

ALANA WALSH

Thank you for an opportunity to comment on the OECD's Framework for Classifying AI systems, it is a welcome contribution. Our comments below are based on the DTA's AI systems experience and other Australian agencies will have their own views.

- Should there be a distinction between core and non-core criteria? In other words, should there be a core classification framework for less-technical audiences plus additional considerations for more technical and informed users?
 - The OECD brings a particular political, economic and social focus to exploring emerging technology. This focus complements technically focused organisations, such as the International Standards Organisation. Designing the framework's criteria around these focuses would prevent unnecessary duplication and confusion.
 - Care should be taken when creating a core criteria structure, so that it doesn't engender an overly reductionist approach to technology. This may hinder a greater understanding of how different AI systems interconnect and how AI systems interconnect to the broader community. Ontologies typically have core criteria, however with emerging technologies these criteria need to be flexible enough to change as the technologies evolve.
- Which characteristics should be core criteria and which 'optional'?
 - If the framework focuses on political, social and economic classifications then technical criteria could be considered as optional.
 - Quantification of criteria should avoid descriptions that are rapidly superseded.
 - ☒ For example the Framework discusses including data scale as a criteria. Size is merely a statement of current technology's ability to easily manage a dataset, and this changes quickly.
 - ☒ If there is a need to maintain this criteria, linking it to current technology's capacity to manage the data will make it an enduring criteria.
 - ☒ Small datasets would be those that are easily managed using current consumer technology, whereas very large datasets would be beyond the capacity of the current technology available to most organisations.
 - Drafters may want to consider the role of AI as a user of another AI system's output. This allows for a greater understanding of how a system operates within an algorithmic ecosystem. It also helps planners understand that some users will automatically ingest the output of a system and this may create complex challenges, such as re-identification of personal information or unconscious bias from one system affecting the output of another system.
 - Data age may need to be considered as a core criteria of data quality. By identifying the age of data points and when the dataset was created, the framework assists users to understand what implicit bias and inappropriate terms may be found in the dataset. This in turn will help mitigate any risks that the AI system's output will have hidden biases and inappropriate categorisations.
- Can AI systems be classified with the core criteria both consistently and reliably?
 - AI's growth over the past 70 years has led to systems being classified by function (Natural Language Processing, Regression, Image classification) or when their use becomes mundane (expert systems). The framework may need some way of recording how a system was previously classified and why that changed.
- Which criteria are useful for a more detailed and technically-oriented framework?

- Model definition and maintenance are the main candidates for more technically oriented frameworks.
- Should there be criteria and classifications that are specific to industries or application domains, e.g. depending on context?
 - While certain types of AI systems are mainly used in specific industries, such as High Frequency Trading algorithms in finance, it does not mean they are the only industry to use that AI system. It may be more accurate to classify AI systems by function and allow industries to identify how they best meet their needs.

Best wishes

Alana