FLI Response to OMB:

Request for Information on Responsible Procurement of Artificial Intelligence in Government

29th April 2024

US Policy Team
policy@futureoflife.org

Request for Information on Responsible Procurement of Artificial Intelligence in Government

Organization

Future of Life Institute

Point of Contact

Isabella Hampton, AI Policy Researcher. isabella@futureoflife.org

About the Organization

The Future of Life Institute (FLI) is an independent nonprofit organization with the goal of reducing large-scale risks and steering transformative technologies to benefit humanity, with a particular focus on artificial intelligence (AI). Since its founding, FLI has taken a leading role in advancing key disciplines such as AI governance, AI safety, and trustworthy and responsible AI, and is widely considered to be among the first civil society actors focused on these issues. FLI was responsible for convening the first major conference on AI safety in Puerto Rico in 2015, and for publishing the Asilomar AI principles, one of the earliest and most influential frameworks for the governance of artificial intelligence, in 2017. FLI is the UN Secretary General's designated civil society organization for recommendations on the governance of AI and has played a central role in deliberations regarding the EU AI Act's treatment of risks from AI. FLI has also worked actively within the United States on legislation and executive directives concerning AI. Members of our team have contributed extensive feedback to the development of the NIST AI Risk Management Framework, testified at Senate AI Insight Forums, participated in the UK AI Summit, and connected leading experts in the policy and technical domains to policymakers across the US government. We thank the Office of Management and Budget (OMB) for the opportunity to respond to this request for information (RfI) regarding the OMB's obligations regarding the responsible procurement of Artificial Intelligence in government, as outlined in the Executive Order on Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence.

FLI's wide-ranging work on artificial intelligence can be found at futureoflife.org.

Executive Summary

We welcome the recent publication of the OMB memorandum on Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence (hereafter known as the 'Memo'), which we see as a promising first step in setting standards for the federal government and beyond for the procurement of advanced AI systems that could threaten the rights and safety of the public. This response to the RfI is intended to help the OMB ensure that AI procurement practices follow this spirit and protect citizens against potential rights- and safety-impacting risks.

This section summarizes our contributions to this RfI, which are as follows:

1. **National security use cases should be within the scope of the Memo's requirements:** At present, any AI system that is part of a 'National Security System' is exempt from the Memo's requirements. We believe bringing these systems within scope is vital to ensuring that the most high-risk cases of AI procurement receive appropriate scrutiny.
2. **AI procurement in rights- and safety-impacting cases should be subject to strict criteria for model developers:** In particular, we lay out the case for the use of Operational Design Domains (ODD) taxonomies relevant to the use of AI technologies. Agencies should be tasked with establishing specific operational conditions for their needs, and should take appropriate steps to ensure that the AI systems they procure are reliable and minimize risks, or can be modified to minimize risks, within these domains.
3. **The granting of extensions and waivers should be strictly limited to appropriate cases:** We ask that OMB only grant these extensions and waivers in exceptional circumstances, such as when temporarily suspending the use of an AI system to achieve compliance would significantly disrupt essential government services, or when the standards themselves would threaten safety and security, and more protective alternative safety standards have been identified and verifiably implemented.
4. **Agencies should only be allowed to opt out of compliance after submitting extensive evidence to OMB and exploring alternative options:** We recommend that the OMB require that agencies submit extensive evidence demonstrating how compliance would increase risks to safety or rights or would pose unacceptable impediments to the agency's operations. This evidence should be reported to the OMB as part of the agency's petition to opt-out of compliance, subject to OMB approval.
5. **OMB should clarify its definition of 'principal basis' for use of AI systems to ensure compliance:** We ask that the OMB clarify this term with a robust set of criteria to ensure that agencies only employ this exception in circumstances where the AI systems concerned do not play a material role in driving important decisions or actions.
6. **An appeals process must be set up to address opt out decisions taken by Chief Artificial Intelligence Officers (CAIO):** We ask that the OMB set up a review and appeals process for decisions made by CAIOs to ensure that they are taken responsibly and within the remit of exceptions outlined in the Memo.
7. **A central office should be created to supervise all CAIOs:** To ensure CAIOs can fulfill their obligations effectively, we recommend establishing a centralized office to supervise CAIOs across the US government. This approach will foster a whole-of-government approach and ensure coordination and consistency across several use cases.
8. **The frequency of reporting should be increased to keep up with rapid AI development:** We recommend that OMB change reporting requirements from at least annually to at least every six months. The publication of agencies' compliance plans should also be updated from every two years to at least annually.

Recommendations

1. National security use cases should be within the scope of the Memo's requirements

At present, much of the Memo applies to all agencies outlined in 44 U.S.C. § 3502(1). Any AI system that is part of a 'National Security System,' however, is exempt from the Memo's requirements, and, in many cases, so are members of the intelligence community as defined in 50 U.S.C. § 3003. We believe bringing these systems within scope is vital for ensuring that the most high-risk cases of AI procurement receive appropriate scrutiny.

AI systems intended for use in national security and military applications present some of the greatest potential for catastrophic risk due to their intended use in critical, often life-or-death circumstances.[1] Considering the sizable impact malfunction, misunderstanding, or misuse of national security AI systems could entail, applicable standards should be at least as rigorous in assessing and mitigating potential risks as those developed for civilian AI systems.

National security AI systems often operate with less transparency and public oversight due to their classified nature. This lack of visibility makes it even more crucial that these systems adhere to rigorous risk assessment, testing, and monitoring standards. Without the Memo's requirements in place, there's a greater risk that flaws in these systems could go undetected or have minimal oversight. The Department of Defense's (DoD) Chief Artificial Intelligence Officers (CAIO) has expressed a need for "more rigor" in AI development, stating that he expects "99.999% accuracy" for LLMs used in military context. These standards could help facilitate the DoD in their pursuit of this goal.

Bringing these systems within scope would ensure they receive the same level of risk assessment and mitigation as civilian AI applications. This is essential for maintaining public trust and ensuring that the benefits of AI in national security are realized without unintended and potentially devastating consequences.

2. AI procurement in rights- and safety-impacting cases should be subject to strict criteria for model developers

Current AI systems operate across a wide variety of domains. To enable the comprehensive evaluation of these diverse use-cases, we propose the development of Operational Design Domains (ODD)[2] taxonomies relevant to the use of AI systems. These taxonomies provide a framework for managing procurement, audits, and evaluations through their operational domain, i.e., the specific functions for which the target system is to be used.

Agencies and departments should be tasked with establishing specific operational domains for their AI procurement needs. This process involves defining the intended operational domain, including, users, vectors, protected characteristics, and assets relevant to their specific use cases. Agencies can better communicate their requirements to model developers or integrators by clearly outlining these conditions and ensure that the procured AI systems are designed to operate safely and effectively within the specified domain.

When evaluating competing bids for AI procurement contracts, agencies should prioritize proposals that demonstrate a strong success rate at minimizing identified risks within the specified operational domain. To prevent misuse or unintended consequences, strict prohibitions should be established on the use of the

[1] Moon, M. (2024, February 27). *The Pentagon used Project Maven-developed AI to identify air strike targets.* Engadget. https://www.engadget.com/the-pentagon-used-project-maven-developed-ai-to-identify-air-strike-targets-103940709.html

[2] This section draws significantly from Heidy Khlaaf's work on AI risk assessments. For more information on ODD taxonomies and their application in AI risk assessment, please see: Khlaaf, H. (n.d.). *Toward Comprehensive Risk Assessments and Assurance of AI-Based Systems.* https://www.trailofbits.com/documents/Toward_comprehensive_risk_assessments.pdf

AI system outside of the application domain, along with appropriate penalties for violations, in accordance with existing legal and regulatory frameworks.[3] Agencies should also implement regular audits and reviews to ensure that procured AI systems are being used responsibly and transparently, without any unauthorized expansion beyond the intended application domain.

Incorporating ODD taxonomies into AI procurement guidelines will allow agencies to ensure a more targeted and risk-informed approach to acquiring AI systems. This not only promotes the safe and responsible use of AI in rights- and safety-impacting cases but also fosters greater transparency and accountability in the procurement process. Model developers will be required to demonstrate that their AI systems can operate safely and effectively within the specified operational conditions, reducing the likelihood of harm to individuals and society as a whole.

3. The granting of extensions and waivers should be strictly limited to appropriate cases

According to current guidance in Section 5(a)(i) of the Memo, agencies must either implement the minimum practices specified in Section 5(c) for safety-impacting AI systems by December 1, 2024, or cease using any AI that is not compliant with these practices, unless they receive a one-year extension or a waiver from the OMB.

While we recognize that one-year extensions may be necessary for the continued use of systems already in place at the time of the Memo's issuance to avoid disruption of essential government services, we believe that these extensions should only be granted in exceptional circumstances where the immediate cessation of the AI system would lead to significant harm or disruption.

To ensure that extensions are only granted when truly necessary, we recommend that OMB require agencies to provide a detailed justification for why the AI system cannot be brought into compliance with the minimum practices within the specified timeframe. This justification should include a comprehensive description of how essential government functions would be disrupted if the system were temporarily taken out of operation while the necessary improvements are made to achieve compliance.

Furthermore, agencies should be required to explain why alternative mechanisms for achieving those essential functions are not feasible in the interim period. This explanation should include an analysis of potential workarounds, manual processes, or other AI systems that could be used to fulfill the essential functions while the non-compliant system is being updated. Only in cases where no reasonable alternatives exist and the disruption to essential services would be severe should extensions be granted.

4. Agencies should only be allowed to opt out of compliance after submitting extensive evidence to OMB and exploring alternative options

While the Memo sets up important requirements for agencies procuring and using AI, it also allows agencies to waive these requirements if "fulfilling the requirement would increase risks to safety or rights overall or would create an unacceptable impediment to critical agency operations." Given the subjectivity of this threshold, we are concerned that these exceptions could transform into a loophole, with requests being granted whenever agency priorities conflict with compliance.

To address this, we recommend that OMB require agencies to submit extensive evidence when seeking a waiver. This evidence should include detailed documentation of the specific AI system, the requirements the agency believes pose a safety risk, and a thorough analysis demonstrating how compliance would directly

[3] For more information on procurement regulations, see Federal Acquisition Regulation (FAR) Subpart 9.1, "Responsible Prospective Contractors," which outlines standards for contractor responsibility and the factors that can lead to suspension or debarment, available at https://www.acquisition.gov/far/part-9#FAR_9_104_1.

cause significant harm or danger. In cases where the agency believes compliance would increase overall risks to safety or rights, they should be required to provide a detailed explanation of these risks, along with alternative practices that are equally or more rigorous than the minimum standards outlined in the Memo.

We also recommend that OMB provide clearer guidance on what constitutes "critical agency operations (as described in section 5(c)(iii) of the Memo). Agencies should be required to justify why the disruption caused by compliance with the Memo's requirements would be unacceptable and demonstrate that they have thoroughly explored alternative options that would be in compliance comply with the standards.

If an agency is granted a waiver, this waiver should be time-limited, and the agency should be required to immediately begin efforts to identify and procure alternative AI systems or other solutions that can comply with the Memo's requirements without causing significant disruption. The waiver should only remain in effect until a compliant alternative can be implemented.

5. OMB should clarify its definition of 'principal basis' for use of AI systems to ensure compliance

The Memo allows agencies to opt out of applying minimum risk management practices to AI systems that are not deemed to be the "principal basis" for any decision or action impacting rights or safety. While this qualifier rightly aims to focus oversight on higher-stakes use cases, the current "principal basis" language is ambiguous and could engender unintended loopholes. The Memo does not provide an explicit definition of "principal basis" or any test for what constitutes a principal basis, leaving room for inconsistent interpretations across agencies.

We recommend that OMB replace the "principal basis" standard for rights- and safety-impacting AI with a clearer "material influence" threshold. The definition of "material influence" should provide a set of criteria to ensure agencies can only use this exception to opt out in circumstances where AI systems do not significantly influence important decisions or actions. An AI system should be considered to materially influence a decision or action when it is a contributing factor that markedly affects the outcome of the decision-making process, even if other information is also considered. Informing or being consulted in a decision or action would not constitute material influence alone, but an AI system need not be the sole or primary basis to meet this standard.

6. An appeals process must be set up to address opt out decisions taken by Chief Artificial Intelligence Officers (CAIO)

We welcome the OMB's 2023 recommendation to set up CAIOs in agencies across the federal government in order to ensure that a dedicated office supervises AI-related functions and decisions, including those taken on procurement. As currently outlined in the Memo, these CAIOs have considerable power in the procurement process - they can unilaterally make the decision to opt out of OMB requirements, and while their decision must be reported to both the OMB and the public, it is final.

Even if CAIOs have nominal independence, they may still act in the perceived interest of their agency's mission, potentially conflicting with responsible, objective decision-making regarding compliance waivers and leading to more liberal use of the opt-out authority than appropriate. We ask the OMB to implement a review and appeals process for these decisions to ensure they are made responsibly and within the scope of exceptions outlined in the Memo.

7. A central office should be created to supervise all CAIOs

As stated previously, we welcome the establishment of CAIOs in each agency to ensure that minimum requirements as set out in the Memo are fulfilled across agencies, while giving different offices within each agency the opportunity to exercise their mandate when it comes to specific use cases. The next step to ensure that the CAIOs can fulfill their obligations is to set up a centralized office which supervises CAIOs across the US government, similar to how the Office of the Director of National Intelligence (ODNI) coordinates all intelligence community activities. This centralized office would foster a whole-of-government approach and ensure coordination and consistency across several use cases. This would complement existing efforts to set up an interagency council to coordinate the development and use of AI in agencies' programs and operations as specified in the Memo.

8. The frequency of reporting should be increased to keep up with rapid AI development

Keeping pace with the rapid advancements in AI technology and its increasing adoption in government agencies is crucial for ensuring that oversight remains relevant. While we commend the Memo for establishing regular reporting requirements for agencies, we believe that the current reporting intervals may not be sufficient to keep up with the pace of AI development and procurement.

To address this issue, we propose that the OMB increase the frequency of information requests from agencies. Specifically, we recommend that, under the existing remit of the OMB, the reporting requirements be updated from at least annually to at least every six months, maintaining the opportunity to further shorten this interval should the pace of AI advancement and procurement continue to accelerate. This more frequent updating will provide the OMB with a clearer picture of how AI systems are deployed and used across government agencies, allowing for more timely identification of potential risks and misuse.

We also recommend that agencies be required to provide updates on their ODD taxonomies as part of their regular reporting, in accordance with our second recommendation. These ODD updates should be accompanied by the results of audits and evaluations, demonstrating that procured AI systems as deployed are operating solely within their intended operational domain. Agencies should also report on any changes to their intended applications, especially those that arise from new releases, security patches, or updates to procured systems. Agencies should be required to further report any expansions to the operational scope of AI systems since the previous reporting period, as well as relevant evaluations that have been performed to ensure the system's reliability, safety, and compliance with OMB standards in the proposed use case. OMB should be granted the authority to reject the expansion of the AI system's operational conditions should evidence of compliance with the standards within the new domain of use prove insufficient. This information will enable the OMB to assess the ongoing compliance of agencies with the Memo and the effectiveness of established risk management practices.

Closing Remarks

We would like to thank the OMB for the opportunity to provide comments on the OMB memorandum on Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence. We believe that the recommendations described above will help ensure the responsible procurement and use of AI systems across the government.