Agile AI Partnerships

A Public-Private FLEXible and SMART Framework for National Security and Competitive Innovation[1]

Karen I. Matthews, PhD MBA

Recanati-Kaplan Fellow
Harvard Kennedy School
Harvard University

MAY 2025

HARVARD Kennedy School
BELFER CENTER
for Science and International Affairs
Intelligence Project

Table of Contents

About the Author

Dr. Karen Matthews is a distinguished leader in emerging technologies and innovation, with over 25 years of experience spanning the Department of Defense (DoD), academia, and industry. She is a DoD member and a Supply Chain and Emerging Technologies professor at the Haslam College of Business, University of Tennessee, Knoxville.

Dr. Matthews holds MEng and PhD degrees in Electrical Engineering, as well as an MBA, all from Cornell University. Her multidisciplinary expertise in engineering and business management has fueled a career dedicated to bridging the gap between emerging technologies and real-world applications.

As a Recanati-Kaplan (RK) Fellow at Harvard University, she focused on Agile AI Partnerships for national security and competitive innovation. Her research introduces the FLEX (Flexible Lifecycle Execution) framework, a groundbreaking approach to AI adoption and implementation that integrates technical, operational, and policy/legal dimensions while embedding cross-cutting themes—Data, Models, Metrics, Visibility, Security, and Compliance—throughout the AI lifecycle. Complementing FLEX is SMART (Systematic Mapping and Reuse Toolkit), a novel methodology that enhances efficiency by enabling the reuse of testing outcomes across similar use cases, reducing redundancy, and accelerating AI deployment. Together, FLEX and SMART ensure AI adoption is both strategic and operationally effective, aligning with national security imperatives while maintaining a technological edge over adversaries.

Before her Harvard fellowship, Dr. Matthews played a pivotal role in Technology Adoption within the DoD, spearheading efforts to transition innovative AI and emerging technology concepts into mission-critical capabilities. Her work at the intersection of policy, governance, and technological innovation has positioned her as a key strategist in AI implementation for national security.

In the private sector, she has led cutting-edge R&D, agile innovation initiatives, and end-to-end technology commercialization, providing her with deep insights into the entire technology lifecycle—from concept and development to deployment and operational impact.

Through her research, leadership, and advisory roles, Dr. Matthews continues to shape the future of AI and emerging technology adoption, fostering a new generation of leaders adept at navigating the complex intersection of technology, policy, and governance.

Acknowledgements

I would like to extend my heartfelt gratitude to all of the interview participants from the academic, public, and private sectors. Although I cannot list all of your names or specific organizations, please know that your candid insights and willingness to share your expertise were essential to the success of this research. Your contributions have enriched this work immeasurably, and I am deeply grateful for the time and effort you dedicated to these interviews.

I am also profoundly thankful to my colleagues in both the Recanati-Kaplan and National Security Fellowships. Your wisdom, experience, and camaraderie made my time at Harvard not only a tremendous growth opportunity but also an immensely enjoyable experience. Special thanks to Maria Robson-Morrow, Michael Miner, Mark Pascale, Hannah Fay, and Graham Allison for your exceptional leadership in the Intelligence Project. Your coordination, guidance, and flexibility have been invaluable, and I truly appreciate your commitment to looking out for our best interests while accommodating the individual needs, aspirations, and goals of each Fellow.

I am deeply indebted to all of my professors—Anil Arora, Leonie Beyrle, Logan McCarty, Jim Waldo, Bruce Schneier, Nicholas Roy, Erik Lin-Greenberg, David Alvarez-Melis, Finale Doshi-Velez, Teddy Svoronos, Sharad Goel, Dan Levy, Eric Rosenbach and Michael McQuade—for sharing your keen insights and experiences. Your willingness to involve esteemed colleagues and subject matter experts in our discussions has greatly enriched my learning experience. I remain in awe of the opportunities I have had here at Harvard.

This experience has been nothing short of transformative. I recognize that to whom much is given, much is required, and I am committed to channeling the wisdom I have gained into my purpose of service. I have tried my utmost to "squeeze" every drop of learning from this incredible opportunity to develop and grow. Thank you all for investing your time and knowledge in me; your impact on my life and career is profound. It has been my honor to learn from you, and I look forward to carrying these lessons forward in my continued service.

Executive Summary

This report offers a practical, actionable roadmap for agile AI adoption and implementation that leverages robust public private partnerships to support national security and sustain competitive innovation. Based on insights from 152 stakeholder interviews and 80 AI-focused sessions across public, private, and academic sectors, best practices from open-source documentation, and insights from AI-related conferences, Harvard courses and MIT courses—including contributions from initiatives such as the EU AI Act and the UN AI Advisory Board—this work reveals that while current opensource documentation provides high-level guidance (under the premise that each agency governs itself), there remains a critical need for detailed, actionable frameworks. Given the evolution of technology leadership and impact of open-source platforms, agile and collaborative frameworks are vital for maintaining a competitive edge and safeguarding national security.

The proposed framework, known as the FLEX (Flexible Lifecycle Execution) framework or FLEX, provides a structured yet agile approach for integrating advanced AI solutions. Acting not only as a comprehensive framework but also as a practical navigation tool, FLEX guides organizations through every phase of AI integration—from defining objectives and identifying risks to designing, testing, deploying, and continuously refining systems. Built on three primary layers (Technology, Operations, and Policy/Legal) and embedding six critical cross-cutting themes (Data, Models, Metrics, Visibility, Security, and Compliance) across its five lifecycle stages (Planning and Assessment, Design and Development, Testing and Validation, Deployment and Monitoring, and Continuous Improvement), FLEX ensures that strategic vision translates into actionable, measurable outcomes.

Complementing FLEX is SMART (Systematic Mapping And Reuse Toolkit), which enables efficient mapping and reuse of testing outcomes across similar use cases, operationally similar scenarios, or shared technical characteristics scenarios - reducing redundancy and accelerating deployment. SMART could be implemented using a simple grid format with axes representing risk level, use case similarity, and testing status, and is updated regularly as part of the continuous improvement cycle.

Together, FLEX and SMART ensure that AI adoption is both efficient and responsive, enabling agencies to implement advanced AI systems that align with national security imperatives while maintaining competitive advantage against adversaries. Unlike existing AI guidelines which remain high-level, FLEX provides concrete steps and a reuse mechanism (SMART) that directly addresses the common pain points such as the lack of standardization and redundant testing.

To integrate AI in a manner that is both innovative and secure, organizations should leverage FLEX and SMART, develop cross-functional teams, implement external review boards, launch pilot projects with iterative refinement, invest in training and capacity building, enhance public-private collaboration, ensure continuous monitoring and improvement, align with international best practices, integrate tailored compliance mechanisms, and address dual-use[2] challenges. These strategic actions empower organizations to translate high-level guidance into practical, actionable steps that drive agile AI adoption. By following these recommendations, agencies can transform abstract principles into concrete outcomes, ensuring sustained competitive advantage and strengthening national security.

Introduction

Rapid advances in artificial intelligence are reshaping sectors as diverse as national security and public service, offering transformative opportunities alongside significant challenges. While traditional open-source documentation provides high-level, nonbinding guidance—operating under the assumption that each agency can effectively self-govern—this approach often leaves practitioners without the practical, step-by-step methodologies required for real-world implementation. This report addresses that gap by proposing an actionable framework for agile AI adoption and implementation.

At the heart of this framework is the FLEX (Flexible Lifecycle Execution) Framework, which translates abstract principles into measurable actions across the entire AI lifecycle. FLEX offers a clear roadmap—from defining objectives and identifying risks to designing, testing, deploying, and continuously refining AI systems—ensuring that each stage of implementation is both rigorous and adaptable. Complementing FLEX is SMART (Systematic Mapping And Reuse Toolkit), a structured tool designed to map and reuse testing outcomes across similar use cases, thereby reducing redundancy and accelerating the deployment of advanced AI technologies.

This dual-tool approach is firmly rooted in robust public–private collaboration and is underpinned by a comprehensive, multi-source methodology. The framework draws on insights from 152 in-depth interviews and 80 AI-focused sessions with key stakeholders from the public, private, and academic sectors and best practices from open-source documentation[3]. Contributions from experts associated with initiatives such as the EU AI Act, the UN AI Advisory Board, and the Department of Defense—augmented by insights from influential conferences like Ergo (2024), The Ash Carter Exchange (2024), and the Hawaii International Conference on System Sciences (2025), as well as advanced coursework at Harvard and MIT—provided a systematic synthesis of diverse perspectives. This rigorous background not only reinforces the need for an agile, actionable framework but also ensures that the proposed approach respects ecosystem variety while emphasizing collaboration.

By clearly distinguishing between guidance—nonbinding recommendations—and governance—binding policies and enforcement mechanisms—FLEX and SMART bridge the gap between abstract principles and on-the-ground implementation. The framework empowers agencies to self-govern while adhering to best practices tailored to their unique challenges. Ultimately, the hypothesis underpinning this work is that an actionable, agile framework will enable U.S. agencies to integrate AI systems that align with national security imperatives, streamline testing processes, reduce redundancy, and sustain a competitive edge in a rapidly evolving technological landscape.

Evolving Leadership in AI: From Public Initiatives to Private Innovation

Artificial intelligence has undergone a profound transformation, shifting from government-led research to a dynamic era of private sector innovation that redefines both technological progress and leadership paradigms. The evolution of AI[4,5]—from early rule-based systems to today's sophisticated deep learning models—illustrates not only its transformative potential but also the inherent risks as the field advances toward Artificial General Intelligence (AGI) and beyond.[6,7,8] This evolution underscores why an agile framework is vital: as AI capabilities grow, so does the complexity of managing dual-use risks and ensuring secure implementation.

In the pre–Cold War era, technological breakthroughs were largely driven by private industry and entrepreneurial ventures, where innovations in communications, automobiles, and industrial machinery emerged from market competition and individual ingenuity. However, with the onset of the Cold War, national security imperatives catalyzed a dramatic shift. During this period, U.S. government agencies such as DARPA and NASA took the lead, channeling significant investments into aerospace, computing, and defense. This era marked a phase where public sector leadership was paramount, transforming the technological landscape through strategic research and development initiatives.

In the post–Cold War era, the balance shifted once again as the private sector reasserted its dominance. Rapid innovation in fields such as information technology, the internet, and digital communications ushered in a new era of private sector leadership. Today, the agile adoption of AI is underpinned by a synergistic blend of public oversight and private innovation. Robust public–private partnerships have emerged as critical for translating advanced AI technologies into secure, scalable, and operational solutions. Lessons learned from these collaborations emphasize iterative innovation, interdisciplinary engagement, and adaptability to varied regulatory environments. This integrated perspective reinforces the need for an agile framework that leverages the strengths of both public oversight and private innovation, ensuring that AI systems are developed in alignment with national security imperatives while sustaining a competitive edge.

Adversarial Acceleration: The Open-Source AI Revolution and the DeepSeek Imperative

The global AI landscape is rapidly transforming as the open-source movement not only democratizes innovation but also empowers adversaries and strategic actors to weaponize technology. Authoritarian regimes and state-aligned actors are harnessing openly accessible AI models to shorten development cycles and outpace traditional, state-controlled research methods. Evidence from recent intelligence assessments indicates that some adversaries are already repurposing these tools[9] for developments that pose serious national security challenges—e.g., enhancing their cyber surveillance, intelligence gathering, and rapid prototyping capabilities.

A striking example is DeepSeek, a Chinese AI company founded in 2023 that launched its first AI model in December 2023. Since its model launch, the company has rapidly expanded its portfolio, including the release of DeepSeek-R1 in January 2025. The DeepSeek platform has quickly emerged as a pivotal tool in the open-source dynamic as its open accessibility and rapid iteration capabilities enable adversaries not only to enhance their existing AI systems at an unprecedented pace but also to strategically weaponize technology. This trend is further compounded by state-driven initiatives such as Made in China 2025[10,11,12] and China Standards 2035[13,14,15] which underscore a determined effort by authoritarian actors to standardize and accelerate AI innovation.

Moreover, while U.S. policies like the CHIPS and Science Act[16] are designed to bolster domestic semiconductor manufacturing and technological resilience, these measures alone may not suffice to counterbalance the competitive pressures arising from adversarial use of open-source AI.

To effectively mitigate these risks, it is imperative that U.S. agencies and companies adopt agile frameworks—such as FLEX and SMART—to integrate secure, cutting-edge AI systems while maintaining rigorous oversight. Proactive public–private collaboration and comprehensive testing protocols are essential to ensure that as adversaries leverage these open tools, U.S. decision-makers can better anticipate their advances, safeguard the nation’s technological edge, and protect national interests.

Best Practices from Current AI Frameworks, Toolkits and Guidelines[17,18]

In today's rapidly evolving AI landscape, a wealth of open-source documentation[19] provides high-level guidance; however, there remains a significant need for detailed, actionable frameworks that enable agile adoption and implementation. This research synthesizes best practices from existing frameworks and toolkits, and provides clear, step-by-step recommendations for practitioners. The focus is on leveraging public-private cooperation, ensuring that agencies can adopt advanced AI models and systems from the private sector while aligning with national security objectives and maintaining a competitive edge. Table 1 provides a high-level overview of the various organizations and their respective frameworks/toolkits/guidelines.

Table 1: Comparative Table of Organizations, intended users, focus areas and gaps

Organization	Document Name	Year	Users	Focus Areas	Key Gaps
DoD	Responsible AI Strategy	2021	Military personnel	Trust, oversight	Real-time systems
NGA	AI Assurance Framework	2023	Geospatial engineers	Reliability, accuracy	Adversarial testing
NSA	AI Ethical Principles	2022	Cryptographers, analysts	Security, ethics	External collaboration
FBI	AI Governance Guidelines	2022	Investigators, developers	Fairness, accountability	Public visibility
CIA	AI Ethics Framework	2023	Analysts, technologists	Security, interpretability	Collaboration
DIA	Responsible AI Guidelines	2023	Intelligence officers	Operational integrity	Unclassified use
Marine Corps Intelligence	Ethical AI Operational Framework	2023	Marines, strategists	Mission-critical AI	Interoperability
NIST	AI Risk Management Framework	2023	Developers, policymakers	Lifecycle risk management	Enforcement
DOE	AI Ethics and Governance Principles	2022	Energy researchers	Sustainability	Scalability
CDAO	Responsible AI Principles	2023	AI leaders, DoD personnel	Visibility, governance	Collaboration

Operationalizing Agile AI: Best Practices, Implementation Roadmap, and Collaborative Strategies

As organizations transition from strategic vision to concrete action in AI adoption, a holistic approach is required—one that integrates practical best practices, a clear implementation roadmap, and robust collaborative strategies. This comprehensive section outlines the core elements necessary for operationalizing agile AI, ensuring that technological innovations are both secure and adaptable while fostering dynamic public–private partnerships.

Agile AI Adoption Best Practices

A successful transition to agile AI begins with establishing solid foundational practices that guide every step of implementation. Organizations are encouraged to:

- **Initiate Agile Pilot Projects and Iterative Innovation:**
  Start with small-scale pilots that serve as controlled environments for experimentation. By adopting an iterative approach, teams can test concepts, gather real-world feedback, and gradually refine their strategies. This minimizes risk while promoting continuous improvement.

- **Foster Interdisciplinary Engagement:**
  Effective AI implementation requires collaboration among technologists, policymakers, legal advisors, and operational experts. Drawing on cross-functional perspectives ensures that technical systems are robust, legally compliant, and operationally viable. Clear communication and shared objectives enhance the integration of AI across different domains.

- **Emphasize Clear Documentation and Standardization:**
  Although widely used tools like model cards[20] and datasheets[21] offer valuable insights, consistent benchmarks, thorough documentation, and standardized reporting protocols promote visibility and facilitate the evaluation and replication of successful practices. This clarity is critical for accountability and for scaling AI solutions across multiple contexts.

- **Integrate Robust Security and Compliance Measures:**
  As AI systems become more integral to operations, embedding state-of-the-art cybersecurity protocols and compliance frameworks is essential. This proactive approach protects sensitive data and ensures that implementations adhere to evolving legal and regulatory standards.

- **Design for Interoperability and Scalability:**
  Future-proofing AI initiatives means building systems that can seamlessly integrate with existing infrastructure and scale to meet expanding operational needs. Flexibility in design allows organizations to adapt to new technologies and regulatory environments over time.

Implementation Roadmap and Next Steps

Translating best practices into actionable strategies requires a well-defined roadmap that guides organizations from initial planning to full-scale deployment:

- **Develop Detailed Implementation Plans:**
  Tailor action plans to your organization's unique mission, outlining clear milestones, responsibilities, and measurable outcomes. Detailed planning bridges the gap between high-level vision and on-the-ground execution.

- **Launch and Refine Pilot Programs:**
  Pilot projects provide invaluable real-world data and offer a low-risk environment for testing AI applications. Use these pilots to gather insights, validate hypotheses, and make iterative refinements before scaling operations.

- **Establish Training and Capacity-Building Initiatives:**
  Equip cross-functional teams with the skills and knowledge necessary for effective AI implementation. Robust training programs, comprehensive user guides, and capacity-building initiatives ensure that all stakeholders—from technical experts to decision-makers—are aligned and informed.

- **Implement Continuous Monitoring and Feedback Loops:**
  Deploy real-time dashboards, automated alerts, and regular review sessions to monitor system performance. These mechanisms enable proactive adjustments, ensuring that AI solutions remain responsive to emerging challenges and opportunities.

- **Foster External Oversight:**
  Engage independent experts and establish external review boards to provide impartial evaluations of AI systems. This external perspective helps identify potential vulnerabilities and validates that implementations adhere to best practices and compliance standards.

Enhancing Public-Private Collaboration and Continuous Adaptation

The rapid pace of AI innovation demands that organizations not only excel internally but also cultivate strong partnerships across sectors:

- **Establish Formal Partnership Channels:**
  Develop structured avenues for collaboration between government agencies, private companies, and academic institutions. Joint research initiatives, shared technology development, and information-sharing platforms foster a vibrant ecosystem that leverages interdisciplinary expertise.

- **Co-Develop Use Cases and Strategic Frameworks:**
  Engage stakeholders collaboratively to design and refine AI applications that address specific operational challenges. By co-developing use cases, organizations can ensure that AI solutions are both innovative and tailored to meet regulatory and practical needs.

- **Align with Global Standards and Best Practices:**
  Monitor and integrate international benchmarks to ensure that domestic practices are competitive on a global scale. This strategic alignment helps maintain interoperability, promotes innovation, and ensures compliance with evolving standards.

- **Embrace Continuous Improvement and Agile Adaptation:**
  The dynamic nature of AI technology requires an ongoing commitment to refinement. Regular system upgrades, feedback-driven adjustments, and long-term strategic planning ensure that AI initiatives remain effective and responsive to both technological advances and regulatory changes.

By integrating these best practices, actionable implementation steps, and collaborative strategies, organizations can operationalize AI in a manner that is secure, adaptive, and forward-thinking. This comprehensive approach not only translates high-level principles into practical outcomes but also establishes a resilient framework for sustaining competitive innovation and meeting national security imperatives in a rapidly evolving technological landscape.

Agile AI Adoption and Implementation Framework: A FLEXible and SMART Roadmap for Action

FLEX (Flexible Lifecycle Execution Framework) is an agile AI adoption and implementation framework that serves as a practical navigation tool designed to guide organizations through every phase of integrating AI solutions with agility. It is not a governance model; rather, it provides clear, actionable steps to help agencies adopt advanced AI models from the private sector while ensuring alignment with national security imperatives and maintaining competitive advantage against adversaries.

To further enhance efficiency and reduce redundancy, this framework integrates SMART (Systematic Mapping And Reuse Toolkit), a tool that maps and categorizes AI use cases based on risk and operational overlap. By leveraging SMART, agencies can reuse testing outcomes from low-risk applications to streamline the evaluation of higher-risk scenarios, operationally similar scenarios, or shared technical characteristics scenarios, ensuring rapid adaptation and continuous improvement.

FLEX (Flexible Lifecycle Execution) Framework[22]

FLEX is structured around three primary layers—Technology, Operations, and Policy/Legal—and embeds six essential crosscutting themes (Data, Models, Metrics, Visibility, Security, and Compliance) across five lifecycle stages. It details activities across each lifecycle stage, ensuring that each element of the framework translates into practical, measurable actions, and draws from Institutional Review Board (IRB) principles to ensure that AI adoption and implementation is conducted appropriately, with ongoing oversight and risk assessment mechanisms. A decision point at the end of each lifecycle stage determines whether the project progresses to the next phase. If all criteria and objectives are met, the project advances; otherwise, unresolved issues must be addressed before proceeding. If the project is no longer viable, termination is a viable outcome to prevent wasted resources. The objective and action steps required for each stage outlined below.

Stage 1: Planning and Assessment

Objectives:

Complete initial scoping, stakeholder engagement, continuous risk assessment, and market research.

Action Steps:

- *Define Objectives and Scope:*
  - Clearly articulate the AI use case, mission needs, and intended outcomes.
  - Identify and document key stakeholders, including internal teams and external partners.

- *Stakeholder Engagement and Market Research:*
  - Convene initial advisory and external review boards to guide planning.
  - Conduct market research to understand competitive benchmarks, user needs, and industry best practices.
  - Identify users in each part of the adoption curve (innovators, early adopters, early majority, late majority, laggards) and their respective access levels. Also identify the cross-functional team.
  - Establish communication channels for ongoing multidisciplinary input.
- *Initial Risk and Compliance Guidance:*
  - Perform a preliminary risk analysis that includes identifying potential technical, operational, and legal risks.
  - Define relevant infrastructure, regulatory, compliance, and security requirements early on.
  - Establish baseline performance metrics (e.g., data quality, model responsiveness) using insights from market research.

Stage 2: Design and Development

Objectives:

Translate planning insights into a concrete design that encompasses both conceptual planning and detailed technical execution.

Action Steps:

- *Conceptual Design:*
  - Draft a high-level design document incorporating stakeholder requirements, early risk insights, market research findings, and operational needs.
  - Outline key architectural principles with an emphasis on security, compliance, and visibility.
- *Detailed Technical Development:*
  - Develop detailed technical specifications, prototypes, and system architecture diagrams.
  - Integrate iterative feedback loops (from technical, operational, and legal perspectives) to continuously refine the design.
  - Build in technical safeguards (e.g., cybersecurity measures, interpretability features) to support operational robustness.
- *Documentation and Standardization:*
  - Create comprehensive design documentation (e.g., design cards) that concisely summarizes key system design elements and facilitates clear communication among stakeholders throughout the development process.

Stage 3: Testing and Validation

Objectives:

Ensure that the system performs reliably under realistic conditions while rigorously evaluating its resilience and risk profile.

Action Steps:

- *Pilot Projects and Simulations:*
  - Launch controlled pilots and simulations to validate performance, accuracy, and security.
  - Use scenario-based testing to assess system robustness under diverse operational
- *Performance Testing and Prompt Engineering:*
  - Conduct detailed performance tests, including prompt engineering evaluations, to measure the system's responsiveness and efficiency.
  - Compare performance metrics against the benchmarks identified during the market research phase. Define criteria if no prior criteria in your space exists.
- *Red Teaming and Adversarial Testing:*
  - Execute red teaming exercises to simulate adversarial scenarios, uncover hidden vulnerabilities, and stress-test the system under attack conditions.
  - Incorporate findings to enhance both technical defenses and operational protocols.
- *Iterative Feedback and Refinement:*
  - Collect real world operational data and feedback from end users.
  - Adjust technical, operational, and compliance aspects based on testing outcomes.
  - Preserve or Pivot: Preserving a technology or pivoting should be a strategic decision driven by continuous learning and adaptability. In an agile process, failing fast enables teams to identify weaknesses early, refine solutions, or redirect efforts toward more viable opportunities. If a technology proves valuable, refining and scaling it is essential; if not, pivoting ensures resources are invested where they will have the greatest impact. The key is to learn quickly, act decisively, and drive innovation forward with purpose.

Stage 4: Deployment and Monitoring

Objectives:

Roll out the system adaptively with continuous oversight to ensure robust performance and prompt issue resolution.

Action Steps:

- *Adaptive Deployment Strategies:*
  - Implement phased rollout approaches (e.g., A/B testing[23], pilot rollouts) to manage risks during deployment.
  - Ensure that deployment aligns with operational readiness and stakeholder expectations.
- *Real-Time Monitoring:*
  - Set up dashboards and automated alert systems to continuously monitor system performance and detect anomalies.
  - Establish mechanisms for immediate issue detection, including regular performance and cybersecurity reviews.

- *Documentation and Accountability:*
  - Maintain comprehensive audit trails and documentation for each deployment step to ensure visibility and accountability.
  - Schedule regular reviews to evaluate deployment effectiveness and system performance.

Stage 5: Continuous Improvement

Objectives:

Support ongoing system evolution through regular updates while establishing a clear process for decommissioning outdated systems.

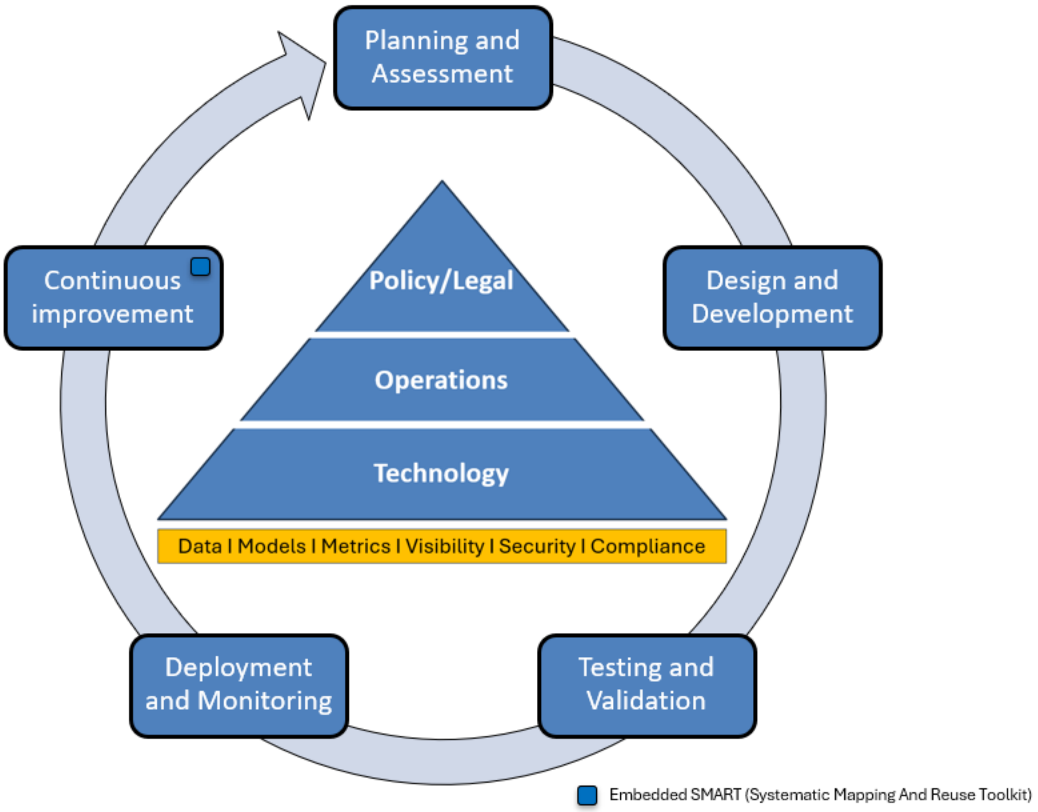
Action Steps:

- *Ongoing Feedback and Updates:*
  - Schedule periodic stakeholder reviews and audits to capture lessons learned and identify areas for enhancement.
  - Update AI models, system software, and technical safeguards in response to new data, technological advances, and regulatory changes, and push results to the SMART dashboard.
- *Process Refinement:*
  - Refine operational procedures and training materials based on real-world experience and continuous feedback loops.
  - Integrate new risk assessments and update compliance measures as standards evolve.
- *System Retirement:*
  - Develop a clear process for decommissioning systems or models that no longer meet performance, security, or compliance standards.
  - Include strategies for knowledge transfer and safe migration of critical functionalities.

This roadmap offers a clear pathway through each lifecycle stage by defining actionable deliverables. Designed with agility at its core, the framework supports adaptive implementation while fostering robust public-private partnerships that align with national security imperatives and competitive innovation. The schematic in Figure 1 graphically illustrates the multi-layered approach and cross-cutting themes, serving as a practical guide for organizations navigating the complex landscape of AI adoption and implementation. Appendix 6: Applied FLEX Framework Use Case Examples further illustrates the application of this framework by providing real-world examples.

This agile process emphasizes rapid learning and adaptation. Identifying challenges early and making necessary adjustments—or halting a phase if needed—enhances efficiency and effectiveness. Flexibility is key to optimizing outcomes and mitigating risks. Sharing insights with the team fosters collective learning, strengthening overall project success.

Figure 1: The FLEX (Flexible Lifecycle Execution) Framework


Planning and Assessment
Design and Development
Testing and Validation
Deployment and Monitoring
Continuous improvement
Policy/Legal
Operations
Technology
Data I Models I Metrics I Visibility I Security I Compliance
🟦 Embedded SMART (Systematic Mapping And Reuse Toolkit)

SMART (Systematic Mapping And Reuse Toolkit)

To enhance agility and reduce redundancy in AI testing and deployment, the SMART (Systematic Mapping And Reuse Toolkit) provides a structured approach for mapping AI use cases based on risk level, operational similarity, and shared technical characteristics. This approach allows agencies to optimize resources, minimize redundant testing, and accelerate AI adoption.

Key elements of SMART include:

- **Agility and Efficiency:** By systematically categorizing AI use cases, organizations can quickly identify overlaps between previously validated models and new applications, reducing unnecessary retesting and speeding up deployment.
- **Risk-Based Testing:** Lessons learned from low-risk AI applications can inform and streamline the evaluation of higher-risk scenarios. This targeted approach ensures that testing efforts focus on areas with the highest impact and security considerations.
- **Data-Driven Decision Making:** SMART provides structured benchmarks and checkpoints, enabling organizations to make informed decisions about AI deployment based on prior testing outcomes and real-world performance metrics.
- **Continuous Improvement:** The toolkit serves as a dynamic feedback mechanism, ensuring that

insights from previous AI implementations are leveraged to refine and accelerate future testing. Regular updates to the SMART mapping ensure that AI systems remain adaptive to emerging risks and technological advancements.

SMART can be visualized as a structured grid or mapping system that categorizes AI use cases along key dimensions such as risk level, similarity to previously tested applications, and current validation status. By embedding SMART within the Continuous Improvement phase of the FLEX framework, organizations can systematically refine their AI adoption processes while maintaining a balance between efficiency and security.

Key Insights and Lessons Learned in Agile AI Adoption[24]

Drawing from extensive stakeholder research and multi-sector engagements, several critical insights and lessons have emerged that inform effective AI adoption. These consolidated learnings underscore both the challenges and opportunities that organizations face as they transition from conceptual frameworks to practical, secure, and scalable AI implementations:

- **Actionable Frameworks are Essential:**
  Detailed, practical steps are needed to translate high-level guidance into tailored implementation strategies.

- **Early Adoption and Iterative Refinement:**
  Many organizations are in the nascent stages of AI integration, relying on pilot programs and iterative development to test and refine their approaches. This early-stage experimentation highlights the need for agile frameworks that can adapt to evolving challenges.

- **Need for Structured and Independent Review Mechanisms:**
  Formal oversight structures—modeled after Institutional Review Board protocols—and independent external assessments are vital for ensuring compliance with national security, legal, and standards. These mechanisms enhance visibility and accountability throughout the AI lifecycle.

- **Bridging Standardization Gaps:**
  While tools such as model cards and dataset datasheets provide valuable insights, the absence of standardized benchmarks complicates cross-comparisons and accountability. Establishing uniform standards is essential for promoting interoperability and clear, consistent documentation.

- **Emphasis on Cross-Functional Collaboration:**
  The successful adoption of AI requires interdisciplinary teams that bring together technologists, policymakers, legal experts, and operational leaders. This collaborative approach fosters robust decision-making and ensures that AI solutions address both technical and regulatory challenges.

- **Future-Oriented and Context-Sensitive Strategies:**
  Anticipating technological advancements—including the dual-use nature of AI—and accommodating regional and sectoral differences are crucial. Flexible frameworks that can adapt to both emerging risks and diverse operational contexts ensure resilience and global competitiveness.

- **Continuous Learning and Knowledge Sharing:**
  Ongoing training initiatives, regular system updates, and effective knowledge-sharing practices are key to fostering a culture of continuous improvement. These efforts help organizations stay ahead of technological changes and refine their AI strategies over time.

- **Building Trust Through Accountability and Visibility:**
  Clear documentation, performance benchmarks, and well-defined metrics[25,26] are critical for ensuring that AI deployments are auditable, reliable, and trustworthy. These practices build confidence among stakeholders and support long-term success.

- **Evolving Public-Private Dynamics and Sectoral Engagement:**
  While early AI governance discussions were predominantly driven by public and academic sectors, the growing role of private-sector innovation necessitates stronger collaborative efforts. Leveraging the strengths of both sectors can balance rapid innovation with robust security and compliance measures.

Together, these insights and lessons form a comprehensive foundation for agile AI adoption. They reinforce the need for actionable, adaptable frameworks that bridge the gap between high-level guidance and real-world implementation, ensuring that AI technologies are developed and deployed in a manner that is both effective and secure.

Pathways to Agile AI Adoption and Implementation

The dual nature of AI—its potential for transformative opportunities as well as significant risks, including its dual-use capabilities for both civilian and military applications—necessitates a strategic and practical approach to its adoption and implementation. As AI evolves from narrow task automation to broader, more advanced applications, adoption frameworks must remain dynamic and continuously adaptable, ensuring alignment with societal values while addressing technical and operational challenges.

Existing open-source documentation on AI for the private sector typically provides high-level guidance rather than prescriptive frameworks, operating on the principle that each organization is best positioned to determine how AI aligns with its unique mission. While this approach offers flexibility, there remains a pressing need for actionable methodologies that practitioners can use to implement AI while maintaining control over their processes. *FLEX* (Flexible Lifecycle Execution Framework) delivers structured, step-by-step methodologies that support agile AI adoption and implementation—without imposing rigid governance constraints. Moreover, to enhance efficiency and reduce redundancy, this paper introduces *SMART* (Systematic Mapping And Reuse Toolkit), a tool that maps and categorizes AI use cases based on risk and operational overlap. By leveraging this matrix, agencies can reuse testing outcomes from low-risk applications to streamline the evaluation of higherrisk scenarios, ensuring rapid adaptation and continuous improvement.

Strategic Recommendations for Agile AI Adoption and Implementation

To integrate AI in a manner that is both innovative and secure, organizations should adopt structured, adaptable frameworks. Key strategic recommendations include:

Table 2: Strategic Recommendations for Agile AI Adoption and Implementation.

Recommendations	Detailed Explanations
Leverage FLEX	Use detailed methodologies to translate high-level principles into practical strategies. Employ SMART at the continuous improvement stage to optimize testing efforts and streamline the adoption process.
Develop Cross-Functional Teams	Assemble multidisciplinary teams combining technical, legal, policy, compliance, and operational expertise.
Implement External Review Boards	Engage independent experts for impartial evaluations to ensure early identification and mitigation of risks.
Launch Pilot Projects with Iterative Refinement	Initiate controlled pilot initiatives to assess risks, gather real-world feedback, and iteratively refine processes before full-scale implementation, using SMART to map and reuse findings across similar use cases.
Invest in Training and Capacity Building	Develop robust training programs and AI literacy initiatives to ensure that all stakeholders understand the risks, benefits, and practical requirements of AI adoption.
Enhance Public-Private Collaboration	Foster strategic partnerships and knowledge-sharing initiatives between government agencies, private companies, and research institutions to drive innovation while aligning with national security imperatives.
Ensure Continuous Monitoring and Improvement	Establish real-time dashboards, automated alert systems, and regular review cycles to continuously evaluate system performance and adapt to emerging risks.
Align with International Best Practices	Benchmark against globally accepted standards and adapt strategies as needed to ensure interoperability and consistency.
Integrate Tailored Compliance Mechanisms	Develop project-specific compliance, security, and accountability guidelines, ensuring that each use case supports both innovation and national security objectives.
Address Dual-Use Challenges	Recognize and plan for the dual-use nature of AI, ensuring that technologies are adopted with clear safeguards to prevent unintended military or non-civilian applications.

By following these strategic actions, organizations can transform high-level principles into practical, actionable steps that drive agile AI adoption, ensuring sustained competitive advantage and national security.

Conclusion: A FLEXible and SMART Approach to Agile AI Adoption and Implementation

The rapid ascent of AI presents transformative opportunities alongside significant challenges—including the dual-use nature of many applications. A proactive, agile approach to AI adoption is essential for navigating this complex landscape. Understanding the historical context strengthens the argument for why contemporary U.S. agencies must actively partner with the private sector. This historical perspective reinforces that, given the evolution of technology leadership, agile and collaborative frameworks—like the proposed framework FLEX (Flexible Lifecycle Execution Framework) and SMART (Systematic Mapping And Reuse Toolkit)—are vital for maintaining a competitive edge and safeguarding national security.

FLEX–offers a comprehensive roadmap that integrates actionable methodologies, continuous monitoring, and iterative improvement. By fostering robust public-private collaboration and leveraging interdisciplinary expertise, FLEX empowers practitioners to deploy AI solutions that are secure, effective, and aligned with both societal values and national security imperatives.

Furthermore, the inclusion of SMART as a mapping tool enhances FLEX's agility by enabling efficient reuse of testing outcomes across similar use cases, thereby reducing redundancy and accelerating deployment. For policymakers, defense analysts, and industry leaders, this paper provides a clear, practical roadmap for integrating advanced AI technologies while safeguarding public trust and ensuring that the United States remains competitive against adversaries. Supplemental insights available in the appendices further enrich this roadmap, offering detailed analyses on global investment trends, emerging large language models, strategic procurement considerations, and historical lessons in public-private collaboration.

Appendices

The following appendices can be found linked to this report on the Belfer Center's Intelligence Project website: https://www.belfercenter.org/programs/intelligence-project. The appendices include four use cases to serve as a guide to using the framework.

Appendix 1: Types of AI, Key Terms and Definitions

Appendix 2: Historical Evolution of AI

Appendix 3: Made in China 2025 (MIC 2025) and China Standards 2035 (CS 2035) Overview

Appendix 4: AI Frameworks, Toolkits and Guidelines

Appendix 5: FLEX Framework Details

Appendix 6: Applied FLEX Framework Use Case Examples

- Use Case #1: Agentic AI for Autonomous Mission Planning
- Use Case #2: Facial Recognition System for Public Safety
- Use Case #3: Customer Support Optimization Using AI-Powered Language Models
- Use Case #4: AI-Driven Emergency Response Coordination

Appendix 7: Interview Detail

Appendix 8: Metrics for Accountability and Transparency

Endnotes

[1] Several iterations of ChatGPT were used to re-write my initial sections and appendices of this paper (and subsequent updates to the sections and appendices), looking for better verbiage, phrasing, flow and consistency of the paper.

[2] Dual-use technologies are technologies that can be used for civilian or military purposes.

[3] Publicly available frameworks guidelines and toolkits- such as AI and Cybersecurity Executive Orders and Government Publications and U.S. National Security and AI Policy Reports and Documentation (such as those from DoD, NGA, NSA, FBI, CIA, DIA, Marine Corps Intelligence, NIST, DOE and CDAO).

[4] See Appendix 1: Types of AI, Key Terms and Definitions for details.

[5] See Appendix 2: Historical Evolution of AI for details.

[6] Cook, Jodie. *Open AI's 5 Levels Of 'Super AI' (AGI To Outperform Human Capability), Forbes (July 16, 2024). This article further contextualizes this progression, providing a structured roadmap for AI development. The framework begins with Level 1, describing today's task-specific Narrow AI, and progresses through specialized AGI (Level 2) and broadly capable AGI (Level 3). It culminates in Level 4, representing Transitional Super AI, and Level 5, which encompasses fully realized ASI capable of transformative societal impacts. https://www.forbes.com/sites/jodiecook/2024/07/16/openais-5-levels-of-super-ai-agi-to-outperform-human-capability/*

[7] While timelines for achieving AGI vary, many experts, including researchers at OpenAI, predict its realization could occur by the 2030s or 2040s. However, uncertainty persists due to technical, ethical, and computational challenges. OpenAI classifies this as Level 3 on its five-level framework for AI progression, describing AGI as capable of outperforming humans in most domains and capabilities while maintaining general adaptability.

[8] Although speculative, some futurists suggest ASI could emerge within decades following the advent of AGI, driven by recursive self-improvement and technological acceleration. OpenAI designates ASI as Level 5 (Super AI) in its framework, describing it as a comprehensive superintelligence with transformative potential for society and existential risks if misaligned with human values. ASI will surpass human capabilities comprehensively, potentially revolutionizing society or presenting significant risks.

[9] See unclassified findings in the NSCAI Final Report, analyses by CISA, and academic research by Brundage et al., 2018.

[10] Kennedy, Scott. *Made in China 2025*. Center for Strategic & International Studies. (June 1, 2015). https://www.csis.org/analysis/made-china-2025

[11] Wubbeke, Jost; Meissner, Mirjam; Zenglein, Max J.; Ives, Jaqueline; and Conrad, Bjorn. *Made in China 2025 The making of a high-tech superpower and consequences for industrial countries. Metrics (August 12, 2016). https://merics.org/en/report/made-china-2025*

[12] McBride, James and Chatzky, Andrew. *Is 'Made in China 2025' a threat to global trade? Council on Foreign Relations (May 13, 2019). https://www.cfr.org/backgrounder/made-china-2025-threat-global-trade*

[13] Murphy, Ben (Editor) and Etcetera Language Group, Inc. (Translator). *China Standards 2035 Center for Security and Emerging Technology (November 8, 2021). https://cset.georgetown.edu/wp-content/uploads/t0406_standardization_outline_EN.pdf*

[14] Koty, Alexander Chipman. *What is the China Standards 2035 plan and how will it impact emerging industries? China Briefing (from Dezan Shira and Associates; July 2, 2020). https://www.china-briefing.com/news/what-is-china-standards-2035-plan-how-will-it-impact-emerging-technologies-what-is-link-made-in-china-2025-goals/*

[15] See Appendix 3: Made in China 2025 (MIC 2025) and China Standards 2035 (CS 2035) Overview for details.

[16] *H.R. 4346 - CHIPS and Science Act (117*th Congress) (2021-2022). https://www.congress.gov/bill/117th-congress/house-bill/4346

[17] See Appendix 4: AI Frameworks, Toolkits and Guidelines for details.

[18] Many of the original webpages are no longer available. In these cases, the closest available reference that could be found is noted.

19 **1. Executive Orders and Government Publications**

- *Executive Order on Advancing United States Leadership in Artificial Intelligence Infrastructure*. U.S. Executive Office (EO 14141, January 17, 2025). https://www.federalregister.gov/documents/2025/01/17/2025-01395/advancing-united-states-leadership-in-artificial-intelligence-infrastructure

- *Executive Order on Strengthening and Promoting Innovation in the Nation's Cybersecurity.* The White House (January 16, 2025). https://bidenwhitehouse.archives.gov/briefing-room/presidential-actions/2025/01/16/executive-order-on-strengthening-and-promoting-innovation-in-the-nations-cybersecurity/

- *FACT SHEET: Ensuring U.S. Security and Economic Strength in the Age of Artificial Intelligence.* The White House (January 13, 2025). https://bidenwhitehouse.archives.gov/briefing-room/statements-releases/2025/01/13/fact-sheet-ensuring-u-s-security-and-economic-strength-in-the-age-of-artificial-intelligence/

- *FACT SHEET: New Executive Order on Strengthening and Promoting Innovation in the Nation's Cybersecurity.* The White House (January 15, 2025). https://bidenwhitehouse.archives.gov/briefing-room/statements-releases/2025/01/15/fact-sheet-new-executive-order-on-strengthening-and-promoting-innovation-in-the-nations-cybersecurity/

- *H.R. 7724 - National Research Act,* Pub. L. No. 93-348. (93rd Congress, 1973-1974). U.S. Government Publishing Office. https://www.govinfo.gov/content/pkg/STATUTE-88/pdf/STATUTE-88-Pg342.pdf https://www.congress.gov/bill/93rd-congress/house-bill/7724

- *US Department of Defense Responsible Artificial Intelligence Strategy and Implementation Strategy.* Department of Defense. (2022) https://media.defense.gov/2022/Jun/22/2003022604/-1/-1/0/Department-of-Defense-Responsible-Artificial-Intelligence-Strategy-and-Implementation-Pathway.PDF

- *GEOINT Artificial Intelligence.* National Geospatial-Intelligence Agency. *https://www.nga.mil/news/GEOINT_Artificial_Intelligence_.html*

- *Principles of AI Ethics for the Intelligence Community.* NSA. https://www.intelligence.gov/images/AI/Principles_of_AI_Ethics_for_the_Intelligence_Community.pdf

- *AI Ethics Framework for the Intelligence Community.* NSA. (June 2020) https://www.intelligence.gov/images/AI/AI_Ethics_Framework_for_the_Intelligence_Community_1.0.pdf

- *FBI Artificial Intelligence.* FBI. (2023) https://www.fbi.gov/investigate/counterintelligence/emerging-and-advanced-technology/artificial-intelligence

- *Artificial Intelligence - Artificial Intelligence (AI) has implications not just for the commercial sector but for national security and law enforcement.* FBI. (2023) https://www.fbi.gov/investigate/counterintelligence/emerging-and-advanced-technology/artificial-intelligence#:~:text=AI%20must%20be%20developed%2C%20acquired,generated%20leads%20with%20human%20experts

- *Audit of the DEA's and FBI's Efforts to Integrate Artificial Intelligence and Other Emerging Technology Within the US Intelligence Community.* Office of the Inspector General (December 2024). https://oig.justice.gov/sites/default/files/reports/25-014.pdf

- Gleeson, Dennis J. Jr. *Artificial Intelligence for Analysis: The Road Ahead.* Studies in Intelligence, Vol 67, No 4, pp 11-15 (extracts, December 2023). https://www.cia.gov/resources/csi/static/88dbcb2b5d4812731b3ff5122e3b6cb5/Article-Artificial-Intelligence-for-Analysis-The-Road-Ahead.pdf

- Brown, Zachery Tyson. *"The Incalculable Element": The Promise and Peril of Artificial intelligence.* National Geospatial-Intelligence Agency (March 2024). *https://www.cia.gov/resources/csi/static/643e18ba5bf779749a14059019db53b2/Article-The-Promise-and-Peril-of-Artificial-Intelligence-Studies-68-1-March-2024.pdf*

- *Government Transformation – Cyber Threat Landscape – CIA's AI Strategy.* CIA (August 25, 2024). *https://www.cia.gov/resources/csi/static/643e18ba5bf779749a14059019db53b2/Article-The-Promise-and-Peril-of-Artificial-Intelligence-Studies-68-1-March-2024.pdf*

- Rosner, Stephanie, Hodosi, Martin and Lim, Rosanna. *Responsible Use of AI in Healthcare Work In Progress.* DIA and Kearney. *https://globalforum.diaglobal.org/issue/october-2024/responsible-use-of-ai-in-healthcare-work-in-progress/*

- *Guiding Principles for the Ethical Use of Artificial Intelligence By Communication Strategy and Operations. US Marine Corps (December 17, 2024)*

https://www.marines.mil/News/Messages/Messages-Display/Article/4001021/guiding-principles-for-the-ethical-use-of-artificial-intelligence-by-communicat/

- *United States Marine Corps Artificial Intelligence Strategy.* US Marine Corps. *https://www.marines.mil/Portals/1/Publications/USMC%20AI%20STRATEGY%20(SECURED).pdf*

- *AI Risk Management Framework.* NIST. (July, 2024); https://www.nist.gov/itl/ai-risk-management-framework

- *CDAO Responsible AI.* CDAO. https://www.ai.mil/Initiatives/Responsible-AI/

- *CDAO Releases Responsible AI (RAI) Toolkit for Ensuring Alignment With RAI Best Practices* US DoD (November 14, 2023). https://www.defense.gov/News/Releases/Release/Article/3588743/cdao-releases-responsible-ai-rai-toolkit-for-ensuring-alignment-with-rai-best-p/

- *Responsible AI. Chief Digital and Artificial Intelligence Office (CDAO)* https://www.ai.mil/Initiatives/Responsible-AI/

- *US Department of Defense Responsible Artificial Intelligence Strategy and Implementation Pathway*. DoD Responsible AI Working Group (June 2022). *https://media.defense.gov/2022/Jun/22/2003022604/-1/-1/0/Department-of-Defense-Responsible-Artificial-Intelligence-Strategy-and-Implementation-Pathway.PDF*

- *Department of Energy Generative Artificial Intelligence Reference Guide.* DOE (2024) https://www.energy.gov/sites/default/files/2024-06/Generative%20AI%20Reference%20Guide%20v2%206-14-24.pdf

- *Artificial Intelligence Guidelines. US Department of Energy. https://www.energy.gov/eere/communicationstandards/artificial-intelligence-ai-usage-guidelines*

- *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*. U.S. Department of Health and Human Services (1979). https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html

2. U.S. National Security and AI Policy Reports

- National Security Commission on Artificial Intelligence. *Final Report: Artificial Intelligence in National Security and Defense*. U.S. Government Publishing Office (2021). https://www.dwt.com/-/media/files/blogs/artificial-intelligence-law-advisor/2021/03/nscai-final-report--2021.pdf

- *2023 Data, Analytics, and Artificial Intelligence Adoption Strategy.* Department of Defense, CDAO (2023). https://media.defense.gov/2023/Nov/02/2003333301/-1/-1/1/DAAIS_FACTSHEET.PDF

- *Artificial Intelligence Ethics Framework for the Intelligence Community*. ODNI (2020). https://www.dni.gov/files/ODNI/documents/AI_Ethics_Framework_for_the_Intelligence_Community_10.pdf

- *Principles of Artificial Intelligence Ethics for the Intelligence Community*. ODNI. https://www.dni.gov/index.php/newsroom/reports-publications/reports-publications-2020/3634-principles-of-artificial-intelligence-ethics-for-the-intelligence-community-1692377385

- *Memorandum on Advancing the Unites States' Leadership in Artificial Intelligence; Harnessing Artificial Intelligence to Fulfill National Security Objectives; and Fostering the Safety, Security, and Trustworthiness of Artificial Intelligence*. The White House (October 24, 2024). https://bidenwhitehouse.archives.gov/briefing-room/presidential-actions/2024/10/24/memorandum-on-advancing-the-united-states-leadership-in-artificial-intelligence-harnessing-artificial-intelligence-to-fulfill-national-security-objectives-and-fostering-the-safety-security/

- *FACT SHEET: Biden-Harris Administration Outlines Coordinated Approach to Harness Power of AI for US National Security*. The White House (October 24, 2024). https://bidenwhitehouse.archives.gov/briefing-room/statements-releases/2024/10/24/fact-sheet-biden-harris-administration-outlines-coordinated-approach-to-harness-power-of-ai-for-u-s-national-security/#:~:text=Today%2C%20President%20Biden%20is%20issuing,policy%20in%20the%20near%20future

- *Framework to Advanced AI Governance and Risk Management in National Security. The* White House. https://ai.gov/wp-content/uploads/2024/10/NSM-Framework-to-Advance-AI-Governance-and-Risk-Management-in-National-Security.pdf

- *Principles of Artificial Intelligence Ethics for the Intelligence Community.* ODNI. https://www.dni.gov/files/ODNI/documents/Principles_of_AI_Ethics_for_the_Intelligence_Community.pdf

- *AI Risk Management Framework.* NIST. (July, 2024); https://www.nist.gov/itl/ai-risk-management-framework

- NIST, Trustworthy and Responsible AI Resource Center (AIRC). *AI Risks and Trustworthiness. NIST, US Department of Commerce.* https://airc.nist.gov/airmf-resources/airmf/3-sec-characteristics/#:~:text=Trustworthy%20AI%20depends%20upon%20accountability,that%20they%20are%20doing%20so

- NIST, Trustworthy and Responsible AI Resource Center (AIRC). *NIST, Trustworthy and Responsible AI Resource Center (AIRC) Webpage. NIST, US Department of Commerce* https://airc.nist.gov

- *NIST AI Symposium: Unleashing AI Innovation, Enabling Trust*. NIST (October 23, 2024). https://www.nist.gov/news-events/events/2024/09/unleashing-ai-innovation-enabling-trust

- NIST, Trustworthy and Responsible AI Resource Center (AIRC). *NIST Risk Management Framework.* https://airc.nist.gov/airmf-resources/airmf/

- NIST, Trustworthy and Responsible AI Resource Center (AIRC). *NIST AI Risk and Trustworthiness* https://airc.nist.gov/airmf-resources/airmf/3-sec-characteristics/

[20] Model cards are documentation for AI models.

[21] Model cards and datasheets are standard AI documentation artifacts

[22] See Appendix 5: FLEX Framework Details for a detailed explanation of the FLEX framework and Appendix 6: Applied Framework Use Case Examples for a detailed use case.

[23] A/B testing is a deployment technique to compare two variants.

[24] See Appendix 7: Interview Details for a quantitative outline of the interview population.

[25] Bommasani, Rishi; Klyman, Kevin; Kapoor, Sayash; Longpre, Shayne; Xiong, Betty; Maslej, Nestor; and Liang, Percy. *The Foundation Model Transparency Index. Center for Research on Foundation Models (May 2024). website: https://crfm.stanford.edu/fmti/May-2024/index.html Paper: https://crfm.stanford.edu/fmti/paper.pdf*

[26] See Appendix 8: Metrics for Accountability and Transparency