Software & Information Industry Association's Feedback on the G7 *International Draft Guiding Principles for Organizations Developing Advanced AI Systems*

Submitted to the U.S. Departments of State and Commerce

October 24, 2023

On behalf of the Software & Information Industry Association (SIIA), the leading trade association for the business of information, we appreciate the opportunity to provide input on the G7's *International Draft Guiding Principles for Organizations Developing Advanced AI Systems* (the "Principles").

We are pleased to see the G7 nations engage in a fulsome effort to align fundamental AI policy internationally through the Principles, the forthcoming code of conduct, and efforts to build best practices with leading international institutions like the OECD and GPAI. These efforts are critical to promote technological innovation in a manner that advances democratic values, harnesses the potential of AI technologies, and appropriately mitigates risk attendant to those technologies. We also appreciate the Principles' recognition of the important role of international technical standards and domestic law that may differ in some ways to reflect differences between jurisdictions.

We hope this input will be helpful as the G7 nations continue to refine the Principles and use the Principles to develop an international code of conduct on AI (the "code of conduct").

Comments on Preamble

Clarify the AI Systems Covered by the Principles

We recommend that the United States and other G7 nations refine the Principles to align with the scope of the White House Voluntary Commitments rather than apply to all "advanced AI systems." While not defining this term, the Principles do include two examples of "advanced AI systems" – foundation models and generative AI – which suggests that the Principles are intended to cover these types of AI systems and not others. The White House Voluntary AI Commitments, in contrast, apply to "generative models that are overall more powerful than the current industry frontier." [1] We suggest that the G7 use the same language to prevent capturing technology that may not present the same level or type of risk, provide clarity to the public, and to further the goal of international policy interoperability.

Define Risk-Based Approach

The Principles call for organizations to take actions "in line with a risk-based approach" as government work to develop more enduring guidance and regulation. We believe there is a lack of harmonization on what a "risk-based approach" means and would recommend that the Code of Conduct explicitly reference the NIST AI Risk Management Framework to provide clarity to the public and organizations alike.

[1] White House, Voluntary AI Commitments (July 2023) (available at https://www.whitehouse.gov/wp-content/uploads/2023/07/Ensuring-Safe-Secure-and-Trustworthy-AI.pdf).

Distinguish Actors Involved in the AI Lifecycle

The preamble states both that the Principles "should apply to AI actors" and that the forthcoming code of conduct will apply to "organizations developing advanced AI systems." Within the text of the Principles themselves, there is ambiguity about which directives would apply to any organization or just to those developing or using AI. We recommend further clarity as some obligations make sense only for those developing AI systems and others for those deploying or using AI systems. We flag certain of these instances in our comments below.

Focus Aspirational Goals

We appreciate the aspirational goals reflected at the top of page 2 of the Principles. We note that the affirmative obligations to respect the rule of law and other values in the design, development, and deployment of AI systems will need to adapt to legal frameworks across the G7 jurisdictions, which may define and advance these values in different ways.

In addition, while we believe the negative obligations in the second sentence of that paragraph are laudable, we recommend revising the phrase "develop or deploy advanced AI systems" to read "develop advanced AI systems that can reasonably be expected to or deploy advanced AI systems". An AI system is a tool, and as with any tool, technological or not – consider, for example, email, which can be used to communicate or for spearfishing, and a hammer, which can be used to repair a broken chair or to assault and individual – is it impossible for a developer to build in controls that insure against any unanticipated improper use. We encourage the G7 to focus instead on mechanisms to limit the use of AI systems in ways that undermine democratic values, human rights, and civil rights.

Comments on Individual Principles

Principle 1

We recommend refinement of Principle 1 to clarify the AI systems to which it refers, consistent with points raised above regarding "advanced AI systems."

In addition, we recommend that Principle 1 incorporate language from the preamble about applying a "risk-based approach" to clarify that "appropriate measures" should be determined based on the risk profile of the AI system at issue and will not be uniform for all AI systems covered by the Principles. For example, independent external testing may be appropriate for certain high-risk use cases but should not be construed as the default for all advanced AI systems.

We also recommend that Principle 1, and any portion of the code of conduct relying on Principle 1, provide clear definitions of terms of art such as "lifecycle" and "traceability." Lifecycle, for example, is defined in the OECD AI Principles.[2] Traceability is among the terms agreed to by the United States and the European Union as part of its effort to align AI taxonomy and terminology in the Trade and

[2] Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449 (May 21, 2019), at I (available at https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449).

Technology Council (TTC).[3] Indeed, to advance the goals of regulatory interoperability across borders, we in general recommend that the G7 nations incorporate definitions agreed upon in the TTC process.

Principle 2

We recommend that Principle 2 provide clarity around the term "organizations." The obligations of organizations with regard to AI risk mitigation and remediation differs by the role they play in the AI lifecycle. This, in turn, necessarily depends on a range of factors, including the intended and actual use of the AI system at issue, the types of vulnerabilities anticipated, and the actual incidents and patterns of misuse observed following deployment. Some steps may be appropriate for developers of AI systems and others for deployers of AI systems.

Principle 3

We reiterate our comments above with respect to "advanced AI systems" and "organizations."

Principle 4

We recommend that Principle 4 be refined to reflect limitations on information sharing due to "privacy laws, regulations and contractual obligations," a critical caveat included in White House Voluntary AI Commitment #2.[4] In addition, while we appreciate the goal for AI organizations to "work towards" information sharing and incident reporting, it is important to recognize that public disclosure obligations must be calibrated against potentially competing needs around national security, trade secrets, and intellectual property protections. There may be situations where sharing is appropriate with some but not all third parties.

Principle 7

We support continued work towards trusted international standards for content authentication and provenance, including watermarking and other techniques to enable users to identify AI-generated content. We recommend that Principle 7 be refined to recognize the need for continued research into digital content provenance and authentication techniques, as they continue to be developed, and to encourage use of tools or APIs that align to interoperable open standards to facilitate trust across the internet. In addition, we recommend focusing Principle 7 on content types, such as audio and visual content (pictures and video), where mitigation research and technology is further along. For example, existing technology does not reliably identify text content generated through natural language processing AI tools. Principle 7 should recognize that applying content authentication and provenance to all AI-generated content may not be appropriate, for example, AI-generated content used in a motion picture should not need to apply such mitigations.

[3] EU-U.S. Terminology and Taxonomy for Artificial Intelligence, First Edition (available at https://www.nist.gov/system/files/documents/noindex/2023/05/31/WG1%20AI%20Taxonomy%20and%20Terminology%20Subgroup%20List%20of%20Terms.pdf).

[4] White House, Voluntary AI Commitments (Sept. 2023), Commitment 2 (available at https://www.whitehouse.gov/wp-content/uploads/2023/09/Voluntary-AI-Commitments-September-2023.pdf).

In general, on content authenticity and provenance, we recommend aligning Principle 7 and the code of conduct as much as possible to the White House Voluntary AI Commitment #5.[5]

Principle 11

We recommend revision of Principle 11 in two respects. First, we agree that developers should incorporate appropriate safeguards around personal data and data protected by IP rights that are contained in training sets. However, the act of ingestion and the permissible scope of permissible output remain the subject of active litigation and should not be prejudged. As drafted, the phrase is subject to overbroad interpretations across and within jurisdictions who may be looking for competitive advantages over AI development. We recommend replacing this category with a statement that it is important to comply with existing IP and copyright legal frameworks.

Second, we support the notion that respect for the proper scope of intellectual property rights is part of the ethical use of artificial intelligence. In particular, exceptions to copyright that would permit commercial entities to ignore license restrictions around the use of intellectual property should not be encouraged. With that said, the scope of copyright protection with respect to data used in AI training sets is being actively litigated and debated in the United States and in foreign jurisdictions. It is important to ensure that the draft principle is not misread as a call to legislate in this area while national conversations around generative AI (as well as the technology itself) are evolving so quickly.

Contact Information

For questions regarding this document, please contact Paul Lekas, Senior Vice President for Global Public Policy & Government Affairs, at plekas@siia.net.

[5] White House, Voluntary AI Commitments (Sept. 2023), Commitment 5 (available at https://www.whitehouse.gov/wp-content/uploads/2023/09/Voluntary-AI-Commitments-September-2023.pdf).