

Comments from Ministry of Economy, Development and Tourism to OECD Framework for “The Classification of AI Systems – Public Consultations on Preliminary Findings”

June 30, 2021

We detail some observations from the document to contribute to discussion of AI systems and its classifications.

1. In Introducing Framework – context section: it could be helpful to state explicitly “direct and indirect” stakeholders, since not only people using/being direct beneficiaries of the AI system are impacted (e.g., schoolmates of a children that is user of an AI system).
2. In Introducing Framework – context section, specifically when named impacts on human rights, and on well-being and its operator, it is not clear if environmental and societal impacts are considered, they could be stated explicitly. Furthermore, environmental impacts are not only important when they affect health and wellbeing, they should be treated independently from the effect on human beings.
3. The environmental impacts should be included in the “Context” box (Figure 2).
4. A decommission stage of the AI lifecycle could be included within “Deployment” and/or “Context” (it should be discussed among experts where they think it is better). How an AI system is taken out from an organization or from people’s lives is relevant when analyzing and classifying them (Figure 3).
5. Page 9, Classification Framework – Context dimension: we also think that environmental impacts should be explicitly included when the effects on human rights and well-being are named as part of “Core characteristics”.
6. Page 10, Table 1: we believe that the relevance of IA principles at least must be classified as “relevant”, because IA could be an important tool for improving the policy/regulation across the techniques that allows to asset the public programs (predict certain behavior of benefited, for instance).
7. Page 11, C. Impacts critical functions / activities [optional criteria]: it could be interesting to lay guidelines on who defines a critical function and how. Should be a self-evaluation? Maybe a common standard? It could be helpful include more details.
8. Page 12, Breadth of deployment characteristics: there is no information in “A pilot project” definition.
9. Page 13, Users of AI systems, we think that a differentiation could be made between “conscious” and “unconscious” users. The first are the ones that interact with AI systems knowingly that they do, while the second interact with them without any knowledge (e.g., people who do not know that AI is used in social network algorithms or recommendation

systems). A differentiation could be made between active and passive users (they interact with the system or just receive its results).

10. Page 13, Impacted stakeholders, optionality and business model: the direct/indirect stakeholders' categories could be included.
11. Page 15, H. Benefits and risks to well-being [optional criteria]: The environmental factor is focused from a human-centered perspective (environmental damage that damages people) not from an environmental-centered point of view (environmental damage that is bad in itself).
12. Page 21, D. Data quality and appropriateness [optional criteria]: An optional criterion could be included with the "Level of encryption" of the data, considering algorithms that can work directly with encrypted data. This could accompany the level of anonymization.
13. Page 23, AI models are not universal: we observed an error due to the reference source is not found.