

Effective AI Management Unlocks Innovation

Purpose-built AI strategy and governance helps federal agencies mitigate risks while amplifying benefits.

Now that artificial intelligence (AI) technologies, including machine learning and generative AI (GenAI), have moved from experimental to mainstream, organizations across the private and public spectrum stand at a technological crossroads as they fight to catch up to the whirlwind speed of these advancements.

“The State of GenAI Today: The Early Stages of a Revolution,” a Guidehouse report produced in partnership with CDO Magazine, reveals that more than three-quarters of data and IT leaders surveyed said their organizations are not fully equipped to harness GenAI today. And 72% reported that GenAI applications are not yet incorporated into their existing data governance and management structures. Respondents from organizations of all sizes said they anticipate that it will be a struggle to do so while acknowledging the transformative nature of capitalizing on the technology’s potential.

Unique Challenges for a Federal Agency

Federal agencies in particular face escalating pressure to leverage AI’s many benefits while managing its inherent risks. Until recently, many of those agencies lacked internal risk



assessment practices suited for the increasingly sophisticated discipline. An executive order issued on October 30, 2023, sought to rectify that by creating a permanent chief artificial intelligence officer (CAIO) position at every agency.¹ Agencies across the federal ecosystem now need a plan for strong AI management—one that simultaneously contains risks and develops a governance approach that maximizes potential benefits.

Guidehouse has been instrumental in helping agencies address these overlapping concerns and develop a sustainable, forward-looking plan for AI management. Our strategy—which is tightly aligned with the National Institute of Standards and Technology’s Artificial Intelligence Risk Management Framework (NIST AI RMF) and President Biden’s executive orders on AI—helps organizations develop safeguards to govern, map, measure, and manage AI in ways that will address the risks of AI systems in practice.²

Developing a Comprehensive AI Strategy

The single most important strategic step in an AI risk management and governance strategy is simply to commit to a need for a cohesive strategy. Without an internal authority, governance framework, and trusted expert advisors, AI governance will inevitably be unfocused. Untethered, policies may prioritize preventing headline-grabbing risks (even if those risks are not germane to the agency's AI needs), prompting agencies to clamp down on potentially beneficial innovations in the name of safety. Inconsistent policies make it challenging to proceed with AI projects in a compliant, approved manner.

Agencies may also lack the internal expertise to differentiate between unacceptable and worthwhile risks. Independent practitioners can provide context and field-tested lessons as AI governance is established and reviewed. Outside experts are also essential to the process of verifying and validating AI models, including validating the models' basic assumptions, data ingestion, and results.

Having adequate internal resources and expertise is critical to AI governance. Before publication of the October 30th executive order, the office of chief data officer (CDO) served as some agencies' clearinghouse for AI management, risk assessment, and guidance. Each agency's new CAIO can now directly lead strategy, including driving AI innovation, managing AI risks, and implementing priorities contained in past and future AI-related executive orders. The CAIO should also continue any existing work to de-silo AI regulations and controls already present in other parts of the agency.

An Executive Roadmap for CAIOs

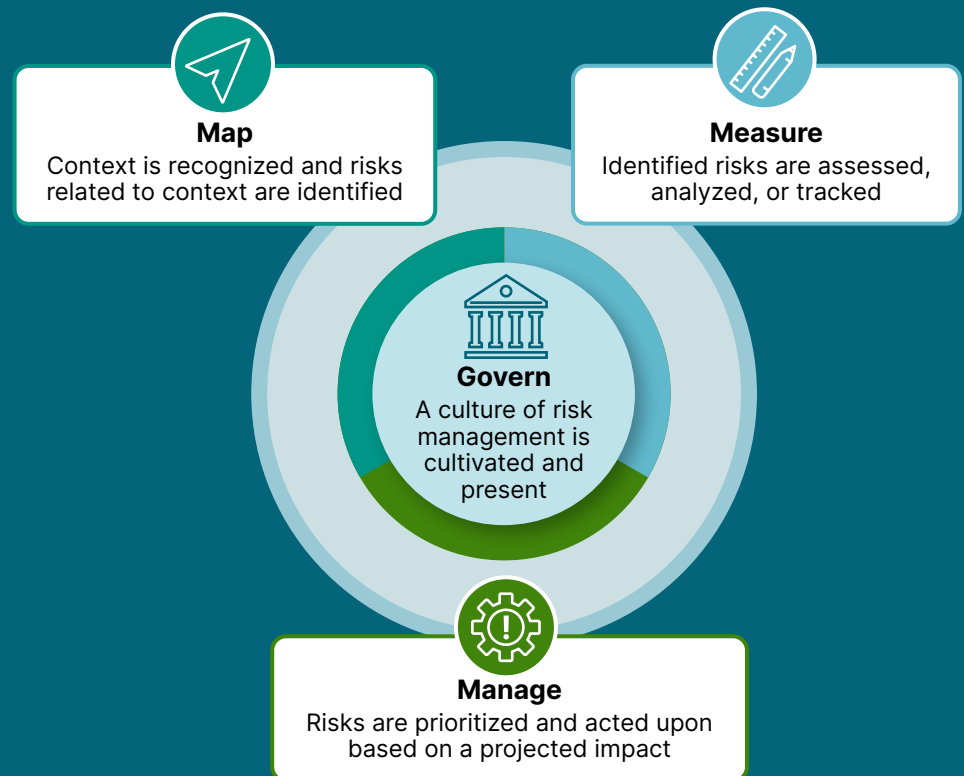
The executive order specifies a number of duties for the CAIO, including:

Promotion of responsible innovation—This includes guiding investment, research, and development efforts as well as addressing human rights and worker training.

Privacy —The CAIO is responsible for keeping the agency's use of AI from infringing on civil liberties and ensuring that data procurement and retention follow applicable laws and policies.

Accountability —This includes watermarking or labeling AI outputs when appropriate, controlling against deceptive or exploitative output from generative AI, and meeting ongoing reporting requirements.

The NIST AI RMF is designed to be comprehensive enough for federal entities while remaining flexible enough to be applied by organizations of many sizes and missions.² The NIST AI RMF organizes AI management operations around three equally weighted functions and one cross-disciplinary core:



MAP —Establish the context for the use of AI, including its intended beneficial uses and potential negative impacts.

MEASURE —Establish metrics to measure risks identified in the mapping stage, rate the trustworthiness of the AI-generated results, and track the collection and application of feedback across AI's field of influence.

MANAGE —Decide how to proceed based on the stated goals and purpose of the AI use weighed against the information gathered in the map and measure stages. Continually assess the risks and benefits introduced by third-party models and data sources.

GOVERN —In the NIST AI RMF model, the govern step overlaps with all three functions. This step includes putting in place formal statements of legal and regulatory requirements along with risk management and accountability practices. It also establishes expectations for a safety-first approach to AI usage.

Advantages of a Comprehensive AI Management Strategy

Taking control of AI governance and putting the right people on task is about more than avoiding inefficiencies, decision paralysis, and siloed operations. The process helps unlock value that might otherwise be hidden or delayed, including:

Effective management and governance processes —Clarity and transparency in the workforce is essential to an effective AI strategy. As new policies, procedures, and technologies are implemented, it's critical to support staff to build their capabilities and understand appropriate and approved AI uses.

Implementation of best practices —It's important to separate proven best practices from popular misconceptions. For example, the most powerful AI is based on neural networks and deep learning—models that are extremely hard to explain but may produce the best results. Expert review helps improve AI model transparency and provide clear guidance on when the model should and should not be used. A sensitive model with obscure inner workings can still have sound predictive value and provide benefit. Discarding all such tools out of hand could lead to missed opportunities or wasted resources spent reinventing something that already exists in a more mature form.

Streamlined handling of complex, high-security tasks —Some models work with extremely sensitive inputs or produce outputs with high-stakes implications for individuals, communities, or nations. The process of reviewing (adjudicating) these models is complex and often time-consuming. Working with experienced AI professionals can marshal the model through the right review process in an expedited, effective manner.

Improved understanding of false positives and false negatives —No model can be 100% accurate. There will always be some risk of both false positives and false negatives. With expert guidance, agencies can understand the tradeoffs of models tuned to accept more of one than the other, then accurately document the potential impact that incorrect predictions could have on the agency, affected individuals, and the broader ecosystem. Establishing appropriate thresholds and tolerance for these errors before an AI model is built can improve the performance of an AI solution and increase stakeholder satisfaction with its results.

Sustaining AI Management Practices

Agencies face significant challenges in articulating and maintaining top-notch AI management practices. Guidehouse works with agency CAIOs, CDOs, and other industry partners, in alignment with evolving guidance from NIST, to navigate these difficulties in a timely manner.

We bring experts in agency operations and emerging AI science to bear toward AI governance challenges. Throughout each stage, we use interpretations, perspective, and field-tested results to address the growing need for agency-specific AI management. As trusted advisors to several federal departments and agencies, we can assess current AI management practices and recommend changes that suit an agency's specific needs. This helps us take existing frameworks, including NIST AI RMF, and quickly adapt them to the unique day-to-day and long-term mission circumstances that each agency faces.

“The journey ahead demands both technological readiness and a vision that comprehensively integrates GenAI into the fabric of organizational strategy and operations.”

- Bob Audet, Partner, Guidehouse
Data & AI Solutions

As AI practitioners, we have developed extensive structure and rules enforcement across internal experiments in generative AI and AI automation. And as experienced modelers, our experts can help refine hypotheses being tested by AI models or explore additional datasets to uncover information that could add value or be a richer source of predictive power than the data already under consideration. We can also assess the potential endurance of a proposed model's predictions and spur development of better models that can iterate over time and be of greater use for a longer period.

As data experts, we can assess the safety and robustness of a model's training set with more refinement than an untrained practitioner. We can also monitor drift and skew found in the datasets that are training an AI model over time, then flag potential issues in performance and bias that those drifts might introduce.

Regulators and other supervisory bodies are often keenly interested in the explainability of AI models. Depending on the chosen algorithm and model design, explainability can be difficult to achieve. We can help ensure that any model approaching production is suitably explainable to meet internal and external stakeholder requirements. This helps people speak credibly to the potential benefits and harms of a model's outputs and ensures that the agency itself clearly understands the risks associated with data source inputs and outputs.

Most importantly, we recognize AI's strengths and weaknesses and the fact that not all problems are best solved with technology.

¹ “Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence,” October 30, 2023 (<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/>).

² “Artificial Intelligence Risk Management Framework” (AI RMF 1.0), January 2023, (nist.gov).

Contacts


April Fordyce, Director
Defense & Security Advanced Analytics
afordyce@guidehousefederal.com

Nong Nai, Director
Cybersecurity
nnai@guidehousefederal.com

About Guidehouse

Guidehouse is a leading global provider of consulting services to the public sector and commercial markets, with broad capabilities in management, technology, and risk consulting. By combining our public and private sector expertise, we help clients address their most complex challenges and navigate significant regulatory pressures focusing on transformational change, business resiliency, and technology-driven innovation. Across a range of advisory, consulting, outsourcing, and digital services, we create scalable, innovative solutions that help our clients outwit complexity and position them for future growth and success. The company has more than 16,000 professionals in over 55 locations globally. Guidehouse is led by seasoned professionals with proven and diverse expertise in traditional and emerging technologies, markets, and agenda-setting issues driving national and global economies. For more information, please visit guidehouse.com.

 guidehouse.com/services/data-analytics-intelligence

 [@GHTechSolutions](https://twitter.com/GHTechSolutions)  [linkedin.com/showcase/guidehouse-technology-solutions/](https://www.linkedin.com/showcase/guidehouse-technology-solutions/)