MITRE's Response to the OMB RFI on Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence Draft Memorandum

December 5, 2023

For additional information about this response, please contact:

Duane Blackburn
Center for Data-Driven Policy
The MITRE Corporation
7596 Colshire Drive
McLean, VA 22102-7539

policy@mitre.org
(434) 964-5023

About MITRE

MITRE is a not-for-profit company that works in the public interest to tackle difficult problems that challenge the safety, stability, security, and well-being of our nation. We operate multiple federally funded research and development centers (FFRDCs), participate in public-private partnerships across national security and civilian agency missions, and maintain an independent technology research program in areas such as artificial intelligence, intuitive data science, quantum information science, health informatics, policy and economic expertise, trustworthy autonomy, cyber threat sharing, and cyber resilience. MITRE's 10,000-plus employees work in the public interest to solve problems for a safer world, with scientific integrity being fundamental to our existence. We are prohibited from lobbying, do not develop or sell products, have no owners or shareholders, and do not compete with industry—allowing MITRE's efforts to be truly objective and data driven. Our multidisciplinary teams (including engineers, scientists, data analysts, organizational change specialists, policy professionals, and more) are thus free to dig into problems from all angles, with no political or commercial pressures to influence our decision making, technical findings, or policy recommendations.

MITRE has a 50-year history of partnering with federal agencies to apply the best elements of artificial intelligence (AI) and machine learning (ML) to advance agency missions while developing and supporting ethical guardrails to protect people and their personal data. Our team's experience with the entirety of the AI/ML adoption and life cycle has strengthened our ability to anticipate and resolve future needs that are vital to the safety, well-being, and success of the public and the country.

Introduction and Overarching Recommendations

MITRE reviewed with interest the recent Executive Order (EO) on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence[1,2] and the Office of Management and Budget's (OMB) draft implementation memorandum. The federal government can serve as a catalyst to accelerate the development of tools and framework to mitigate risks and enhance the trustworthiness and security of AI systems, and the approaches outlined in the EO will help advance that outcome considerably. Our overall reactions are provided below, with answers to the OMB request for information's (RFI's) specific questions in the section that follows.

Level of selected governance model. Managing and regulating emerging technologies like AI is challenging due to the field's rapid evolution, wide-ranging adoption possibilities, and the community's developing understanding of AI's risks. The EO and draft memorandum take a balanced governance approach, allocating complementary responsibilities across both the Executive Office of the President (EOP) and individual federal agencies. This approach avoids the pitfalls of a single entity approach, which can result in decisions lacking context and

[1] Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. 2023. The White House, https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/. Last accessed December 1, 2023.

[2] Please note that any unspecified references to an EO within this document are explicitly referring to this particular EO (14110 of October 30, 2023).

regulatory confusion. It also avoids the pitfalls of a completely distributed approach, which demands duplicative expertise and is nearly impossible to implement consistently or analyze effectively at the whole-of-government level.

However, this balanced governance approach necessitates substantial effort in bridging the gap between EOP policymakers and agency implementation, as well as between government and industry. Activities such as establishing a common understanding, standardized frameworks, and shared infrastructure are crucial. While EOP staff may be able to prioritize these tasks initially, such coordination efforts run the risk of being sidelined over time (especially across presidential administrations), jeopardizing the long-term success of the initiative. Therefore, MITRE recommends that OMB incorporate this bridging function into its implementation plan for the EO. Various coordination approaches can be employed, including a federal coordination office, an Information Sharing and Analysis Center, a Center of Excellence, a consortium, and/or an FFRDC. MITRE has experience with each of these approaches and would be pleased to discuss their advantages and disadvantages in helping OMB and federal agencies achieve their AI goals.

EOP messaging. It is essential for the EOP to strike a better balance between focusing on managing risks posed by and ensuring agencies feel encouraged to leverage AI to enhance their missions. By fostering a supportive environment for responsible and impactful AI use, agencies will be more inclined to adopt AI technologies and contribute to the overall advancement of AI in the federal government.

The EOP should also openly lead by example by initiating projects to responsibly leverage AI in meeting their own missions. This would not only set an example for agencies to follow but also provide the EOP with insights and experiences that will be necessary in fulfilling their oversight and interagency coordination responsibilities.

Flexibility and adaptability in AI governance. The importance of flexibility and adaptability in AI governance cannot be overstated, as each agency will have different needs and requirements based on their size, organization, budget, mission, and internal AI talent. Common best practices, guidance, training, and a list of priority risks of common concern[3] will help agencies navigate the AI landscape without OMB direction becoming overly prescriptive.

Clear governance inputs and processes. Based on our extensive experience in cyber governance, we understand that efficient governance necessitates well-defined inputs, processes, outputs, roles, responsibilities, information sharing, and decision-making authorities. The same will hold true for AI governance. This clarity is crucial regardless of whether a new AI governance body is established or AI governance is integrated into existing governance bodies. A well-defined governance structure will facilitate the responsible planning, development, deployment, and use of AI across federal agencies.

Collaboration and coordination. The advancement of responsible AI innovation and the achievement of EO objectives rely heavily on collaboration and coordination between agencies, as well as with external stakeholders and subject matter experts (SMEs). By fostering a culture of cooperation, the federal government can harness the collective knowledge and expertise of various stakeholders to ensure the safe, secure, and trustworthy development and use of AI

[3] Developing something similar to Appendix 5 within the Capital Programming Guide (https://www.whitehouse.gov/wp-content/uploads/2021/01/capital_programming_guide.pdf) supplement to OMB Circular A-11 could be useful.

technologies. The EO and this draft OMB memorandum naturally focus on internal federal government activities, but efforts to ensure the federal government collaborates with and leverages insights from nongovernmental organizations will also be required.

Requirements for agencies that are part of a CFO Act agency. The draft memorandum prescribes requirements for agencies that are considered a CFO Act agency[4]. A CFO Act agency must designate a Chief Artificial Intelligence Officer (CAIO) responsible for the requirements specifically called out for CFO Act CAIOs, as well as the broader CAIO requirements. Agencies that are not CFO Act agencies must also establish a CAIO. However, the current memo is confusing regarding whether a CFO Act agency that has one CAIO at the top level (e.g., HHS) obviates the requirement for (subordinate) agencies within that CFO Act agency to establish a CAIO (e.g., for FDA, CMS, CDC, etc.). Several sections of the draft memorandum call out different requirements for all CAIOs versus CFO Act CAIOs specifically (e.g., Section 3(a)(i), Section 3(a)(ii), Section 3(b)(i), Section 3(b)(iii)), but the guidance can be interpreted as requiring a "top-level" CAIO only for CFO Act agencies. MITRE recommends that OMB better clarify these requirements and their expectations on different categories of agencies in the memo.

Resourcing. The requirements outlined in the EO and this draft OMB memorandum are extensive. Moreover, our discussions throughout the federal AI community suggest that the community has likely been underestimating the resources necessary to properly assure AI. We recommend that OMB proactively work with agencies (in the near term) as well as with Congress (on future budgets) to find appropriate ways to ensure resources are available to meet these requirements.

In conclusion, MITRE emphasizes the importance of a balanced governance model, flexibility and adaptability in AI governance, clear governance inputs and processes, and collaboration and coordination between agencies and external stakeholders. By addressing these key points, the federal government can more effectively advance responsible AI innovation and meet the objectives outlined in the Executive Order and OMB memo.

Questions Posed in the RFI

1. The composition of Federal agencies varies significantly in ways that will shape the way they approach governance. An overarching Federal policy must account for differences in an agency's size, organization, budget, mission, organic AI talent, and more. Are the roles, responsibilities, seniority, position, and reporting structures outlined for Chief AI Officers sufficiently flexible and achievable for the breadth of covered agencies?

The roles, responsibilities, seniority, position, and reporting structures outlined for CAIOs are sufficiently flexible and achievable from an overall perspective. Maximizing flexibility in meeting these requirements will be critical given the wide variance of agency sizes, operating functions, and management approaches taken. However, there are areas for recommended improvement, as listed below.

[4] As defined under the Chief Financial Officers (CFO) Act of 1990.

Authorities and accountability. One major gap in this memorandum is specificity on authorities and accountability of the CAIOs within the agency's management and decision-making processes. Effectively describing these management and decision-making authorities and accountabilities will benefit not only the CAIOs, but also other leaders and staff throughout the agency. These authorities and accountabilities should not be fully specified within the OMB memorandum, but OMB should select a minimal core set that agencies could fine-tune (and expand upon) for their unique situations. Without doing so, the chances of confusion within agencies and uniformity across agencies are high, which could impede both AI oversight and the agencies' business operations.

The clarification of management and decision-making authorities and accountabilities is crucial for individuals in the CAIO role, as it ensures they have the necessary authorities outlined in Section 3.b.H of the draft memo to eliminate barriers effectively. However, mandating these authorities may create issues for some agencies, as there could be valid existing reasons or processes that need to be addressed before removing the barriers appropriately. OMB must exercise caution when defining expectations for CAIO authorities and roles within agencies to avoid creating unnecessary bureaucratic layers or circumventing critical decision-making processes.

A crucial aspect of AI governance is the review and management of investments, which ideally should be integrated with an agency's existing investment management process. AI leadership should have a say in investment and resource allocation decisions. However, it is not advisable to designate funding for AI exclusively under the management of the CAIO. Instead, the CAIO role should function as a collaborative intermediary between the acquisition process and the mission-oriented stakeholders (with intended users) who are funding the development. This approach fosters better coordination and alignment of AI investments with the agency's overall objectives.

Resourcing. For large agencies with multiple operating divisions, the work required to perform the tasking in this EO and memorandum will be quite extensive, necessitating more than simply designating a single individual as a CAIO. These large agencies will likely need a dedicated staff office to perform CAIO functions. These agencies should signal this need to OMB and work collaboratively with them to determine how to adequately resource the staff to ensure success.

Creating new bodies versus leveraging existing ones. Agencies should have the flexibility to choose the most suitable approach for both the CAIO position and the internal agency governance body. However, MITRE recommends that OMB provide some baseline guidance. For instance, an agency should not automatically create new AI governance positions and bodies just because the EO calls for those functions. Instead, new AI governance positions and/or bodies should be established only if the agency believes doing so will add significant value compared with utilizing existing governance positions and/or bodies. In the short term, agencies must also consider whether they possess adequate AI expertise to staff their desired approach, or if temporary alternative approaches need to be explored.

The requirement in 3.c.i of the draft memo that agency AI Governance Boards must be vice-chaired by the CAIO could in many cases effectively preclude agencies from using existing boards to fulfill this function, which conflicts with providing this as an option. OMB should rethink this requirement to ensure that use of existing boards truly is an option for agencies to consider.

CAIO position qualifications. The minimum qualifications of a CAIO provided in 3.b.i of the draft memo are vague. While that vagueness may be necessary in the early stages of implementing this memorandum, OMB should also initiate work to more fully define minimum requirements (or at least best practices) and initiate training programs to prepare the selected officials for success. The MITRE paper *Interagency S&T Leadership*[5] provides experience-driven insights on how to effectively provide such leadership from both inside a federal department/agency and from an EOP/interagency perspective, and can be leveraged to help in fulfilling this task.

The requirement in 3.b.iii of the draft memo for CAIOs to be senior executive service (SES) equivalents in CFO agencies could create near-term issues given the short timelines provided in this memorandum. Finding individuals who both are already SES certified and have the required AI experiences to succeed in this role will be extremely difficult. Requisite AI expertise, relationships with agency senior leaders, and formally dictated authorities and accountabilities will be much more important than "checking an SES box" upfront. Rather, OMB should consider this SES requirement to be an aspirational goal for agencies to reach within a defined timeframe rather than an upfront requirement that could create more issues than it solves.

Communications. In addition to formal reports to OMB on implementing this memorandum, agency CAIOs should also be expected to (1) document and exchange lessons learned and best practices with their peers in other agencies and (2) collaborate with their agency's external communications teams to explain how they are leveraging and assuring AI, as well as regularly highlighting the benefits that their use of AI has provided.

2. What types of coordination mechanisms, either in the public or private sector, would be particularly effective for agencies to model in their establishment of an AI Governance Body? What are the benefits or drawbacks to having agencies establishing a new body to perform AI governance versus updating the scope of an existing group (for example, agency bodies focused on privacy, IT, or data)?

There are many potential governance coordination mechanisms and approaches, the best fit of which will vary based on an agency's AI goals and status of existent governance processes and mechanisms. Before making decisions on those, however, MITRE feels that a more definitive description from OMB is needed about what AI governance entails.

AI governance. First, agencies will benefit from defining "AI governance" to ensure clarity in their approach to establishing AI governance mechanisms. The following definition of AI governance is offered based on MITRE work: *AI governance is the set of processes, practices, principles, and policies that an agency uses (within and adjunct to existing governance structures) to direct and control its planning, development, deployment, and use of AI to meet mission needs. The purpose of AI governance is to manage and address potential risks, harms, and challenges posed by AI while maximizing the benefits of AI and ensuring alignment with trustworthy AI principles (e.g., safety and security, human rights protection, privacy,*

[5] D. Blackburn. Interagency S&T Leadership. 2023/2016. MITRE, https://www.mitre.org/sites/default/files/2023-08/PR-16-0916-interagency-s-and-t-leadership-full-paper.pdf.

fairness/equity, responsible and ethical use of AI (including accountability), transparency, explainability, robustness and reliability, and traceability).

Second, more definitive descriptions of the types of processes included in AI governance are also needed. MITRE recommends that an agency's AI governance—whether integrated into or established alongside existing governance processes—include the following processes and coordination mechanisms:

- AI strategy development, including defining the agency's AI strategy (direction), monitoring progress of strategy implementation, and refreshing the AI strategy when needed. The AI strategy should align with the agency's overall strategic goals and objectives.
- A definition of Trustworthy AI and operating principles, aligned with the agency's mission.
- AI investment management, integrated into the agency's existing investment management process. This activity includes the regular maintenance of the agency's AI Use Case Inventory (required by EO 13960) as well as overall awareness and tracking of the agency's current AI efforts inclusive of AI prototypes, pilots, operationalization projects, and (when appliable) programs of record. Including AI efforts in existing investment management reviews ensures that AI investments align with the agency's strategic objectives, ethical principles, and risk management practices. In addition to common elements captured in investment reviews (e.g., strategic alignment, risk, resource allocation, performance measurement), additional investment review elements to capture for AI efforts include (but are not necessarily limited to) ethical/responsible use, impacted users, safety impacts, rights impacts, and regulation compliance. Capturing and tracking the pipeline of AI efforts helps the agency make more effective decisions on whether to move an AI effort forward to the next stage based on measured value demonstrated.
- AI risk and harm management approach and procedures as part of existing agency investment and risk management processes, including identification of potential risks and harms (especially safety-impacting and rights-impacting risks), mitigation strategies, and monitoring.
- AI assurance, a lifecycle process that provides justified confidence in an AI system's ability to operate effectively with acceptable levels of risk to stakeholders. AI assurance provides the operational framework for risk management and procedures, as part of existing agency investment and risk management processes, including identification of potential risks and harms (especially safety-impacting and rights-impacting risks), mitigation strategies, and monitoring. The AI assurance approach should provide AI capability development oversight, which may entail an independent internal-agency SME committee reviewing AI development efforts throughout their lifecycle and, when needed, providing recommendations to AI governance decision-making body(ies). An AI capability development oversight process may also include algorithm "registration" and SME committee review checkpoints to ensure AI-enabled capabilities are functioning as intended, will integrate into agency business processes and workflows, will integrate into agency IT systems, will be accepted by business owners and users, and adhere to the

agency's trustworthy AI principles. If a similar oversight process exists for IT or analytics projects, the agency should identify whether AI efforts can be addressed in the existing process or whether a special committee is needed to review the unique aspects of AI.

- AI policy creation and response, which includes creation and dissemination of agency policies related to AI planning, experimentation, prototyping, piloting, and operationalization, as well as agency response to government oversight policies and regulations.
- External AI stakeholder collaboration and partnerships to share best practices, tools, and data; obtain feedback; and stay informed on emerging trends and regulatory developments. Public-private partnerships are an effective way for government, industry, and public interest groups to develop solutions that address many stakeholder perspectives. These entities provide a voice to often underrepresented groups and allow diverse opinions to be heard in a moderated and safe environment. Considering the viewpoints of the public and community representatives is especially important for safety-impacting and rights-impacting AI. Public-private partnerships can offer valuable insights to government agencies, leading to better policies, practices, and AI-enabled systems.

Of the above processes and practices that are part of AI governance, two should be integrated into existing agency processes: risk management and investment management. An AI capability development oversight process may be able to be integrated into existing project reviews and checkpoints for analytics or IT, depending on scope of and participation in the existing process.

Coordination mechanisms. Additional coordination mechanism recommendations for effective AI governance include:

- Clearly defined and communicated AI governance processes, featuring well-defined inputs (e.g., identifying AI efforts by risk level and life-cycle stage), procedures, outputs (decisions), and decision-making roles, authorities, and responsibilities. This includes identifying who has decision-making authority and who has responsibility for acting on decisions. The agency's AI Use Case Inventory serves as a valuable mechanism for clarifying inputs to AI governance processes.
- Involvement of stakeholders and subject matter experts in governance forums to inform decision makers. These stakeholders and SMEs may include acquisition managers, project/program managers, AI developers, data managers, ethicists, business owners, end-user representatives, business process specialists, attorneys, and others as relevant to the AI use case, type of AI effort life-cycle stage, and decisions being made.
- Allocation of sufficient resources to organize AI governance activities, such as creating governance decision-making schedules, providing quality assurance of incoming materials, prioritizing topics for review, tracking action items to completion, managing risk status, and communicating with stakeholders. These resources may report to the CAIO or other senior leader role fulfilling AI governance duties.
- Integration of the AI governance decision workflow into a workflow system, as part of existing governance process workflows where applicable and to the extent possible.

Updating existing governance or establishing new AI governance. As mentioned earlier, governance best practices recommend incorporating AI governance into existing agency governance processes, practices, and bodies wherever possible. When discussing AI investments/projects in governance processes, additional unique factors must be addressed, such as responsible use of AI, visibility to vendor data and models, and impacts to safety and rights. Addressing these unique AI factors can be achieved by bringing AI SMEs (e.g., acquisition managers, developers, data scientists, program/project managers, ethicists, attorneys, business/mission representatives) into governance forums to provide guidance and recommendations to decision makers.

- Clear objectives, processes, decision-making authorities, roles, and responsibilities are crucial for effective governance. The agency must precisely outline the decisions that need to be made concerning various AI initiatives, such as prototypes, pilots, operationalization projects, and enterprise programs. It is essential to determine when and where these decisions should be made (within specific governance processes and forums), who is responsible for making them (decision-maker), who will be affected by them (the recipient of the decision), and who must take action based on the decisions. This clarity ensures a well-structured and efficient governance system for AI efforts within the agency.If existing governance bodies and governance processes are used for AI governance, the agency must be very clear about when and how AI decisions are integrated into existing bodies and decision-making processes, and who is accountable and responsible for decisions related to the planning, development, deployment, and use of AI.
- If a new AI governance body and separate AI governance processes are established, it is essential to clearly define the roles, responsibilities, decision-making authorities, and the inputs and outputs that interact with other (preexisting) governance processes and bodies. Furthermore, it is necessary to develop criteria that indicate when a topic should be escalated from AI governance to the next higher level of authority. This clarity ensures a well-structured and efficient governance system that effectively integrates AI governance with the broader organizational framework..

Recall that AI governance set up separately from existing governance structures should primarily focus on agency AI strategy development, internal AI standards (e.g., trustworthy AI principles and AI-specific policies), AI partnerships, and AI assurance. AI capability development oversight may be integrated into existing project review processes (e.g., for IT or analytics projects) or may need to be set up separately. Existing governance authorities and processes should add AI representation to their decision-making bodies to inject subject matter expertise, provide risk management, and conduct investment reviews and make decisions on possible AI-enabled solutions to mission-focused needs.

The benefits and drawbacks of establishing a new AI governance body versus updating existing governance bodies are outlined below:

- Benefits of establishing a new (and dedicated) AI governance body:
    - Dedicated focus on AI governance, which can elevate AI into mainstream decision-making processes and lessen financial and mission risk

  - Increased opportunities for AI adoption in the agency given increased communication and awareness of AI's potential across programs and mission areas
  - Increased focus on attracting and retaining AI-specific expertise, which may not be present in existing governance bodies (such as those focused on privacy, IT, or data)
  - Greater visibility and prominence for AI issues, which can help in raising awareness and promoting accountability
  - Potentially increased adoption of agency AI standards that align with national/international AI regulations and policies
  - Potential for lessened AI-related risks
- Drawbacks of establishing a new AI governance body:
  - Potential duplication of efforts and resources, as existing governance groups (especially data and analytics governance) may already be working on AI-related investment decisions, risk management, oversight, and other AI topics
  - Potential duplication of existing governance processes that may not be AI specific but AI relevant.
  - The need for additional resources, such as funding and personnel, to establish and maintain a new AI governance body
  - Probable challenges in coordinating and integrating the work of the new AI governance body with existing groups and initiatives, which may create conflict and/or confusion and may minimize/overtake existing responsibilities
  - Possibility of governance fatigue if membership overlaps too much across the new and existing bodies
- Benefits of updating the scope of an existing governance body(ies):
  - Leveraging existing resources, expertise, and relationships, which can help establish efficient and effective AI governance that is integrated with other investment management and risk management
  - Leveraging existing governance processes that may not be AI specific but AI relevant.
  - Avoiding duplication of efforts and resources, as existing groups can build on their current work to address AI governance issues
  - Easier integration and coordination with other existing groups and initiatives, as the updated group would already be part of the existing governance ecosystem
- Drawbacks of updating the scope of existing governance body(ies):
  - The potential for AI governance issues to be overshadowed by other priorities and concerns within the existing group
  - The possible lack of AI-specific expertise within the existing group, which may limit its effectiveness in addressing AI governance issues

- Challenges in adapting the existing group's processes and structures to accommodate the unique complexities and challenges of AI governance

3. How can OMB best advance responsible AI innovation?

In addition to the guidance OMB provides in its draft memo, the following are other ways for OMB to advance responsible AI innovation:

First and foremost, OMB should lead by example by incorporating AI into its own core functions, many of which stand to benefit significantly from AI integration. By doing so, OMB can demonstrate the potential benefits AI can bring in terms of enhancing mission effectiveness, reducing time and costs, and ensuring security and privacy, thus encouraging agencies to follow OMB's example in their own missions and use cases. For example, OMB could leverage AI-driven analytics and decision-making tools to optimize budget allocation and resource management by analyzing historical data, predicting future trends, and identifying inefficiencies.

Relatedly, OMB could initiate a new and substantial activity that encourages agencies to leverage AI while meeting the current President's Management Agenda. This could involve, for instance, working with federal agencies to identify areas where AI can be used to improve service delivery, reduce wait times, and enhance the overall customer experience.

Such activities will not only enhance OMB's own mission but also provide OMB with hands-on insights and experiences that are necessary to effectively oversee and guide federal department and agency activities, as well as earn the respect of other federal agencies in the process. By actively embracing AI and showcasing its potential, OMB can effectively promote responsible AI innovation across the federal government and its various sectors.

Address resource requirements for AI capabilities. The EO and this draft memorandum mandate numerous additional activities that federal agencies must undertake using their existing funding. While this is a standard aspect of EOP leadership, the size and scope of activities called for within this EO and draft OMB memo are unprecedented. The resources required to meet these requirements are substantial. Moreover, our discussions throughout the federal AI community suggest that the community is likely underestimating the resources necessary to properly assure AI as they use it to enhance agency missions. Pilots to better characterize cost requirements are needed. Operating in a level budget environment while addressing these factors necessitates the reallocation of resources from other activities to meet these demands. OMB has a critical role to play in this regard and should proactively work with agencies (in the near term) as well as with Congress (on future budgets) to find appropriate ways to address this resource issue.

As a first step, OMB should proactively identify existing requirements (e.g., Privacy Act implementation, Federal Data Strategy) that are similar or closely aligned to the new AI requirements and then develop strategies to meet common requirements with a single work product.

Foster public-private partnerships to help agencies meet AI objectives (from the EO and OMB's draft memo) and accelerate the advancement of priority initiatives. The federal government recognizes the value of AI knowledge and use cases from various sources, including nongovernmental entities. These entities will also be developing insights and lessons learned, which the federal government should proactively seek to leverage. As we previously mentioned in our discussion on AI governance, engaging a broad set of stakeholders is critical not only for knowledge-sharing but also for risk management; different stakeholders can perceive and articulate diverse risks in AI utilization. Likewise, OMB should ensure that agencies engage a diverse range of stakeholders throughout the AI life cycle, including business/mission owners, user representatives, ethicists, legal SMEs, business process designers, organizational change management SMEs, and human-machine teaming SMEs. By fostering collaboration and knowledge-sharing, OMB can further promote responsible AI innovation across various sectors.

Better balance OMB's focus on safety/security with encouraging agencies to leverage AI. The EO and this draft OMB memorandum both place significant emphasis on ensuring safety and equity in AI development and deployment. While these aspects are undoubtedly crucial, there is comparatively less focus on mandating and ensuring that agencies proactively work to leverage AI appropriately to enhance their missions. To address this imbalance, OMB should allocate more attention to identifying, characterizing, and promoting the practical application of AI within agencies and emphasize the importance of harnessing AI's potential to improve operational efficiency and effectiveness. In doing so, OMB should ensure that its communications more fully promote both connected objectives – the responsible development and deployment of AI, as well as its practical application to enhance agency missions. By striking a better balance between these two aspects, OMB can foster a more comprehensive approach to AI innovation that not only ensures safety and equity but also actively encourages agencies to harness the transformative power of AI to improve their operations and better serve the public.

Additionally, OMB should explicitly state that decisions on whether to use AI should be based on its improvement over the status quo, rather than how close it comes to reaching an idealized state of accuracy and bias. Current (non-AI) practices are not, by any means, free from error or bias. The pursuit of "perfection" should not overshadow solutions that are simply "significantly better" than current practices and can still provide substantial benefits.[6]

Refine the concept of "responsible use" and establish clear expectations. Currently, the term "responsible use" lacks a precise definition, leading to varying interpretations and potential inconsistencies in its application. To address this issue, OMB should develop a consistent definition and baseline expectations. OMB and agencies can incorporate this understanding into their AI governance processes, such as AI strategy development, mission need/problem identification, the AI use case selection process, the AI algorithm oversight process, the AI assurance process, and the AI risk management process. By embedding responsible use principles in these processes, agencies can ensure that AI systems are developed and deployed in a manner that aligns with ethical, legal, and societal considerations.

[6] This is discussed in further detail in *MITRE's Response to the OSTP RFI on a National Artificial Intelligence Strategy*, available at https://www.mitre.org/sites/default/files/2023-07/PR-22-01891-22-MITRE-National-AI-Strategy.pdf.

Missing or minimized elements:

- ***AI assurance***. While the EO references AI assurance a few times, this draft OMB memo does not explicitly define or address AI assurance as a distinct concept, although some aspects are scattered throughout the document. AI assurance should be a clearly visible building block for the EO's implementation. It is a key part of AI governance that informs and is informed by other AI governance processes; governing AI systems without assurance is not feasible.

  MITRE defines "AI assurance" as *a life-cycle process that provides justified confidence in an AI system's ability to operate effectively with acceptable levels of risk to its stakeholders*.

  - Operating effectively entails meeting functional requirements with valid outputs.
  - Assurance risks may be associated with or stemming from a variety of factors depending on use context, including but not limited to AI system safety, security, equity, reliability, interpretability, robustness, directability, privacy, and governability.

- ***Acquisition considerations***. The memo currently provides limited guidance on acquisition considerations, even though they will significantly influence the federal government's use of AI. Therefore, acquisition should be addressed more comprehensively. It can be helpful to consider AI-specific aspects for acquisition in much the same way the community has considered cybersecurity throughout the system's life cycle: establishing clear business and technical requirements, ensuring fielded solutions comply with appropriate AI policies and regulations, verifying and validating algorithms and data relative to the established requirements and to criteria defined to assure trustworthy AI, and validating to an Authority to Operate, with ongoing measurement and monitoring thereafter.

- ***Accountability and protections for agencies and federal staff***. The EO and memo currently do not provide protections for agencies and federal staff when AI-enabled outcomes are not ideal, even if they exhibited due diligence in their AI acquisition, deployment, and governance decisions. Addressing this aspect is essential to encourage responsible decision making; otherwise, agencies and federal staff may remain too risk averse.

- ***Enhancing agency AI workforce endeavors***. The OMB memo's inclusion of non-practitioners in its AI workforce planning requirements for agencies is positive. The internal agency AI workforce planning and development requirements in the OMB memo can be enriched by identifying four workforce categories that agencies need to address: AI end-users (such as employees whose work is directly impacted by AI), technical AI practitioners (e.g., AI developers, AI engineers, data scientists), non-technical AI practitioners (e.g., acquisition managers, project/program managers, business process designers, governance actors, risk managers, investment managers), and the general workforce (for overall agency workforce AI literacy). For each of these four workforce categories, agencies can use specific strategies to build workforce capacity for AI.

- ***Whole-of-government workforce endeavors***. The memo does not identify any whole-of-government AI workforce initiatives. This is a notable gap, as there are likely many

similarities in workforce development activities across federal agencies that could be more effectively and consistently addressed through a holistic approach. In a recent MITRE project, we investigated whole-of-government workforce investment needs for AI and a handful of other critical and emerging technologies, which are summarized in Figure 1 below. This can be a good starting point for subsequent OMB analysis.

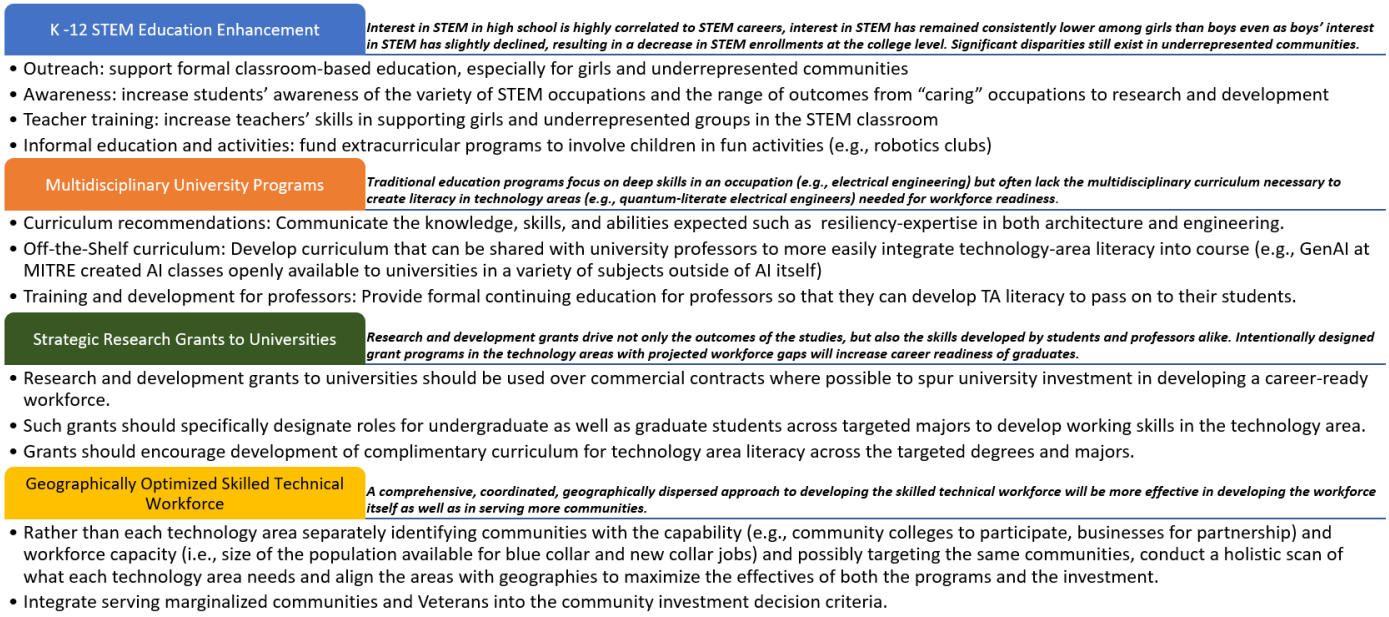

Figure 1 - Workforce Needs for Critical & Emerging Technologies

4. With adequate safeguards in place, how should agencies take advantage of generative AI to improve agency missions or business operations?

Agencies should take a phased approach to adopting generative AI. Adopting an enterprise view of generative AI (and AI in general) provides guardrails that increase trustworthiness and enable increasing sophistication, while managing and mitigating risk. A principled approach to generative AI adoption will enable more success through tools, processes, and frameworks that adapt to evolving risks and requirements. Agencies will need to leverage third-party foundation models for a variety of use cases, ranging from relatively low-risk situations to those with very high consequences, spanning diverse business needs.. Such adoption should follow a healthy prescriptive of "crawl, walk, run," where an agency should:

1. Initially focus on value-added generative AI capabilities (and AI capabilities in general) that have a high probability of end-user adoption success and low risk.
2. Learn from that experience and evolve AI efforts and supporting processes and practices that can inform the adoption of additional generative AI capabilities.
3. Advance generative AI adoption until the desired improvements and transformational changes in mission operations and/or service delivery are achieved.

Preparing an agency for adopting responsible generative AI (and AI in general) can be addressed by the agency executing the three phases described below. Each phase will yield products that advance the agency's ability to adopt generative AI more efficiently and with more assured systems.

1. ***Assess***. Perform an assessment of the agency's current and desired states of generative AI maturity using the AI Maturity Model.[7] A guided AI maturity assessment will yield current and desired states, potential blockers, and priorities for building the agency's capacity for AI implementation in general, as well as progress measures. Aspects to assess are governance and strategy; ethical, equitable, and responsible use; organization; technology enablers; data; and performance and application. As part of this assessment, examine acquisition practices as well as business processes and workforce from enterprise and organizational change management perspectives.
2. ***Strategize***. Define an overarching AI strategy that articulates the agency's goals and objectives for using AI to improve mission operations and service delivery. Identify current and near-term planned AI use cases to solve important mission problems and needs. For each use case, identify the business and end-user impacts and corresponding business process and organizational (e.g., role, responsibility) changes needed for successful use case implementation. Demonstrated mission value, end-user and business owner acceptance, and fit within business processes and workflows will all be required for successful AI adoption. Also identify infrastructure changes (e.g., data pipelines, computing resources, tools, software) needed to support the agency's pursuit of the AI use case. Use the results of the assessment; the selected AI use case(s); and the identified infrastructure, business process, and organizational changes to develop a roadmap with checkpoints, measures of progress, and sequenced generative AI efforts.
3. ***Enact***. Perform generative AI projects in sequence, expanding scope from experiments to proofs-of-concept, prototypes, and operationalization projects when value is shown. Measure agency AI maturity progress at checkpoints, adopt beneficial practices, and iterate learning. The sequences for projects will follow the "crawl, walk, run" strategy mentioned above that starts with straightforward, mature AI technologies and advances in complexity as AI project experiences build the organization's acceptance of and capacity for AI over time. Progress measures in the roadmap are assessed at checkpoints and will identify beneficial changes and enable learning for course correction.

With these guardrails in place, agencies should explore and leverage generative AI for a wide range of use cases. To begin, an agency can pursue low-hanging fruit with generative AI applications that are immediately possible and low risk. Success here builds workforce acceptance of and citizen confidence in leveraging generative AI, which in turn builds agency momentum to focus on more advanced use cases involving more risk but greater rewards from positive outcomes. Below are a few examples (from simple and lower risk to more complex) of the many evolving opportunities agencies might take to responsibly use generative AI to improve service delivery and mission operations:

[7] E. Bloedorn et al. The MITRE AI Maturity Model and Organizational Assessment Tool Guide. 2023. MITRE, https://www.mitre.org/sites/default/files/2023-11/PR-22-1879-MITRE-AI-Maturity-Model-and-Organizational-Assessment-Tool-Guide.pdf.

- Use "out-of-the-box" third-party generative AI for basic functions such as summarizing meeting notes, drafting internal communications, and drafting documents and emails, etc.
- Use a conversational large language model (LLM), also known as a chatbot, to help with information gathering applications, such as helping a person understand what information may be necessary for filling out a form.
- Take a third-party LLM used for information retrieval and summarization and tune/augment it on agency-specific data to enhance the agency's ability to discover information from its large repository of unstructured documents.
- Use an LLM to train an agency-supporting chatbot that in turn would have expertise in answering questions (e.g., about employee benefits or helping connect an employee to training or new work opportunities).
- Have non-technical workers use an LLM "agent" that can take a natural language request for information and translate it into a technical (e.g., SQL) database query to access corporate records.
- Use generative AI to create future "trajectories" of agency services, resource consumption, or behavioral patterns of workers or clients for the purpose of robustness and resilience planning for possible future events and circumstances.
- Use an agency-augmented LLM to assist in course of action planning, where the user would tell the AI model the objective, and the AI would help a human develop a plan, drawing on both internal agency information and external knowledge.

5. Are there use cases for presumed safety-impacting and rights-impacting AI (Section 5 (b)) that should be included, removed, or revised? If so, why?

The use cases outlined in Section 5(b) of the draft OMB memo appear suitable, and we do not have substantive revisions to suggest. However, we would like to offer two related recommendations that OMB will find beneficial as further detailed and mature work in this area progresses: tool qualification and synthetic data.

Tool qualification. The example use cases for safety-impacting and rights-impacting AI primarily focus on the deployed system itself—the running code in use. While this is undoubtedly the first priority, it may be beneficial to also consider the concept of "tool qualification" to increase rigor and reduce potential rework. Tool qualification refers to a process that promotes the safety of a system by minimizing the risk of tool errors, either by reducing the number of errors or ensuring that they do not impact safety. Although establishing justified confidence in the deployed and to-be deployed systems that are rights-impacting and safety-impacting is of utmost importance, requirements for tool qualification should likely be considered a secondary priority. Introducing the topic now could prompt valuable efforts to include this aspect in the justification process.

In safety-critical sectors such as transportation, tool qualification is a mandatory safety procedure, as outlined in standards DO-178C and DO-330 for aviation[8] and ISO-26262[9] for automotive safety. These standards typically define three levels of tool qualification:

1. Development tools that could introduce errors into operating safety-critical software
2. Verification tools that might fail to detect errors but are used to reduce other development or verification activities
3. Verification tools that might fail to detect errors but are not used to reduce other development or verification activities

The Tool Qualification Level (TQL) required for a tool depends on its criteria and the criticality of the software for which it is used. DO-330 provides detailed guidance on tool qualification, acknowledging the differences between the execution environments for airborne software and tools.

Errors introduced into safety-impacting or rights-impacting AI are well covered. However, some tools may be used to verify or characterize rights-impacting or safety-impacting AI, and they should also be considered based on their role. A machine learning-based test case generator and automated test harness could, for example, miss failure conditions if it failed in execution. This would increase the risk that latent failures in the AI system to be deployed are not detected and corrected, reducing confidence in the final system.

A detailed set of verification tool use cases is not required, but it would be useful to introduce the concept of tool qualification for verification tools in the rights-impacting and safety-impacting development chain.

DO-330 also addresses the use of previously qualified tools. In summary, reusing a previously qualified tool is permitted if the developer can demonstrate, through a change impact analysis, that the tool still meets its TQL requirements despite any changes in the operational environment or the tool itself.

Synthetic data. A second related recommendation is to focus on the application of AI and/or synthetic data in the design, development, or assessment of other AI capabilities that are considered safety-impacting or rights-impacting. This use case is important to consider because AI systems and synthetic data are increasingly being utilized in the development and evaluation of other AI technologies. Incorporating AI and synthetic data in the design and assessment process can potentially introduce biases, inaccuracies, or other unintended consequences that may impact the safety and rights of individuals. Therefore, it is crucial to ensure that AI systems and synthetic data used in these contexts adhere to the same standards of safety, security, and trustworthiness as the AI capabilities they are helping create or evaluate.

To address this use case, this memo should provide guidance on the following aspects:

---

[8] DO-330 Software tool qualification considerations: When, where, and how it applies. 2023. LDRA, https://ldra.com/do-330/. Last accessed December 1, 2023.

[9] ISO 26262-1:2018 - Road vehicles — Functional safety — Part 1: Vocabulary. 2018. International Organization for Standardization, https://www.iso.org/standard/68383.html. Last accessed December 1, 2023.

- Ensuring that AI systems and synthetic data used in the design, development, or assessment of safety-impacting or rights-impacting AI capabilities are subject to the same requirements and oversight as the AI capabilities themselves.
- Establishing best practices for the use of AI and synthetic data in the design and assessment process, including transparency, documentation, and validation of the AI systems and synthetic data used.
- Encouraging collaboration and information sharing among agencies and stakeholders to identify and mitigate potential risks and challenges associated with the use of AI and synthetic data in the development and evaluation of safety-impacting or rights-impacting AI capabilities. This collaboration can also share best practices and opportunities to use synthetic data to enhance confidence in the safety or rights-preserving assurances for an application.

By including this consideration in the memo, the government can ensure that AI systems and synthetic data used in the design and assessment of safety-impacting or rights-impacting AI capabilities are held to the same high standards of safety, security, and trustworthiness, ultimately leading to more reliable and responsible AI applications in the public sector.

6. Do the minimum practices identified for safety-impacting and rights-impacting AI set an appropriate baseline that is applicable across all agencies and all such uses of AI? How can the minimum practices be improved, recognizing that agencies will need to apply context-specific risk mitigations in addition to what is listed?

MITRE has the following recommendations to enhance these minimum practices:

- Section 4.D, pg. 17. We suggest adding a requirement that the monitoring for performance degradation must not focus only on aggregate performance measures that can mask disparate outcomes for various population cohorts. For example, some facial recognition systems may have reasonable overall precision and recall averaged across the entire population but show significant underperformance for women of color compared with the aggregate and over-performance for white males compared with the aggregate. An explanation of the population cohorts used to analyze errors and the results by cohort should be included in the documentation of the monitoring process and model performance.
- Appendix I; pg. 26; Consolidated Table of Actions. We recommend clarifying with an additional table footnote that rows 6 (Section 5(a)(i)), 8 (Section 5(c)(iv)(D)), and 9 (Sections 5(b) and 5(c)(iii)) apply to AI used by each agency, including DoD agencies, on information systems other than National Security Systems.
- Section 4(b)(ii); pg. 9; Data. Making datasets publicly accessible should be weighed against the potential for misuse by domestic or foreign actors who seek to thwart U.S. interests.

- Section 5(c)(iv)(A)(3); pg. 15. Cases where agencies cannot access data and must obtain sufficient descriptive information require establishing requirements for AI or data providers to follow.
- Missing definitions:
  - Section 5(c)(v)(A)(3); pg. 19; "Communities"
  - Section 6; pg. 24; Recommend adding an item (8) that explicitly speaks to AI security more comprehensively.
  - Section 6; pg. 25; #2 should be clarified such that it refers to climate or environment outcomes resulting from actions taken on the basis of AI outputs and not to climate or environment outcomes resulting from the training of AI.

7. What types of materials or resources would be most valuable to help agencies, as appropriate, incorporate the requirements and recommendations of this memorandum into relevant contracts?

Create a federal AI acquisition best practices guidebook. The Department of the Air Force and Massachusetts Institute of Technology's AI Accelerator has developed an AI Acquisition Guidebook[10] focused on providing a basic understanding of the AI acquisition life cycle related to data, finance, and legal considerations, which can be leveraged to create a federal AI acquisition best practices guidebook.

Develop and share guidelines for market research, RFIs, solicitations, and contracts/agreements for agencies to use. These guidelines should address aspects such as how AI could be applied as a full or partial solution; what (if any) embedded AI solutions a vendor plans to propose (e.g., commercial solutions); and language that provides the acquiring government agency visibility into the training, testing, and validation of datasets used by the vendor as well as the vendor's AI model. In addition to guidance offered in Section 5.2 of the DoD Office of Developmental Test, Evaluation & Analysis's Systems Engineering Processes to Test AI Right,[11] the following may be considered for acquiring AI solutions:

- Treat the AI model as a subcomponent with its own specific test and evaluation (T&E) requirements.
- Ensure contracts have access to test the AI model subcomponent and test infrastructure resources.
- Plan for periodic model retraining post-deployment by monitoring the system in operation to ensure that the periodic retraining is actually improving or at least maintaining performance.

[10] Artificial Intelligence Acquisition Guidebook. 2022. Massachusetts Institute of Technology, https://aia.mit.edu/wp-content/uploads/2022/02/AI-Acquisition-Guidebook_CAO-14-Feb-2022.pdf.

[11] C. Balhana et al. Systems Engineering Processes to Test AI Right (SEPTAR) Release 1. 2023. MITRE, https://apps.dtic.mil/sti/trecms/pdf/AD1211716.pdf.

- Determine and consider model training data approaches in contracts, such as buying off-the-shelf pre-trained models, providing government-furnished information, or paying contractors to gather data.
- Adopt Data Cards and Model Cards to document dataset contents and trained ML model specifications.
- Establish agreements to document the AI method (e.g., Convolutional Neural Network) and, possibly, its training parameters to enable effective T&E and risk assessment.
- Address training data, test data, and (potentially) AI-generated data ownership, management, location, access, security, rights, retention, and provenance in agreements and memorandums.
- Communicate the relevance of the AI threat during contracting and use resources like DoD's Validated Online Lifecycle Threat to document known threats.
- Include clauses in contracts to address additional or added scenarios where the AI-enabled system is expected to be deployed in conditions beyond its originally intended use.
- In requests for proposals, require respondents to advise on the presence of AI and the nature of its implementation, and consider engaging T&E SMEs to support the evaluation process.

Train the federal acquisition workforce on AI. Topics should include AI "101" fundamentals (how AI differs from traditional software), best practices for acquiring AI-enabled systems, and program and contract considerations for AI acquisitions. Deploy this training through a variety of methods, including through the Federal Acquisition Institute.[12] Additional considerations include:

- Engage with the Defense Acquisition University (DAU) to collaborate on key learning objectives for AI/ML training. DAU has developed an introductory Machine Learning course video suitable for technical and non-technical audiences.[13]
- Engage with the OPM-moderated Chief Learning Officers Council[14] to share leading practices and learn from each other to train the acquisition and procurement workforce at federal agencies. Identifying AI career paths and succession planning may also be an important contribution that Chief Learning Officers can provide to their respective agencies.

Develop plans to cultivate an organizational culture that embraces AI as part of doing business as well as acquiring AI solutions. Evolve organizational mindsets from risk averse and compliance-focused to embracing new technology that enables processes and solutions that

[12] Federal Acquisition Institute. 2023. General Services Administration, https://fai.gov/. Last accessed December 1, 2023.

[13] R. Maus. SWE 0057 What is Machine Learning?. 2023. Defense Acquisition University, https://media.dau.edu/media/t/1_hsij16av. Last accessed December 1, 2023.

[14] Chief Learning Officers Council. 2023. Office of Personnel Management, https://www.chcoc.gov/content/chief-learning-officers-council-cloc. Last accessed December 1, 2023.

meet standards and requirements. Focus on recruiting an AI NextGen workforce to enable AI-driven solutions and processes.

Establish auditing standards to ensure that vendors of AI tools are honoring agreements in contracts. For example, maintain government awareness of the evolution of any "black boxes" post-testing/acceptance and constant monitoring and correction as more datasets are added and outputs may drift unexpectedly.

8. What kind of information should be made public about agencies' use of AI in their annual use case inventory?

The Advancing American AI Act and Executive Order 13960 require agencies to "prepare and maintain an inventory of the AI use cases of the agency, including current and planned uses." This information is to be shared across agencies and with the public. To date, 21 agencies have listed about 500 use cases on inventories available via ai.gov, with six agencies declaring no AI use cases. These inventories vary widely in terms of the fields provided, depth of description, and completeness of the information. Many of the 21 inventories reuse fields from other inventories.

MITRE supports several sponsors that have created systems to track AI capabilities and how they are used. These systems organize information into the following categories:

1. Description: Information about who developed the capability, how it was developed, and what it is intended to do
2. Performance: Information about how the capability was tested and its performance, resilience, and usability
3. Use: Information about how the capability is used (more specific than its intended function), who is using it, what it is approved for, and how it is performing in operation

Description information should include known stakeholders (beyond the users of the capability) and information about whether those stakeholders have been or can be involved in the development and/or monitoring of the capability. Stakeholders who may be impacted include entities whose data is used in the use case and people whose jobs may be affected by the use case.

Information about the development of the capability should include a description of any training and test data used and the AI method (e.g., natural language processing) used to create the capability. This information, combined with performance information, can be used by others to validate the correctness of the approach.

Section 4.b of the RFI says, "Where an agency currently uses or plans to use AI for a purpose described below, the CAIO, in coordination with other relevant officials as specified by the agency, may make a determination (or reverse a prior determination) that the AI application or component does not match the definitions of 'safety-impacting AI' or 'rights-impacting AI' and is therefore not subject to the minimum practices." Description information should include the results of this determination for the capability and its intended use, with supporting information about the reasons for this determination. The benefits of doing this include greater public

transparency, greater ease for auditing, and clearer and more consistent criteria for third-party developers to appreciate earlier in their modeling process.

Performance information should include specific details about metrics used to determine whether to approve the use case. Typically, precision and recall are important metrics, but the use case may weigh one more than the other and may have additional metrics, including operating factors like size of the trained model and how long it takes to process a new input. Performance information should include information about the diversity of test conditions and variation in performance across them.

Use information should include the current status of the capability, whether planned, under development, in pilot, in full operation, and so forth. Many of the existing inventories include fields for this information, but currently list only capabilities that are in operation or that are planned. It is important for agencies to maintain inventories of both planned and operational AI use cases.

Use information should include what the agency has approved as responsible uses of the capability. The memorandum should elaborate the definition of what must be considered when determining responsible uses, including impacts to safety, rights, and jobs. This information should address both uses of the capability (including potential uses outside the agency and reuse within the agency) and uses of the information produced by and decisions made by the capability. Finally, when possible, use information should reference mission efficacy of AI that has been deployed. Although it may not be possible to establish cause and effect, it may still be beneficial to benchmark relevant mission performance indicators before AI deployment against post AI deployment.

An important consideration for the memorandum is which aspects of this information should be made available to the public, and how often. While transparency is important for ensuring equity and enhancing quality, it also can enable adversarial attacks on AI capabilities by providing insight into how those capabilities work and how they are used. Disclosure may also impact the privacy of people whose data were used in the development of a capability, because some attacks can reverse engineer training data from an operational model. OMB may need to provide for third-party technical review of inventory information to identify risks of disclosure and maximize transparency.

AI incident reporting should also be considered in the context of any published, deployed AI use cases. Sharing AI incidents through collaboration frameworks such as MITRE ATLAS[15] helps AI system developers and owners mitigate risks for the government and public. In cases where publicly disclosing such information may not be permissible, anonymized disclosure approaches can be used through trusted parties.

[15] MITRE ATLAS™ (Adversarial Threat Landscape for Artificial-Intelligence Systems). 2023. MITRE, https://atlas.mitre.org/. Last accessed December 1, 2023.