30 June 2021

Via email to ai@oecd.org

BSA Comments on the OCED Framework to Classify AI Systems

BSA | The Software Alliance (BSA), the leading advocate for the global software industry, appreciates the opportunity to provide input on the OECD Framework to Classify AI Systems ("Classification Framework"). BSA's members[1] are among the world's most innovative companies, creating software solutions that spark the economy. Our members are at the forefront of software-enabled innovation that is fueling global economic growth, including cloud computing and AI products and services. BSA members have made significant investments in developing innovative AI solutions for use across a range of applications. As leaders in AI development, BSA members have unique insights into both the tremendous potential that AI holds to address a variety of social challenges and the governmental policies that can best support the responsible use of AI and ensure continued innovation.

At the outset, we would like to applaud the OECD for pursuing this important project. As policymakers around the world evaluate whether they have the right policies in place to promote the benefits and address the unique risks of AI, the Classification Framework can serve as an important resource to help them tailor policies to address risks that are implicated by specific applications of AI. By facilitating the development of targeted policies, the Classification Framework can help reduce the uncertainty that can be created by broad or vaguely worded AI regulations. BSA has long focused on the benefits of risk-based evaluation of AI, and we are pleased to see that holistic approach is utilized in the Classification Framework. Risk-based approaches help establish foundational trust in AI by creating an agreed-upon set of features and elements of AI systems that can inform meaningful understanding by a variety of stakeholders, precisely the intended use of this Framework.

The Classification Framework has the potential to be an invaluable tool for policymakers that will help them identify the key technical characteristics of AI systems that should be taken into account when developing policies aimed at advancing the OECD AI Principles. To that end, the Classification Framework organizes the analysis into four "dimensions," focusing on a system's "Context" of use, the "Data and Input" used to train the system, the nature of the "Model" that undergirds the system, and the "Tasks and Output" the system generates. While there are many ways that the Classification Framework could be organized, we agree that these four dimensions provide a helpful heuristic for understanding the key technical characteristics of AI systems and appreciate the effort to organize the Framework around the AI lifecycle.

---

[1] *BSA's members include: Adobe, Atlassian, Autodesk, Bentley Systems, Box, CNC/Mastercam, DocuSign, IBM, Informatica, MathWorks, Microsoft, Okta, Oracle, PTC, Salesforce, ServiceNow, Siemens Industry Software Inc., Slack, Splunk, Trend Micro, Trimble Solutions Corporation, Twilio, Workday, and Zoom.*

The specific technical characteristics (i.e., "sub-dimensions") identified in the Framework are likewise commendable, as is the effort to explicitly explain how they relate to specific OECD AI Principles. As you continue to develop the Framework, we would urge you to consider the inclusion of an additional sub-dimension within the Model category. Specifically, we recommend that you include a sub-dimension that is focused on the manner in which an AI system's underlying model was developed and/or maintained. In our recent whitepaper on AI bias risk management[2], we characterized AI model development as occurring along a spectrum:

- Universal Model: The term "universal model" is used to describe circumstances in which an AI developer provides multiple customers (i.e., AI deployers/users) with access to a single pre-trained model.
- Customizable Model: The term "customizable model" is used to describe circumstances in which an AI Developer provides a pre-trained model to AI Deployers who can customize and/or retrain the model using their own data.
- Bespoke Model: The term "bespoke model" is used to describe circumstances in which The AI Developer trains a bespoke AI model on behalf of an AI Deployer using the AI Deployer's data.

Understanding where an AI model fits along this spectrum will be critical as policymakers try to develop policies to advance several core OECD AI Principles. For instance, effectively promoting AI Robustness, Security, and Safety" (OECD Principle 1.4) will require policies that that encourage stakeholders "based on their roles, the context, and their ability to act, [to] apply a systematic risk management approach to each phase of the AI system lifecycle." As policymakers try to craft such policies, understanding how an AI system's model was developed and/or maintained will be a key consideration. While it is impossible in the abstract to assign roles or delegate specific responsibilities for AI risk management functions, understanding how a model was developed will be a key determining characteristic. For instance, in circumstances involving Universal Models, the AI Developer that trained and maintained the model on behalf of its customers will generally be best positioned to address most aspects of model risk management throughout the system's lifecycle. In contrast, in circumstances involving Customizable Models, many key risk management responsibilities will likely shift to the organization that re-trained and/or customized the model. And in circumstances involving Bespoke Models, the bulk of risk management responsibilities will necessarily fall on the entity that contracted for the system's development. Understanding the manner in which a model was developed and/or maintained will be similarly important in the context of policies related to "Transparency and Explainability" (OECD Principle 1.3) and "Accountability" (OECD Principle 1.5).

BSA appreciates the OECD's investment of time and resources into the development of the AI Classification Framework. We agree that the project can play a key role in helping policymakers develop risk-based and context-specific AI policies. Given the somewhat abstract nature of our recommendation, we would be welcome an opportunity to discuss these issues with you at your convenience. We look forward to continued collaboration with the OECD.AI network.

---

[2] See *Confronting Bias: BSA's Framework to Build Trust in AI,* available at
https://ai.bsa.org/confronting-bias-bsas-framework-to-build-trust-in-ai