15 January 2025

COMMENTS IN RESPONSE TO EUROPEAN COMMISSION CALL FOR EVIDENCE SURVEY ON "SECOND DRAFT GENERAL PURPOSE AI CODE OF PRACTICE FOR PROVIDERS OF MODELS WITH SYSTEMIC RISK"

The Association for Computing Machinery (ACM) is the world's longest established professional society of individuals involved in all aspects of Computing. It annually bestows the ACM A.M. Turing Award, often popularly referred to as the "Nobel Prize of Computing." ACM's Europe Technology Policy Committee ("Europe TPC") is charged with and committed to providing sound **technical information** to policy makers and the general public in the service of sound public policymaking. Europe TPC has responded to the European Union stakeholder's consultations in the past in the context of the AI Act[1], the Data Act[2], the Digital Services Act[3][4], the Digital Citizen Principles[5], the Cyber Resilience Act[6], amongst others[7]. ACM and Europe TPC are non-profit, non-political, and non-lobbying organisations.

Europe TPC is pleased to respond to the European Commission's call for feedback launched on 19 December 2024 on the European Union's **"Second Draft General Purpose AI Code of Practice for Providers of Models with Systemic Risks"**. Europe TPC supports the European Commission's aim to establish a guiding document for providers of general-purpose AI models when demonstrating compliance with the AI Act along the full life cycle of the models with systemic risk, through four Working Groups working in close collaboration with a pool of experts.

Europe TPC is gratified that eight (8) out of twelve (12) initial recommendations made for the European Commission's **First Draft** are integrated in the main text of the **Second Draft**. This response extends the previous twelve (12) recommendations[8] to reflect the latest changes in the second draft, and proposes two (2) new recommendations[9].

[1] https://www.acm.org/binaries/content/assets/public-policy/europe-tpc-comments-ai-consultation.pdf
[2] https://www.acm.org/binaries/content/assets/public-policy/acm-eur-tpc-data-act-comments-13may22a.pdf
[3] https://www.acm.org/binaries/content/assets/public-policy/europetpc-digital-services-act-comments.pdf
[4] https://www.acm.org/binaries/content/assets/public-policy/acm-europe-tpc-dsa-comments.pdf
[5] https://www.acm.org/binaries/content/assets/public-policy/europetpc-comments-digital-principles.pdf
[6] https://www.acm.org/binaries/content/assets/public-policy/acm-europe-tpc-cyber-reslience-comments-pdf
[7] https://www.acm.org/public-policy/public-policy-statements
[8] Europe TPC has extended the eight (8) recommendations that were integrated to reflect the latest changes in the second draft and reiterates the four (4) recommendations that are not integrated in the second draft, requesting for these to re-considered for the third draft (including recommendations 1, 2, 6 and 11).
[9] The net-new recommendations are recommendation 13 and recommendation 14.

Global Recommendations

- **Recommendation 1** - Although the document aims to introduce guidance for model *providers* with systemic risks only, a significant portion of the document (Measures 6+) seems to directly depend on the *deployment* and use-cases of the models in the context of encompassing AI systems. Many measures reference expectations for model *deployers*, but provide them under a Code of Practice that is targeted at model *providers*. The main recommendation from Europe TPC is that the code of practice should stay within the limits of its title "General Purpose AI Code of Practice for Model Providers with Systemic Risks". Europe TPC acknowledges that the European Commission and the Chairs have taken explicit action in the second draft to further focus the scope of the document towards Model Providers; however, Europe TPC would still call out that there is still a significant direct and indirect scope providing expectations towards Model Deployers. The Commission should therefore revisit the content of the Code of Practice to ensure that it encompasses solely the development and release stages of models by GenAI model providers, as opposed to encompassing both GPAI model providers and GPAI model deployers. Europe TPC recognises that it may be challenging to separate the two scopes from a technical and domain perspective; hence an alternative would be to expand the scope of the Code of Practice to the deployment of models and to re-calibrate the participation in the working groups to ensure appropriate representation from model deployer organisations as well.

- **Recommendation 2** - EuropeTPC recommends that the European Commission makes clear that the processes and mitigations in place for models with systemic risks are a superset of the mechanisms that would be required for models with high and de-minimis risk. If possible the commission should provide guidance on which processes are a MUST/SHOULD/COULD for models with systemic, high and de-minimis risk respectively. Furthermore, such guidance should consider the probability of the risk[10] and the impact of the outcome(s)[11].

- **Recommendation 3** - The European Commission has made a clear effort to clarify that the guidance that the Code of Practice extends across modalities beyond only text and image models. Europe TPC applauds further emphasis in the latest draft and recommends that this disambiguation be made explicitly in the main text.

- **Recommendation 4 -** The European Commission has made a clear effort to clarify that the guidance of the Code of Practice extends across ML-types beyond only text

[10] To better align with state of the art procedures in (operational or cyber) risk management, risk tiers should be defined in terms of frequency and severity. Europe TPC suggests adding the following clarification after the underlined text: "risk tiers should be defined by taking into account both the frequency (likelihood) and the expected impact (severity) of the identified risks"

[11] Any mitigations would need to be aligned with the probability and impact of the risk, encompassing a quadrant that can be suggested as part of the Code of Practice.

and image models. Europe TPC applauds further emphasis in the latest draft, and recommends that this disambiguation be made explicitly in the main text[12].

Measure-specific recommendations

- **Recommendation 5 -** In the context of *Measure 1*, Europe TPC acknowledges that the European Commission has extended these sections to provide guidance to model providers on defining the repeatable process required to reduce overhead when registering models and their respective metadata[13], particularly when multiple versions of a model may be released in a relatively agile manner.

- **Recommendation 6 -** In the context of *Commitment 3.1*, the General-Purpose AI Code of Practice Draft highlights that one use case treated as systemic risk is "Automated use of models for AI Research and Development". EuropeTPC recommends that this is reconsidered and removed as a systemic risk, as developments that "could greatly accelerate AI research and development" should not be seen as detrimental to the European scientific community and ecosystem. Furthermore, in *sub-measure 6.3.1. Dangerous model capabilities***,** Europe TPC points out that "Long horizon planning, forecasting and strategising" is a generic term. A significant subset of machine learning use cases, such as demand forecasting, are leveraged commonly in industries for low-risk use-cases. The Code of Practice should better qualify when these capabilities should be considered a dangerous model capability[14]. Europe TPC calls out that this recommendation was not considered in the revised version, and no further changes have been added; hence Europe TPC urges the commission to consider this recommendation.

- **Recommendation 7 -** In the context of *Measure 10*, EuropeTPC recommends that when it comes to robust evaluation methods, these should, at the very minimum, ensure that relevant technical and non-technical domain experts are involved to ensure fit-for-purpose evaluation. Europe TPC acknowledges that the latest draft makes an indirect reference through the term "Experts with relevant expertise", however, it is still recommended that this is made explicit in the main text as opposed to the current implicit statement.

- **Recommendation 8 -** In the context of *Sub-measure 10.3*, Europe TPC suggests that rather than purely seeking to establish a gold standard for operationalising high scientific rigour, the Commission identifies a set of parameters or metrics that can be examined for effectiveness on a regular basis. For example, such parameters or metrics could include (but not be limited to) verifiability of the reported findings, adherence to open data principles, declaration of competing interests, type of data

[12] Such as classification or regression.
[13] These templates could encompass model-fair-use templates which are compliant with AI Act requirements, similar to how standardised licenses exist for open source code (e.g. Apache, MIT, etc)
[14] The role of the Human In The Loop (HITL) should be a key consideration in any assessment of model capabilities deemed dangerous.

used and its relevance to the scientific endeavour, correlation between the inputs and outputs, appropriateness and relevance of methods used, approaches to classification and clustering of data, justifications of the research goals, and roles and responsibilities of those engaged in the scientific endeavour.

- **Recommendation 9 -** In the context of *Sub-measure 10.8*, Europe TPC acknowledges that the European Commission has made an explicit reference to ensuring alignment and compatibility with independent organisations such as the UK AI Safety Institute[15], the US AI Safety Institute[16], and others as channels, organisations and methods that would facilitate the sharing of evaluations, tools and best practices. These organisations can support the testing of high-risk models, as well as define standard frameworks that can be adopted to ensure robust and consistent testing. Europe TPC highlights the importance of the link and alignment with such organisations and recommends extending the scope to provide further references of such organisations as examples, as well as the type of expectations with such organisations.

- **Recommendation 10 -** In the context of *Sub-measure 12.2*, Europe TPC acknowledges that the European Commission has highlighted the importance of collaboration with standardisation bodies and organisations that publish and drive forward industry standards in AI & ML. Europe TPC would like to once again highlight a broad set of initiatives that are standardising taxonomies and risks within cybersecurity, such as the Institute for Ethical ML's MLSecOps framework[17], the MITRE Attack Framework[18], the OWASP Top 10 ML[19], and the UK National Cyber-security Centre Machine Learning Security Principles[20]; furthermore Europe TPC recommends that the European Commission set up respective initiatives to develop EU cybersecurity-specific standards by CEN, CENELEC or ETSI in support of legislation (AI Act) and the associated Code of Practice.

- **Recommendation 11 -** In the context of *Measure 14*, EuropeTPC recommends that the term "deployment" is revisited to disambiguate it in this context, as it seems to refer in this section to "making the model available for use". However, the term "deployment" in the AI Act is used in the context of "model deployers" who integrate a model into a production AI system. Based on this, a more accurate term would be to "release the model", as the section specifically suggests mitigating the "release" of models based on the constraints provided. Once a model is released, model deployers will be able to deploy and integrate it into their AI systems.

- **Recommendation 12** - In the context of *Sub-measure 18.2*, Europe TPC acknowledges that the European Commission has reflected that the consequences of any serious incidents need to be aligned with the type of capital (i.e. human, natural,

[15] https://www.aisi.gov.uk/
[16] https://www.nist.gov/aisi
[17] https://ethical.institute/security.html
[18] https://attack.mitre.org/
[19] https://owasp.org/www-project-machine-learning-security-top-10/
[20] https://www.ncsc.gov.uk/collection/machine-learning-principles

social, manufactured, or financial) impacted. Europe TPC recommends that the European Commission continues to emphasise that any identified corrective measures should be proportionate to the consequences of the serious incident on the impacted capital(s).

- **Recommendation 13** - In the context of *Measure 2*, Europe TPC recommends the creation of an "incident register" that can be used to record, track, and resolve known violations across dimensions such as copyright and security, among many others. This "incident register" can be used to extend the existing KPIs to ensure that documentation and auditability are in place and should be maintained for a period of time (e.g. 5 years). This incident register can be managed at a European Commission level and/or at an organisational level; this is a consideration that the expert working group should define, including the responsibility and accountability of the register being kept up to date.

- **Recommendation 14** - In the context of *Measure 12.1*, Europe TPC supports the European Commission's intent to encompass important security best practices and recommends extending this measure to include the need to establish a plan for incident response and handling of security breaches.