Ernst & Young Global Limited
6 More London Place,
London,
SE1 2DA,
United Kingdom

Building a better working world

29 June 2021

Public consultation – EY comments on the OECD Framework for Classifying Artificial Intelligence (AI)

Ernst & Young Global Limited, the central coordinating entity of the Ernst & Young organization, welcomes the opportunity to participate in the public consultation on the proposed OECD framework for the classification of AI systems. We have extensive global experience helping our clients from a variety of industries implement trustworthy AI. We also work with regulators, academia, industry bodies and international organizations such as the OECD to understand the policy risks and opportunities associated with AI.

We agree with the intended purpose of the proposed framework for AI system classification, which has the potential to allow policymakers to "distinguish the different types of trees within the forest" of systems that are commonly grouped within the catch-all term of AI. We hope that the final iteration of the classification framework will facilitate a nuanced and precise policy debate, and if needed, the development of regulations that are appropriate and proportionate.

We note that the framework is positioned as classifying AI systems, for the purpose of identifying links to policy implications, while using an approach that is largely focused on the technical characteristics of AI systems. Technical characteristics of AI systems undoubtedly impact on the technical procedures that will be required for appropriate implementation of future regulations. EY believes the technical characteristics and procedures for achieving compliance with future regulations is best addressed through globally applicable standards, which can more easily be adapted as the technology develops over time. We therefore recommend that the assessment of policy implications should focus largely on the (intended) use and outcomes of the AI systems, as is the case in the proposed AI regulation that was published by the European Commission on 21 April 2021.

With regards to "Operators of AI Systems", EY recommends greater emphasis on the risks inherent in the design or problem specification phase. Many of the major public failures and issues raised in other sections of the OECD's framework document could have been avoided if first addressed at the specification phase. Emphasizing the technical aspects rather than the specification (or business) aspects places the burden of responsibility with technical teams who may not have sufficient context or expertise to mitigate mis-specification risks.

EY has submitted an annotated version of the draft report with more detailed comments and observations. We would be pleased to discuss our views in more detail and thank the OECD for the opportunity to provide detailed comments.

Yours sincerely,
Ansgar Koene/ Global AI Ethics and Regulatory Leader at EY / Ansgar.Koene@uk.ey.com

Appendix – EY responses to five key questions from the OECD's Framework for Classifying Artificial Intelligence (AI)

1. Should there be a core classification framework for less-technical audiences and additional considerations for more technical or informed users?

   This depends on the intended use of the classification system. If the classifications are meant to support policy related actions such as the risk assessment presented towards the end of this document, then all AI systems should be provided with a full set of classifiers. If the classification is performed by the organization that is operating the AI system, they should have the required technical knowledge to provide a full classification.

2. Which characteristics should constitute core criteria for a user-friendly policy-oriented classification framework? Could you please comment on the tentative suggestions by the expert group for core criteria? (the criteria that are not marked as "optional criteria")

   We recommend that the distinction between "core" and "optional" criteria should be removed. If the more technical classification questions are relevant for something like a risk assessment, or anything else related to AI policy, then they should be included in the policy-oriented classification. If they are not relevant, then there is no need to include them as optional.

3. Can users consistently and reliably classify specific AI systems using these core criteria?

   The "core criteria" are generally clearly defined. However, the "core application areas" requires some additional options. Additional "core application areas" could be: Anomaly detection (as in fraud detection, cybersecurity, device failure prediction etc.); Prediction (as in credit assessment, weather forecasting, etc.).

4. Which characteristics may be useful for a more detailed and technically oriented framework? Could you please comment on the tentative suggestions based on the expert group's feedback for additional, more technical, "optional criteria".

   Some of the "optional criteria" are not well defined yet, especially the "business model" criterium does not seem very useful in its current form. Such disagreements suggest a need for further refinement of the relevant definitions. For example, distinguishing between any of the "for-profit use" business models or the "non-profit use" or "public service" does not directly indicate anything about the AI system. It merely provides an approximate proxy for the optimization goal. For the purposes of classifying the AI system, it would be less ambiguous if the "business model" criterium were replaced by the "fairness" criterium (e.g., maximizing cumulative outcomes vs. minimizing differences in outcomes). Either of these could apply equally to any of the "for-profit use" models or the "non-profit use" or public service.

5. Should there be industry or application domain specific criteria and classifications, e.g. depending on context?

   Yes, application domain specific criteria could follow the application of domain specific regulatory regimes allowing greater cross-border harmonization with existing policy frameworks like domain specific safety requirements.