Comments from BSA | The Software Alliance on AI Guidelines for Business

February 19, 2024

General Comments

BSA | The Software Alliance (**BSA**)[1] appreciates the opportunity to provide comments to the Ministry of Internal Affairs and Communications (**MIC**) and Ministry of Economy, Trade and Industry (**METI)** in response to public consultation on Draft AI Business Guidelines for Business (**Draft Guidelines**).[2] We commend the leadership of MIC and METI in compiling the draft Guidelines to support artificial intelligence (**AI**) business operators to develop, deploy, and use AI responsibly. We are encouraged that the draft Guidelines[3] take a risk-based, life-cycle approach that supports industry's voluntary efforts to promote innovation and the utilization of AI while providing appropriate safeguards to minimize associated risks. BSA and its members are eager to work with the Government of Japan to support this effort.

BSA is the leading advocate for the global software industry. BSA members are at the forefront of developing cutting-edge services — including AI — and their products are used by businesses across every sector of the economy.[4] For example, BSA members provide tools including cloud storage and data processing services, customer relationship management software, human resource management programs, identity management services, cybersecurity services, and collaboration software. BSA members are on the leading edge of providing AI-enabled products and services. As a result, they have unique insights into the technology's tremendous potential to spur digital transformation and the policies that can best support the responsible use of AI. BSA's views are informed by our experience working with member companies to develop the BSA Framework to Build Trust in AI,[5] a risk management

[1] BSA's members include: Adobe, Alteryx, Altium, Amazon Web Services, Asana, Atlassian, Autodesk, Bentley Systems, Box, Cisco, Cloudflare, CNC/Mastercam, Dassault, Databricks, DocuSign, Dropbox, Elastic, Graphisoft, Hubspot, IBM, Informatica, Kyndryl, MathWorks, Microsoft, Nikon, Okta, Oracle, PagerDuty, Palo Alto Networks, Prokon, Rockwell, Rubrik, Salesforce, SAP, ServiceNow, Shopify Inc., Siemens Industry Software Inc., Splunk, Trend Micro, Trimble Solutions Corporation, TriNet, Twilio, Workday, Zendesk, and Zoom Video Communications, Inc.

[2] *Call for Opinions on the "Draft Guidelines for AI Business Operators"*, January 20, 2024 at https://public-comment.e-gov.go.jp/servlet/Public?CLASSNAME=PCMMSTDETAIL&id=145210224&Mode=0 (Japanese)

[3] *Draft AI Guidelines for Business*, January 2024, Main Body at https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20240119_4.pdf (English) and https://public-comment.e-gov.go.jp/servlet/PcmFileDownload?seqNo=0000267013 (Japanese)

Draft AI Guidelines for Business, January 2024, Appendix at https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20240119_5.pdf (English) and https://public-comment.e-gov.go.jp/servlet/PcmFileDownload?seqNo=0000267014 (Japanese)

[4] BSA | The Software Alliance, *Artificial Intelligence in Every Sector,* June 13, 2022 at https://www.bsa.org/files/policy-filings/06132022bsaaieverysector.pdf

[5] *Confronting Bias: BSA's Framework to Build Trust in AI*, June 8, 2021 at https://www.bsa.org/reports/confronting-bias-bsas-framework-to-build-trust-in-ai

22F Shibuya Mark City
1-12-1 Dogenzaka, Shibuyaku,
Tokyo 150-0043

P +81 3 4360 5473
F +81 3 4360 5301
W bsa.org

Japan Representative Office

framework we published more than two years ago to help companies mitigate the potential for unintended bias in AI systems. Built on a vast body of research and informed by the experience of leading AI developers, the BSA Framework outlines a lifecycle-based approach for performing impact assessments and highlights corresponding best practices.

Our experience with these issues informs our recommendations below and builds on our earlier submission responding to the Skeleton of the Draft Guidelines ("Skeleton Draft).[6]

Global Harmonization

[Main Body, Introduction, page 2 and throughout the document]

As policymakers around the world are developing regulatory approaches to AI, the global nature of today's technology ecosystem demands coordinated policy responses to foster innovation. BSA supports Japan's leadership in driving international discussions, as demonstrated through the Hiroshima AI Process. We encourage countries to pursue interoperability through multistakeholder dialogue, developing a shared vision for a risk-based policy approach for addressing common AI challenges and advancing norms around responsible AI governance (e.g., risk-based approach, proportionate and role-based responsibilities along the AI value chain). Global partners should also agree on common AI terminology and taxonomy that enable innovators to confidently and flexibly adopt the technology for beneficial applications. We recommend the Draft Guidelines to reflect such a harmonized approach.

Definitions

[Main Body, Part 1 Definitions, page 8]

In the Draft Guidelines, an AI system is defined as "a system that includes software as an element that has the ability to operate and learn with various levels of autonomy through the process of utilization (machines, robots, cloud systems, etc.)." Given that AI systems are developed and deployed in an international context, definitions that apply to AI should operate across different jurisdictions to facilitate and promote further widescale adoption and use of AI technologies. We recommend that Japan adopts the OECD's definition of AI.[7] Using an accepted and internationally recognized definition of AI, such as the OECD's, will facilitate the international alignment of Japan's policies, promoting dialogue, adoption, and compliance with the Guidelines.

[6] *Recommendations from BSA | The Software Alliance on the Skeleton Draft of the New AI Business Operator Guidelines*, October 27, 2023 at https://www.bsa.org/policy-filings/recommendations-from-bsa-the-software-alliance-on-the-skeleton-draft-of-the-new-ai-business-operator-guidelines

[7] *Updates to the OECD's definition of an AI system explained*, November 29, 2023 at https://oecd.ai/en/wonk/ai-system-definition-update

> "An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment."

Balanced and Proportionate Allocation of Responsibilities Among AI Actors in the AI Ecosystem

[Main Body - Introduction, page 5, Part 5 Matters Related to AI Business Users page 34-35]

Given the diverse AI value chain, we support the Draft Guideline's recognition of the importance of allocating responsibilities to the entities best placed to comply with them. We also welcome the Draft Guideline's streamlining of the categories of AI actors from the previous Skeleton Draft.[8]

However, the Draft Guidelines could be improved further by providing examples of the types of AI actors that fall under each of these categories. It would also be helpful to provide additional explanations to ensure that the responsibilities in the AI value chain are balanced and proportionate by noting, for example, that AI business users have the ability to alter the AI system if it is customizable.

The current definition and guidance do not account for customizable AI, which is a product that many enterprise companies offer to AI business users. AI Developers often create general customizable AI tools, the intended purpose of which is low risk. It is then up to the AI business users/customers to decide how these tools are employed. In the business-to-business (B2B) context, it is often the customers that ultimately control the data that is submitted to the AI, direct how the AI is configured, decide on when the AI system is used and in which context, and, most critically, determine how the results are used. As AI business users are generally best positioned to provide details on how the system is being used and have more insight into the data inputs and the resulting outputs and other real-world factors affecting the system's performance, we recommend including customizable AI in "Part 5 Matters Related to AI Business Users".

Risk-Based Approach

[Main Body, Part 2, C. Common Guiding Principles, 3) Fairness, page 14, 15/Appendix, 1, B. Benefits and Risks of AI, Risk of AI, page 14-17]

BSA supports the implementation of risk management programs and supports the Draft Guidelines' recommendations to encourage AI business operators to set governance goals and implement "AI management system(s)" to achieve such goals. Effective risk management programs enable organizations to identify the personnel, policies, and processes necessary to manage AI risks. Elements of a risk management program may include clearly assigning roles and responsibilities, establishing formal policies, using evaluation mechanisms, ensuring executive oversight, performing impact assessments for high-risk AI, and having internal independent review mechanisms, such as interdepartmental governance or ethics committees, to evaluate and address AI issues that pose high risks. Organizations can incorporate these practices into a broader corporate risk management program or establish them in a separate AI program.

We strongly support a risk-based approach to AI policies that focuses on use cases that create high risks to individuals. To adopt this risk-based approach, policies should identify a subset of AI uses as high-risk AI uses. These include AI systems that determine an individual's eligibility for housing, employment, credit, education, healthcare, or insurance. While examples of risks are listed in the Appendix under the "Risk of AI" section, the Draft Guidelines do not clearly

[8] *New AI Business Guidelines Skeleton (Draft)* at https://www8.cao.go.jp/cstp/ai/ai_senryaku/5kai/gaidorain.pdf

differentiate between high and low risk applications. BSA recommends that the Guidelines include a clear distinction or a definition of what constitutes high risk uses of AI and a recommendation that organizations that develop or deploy high-risk AI conduct impact assessments and publicly affirm that they have done so.

External Audit

[Appendix2., Part 2 E. Building AI Governance, 5. Evaluation, page 53-56]

The Draft Guidelines suggest using either in-house resources or external audit entities to determine whether AI management system are appropriately implemented in line with the AI governance goals set by individual organizations. While we understand that external audits are presented as one option, as noted in our earlier submission, we caution against the use of external audits at this point, as currently auditable standards for AI are not mature. There are few existing procedures or best practices for companies to either: (1) choose a reputable company capable of auditing an AI system, or (2) determine what standards any such auditing company should apply. Although the International Organization for Standardization (ISO) has issued a number of AI-related standards, many standards remain under development. Consequently, currently there is a lack of sufficient voluntary consensus-based standards addressing AI systems. Without common standards, the quality of audits will vary significantly because different audits may measure against different benchmarks, undermining the goal of obtaining an evaluation based on an objective benchmark.

Furthermore, while BSA understands the need to promote transparency, we recommend against requiring business operators to publish audited results, which include confidential or proprietary information. This could disincentivize companies from voluntarily undertaking a rigorous review of their AI systems. For these reasons, external audits are not an appropriate solution to achieving transparency over AI governance, and we recommend removing this recommendation from the Draft Guidelines.

Common Guiding Principles for AI Business Actors Involved in Advanced AI Systems and Code of Conduct for Organizations Developing Advanced AI Systems

[Main Body, Part 2/D. Common Guiding Principles for AI Business Actors Involved in Advanced AI Systems, page 22-24/ Additional Entries in the "International Code of Conduct for the Hiroshima AI Process for Organizations Developing Advanced AI Systems", page 28-30]

The Draft Guidelines references International Guiding Principles for All AI Actors[9] and for Organizations Developing Advanced AI Systems[10], as well as the International Code of Conduct for Organizations Developing Advanced AI Systems [11] developed under the Hiroshima AI Process. While we support the objectives of the Principles and Code of Conduct, which aim to promote safe, secure, and trustworthy AI, they will benefit from further improvement as articulated below:

[9] *Hiroshima Process International Guiding Principles for All AI Actors* at https://www.soumu.go.jp/hiroshimaaiprocess/pdf/document03_en.pdf
[10] *Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI Systems at* https://www.soumu.go.jp/hiroshimaaiprocess/pdf/document04_en.pdf
[11] *Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems* at *https://www.soumu.go.jp/hiroshimaaiprocess/pdf/document05_en.pdf*

Clarifying the Scope of the Principles and Code of Conduct

We recommend clarifying the scope of the Principle and Code of Conduct. The Draft Guidelines state that they apply to "advanced AI systems." However, the Draft Guidelines do not define "advanced AI systems". We recommend amending the Draft Guidelines to explain this term, making clear that "advanced AI systems" only encompass the most capable models that pose a high risk of harm. This approach avoids placing responsibilities on AI systems that may be used in low-risk scenarios and instead focuses resources on areas that have the most significant impact on individuals. The term "advanced AI systems" should also be used consistently. For example, Section II) currently refers to AI systems more broadly, without clearly focusing on advanced AI systems.

I) Measures to Identify, Evaluate, and Mitigate Risks Across the AI Lifecycle

First, this section should promote the use of internal testing and avoid suggesting that external tests always should be conducted. This section addresses the identification and mitigation of risks throughout the AI lifecycle, describing both internal and independent external testing as measures that organizations should perform. We agree that testing is a key part of identifying risks but advise against suggesting that organizations always should conduct external testing. There are circumstances where an organization may elect to perform external testing. However, internal testing — which can be performed by a team of employees that is independent from the team tasked with developing an AI system — can identify and mitigate risks without creating concerns about sharing trade secrets, information that could jeopardize information or network security, and other proprietary information that will arise in external testing. As a result, we recommend focusing this section on internal testing and removing the reference to independent external testing.

Second, this section should be updated to reflect the different roles in the AI value chain such as those of developers of AI systems and deployers of AI systems. Guidance in this section should recognize these different roles, because developers, deployers, and other parties within the value chain will each have access to different types of information and will be able to take different actions to mitigate risks. In its current form, Section I) can be read to assume that the developer of an AI system can identify, evaluate, and mitigate risks associated with that AI system — even after the AI system has been acquired and deployed by another organization. That often is not the case. As noted above, instead of assuming that all AI actors have access to information created at all stages of the AI lifecycle, Section I) should promote the identification, evaluation, and mitigation of risks by different organizations based on their role in the lifecycle. The relevant responsibility and accountability should be assigned to the most appropriate role based on the organization's knowledge, control, and position in the AI value chain that makes it possible to address specific risks.

II) Identify and Mitigate Vulnerabilities and Misuse After Deployment

Section II) should be updated to reflect the different roles of developers of AI systems and deployers of AI systems. Like Section I), Section II) appears to assume that all actors in the AI value chain have access to all information about an AI system throughout its lifecycle. Because Section II) focuses on vulnerabilities occurring *after* deployment, we strongly recommend that it be revised to apply to the appropriate role — the deployer (e.g. AI business users) using the AI system — and not impose such responsibilities on organizations that are not in a position to address the concern, such as the developer of that system. The importance of distinguishing responsibilities for organizations with different roles such as developers and deployers (e.g., AI business users) becomes clear when considering the types of information available to

organizations in each role. For example, the developer of an AI system is well positioned to describe features of the data used to train that system, the system's known limitations, and its intended uses, but generally will not have insight into how the system is used after it is acquired by another organization and deployed. In contrast, a deployer (e.g., AI business users) is well positioned to understand how the system is actually being used, what type of human oversight is in place, and whether there are complaints about how the system works in practice. Organizations may also take on other roles, such as integrating an existing AI model into the organization's products and services. Any responsibilities placed on these organizations should similarly reflect their role in integrating the AI system into the organization's products and services.

Creating role-based responsibilities are not unique to AI; they are considered best practice in privacy and security legislation worldwide.[12]

Also, Section II) should clarify its approach to vulnerabilities. Section II) focuses on identifying and mitigating vulnerabilities and, where appropriate, incidents and patterns of misuse. It also refers to other stakeholders in connection with these efforts. Importantly, vulnerability reporting should be handled confidentially, as it may otherwise interfere with customer contractual agreements or raise concerns about proprietary information. Further, other security implications should be considered when addressing vulnerabilities. It is a prevailing practice in the industry that a vulnerability is not publicly disclosed until a patch or other mitigation measures are in place to limit further harm and prevent potential exploitation by malicious actors. Laws and policies related to vulnerability reporting should be risk-based and in line with internationally recognized standards and best practices. The Draft Guidelines should recognize the importance of confidentiality in responding to these security incidents.

III) Public Reporting of Advanced AI Systems' Capabilities, Limitations, and Domains of Appropriate and Inappropriate Use

We support Section III). We recommend further clarifying how the transparency responsibilities in Section III) should be allocated among the different organizations that play different roles in the AI value chain. Section III) calls for publicly reporting key information about advanced AI systems, including capabilities, limitations, and domains for appropriate or inappropriate uses. We recognize that developers of AI systems are creating a range of new resources to provide transparency to their customers about those AI systems, such as documentation that provides information on responsible AI design choices, as well as best practices for deploying and optimizing the performance of a particular AI service. The Government should support efforts to provide deployers with this type of information, while avoiding requirements to disclose underlying training data or other information for which disclosure would create trade secret, confidentiality, cybersecurity, and privacy concerns.

IV) Work Toward Responsible Information Sharing and Reporting of Incidents

We recommend clarifying aspects of the vulnerability reporting referred to in Section IV). Vulnerability reporting and incident response are key components of an effective security program. The public incident reporting recommended in Section IV) could interfere with customer contractual agreements and measures to safely address vulnerabilities. As discussed above, a company generally should not report a vulnerability until it has developed a patch or implemented other mitigation measures. Laws and policies related to vulnerability reporting should be risk-based and in line with internationally recognized standards and best practices.

[12] *AI Developers and Deployers: An Important Distinction*, March 16, 2023, at https://www.bsa.org/files/policy-filings/03162023aidevdep.pdf

We encourage the Government to recognize the importance of confidentiality in responding to these security incidents.

V) Develop, Implement, and Disclose AI Governance and Risk Management Policies, Grounded in a Risk-Based Approach

We support Section V), which recognizes the importance of risk management policies and practices to enhancing organizational accountability and ensuring responsible AI.

Section V) recognizes that organizations should develop and implement risk management programs to help them evaluate and mitigate risks throughout the AI lifecycle. We encourage the Draft Guidelines to recognize that a key part of an effective risk management program is conducting impact assessments. Impact assessments enable organizations to identify and mitigate risks and should be conducted by developers and deployers for high-risk uses of AI systems. By allowing personnel across the organization to examine the objectives, data preparation, design choices, and testing results, these assessments help refine AI products and services and drive internal changes to an organization's risk management program. Implementing these changes enables organizations to better address existing concerns and adapt to new risks as they emerge.

Section V) also refers to disclosing AI governance policies and organizational mechanisms to implement policies. We encourage the Draft Guidelines to recognize that impact assessments should be treated as confidential to preserve the incentives for organizations to implement them through rigorous processes that identify and mitigate a wide range of potential risks. The fact that assessments are being performed for high-risk uses of AI systems promotes trust for external stakeholders because they will know that an organization is conducting a thorough examination of AI systems; those assessments should also be available to regulators in the course of an investigation, under existing domestic laws. We support the aim of Section V) in ensuring the implementation of risk management policies. Section V) should refer to impact assessments as an important accountability tool that can help achieve this goal.

VII) Develop and Deploy Reliable Content Authentication and Provenance Mechanisms

We support Section VII). The development and deployment of reliable content authentication and provenance mechanisms (e.g., watermarking) that can help users identify AI-generated content is an important focus for AI policies. Any content provenance requirements for AI-generated content should focus on images, audio, and video content, since it is unlikely that tools developed for labeling image and audio-visual content would be effective for text. To ensure provenance of text-based AI generated content, we recommend focusing on other transparency mechanisms that can help individuals know when they interact with an AI system.

We encourage the Government to leverage the work of the Content Authenticity Initiative[13] and the Coalition for Content Provenance and Authenticity,[14] which promote the adoption of an open industry standard for content authenticity and provenance. This work can enable viewers to identify information about the origins of an image or video, such as the photographer, the location where the image was generated, and if it was edited using software, assisting viewers in determining the content's authenticity.

[13] *Content Authenticity Initiative* at https://contentauthenticity.org/

[14] *Coalition for Content Provenance and Authenticity* at https://c2pa.org/

XI) Implement Appropriate Data Input Measures and Protections for Personal Data and Intellectual Property

This Section is unnecessary. Unlike Sections I) to X), which address system-level risks that are not covered in existing regulatory frameworks, Section XI) implicates issues for which existing regulations already are in force. We agree with the importance of implementing appropriate safeguards regarding input data, but including a new AI principle and code of conduct implies organizations are not already required to maintain proper data governance. In addition, the description of Section XI) refers to transparency of datasets, without recognizing that those datasets may be confidential and contain a range of proprietary information and, therefore, should not be disclosed.

Specific Methods to Implement the Approaches in the Guidelines

[Appendix]

We appreciate that the Draft Guidelines provide an Appendix presenting specific methods -- together with “practical examples” -- to implement what is presented in the main body of the Draft Guidelines. We recommend clarifying that these specific methods are presented as mere examples and are not the only means to achieve the basic philosophy, principles, guiding principles, and code of conduct articulated in the main body of the Draft Guidelines. This will enable businesses to clearly understand that there are additional implementation options.

Conclusion

BSA and our members appreciate METI and MIC preparing English translation at the time of consultation, which has been very helpful, particularly for a document of this length. We look forward to supporting the goal of Government of Japan to develop effective AI governance policies and look forward to the opportunity for a continued dialogue on how we can further assist in the effort.