Consultation Response

Stakeholder Consultation on the Draft Guidelines on the Classification of High-risk AI Systems under Article 6 of the AI Act

23 July 2026

The Association for Financial Markets in Europe (AFME) welcomes the opportunity to comment on the **Stakeholder consultation on the Draft Guidelines on the classification of high-risk AI systems under Article 6 of the AI Act** (the "Draft GLs"). The Association for Financial Markets in Europe (AFME) is the voice of the leading banks in Europe's financial markets, providing expertise across a broad range of regulatory and capital markets issues. We represent over 150 leading global and European banks and other significant market players. Our members play a vital role in Europe's financial ecosystem, underwriting around 90% of European corporate and sovereign debt, and 85% of European listed equity capital issuances. Importantly, AFME members are market makers, providing liquidity, which is essential for ensuring financial markets can function efficiently. We also represent law firms and other associate members which advise market participants and support AFME's legal and regulatory initiatives.

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

General principles for classification of high-risk AI systems under Article 6 (Section II of the draft Guidelines)

II.1 The system must be an AI system

Question A. Are there any aspects / paragraphs of this Section of the draft Guidelines that you believe require further clarification, and if so, how would you suggest they be improved to ensure effective implementation and compliance? Please indicate specific parts of the draft Guidelines (i.e., paragraph(s) number).

We reiterate that traditional statistical models, i.e. static logistic regression models, should not qualify as AI systems, as they are deterministic, non-adaptive, rely on manually selected variables, have high explainability, are not autonomous and are under full human control. This is consistent with ECB Opinion CON/2026/10, which says that highly explainable models lacking black-box characteristics should not be subject to the high-risk regime solely due to their use case. The definition needs to be precise and clear enough to prevent divergent interpretation across Member States, which would run counter to the aims of the Act and the broader goals of the

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

Company Registration No: 6996678 Registered Office: Level 10, 20 Churchill Place, London E14 5HJ
AFME is registered on the EU Transparency Register, registration number 65110063986-76

digital single market. If traditional statistical models are to be included in the definition, we suggest that they should at least not qualify as high-risk.

However, the Draft GLs appear to apply the Article 6(3) exemption filter to systems that may not qualify as AI systems under Article 3(1) AI Act in the first place. For instance, the grade calculator in paragraph 225(c) performs a deterministic weighted average: it infers nothing, learns nothing, and produces a fully transparent arithmetic output.

Finally, will the forthcoming Art. 25 GLs address group-internal scenarios specifically?

II.2 Intended purpose(s) of the AI system

Question A. Are there any aspects / paragraphs of this Section of the draft Guidelines that you believe require further clarification, and if so, how would you suggest they be improved to ensure effective implementation and compliance? Please indicate specific parts of the draft Guidelines (i.e., paragraph(s) number).

Getting the intended purpose and associated responsibilities right is a key priority for this consultation, particularly for financial institution who are likely to be deployers of AI systems (rather than providers) in most instances.

We suggest it would be preferable if a system is not suitable for high-risk use cases unless expressly indicated. More fundamentally, the GLs should reflect the distinction deliberately established by the AI Act between (1) GPAI models or systems and (2) high-risk AI systems, which are subject to separate compliance regimes under Chapter V and Chapter III, respectively. To deem all GPAI systems as capable of high-risk use contradicts the Level 1 text approach and would mean that deployers must comply with verification requirements and report where a high-risk AI system is non-compliant. In many cases, this could be simply down to oversight of the GPAI System provider and would lead to unnecessary notifications and obligations on the provider. Furthermore, in line with the intended purpose definition, there should be a level of intentionality by the provider to invoke high risk requirements.

At the same time, where high-risk deployment is reasonably foreseeable, or where a downstream deployer is deemed to assume provider-like responsibilities because it configures, integrates or deploys a GPAI model or system for a high-risk use case, the scope of those responsibilities should be aligned with what is established in the AI Act and calibrated to the deployer's effective ability to comply. High-risk obligations should not be imposed on downstream deployers beyond what they can reasonably fulfil. This should depend, in particular, on the degree of control exercised by each party over the relevant elements of the system, the technical information available to the deployer, the safeguards implemented or supported by the upstream provider, and the practical means available to the deployer to manage the associated risks. Such obligations should also be limited to the specific substantial modifications implemented by downstream operators or those aspects of the AI system over which they have effective control.

This is particularly important for regulated entities relying on externally sourced GPAI models or systems, where key elements such as model design, training data, testing, robustness, updates and technical limitations remain under the control of the upstream provider.

Therefore, the current approach should be aligned with the specific treatment of GPAI models in the Level 1 text. Additionally, the GLs should at least clarify:

- That the absence of an express exclusion of high-risk uses should not, by itself, lead to a presumption that such uses form part of the intended purpose. Classification should be based on the provider's intended use and not on documentary omissions alone. The presumption should apply only where there is positive evidence that the provider actively promotes or facilitates a high-risk use case. In addition, where access to a system is technically restricted to non-high risk use cases through appropriate controls, segmentation and governance measures, such restrictions should be considered when assessing the intended purpose
- How *"intended purpose"* vs *"reasonably foreseeable misuse"* applies to para 12, which treats uses as part of purpose - classification by capability.
- What steps a downstream provider who puts a broadly applicable AI system into service for a non-high-risk purpose should take to avoid automatic high-risk classification, particularly where para 12 doesn't apply.
- How a deployer may ensure that it does not become subject to provider obligations under Article 25 where it white labels (i.e. puts name or trademark) a GPAI system already placed on market, where it doesn't intend to use the AI system for high-risk purpose, but where the provider has not done "enough" to exclude it from being high-risk.
- How "feasible and reasonably foreseeable" high-risk use interacts with:
    - a provider's ability to legitimately limit the intended purpose; and
    - the need for downstream deployers to receive sufficient documentation, support and control capabilities where they are expected to assume provider-like obligations. In the absence of clear and objective criteria, this creates significant legal uncertainty for deployers of externally sourced AI systems, who typically lack access to training datasets, model design choices, or robustness and fairness testing results - a structural information asymmetry that makes independent high-risk assessment extremely difficult.
    - How para 12 aligns with Recital 52 proportionality, and whether high-risk classification should reflect the actual likelihood of high-risk use rather than theoretical capability.

High-risk classification according to Article 6(1) – Annex I (Section III of the draft Guidelines)

AFME has no comments in response to these questions.

High-risk classification according to Article 6(2) – Annex I (Section IV of the draft Guidelines)

IV.1 Rationale and approach

Question A. Are there any aspects / sections / paragraphs of this Section of the draft Guidelines that you believe require further clarification, and if so, how would you suggest they be improved to ensure effective implementation and compliance?

Para 66
Para 66 could further clarify that the purpose of Article 6(3) is to ensure a proportionate application of the high-risk regime by excluding systems that do not materially influence decision-making outcomes. Additional guidance on the distinction between genuinely low risk uses and systems that materially affect decisions would improve legal certainty and reduce the risk of over-classification across Annex III use cases.

Question B. Do you think there are additional examples or use-cases, relevant for this Section of the draft Guidelines, that you find important that the Guidelines include or exclude from the scope of high-risk classification, taking into account the legal provisions in the AI Act?
a) please explain why do you think they should be treated as such and
b) if your suggested example(s) or use case(s) are to be considered as in or out of scope.

Additional examples that should be treated as non-high risk include transaction monitoring systems analysing structural transaction data without assigning personal risk scores

IV.2 Horizontal issues applicable to Annex III use cases

Question A. Are there any aspects / sections / paragraphs of this Section of the draft Guidelines that you believe require further clarification, and if so, how would you suggest they be improved to ensure effective implementation and compliance?

Paragraph 71
Further clarification would be helpful regarding the role of human review when assessing whether a system performs preparatory, procedural or support functions under Article 6(3). Additional examples illustrating when human involvement demonstrates the absence of material influence would improve consistency across sectors.

Para 75-76
Adding generally exempt modules to a high-risk AI system unduly broadens the scope of application of high-risk AI obligations, creating significant operational burden not intended by the legislator.

Additionally, this interpretation of the scope of an “AI system” must be confined to high risk only. There is no legal basis in the primary regulation for expansion, or indeed in the AI system definition guidance. E.g. if workflow components 1, 2 and 3

are from three different suppliers – this would make the *workflow* developer*,* a provider of all composite systems even if the intended purpose of the workflow is consistent with the intended purpose of each individual AI system. Art 25 does not apply to non-high risk so we do not believe there would be legal basis for this.

Regarding the concept of *"material influence"* referred to in paragraphs 75-76, we suggest an amendment such as *"…complex, interconnected setups, like agentic AI systems, where the interconnected systems materially influence an individual decision in the context of the use cases under Annex III"*. Without clearer criteria, there is a risk that components performing low-risk support functions become automatically subject to high-risk classification solely because their outputs are technically available to a broader system.

The GLs should also recognise that effective functional separability may, in appropriate circumstances, be sufficient even where complete technical separation is not feasible. In particular, where a component can be independently identified, documented, audited and governed, and where its outputs do not directly determine the high-risk outcome, as with preparatory tasks. High-risk AI requirements should apply only to the relevant high-risk AI subsystem and to the interfaces necessary to ensure compliance, not to the entire surrounding IT environment.

The GLs should clarify that non-AI components, deterministic rules engines, workflow tools or ordinary IT modules should not automatically be absorbed into the high-risk AI system boundary merely because they are technically integrated with an AI component, unless their joint outputs materially influence the decision. Nor should structuring outputs make an AI system ineligible for Art 6(3) filter.

This issue is particularly relevant for agentic systems, where interaction between specialised functions is an inherent feature of the architecture. The GLs could usefully clarify how the concepts of material influence and genuine separability apply in agentic environments, including whether traceability, governance controls and selective human oversight of high-risk functions may be taken into account when assessing system boundaries.

For example, an agentic assistant used by a bank to prepare a credit file may autonomously retrieve documents, extract factual information, identify missing items, query internal policies and draft a non-binding credit memo for human review. Such an assistant may interact with multiple tools at runtime, but, by design and controls, it does not score the applicant, rank applications, recommend approval or refusal, or set credit terms. Any creditworthiness model remains separately governed and auditable. This example would help clarify that, for agentic systems, high-risk classification should depend on whether the agentic component materially influences the individual credit decision, rather than on the mere fact that it operates within a broader credit process or can call multiple tools. Traceability of tool calls, permissioning, audit logs, governance controls and human approval gates should be relevant to assessing genuine separability and system boundaries.

Para 84
We consider the application of the Article 6(3) filter to be a key priority under this consultation. Our specific comments are set out under subsequent paragraphs but focus on the need to ensure that the key criterion remains whether an AI system directly and materially influences the outcome of a decision.

Para 85
The principle of linking the filter to the absence of material influence on decision-making should be explicitly reaffirmed as the guiding criterion for the application of Article 6(3). In particular, where an AI system does not materially influence the outcome of a decision, it should not be classified as high-risk, even if it is used within a broader decision-making process.

Para 86
The conditions for the filter should be applied in a substantive rather than formalistic manner. Systems that effectively perform narrow procedural, preparatory or supportive tasks should remain eligible for the filter where they do not materially influence decision outcomes, even if they operate within complex or integrated workflows.

Para 90
Excluding use of the Article 6(3) filter where an AI system forms part of a complex or interconnected configuration whose combined intended purpose or outputs materially influence an individual decision risks exceeding the Article 6(3) scope. This would capture e.g. OCR tools that extract information from document scans or translation tools. It should be deleted or clarified that the assessment must focus on whether the specific component makes a direct and material contribution to the decision.

Para 92
The GLs should clarify whether systems applying simple, objective and non-interpretative criteria may be excluded where the output follows directly from predefined requirements and does not involve an evaluative judgement by the AI system. E.g. a system that identifies whether an objective prerequisite is not met etc. should not be treated in the same way as a system that scores suitability, assesses eligibility, ranks candidates or predicts creditworthiness.

Para 93
This risks being interpreted too restrictively by suggesting that categorisation involving existing scores, rankings or labels may fall outside the scope of *"narrow procedural tasks"*. It should be clarified that the decisive criterion should not be the form of the output, but whether the system performs a substantive evaluative function or materially shapes the decision, rather than merely organising or supporting information processing. In addition, not all profiling has the same impact on the individual's rights and this should allow for context-specific considerations. E.g. using labels that refer to the likelihood of someone being interested in certain advertisements or as being more prone to use digital contact methods vs. phone lines should not be considered high-risk. Also, sorting information based on existing scores should be explicitly excluded here. This also seems at odds with the interpretation by the CJEU (Case C-634/21) which differentiates between different types of profiling rather than treating them uniformly.

Para 94
This would benefit from clarification of the notion of a *"completed"* human activity. This should be understood as referring to the completion of the substantive human assessment, even where subsequent procedural, formalisation, verification or quality-assurance steps remain. A narrower interpretation risks excluding typical ex post support functions that do not affect the underlying decision.

Para 95
This may be interpreted too narrowly by describing AI systems as providing only an *"additional layer"* to a completed human activity. It should be clarified that this includes typical ex post support functions, such as consistency checks, error flagging, traceability enhancement or the retrieval of relevant information, provided that such systems do not replace or autonomously perform the human assessment.

Para 96
Please further clarify the notion of *"improve"*, ensuring that it does not exclude systems that support validation or quality assurance. The key criterion should be whether the AI system produces a materially different substantive outcome or effectively replaces the human assessment. A verification of the feasibility of process steps and reasons for drop-outs / interruptions or recommendation to review / verify an assessment taken by a human should be captured by the interpretation of "improving".

Para 100
We understand the notion of a *"completed"* human assessment as referring to the point at which the substantive human judgment has been made, even if subsequent verification, documentation or validation steps are still ongoing, in order to avoid excluding typical ex post analytical tools.

Para 101
The distinction between comparative analysis and proposing a new assessment should be made more explicit. Presenting discrepancies, historical benchmarks or comparative information should not be considered as proposing a new assessment unless the system assigns scores, ranks alternatives or steers the decision-maker towards a specific conclusion. *"propose a new assessment"* should be limited to where a system would propose a new assessment result (not only a review).

Para 102
This would benefit from further operational clarification of *"proper human review"*. In particular, it should require that the reviewer retains full autonomy, has access to the underlying case material, can disregard the AI output and is not constrained by it, ensuring that the system merely informs the assessment rather than influencing or determining it.

Para 103
Para 103 introduces preparatory tasks as a distinct category under Article 6(3)(d), but further clarification is needed to ensure that this notion is not interpreted too narrowly. In particular, it should expressly cover systems that provide informational support to a subsequent assessment, without performing any substantive evaluation of the case.

Para 104
The Draft GLs should clarify that preparatory tasks and procedural tasks may overlap in practice. Systems performing data structuring, filtering or organisation prior to an assessment should remain within scope of the definition of preparatory tasks where their role is limited to preparing information and does not involve any assessment of the merits of the case.

Para 105
A *"very low"* level of impact should be interpreted in line with the absence of material influence. It should be clarified that improving the availability, organisation or clarity

of information does not constitute material influence, provided that the system does not shape the substantive evaluation or decision outcome.

Para 106
In addition to the examples provided in para 106, we suggest the inclusion of functions such as data enrichment, completeness checks or evidence retrieval, provided that they do not perform evaluative judgments or determine outcomes.

Para 107
Para 107 correctly identifies the role of the system in the decision-making process as the decisive factor, but this should be applied substantively. The assessment should focus on whether the system contributes to the substantive evaluation, rather than on its proximity to the final decision point, ensuring that preparatory functions remain within the filter.

Para 108
This risks being interpreted too restrictively by limiting preparatory tasks to outputs that are *"general inputs or factors"*. It should be clarified that preparatory systems may provide case-specific information - such as retrieving relevant data, linking evidence or identifying missing element - provided they do not evaluate the case, rank alternatives or recommend a decision.

Para 110
The Draft GLs note that not all personal data processing or pattern analysis is profiling. We suggest that the key question should be whether the system performs automated processing of personal data to evaluate personal aspects of a natural person in a way that materially influences an individual decision within an Annex III use case. The ambiguity also seems at odds with the CJEU interpretation (Case C-634/21) which differentiates between profiling types rather than treating them uniformly.

Para 112
Please confirm that there is no automatic exclusion from the Art 6(3) filter for AI-enabled tools that perform only narrow procedural/preparatory functions, where they process personal financial data (and so profile), but do not evaluate creditworthiness or establish a credit score. The criterion should be whether the AI system performs an evaluative function involving prediction/assessment/judgement about the natural person, or whether it only extracts/validates/organises factual information.

Paragraphs 113–115
The GLs could clarify the type of evidence and documentation that providers may rely on when demonstrating that a system does not materially influence decision-making. Additional examples of acceptable technical, organisational and governance controls would improve consistency and legal certainty in the application of Article 6(3).

Question B. Do you think there are additional examples or use-cases , relevant for this Section of the draft Guidelines, that you find important that the Guidelines include or exclude from the scope of high-risk classification, taking into account the legal provisions in the AI Act?
a) please explain why do you think they should be treated as such and
b) if your suggested example(s) or use case(s) are to be considered as in or out of scope.

Re our comments on para 110-112, a credit-specific example under Annex III point 5(b) (creditworthiness) would be useful on where tools may process personal financial data, but they do not themselves evaluate the applicant's creditworthiness or establish a credit score.

IV.3 High-risk AI systems in the areas of Annex III

3.1 Biometrics

Question A. Are there any aspects / sections / paragraphs of this Section of the draft Guidelines that you believe require further clarification, and if so, how would you suggest they be improved to ensure effective implementation and compliance?

Paras 119-177
Similar to the chapter on the Art 6(3) filters, this section of the GLs does not seem to recognise that profiling is not per se negative. The example for call centre routing listed under para 177 as high-risk is also inconsistent with the example of improved customer experience listed under para 170. We recommend either deleting the call centre example under para 177 or moving it to the examples for AI systems falling outside the high-risk use cases.

Para 129
In general, regarding remote identification system, we would like to stress the importance of harmonise the content of the present Guidelines with:

- The notion of biometric data as revised under the Digital Omnibus on Data Acquis, GDPR, and
- EDPB Guidelines 5/2022.

Para 129–142
Further clarification would be beneficial regarding liveness detection systems used in consented digital onboarding and eKYC processes. The GLs should expressly confirm that systems used solely to verify the presence of a genuine user, including where comparisons are performed against internal fraud databases, do not constitute remote biometric identification where their purpose is identity verification rather than identification.

Para 171-177
The boundary between speech analytics, transcription, sentiment analysis and emotion recognition based on vocal biometrics is not sufficiently clear. The GLs should clarify which specific features transform speech analytics or call-centre QA tools into high-risk emotion recognition systems.

Para 177
Please clarify borderline use cases involving voice-derived information, e.g. AI systems used to analyse recorded calls where:

- the system processes transcripts, including derived/embedded information on vocal characteristics.
- the aim is to assess client sentiment or compliance with service standards, rather than to infer emotional states.

Would inclusion and analysis of vocal features, even via transcripts or metadata, mean the system uses biometric data?

Question B. Do you think there are additional examples or use-cases, relevant for this Section of the draft Guidelines, that you find important that the Guidelines include or exclude from the scope of high-risk classification, taking into account the legal provisions in the AI Act?
a) please explain why do you think they should be treated as such and
b) if your suggested example(s) or use case(s) are to be considered as in or out of scope.

We would appreciate illustrative examples clarifying the boundary between emotion recognition systems (as described in the case of voice-based emotion inference) and non-biometric sentiment analysis (as in text-based analysis), particularly where hybrid approaches combine transcript analysis with voice-derived indicators.

For example, liveness detection and selfie verification systems used in consented eKYC onboarding processes constitute biometric verification rather than remote biometric identification, including where such systems are used in conjunction with internal fraud-prevention controls.

A banking-specific onboarding example would improve legal certainty and support consistent implementation across the financial sector.

In addition, under paragraph 177, we would appreciate if the GLs could specify that emotion recognition systems using recordings that are in the public domain should not be categorised as high-risk under point 1c.

3.2 Education and Vocational Training

Question A. Are there any aspects / sections / paragraphs of this Section of the draft Guidelines that you believe require further clarification, and if so, how would you suggest they be improved to ensure effective implementation and compliance?

Para 209-215
We request confirmation that the scope of *"educational and vocational training institutions"* both here and in the examples below is limited to institutions where education/training is the primary purpose (e.g. schools) and does not extend to private actors such as companies providing internal or mandatory training.

Question B. Do you think there are additional examples or use-cases , relevant for this Section of the draft Guidelines, that you find important that the Guidelines include or exclude from the scope of high-risk classification, taking into account the legal provisions in the AI Act?
a) please explain why do you think they should be treated as such and
b) if your suggested example(s) or use case(s) are to be considered as in or out of scope.

AFME has no comments in response to this question.

3.4 Employment, workers' management and access to self-employment

Question A. Are there any aspects / sections / paragraphs of this Section of the draft Guidelines that you believe require further clarification, and if so, how would you suggest they be improved to ensure effective implementation and compliance?

Para 236-278
We would welcome clarification of the geographical scope of the Act, namely Article 2(1)(c), in relation to employment. We suggest that it should not cover, for example, an EU-based candidate looking for a job in the UK who may be subject to AI assessment.

Paragraph 247
The GLs should clarify job descriptions/qualifications/skills/etc. generation is a preparatory activity before any candidate assessment. The consideration should not be whether those elements were initially drafted by a human or an AI system, but whether the AI system then evaluates identifiable natural persons or otherwise materially influences specific recruitment decisions. The human review requirement may also ensure a more proportionate interpretation, in line with paras 70-71.

Para 250
"Meaningfully conditions access to employment opportunities" should be understood as covering situations where an AI system materially restricts, prioritises or excludes the visibility of specific job opportunities for certain individuals or groups, thereby significantly affecting their realistic ability to access those opportunities. A distinction should be drawn between such cases and systems that merely optimise reach or engagement.

Para 252
This correctly excludes targeting based on objective, reasonable and job-related criteria, as well as contextual targeting. It should be further clarified that such criteria include elements such as location, licensing, language skills or profession-specific channels. Conversely, systems that selectively limit the visibility of job opportunities based on behavioural or personal profiling may fall within scope where they materially affect access to employment.

Para 258
This would benefit from clarification of the notion of *"significant reliance"*, which is currently broad and risks capturing ordinary decision-support tools. It should be clarified that this concept applies only where the AI output effectively determines or materially constrains the decision outcome, rather than merely supporting or informing human judgment.

Para 267
While task allocation may materially affect career prospects and access to opportunities, we request clarification that only systems that effectively determine access to more or less favourable assignments fall within scope, and not ordinary operational allocation tools that do not materially affect career progression or income.

Clear examples should distinguish between decision-making systems and decision-support tools, e.g. in areas such as task allocation, scheduling or workload management.

Para 268
This appropriately identifies risks of discrimination linked to behavioural or personal traits. This should be clarified to ensure that such risks are not presumed in all cases where individual data is used, but only where such characteristics are used to structure allocation decisions in a way that may lead to systematic disadvantage or unequal access to opportunities.

Para 269
This defines the scope broadly, potentially capturing a wide range of systems using behavioural or inferred traits. It should be clarified that point 4(b) applies only where such data is used to rank, prioritise or exclude workers in a way that materially affects access to work, income or progression, rather than where it is used in a limited or purely operational manner.

Para 270
Para 270 includes common operational indicators such as punctuality, responsiveness or performance ratings, which risks over-classification if not further clarified. It should be specified that such indicators trigger high-risk classification only where they are used to materially condition access to work opportunities or career progression, rather than for minor or non-decisive operational adjustments.

Para 272
This correctly excludes allocation based on objective, neutral and external criteria. This approach should be reinforced and expanded with additional examples, including hybrid systems, to clarify the boundary between behavioural allocation leading to high-risk classification and neutral task distribution based on availability, location or required qualifications.

Para 277
Please clarify for situations in which systems generate alerts/“hits” that may lead to human-led investigations or employment decisions, e.g. systems used for:
AML and anti-fraud checks involving employees (e.g. transaction monitoring or internal investigations). The current drafting does not make clear whether the 'exclusively' condition is lost where the system's alerts or outputs are, as a practical matter, used as inputs into employment-related decisions.

Question B. Do you think there are additional examples or use-cases , relevant for this Section of the draft Guidelines, that you find important that the Guidelines include or exclude from the scope of high-risk classification, taking into account the legal provisions in the AI Act?
a) please explain why do you think they should be treated as such and
b) if your suggested example(s) or use case(s) are to be considered as in or out of scope.

Please include examples in the GLs on the boundary between regulatory compliance uses and employment monitoring, particularly in financial services contexts, to support consistent interpretation and implementation. For example, trader surveillance systems are mandatory in financial services contexts for the purposes of

detecting market abuse and individual performance monitoring per se. However, alerts may then lead to human-led investigations or employment decisions.

On the example below 247: where an AI system is solely intended to generate or suggest job descriptions, qualifications, skills or recruitment requirements, and does not evaluate, rank, score or filter candidates, it should be considered to perform a preparatory or procedural task within the meaning of Article 6(3)(a), regardless of whether those elements are initially provided by a human or generated by the AI system itself.

3.5 Access to and enjoyment of essential private services and essential public services and benefits

Question A. Are there any aspects / sections / paragraphs of this Section of the draft Guidelines that you believe require further clarification, and if so, how would you suggest they be improved to ensure effective implementation and compliance?

Para 292
It would be helpful to make more explicit that the final decision remains with a human, particularly in private sector use cases. In addition, §3.5 currently includes limited examples from the banking and private sector, with most examples referring to public-sector deployments. Additional private-sector examples could improve clarity and practical relevance.

Para 295
We request clarification to ensure that creditworthiness evaluation / the establishment of a credit score are not interpreted so broadly as to capture all AI systems used within a credit process, rather than only those specifically intended to perform these functions.

Also, AI systems that merely digitize application forms, i.e. whose sole function is to accurately reproduce handwritten information, are not *"evaluating creditworthiness"* when they are used in the processing of credit applications.

Para 296
This broad definition of creditworthiness evaluation risks being interpreted as covering upstream, downstream or purely supportive tools used within the credit process. This would be inconsistent with examples excluding systems such as customer segmentation for marketing purposes. It should be clarified that only systems that directly assess a person's ability or willingness to meet financial obligations fall within scope.

Para 297
Please clarify that a *"credit score"* refers only to outputs intended to serve as a primary or decisive input in determining access to financial resources or essential services. Scores generated for analytical, internal or support purposes should not be sufficient to trigger high-risk classification where they do not materially influence the decision outcome.

Para 309
The GLs should clarify that AML/CFT systems whose outputs are merely available to, but not determinative of, a subsequent creditworthiness assessment are not

automatically covered by point 5(b), provided that the relevant AML/CFT and credit-scoring functionalities remain functionally and technically separable.

The statement that AML/CFT controls are not covered by the fraud exemption seems misaligned with the legal reality that fraud is a criminal activity under Article 2(1) of Directive (EU) 2018/1673 and is a predicate offence for the purposes of Regulation (EU) 2024/1624. Such a contradiction should be removed allowing any AML/CFT control benefit for the fraud exception under point 5b. Moreover, this assessment is not our own, but the own co-legislators' understanding as it is part of the content of recital (86b) in future Payment Services Regulation, currently undergoing approval.

Additional guidance would be welcome on the concept of *"functionally linked"* systems, including clarification that this should be limited to situations where the output of the AML/CFT module constitutes a determining input to the creditworthiness assessment or scoring outcome.

Finally, we understand that post-grant monitoring and periodic reviews of existing credit limits are out of scope, particularly where no new creditworthiness assessment is performed and customers retain access to appropriate review mechanisms.

Para 321
The absence of a fraud-detection exception under point 5(c) creates a potential inconsistency with point 5(b), which is difficult to justify in practice, particularly where similar AI systems are used across insurance and lending contexts. Please clarify whether the exception could apply where fraud detection is the main intended use or, alternatively, how to assess integrated systems (e.g. in bancassurance models) combining fraud detection with risk assessment or pricing functions.

Para 323
Para 323 correctly clarifies that AI systems used solely for prudential purposes fall outside the scope of high-risk classification where they are not intended for creditworthiness assessment or scoring. This reflects a function-based approach, which should be consistently applied across the section.

Para 324
We understand that the underlying reason why credit scoring is a high-risk practice lies in its close association with the individual assessment of people. In this sense, it could be clarified that all prudential models (and even non-prudential ones) that do not assess people individually or anonymously should not be considered high-risk (e.g., IRB standardization, IFRS 9 provisions, stress testing, economic capital, liquidity risk, market risk, operational risk).

The GLs should also provide additional clarity on system boundaries in banking environments. In particular, the mere fact that IRB and credit-scoring processes share data, infrastructure or modelling components should not automatically result in their classification as a single AI system. Where a common engine generates both prudential and credit-scoring outputs, only the functionality intended for creditworthiness assessment should be relevant for point 5(b).

Para 325
That separate AI systems may coexist and should be assessed independently should be reinforced to confirm that system interaction or shared inputs does not, per se, extend high-risk classification to systems performing purely prudential functions.

Also, where an AI system feeds its outputs into a non-AI credit-scoring model (but is genuinely separable and does not materially influence the decision), this should not be treated as a single AI system and the AI element should not be in scope of high-risk classification.

Para 326
More concrete criteria should be developed to identify system boundaries, particularly where models interact or share components, to ensure consistent application. We cover this further under paragraphs 75-6 above.

Para 327
The GLs should clarify that prudentially required recalibrations, parameter updates and model re-estimations do not, in themselves, constitute a significant change in design under Article 111(2). In the banking sector, such changes are subject to established supervisory governance and are frequently required by prudential regulation.

We encourage alignment of the interpretation of *"significant change in design"* with the prudential concept of material model change. Changes considered non-material by the competent prudential supervisor should not automatically be treated as significant changes for the purposes of the AI Act. Significant changes should be limited to modifications that materially alter the model's core logic, architecture or intended purpose. It is also unclear which elements of an IRB model are part of the *"design"* relevant to article 111(2).

This would also be consistent with the concerns expressed by the ECB in Opinion CON/2026/10 regarding the need to avoid disproportionate compliance burdens and ensure coordination between the AI Act framework and prudential supervision.

Please also ensure that testing, validation and model development environments do not inadvertently trigger classification or compliance obligations prior to actual deployment.

Para 328
We suggest clarification that it is the go-live date of the AI system that marks the end of grandfathering and the start of AI Act obligations (as opposed to any intermediate steps in the deployment process).

Question B. Do you think there are additional examples or use-cases, relevant for this Section of the draft Guidelines, that you find important that the Guidelines include or exclude from the scope of high-risk classification, taking into account the legal provisions in the AI Act?
a) please explain why do you think they should be treated as such and
b) if your suggested example(s) or use case(s) are to be considered as in or out of scope.

We recommend adding cases that fall under the filter exemption, particularly in relation to narrow procedural and preparatory tasks. Modern credit processes rely on a range of digital support tools. The key question should be whether the intended purpose is a genuinely separable procedural or preparatory function that does not materially influence the actual assessment.

E.g.

- Systems that request/gather information before a creditworthiness assessment.
- Text recognition/document analytics: a system receives customer documents, extracts text (including OCR), classifies files into technical categories, identifies entities, structures information, and returns content and metadata in a standardised format for downstream processing - without itself scoring/ranking applicants or recommending a credit outcome.
- Electronic self-disclosure workflows, limited to organising and standardising client-provided or account-based information, mapping into predefined disclosure fields etc. before the actual assessment. Especially where there are plausibility checks, documented correction, and human/downstream assessment before any formal decision.
- Systems supporting a human decision-maker without shaping, prioritising, or recommending an outcome (e.g. Copilot used to draft summaries or retrieve information).
- Transaction Tagger systems limited to reducing manual errors, translating transaction descriptions into predefined technical categories and feeding standardised information into broader client disclosures. (Not where the same system begins to generate applicant-level risk signals, prioritisation outputs or decision-oriented recommendations)

Please also include an example of out-of-scope client segmentation, such as a system used to group clients by preferred communication channel, product interest or likelihood of responding to a marketing campaign (provided that all clients remain able to apply for the product and any subsequent creditworthiness assessment is done separately and independently). Such segmentation is not limiting access to essential services where solely aimed at improving client engagement, accessibility and service delivery. E.g. segmenting clients according to preferred communication channel does not exclude any population group but rather helps ensure that all clients can interact through the channels most appropriate to their needs and behaviours, thereby supporting broader and more effective access to the relevant products and services.

Questions in relation to Article 112 paragraph 1 (Evaluation and Review)

AFME has no comments in response to these questions.

AFME Contacts

Coen ter Wal
Managing Director, Technology & Operations
coen.terwal@afme.eu
+44 (0)203 828 2727

Stefano Mazzocchi
Managing Director, Advocacy
stefano.mazzocchi@afme.eu
+32 2 883 55 46

Fiona Willis
Director, Technology & Operations
fiona.willis@afme.eu
+44 (0)203 828 2739