Consultation Response

Consultation: Commission Guidelines to Clarify the Scope of the General-purpose AI Rules in the AI Act

15 May 2025

The Association for Financial Markets in Europe (AFME) welcomes the opportunity to comment on the **European Commission Guidelines to Clarify the Scope of the General-purpose AI Rules in the AI Act**. AFME represents a broad array of European and global participants in the wholesale financial markets. Its members comprise pan-EU and global banks as well as key regional banks, brokers, law firms, investors and other financial market participants. We advocate stable, competitive, sustainable European financial markets that support economic growth and benefit society.

AFME is the European member of the Global Financial Markets Association (GFMA) a global alliance with the Securities Industry and Financial Markets Association (SIFMA) in the US, and the Asia Securities Industry and Financial Markets Association (ASIFMA) in Asia.

AFME is registered on the EU Transparency Register, registration number 65110063986-76.

We summarise below our high-level response to the consultation, which is followed by answers to the individual questions raised.

Executive Summary

This response outlines the concerns of members regarding the use of compute as the sole determinant for a GPAI designation. It also clarifies that, while the 10^22 threshold is sufficient at the current time, it is desirable to refine this metric to ensure it remains futureproof as technology develops.

Additionally, members request clarification that modification, eg. fine-tuning, of a GPAI foundation model would be deemed a business-as-usual Deployer activity and would not give rise to a designation of a GPAI Provider.

Lastly, it highlights the need for clear definitions of open-source and clarifications of what constitutes placing a GPAI model "on the market", especially within the context of deployments within group entities under the same corporate umbrella.

Questions

1) Many entities will have to assess the general-purpose nature of their models to determine whether they need to follow the obligations for providers of general-purpose AI models. A pragmatic metric is thus highly desirable to limit the burden, especially on smaller entities. Do you agree that training compute is currently the best metric for assessing generality and capabilities, despite its various shortcomings?

No

Please explain why and which alternatives may be preferable.

Association for Financial Markets in Europe
London Office: Level 10, 20 Churchill Place, London E14 5HJ, United Kingdom T: +44 (0)20 3828 2700
Brussels Office: Rue de la Loi 82, 1040 Brussels, Belgium T: +32 (0)2 883 5540
Frankfurt Office: c/o SPACES - Regus First Floor Reception Große Gallusstraße 16-18 60312 Frankfurt am Main, Germany T:+ 49 (0)69 710 456 660
www.afme.eu

Compute may not be a reliable and sustainable proxy for assessing model generality or capabilities when considered in isolation. Our members have concerns regarding its long-term relevance given rapid hardware and software developments. Suggestions for complementary criteria include model architecture (specifically the number of input sources e.g. text/videos/images) and task range to ensure a more accurate and future-proof assessment framework.
From a practical perspective, while compute can be estimated, precise measurement is not always possible. It is also likely that hardware (i.e. GPU chip providers) and cloud providers may need to update their service terms to provide this additional data (the way it is done for CO2 emissions, as an example).
Additionally, in certain systemic GPAI model modification contexts, a compute-based threshold creates compliance dependencies across the value chain as the modifiers of models must rely on the representations of and classification by upstream providers.

It should also be noted that compute is not always a metric that model providers are willing to disclose which causes challenges for members to comply with their obligations under the AI Act.

2) Is 10^22 FLOP a reasonable threshold for presuming that a model is a general-purpose AI model?
Yes

3) With the proposed threshold of 10^22 FLOP, or your alternative threshold suggested above, how many models and how many entities do you expect to be in scope of the AI Act, and why?
While the currently proposed threshold is suitable and sufficiently high to capture significant models or modifications at the current time, we reiterate our concerns in Q1 re using compute as a sole measure and the consequent futureproofing of such a metric.

Additionally, as stated above, it should be noted that many model providers do not disclose underlying compute metrics, hence it can be challenging to precisely determine compute usage.

It is also the view of our members that downstream users of GPAI models should not be classed as "Providers" based solely on compute levels. This is also important, to ensure that supervision is limited to companies that genuinely play the role of providers.

4) In addition to the examples presented in section 3.1.1 of the working document, are there other examples for which it would be important to clarify whether the presumption of being a general-purpose AI model based on the training compute threshold may be rebutted?
Regulated, well-defined use cases (e.g. credit assessment or compliance tasks) should not be presumed to be GPAI solely based on compute levels.
It is also the view of our members that downstream users of GPAI models should not be classed as "Providers" based solely on compute levels. This is also important, to ensure that supervision is limited to companies that genuinely play the role of providers.

5) Besides the criteria presented in Section 3.1.2 of the working document, are there other criteria that can be used to determine whether iterations, instances, or derivates of a model constitute distinct models for the purposes of the AI Act?
Fine-tuned or modified models should not be considered a new or distinct model from the original.

Beyond compute-based criteria, it is suggested that modifications such as fine tuning should significantly expand the range of tasks the model can perform or the modalities it supports, in order to be considered a distinct model.

8) Many downstream modifiers will have to assess whether they need to comply with the obligations for all providers of general-purpose AI models and the obligations for providers of general-purpose AI models with systemic risk. A pragmatic metric is thus highly desirable to limit the burden on downstream modifiers having to make this assessment, especially on smaller entities. Do you agree that training compute is currently the best metric for quantifying the amount of modification, despite its various shortcomings?

YES

9) Are there examples of modifications of general-purpose AI models that meet the proposed training compute threshold of 3x10^21 FLOP, yet which should not result in the downstream modifier being considered a provider?

- Fine-tuned or adapted models performing narrow, domain-specific tasks should be excluded from GPAI obligations even if based on larger underlying foundation model. This would be viewed as a business-as-usual Deployer activity not a Provider activity.
- Internal-facing tools, such as chatbots accessed via internal web interfaces, should not be classified as GPAI merely due to how they are accessed.
- For the definition of 'placing on the market', we would request clarification of whether this would include providing a GPAI service to another entity under the same corporate umbrella. In our members' view, that should not be the case, as a firm would not be placing a service on the 'EU marketplace', but rather for internal/ intragroup purposes.
- In addition to the above, it is our members' view that modifications such as fine tuning should significantly expand the range of tasks the model can perform or the modalities it supports, before the modifying entity could be considered a provider.

11) In addition to the examples presented in section 3.3.1 of the working document, are there other examples of when a general-purpose AI model should be considered as being placed on the market?

The examples given currently are too vague – for example: 'A general-purpose AI model is integrated into a chatbot made available via a web interface'. These examples are somewhat simplistic and hence their value is questionable. Additionally, examples of what is definitely not considered placing on the market of general-purpose AI models would be of value to aid clarity.

We would also welcome clarity that the definition for 'free and open-source software' in the AI Act will be aligned to the definition provided in the Cyber Resilience Act and the further clarity being provided in expected Guidance (Article 26) regarding the application of the Act to open-source software. There are instances of manufacturers claiming to provide open-source software, however, restrictive licensing practices can significantly limit how firms can utilise their models. This undermines the open-source attributes of the model and contravenes the definition of open-source software used in the Cyber Resilience Act. These manufacturers who do not align with the Cyber Resilience Act should not benefit from an exemption.

Lastly, for the definition of 'placing on the market', we would request clarification of whether this would include providing a GPAI service to another entity under the same corporate umbrella. In our

members' view, that should not be the case, as a firm would not be placing a service on the 'EU marketplace', but rather for internal/ intragroup purposes. Further, the Consultation Document references the Blue Guide. Section 2.3 (Placing on the Market) of the Blue Guide states that a manufacturer is not considered to be placing a product on the market where a product is manufactured for one's own use. Assuming "manufacturer" in this context is equivalent to "provider", this would support our view above.

AFME Contacts
Amandeep Luther
AFME AI Lead
Amandeep.luther@afme.eu