types of models, of the new term of the scientific panel following the appointment of experts;
> - the assigned responsibilities and duties of the rapporteurs and contributing experts when they have been appointed to prepare recommendations, opinions and qualified alerts."

Timely public disclosure of the panel's activities allows stakeholders to validate and build upon the panel's systemic risk assessments, which will support broader AI safety research. This scrutiny encourages evidence-based recommendations, opinions and qualified alerts, while providing assurance to civil society, industry and GPAI providers alike that assessments are independent, rather than arbitrary.

Article 16: Conditions for granting access to received documentation or information

> **ARTICLE 16(2)**
>
> Access to the requested information shall be restricted to the rapporteur and designated contributing experts working on the specific task and shall be limited in time to an appropriate duration for the purpose of fulfilling the relevant task plus 30 days with the possibility for extension upon a duly justified request."

This enables effective collaboration and sufficient time for thorough analysis, while maintaining security and providing clear temporal boundaries for information access.

Article 17: Procedure to issue qualified alerts

ARTICLE 17(1)
"A decision by at least a ~~qualified~~ simple majority of the members of the scientific panel shall be necessary to issue a qualified alert to the AI Office..."

Since qualified alerts don't automatically trigger systemic risk designation for a model by the Commission, a qualified majority is an unnecessarily high bar, preventing the scientific panel from meaningfully supporting the Commission's monitoring. Alerts will provide concrete evidence of Union-level risks. This critical information, necessary to protect health, safety, and fundamental rights, may be missed due to the difficulty of achieving a weighted majority. Consensus can still be reflected via a simple majority, which is less likely to lead to deadlock and protracted deliberations, and better reflects the precautionary principle needed for emerging technologies. Faster alert mechanisms are essential to address the rapid scaling of GPAI models and a fast-evolving AI market.

Article 18: Handling of qualified alerts

ARTICLE 18(1)
"The AI Office shall evaluate the qualified alerts issued pursuant to Article 90(1) of Regulation (EU) 2024/1689 and take a decision whether to launch any measures as provided for in Article 91 to 93 of that Regulation or not to follow-up on the alert. The AI Office shall make publicly available a summary of their decision on whether to launch any measures or not following a qualified alert, taking into account the need to protect trade secrets and business confidentiality."

Beyond greater transparency, accountability, and public safety awareness, informing the public about a planned designation would allow citizens and businesses alike to plan accordingly, particularly as GPAI models form the basis for many different downstream applications.