AI Oversight in US Government: From Formal Policies to Functional Accountability

Kwame Amponsah[1], Kelvin Gyimah Agyei[2], Frank Boakye[3], Usman Nasiru[4], Manso Frimpong[5]

[1]College of Business, Westcliff University, USA

[2]Management Information Systems, Fogelman College, University of Memphis

[3]Department of Management Information Systems, University of Memphis.

[4]Department of Statistics, Ohio State University, Columbus, Ohio, USA

[5]Fogleman College of Business and Economics, University of Memphis, Memphis TN, USA

DOI: https://doi.org/10.51244/IJRSI.2026.1304000011

Received: 24 March 2026; Accepted: 30 March 2026; Published: 23 April 2026

ABSTRACT

When artificial intelligence (AI) gained attention about ten years ago, its application in federal agencies has increased dramatically. Today, it is a vital tool for efficiency, security, and creativity across federal agencies, having begun as an experimental technology with few government applications. However, a big disconnect exists between stated policies and real supervision procedures due to the US government agencies' expanding usage of AI systems. By examining federal regulations, auditing instruments, and agency-specific procedures, this study seeks to determine how successful the US government's present AI supervision procedures are. The findings of this paper show that, although some agencies have been able to put the overview of artificial intelligence in place, other small agencies that do not have enough resources to fund the project or lack trained personnel to carry out the oversight lag behind. Drawing from case studies, policy documents, and academic research, this study spotlights challenges in guaranteeing accountability, weaknesses in enforcement, and limitations of the existing auditing processes. The research aims to support the creation and development of functional oversight mechanisms that fully preserve and safeguard the interests of the public while maintaining democratic ideals. This has been accomplished by developing tactics and best practices to increase accountability. By informing stakeholders and policymakers about the necessary adjustments to bridge the gap between AI governance practice and policy, the findings of this research will ultimately boost public trust in government AI applications.

Keywords: artificial intelligence, government oversight, algorithmic accountability, federal policy, administrative law, bias detection

INTRODUCTION

The US government promotes research and development and fosters an open regulatory environment to safeguard its position as the world's leading AI technology and guarantee a sizeable portion of the anticipated global AI investment. Healthcare, public transport, and public safety have greatly utilized artificial intelligence in the United States. Incorporating artificial intelligence systems in the United States government has been one of the most transformative developments because it has brought more efficiency to various federal agencies regarding decision-making and operational effectiveness. However, despite all the positive impact it has had, it has had its shortcomings, especially when it comes to ethical considerations and accountability. Executive Order 14110 on Safe, Secure, and Trustworthy Artificial Intelligence, which established detailed rules for federal AI deployment and monitoring, results from substantial policy discussion spurred by this technological revolution (Biden, 2023; The White House, 2025). To guarantee AI systems' efficacy and equity, ongoing assessment is necessary. Creating increasingly complex auditing instruments and encouraging multidisciplinary cooperation between ethicists, technologists, and legislators are two examples of this. Nonetheless, there is still inconsistency and incompleteness in the federal bureaucracy's implementation of these ambitious policy frameworks into effective supervision systems.

Since official policies need, or rather require, accountability measures, strong human oversight, and transparency, one finds that there are major gaps in how the artificial intelligence systems are governed due to how different agencies implement these requirements (Coglianese & Lehr, 2019). Resultantly, those differences raise many questions on the effectiveness of the control mechanisms in place and their ability to guarantee the proper use of AI in government operations. O'Neil (2016) and Angwin et. al. (2022) state that these systems have a far higher risk of bias, harm, and exclusion when run without proper monitoring, threatening individual liberties and public confidence in governmental institutions. Fundamental facets of individuals' lives are impacted by government AI systems, including choices about criminal justice, healthcare access, immigration, and social service eligibility, and therefore, it is impossible to overestimate the importance of bridging these disparities.

This article aims to answer three major questions comprehensively. First, to what extent do existing federal AI policies translate into measurable oversight mechanisms across key government agencies? Second, which auditing and monitoring tools prove most effective in detecting bias and ensuring accountability in government-deployed AI systems? Third, what legal and procedural mechanisms exist for holding government agencies accountable when algorithmic decisions cause harm?

METHODOLOGY

This study employs a qualitative research design grounded in case study review and policy analysis. This includes case studies from 2019 to 2024 and government reports from 2016 to 2024. This study uses database documents from government portals (e.g., agency websites, federal register), for example, the Government Accountability Office (GAO), and academic databases (including IEEE Xplore, JSTOR, and Google Scholar). The Department of Homeland Security, the Department of Defense, the Internal Revenue Service, and the Social Security Administration were prominent government agencies whose policy documents were examined.

The paper applied thematic analysis to single out repetitive oversight gaps and patterns across data to comprehend practical issues. Policies were analyzed for translation into measurable oversight approaches, auditing procedures were assessed for adoption and effectiveness, and accountability mechanisms were screened for procedural and legal sufficiency.

The reliance on publicly available data limits this study, given that publicly available data can omit classified AI utilization in the context of national security. In addition, the case studies utilized may mirror high-profile examples and may not expose the full range of AI use across the government agencies. Besides these limitations, the methodology offers a robust foundation for evaluating the gap between AI policy and practice in the US government.

LITERATURE REVIEW AND THEORETICAL FRAMEWORK

Overview of AI in Government

Machine learning algorithms have been utilized to analyze data and identify patterns; for instance, they can analyze health data to spot patterns and enhance reaction tactics in medical emergencies (Challen et al., 2019). In the healthcare sector, artificial intelligence supports areas such as predicting a disease outbreak, diagnosing patients, and allocating funds in various healthcare sectors. However, applying artificial intelligence to generate insights raises ethical concerns regarding data privacy and algorithmic bias and requires much oversight (Angwin et. al., 2022; Burrell, 2016). In administrative processes, artificial intelligence has been greatly incorporated in tasks such as data entry and documentation procedures. Deloitte (2020) implies that artificial intelligence can cut costs that are used for operations while enhancing effectiveness and efficiency (Deloitte, 2020). In the long run, Kroll et al. (2017) argue that the decision-making process may become more difficult to sustain regarding transparency and responsibility if robotic techniques are used.

The public safety and surety docket has also massively integrated AI for prospective law enforcement, which employs statistical data analysis to foresee criminal activities. This also has its shortcomings, as Garvie (2016) explains that as facial recognition and behavioral analysis-based surveillance technologies have become more

popular, discussions about privacy rights and the possibility of discriminatory practices have arisen. Lastly, in transportation, AI can improve the pace of traffic and lessen gridlock, according to research, which will help make transportation networks more effective, thus increasing urban mobility. Self-driving vehicles and AI-powered traffic control systems are being considered. On the contrary, accountability in case of accidents brought about by these technologies is still in question (Government Accountability Office, 2024).

THEORETICAL FRAMEWORK

Algorithmic Governance Theory: This theory mainly focuses on the challenges accompanying algorithmic systems in government settings. Algorithmic governance refers to organizing and managing public functions through automated systems, which are intricate collections of digital instructions. It mainly focuses on fairness, explainability, auditability, and responsibility (Kroll et al., 2017). Traditionally, governance depended on human discretion, laws, and policies to guarantee legitimacy and accountability (Burrell, 2016). However, with the rapid adoption of AI, these procedures are now performed by automated systems that work with limited transparency. In the United States, rules regarding the use of AI in the government require human oversight of the artificial intelligence systems that make decisions about citizens. However, research shows a huge gap between what happens and what is required. Human oversight is non-functional, which may worsen things (Coglianese & Lehr, 2019). According to the Department of Homeland Security policy, a person must review decisions made by AI before they become final. Before making decisions, personnel responsible for the AI systems must fully understand why an AI system has recommended and why, be fully responsible for the final decisions, as well as regularly check how well the system is working (Department of Homeland Security, 2023). Some of the gaps between mandated human oversight and actual practice in algorithmic governance include:

Inadequately skilled personnel: The administrative rule states that humans with technical knowledge should oversee AI systems. In reality, most administrative employees are unaware of how artificial intelligence operates (Salesforce, 2024). Therefore, they are uncertain of what inquiries to make regarding AI suggestions, lack knowledge about the constraints of AI systems, have more confidence that AI is "objective" and devoid of prejudice, and are unable to identify biased or incorrect conclusions made by AI (Angwin et. al., 2022).

Excessive Dependence on AI: The rules offer balanced human-AI collaboration, but people grow overly reliant on AI suggestions. You find that as time goes by, human judgmental abilities deteriorate. Also, due to over-independence, human choices have a major impact on AI's initial suggestion. This makes the set policies inefficient in achieving transparent accountability and transparency.

Fake Review Process: Rules state that there should be a relevant human evaluation of AI suggestions. However, people "rubber-stamp" their agreement with AI without actually considering the decision (O'Neil, 2016). People assigned oversight duties approve AI recommendations without oversight mainly because they assume that AI is always correct. In contrast, others may lack enough time to look closely at the reviews due to heavy workloads. Others do not even know how to filter and evaluate decisions made by AI. This deteriorates the efficiency of the set policies.

Workplace Pressure: Instead of integrating a careful oversight of AI, corporations opt to forego a thorough review in favor of speed and cost savings. The pressure to cut costs by reducing human involvement and the fear of criticism for superseding "objective" artificial intelligence systems are just some of the reasons for the recklessness. Employees are judged on how quickly they handle issues, not how well they make decisions. This pressure due to time constraints also precludes thorough analysis, deteriorating the results of the oversight.

Furthermore, procedural justice theory and administrative law are common frameworks emphasizing transparency and equitable procedures. For instance, the primary goal of Democratic Theory and Participatory Governance is to examine how AI affects public participation and political credibility thoroughly. On the other hand, the goal of network governance theory is to supervise AI administration for multiple stakeholders. The theory of precautionary governance and the risk society focuses on managing emerging hazards related to artificial intelligence through anticipatory governance. Sociotechnical systems theory aims to view artificial intelligence as a complex technology-social connection.

Accountability and Oversight

According to Corbett-Davies & Goel (2018), accountability comprises transparency during decision-making processes and holding corporates accountable in case of negative algorithmic results. Therefore, accountability in governance can be defined in simpler terms as the need of government officials to take accountability and be able to answer questions regarding their actions. For algorithmic governance to be fair, identify and reduce biases, and preserve public confidence, oversight procedures are crucial; therefore, in order to guarantee that laws, rules, and moral principles are followed, oversight refers to the procedures and frameworks that keep an eye on and assess the activities of governmental bodies.

In the United States, accountability and oversight policies have developed significantly since 2016, when government agencies started understanding the benefits and importance, as well as the risks, of artificial intelligence. Leaders such as Barack Obama started focusing on and funding research boards such as the National AI Research and Development Strategic Plan in 2016 to coordinate AI across government agencies (AIRDIWG, 2019).

This progressed when the Biden administration took charge, since he signed Executive Order 14110 (The White House, 2025). This order focused on more detailed regulations that govern and shape AI safety and transparency across all federal agencies. For example, Executive Order 14110 held that agencies must assess AI systems thoroughly, test them for bias, and ensure humans continue to supervise high-risk AI applications per this mandate (Biden, 2023; The White House, 2025). This took a stronger approach in regulating the use of AI in the government (Corbett-Davies & Goel, 2018). This gradual evolution is clear evidence that the government's thinking on AI has changed since the Obama regime encouraged this technology. In contrast, the Biden regime focuses on controlling and closely managing AI and the risks that come along.

Historical Evolution

Early Federal AI Initiatives (2016-2020).

Artificial intelligence has made its debut and is revolutionizing a wide range of government operations, including in sectors such as production, transportation, and hospitals. Due to its robust innovation environment, which allows the government to leverage AI tools, the US has positioned itself as a world leader in AI applications in its major sectors (Trump White House Archives, 2019). Although it started during the tenure of Barack Obama, the United States government's official involvement with artificial intelligence (AI) legislation took off under President Trump.

Some key policies during this time included the National AI Research and Development Strategic Plan (2016). This policy focused on ensuring clearly defined scientific agendas along with techniques for overseeing (AIRDIWG, 2019). This policy was then topped by Executive Order 13859 (2019). The Executive Order 13859 focused on research, standardization, removing regulatory obstacles, and preserving American dominance in Artificial Intelligence (The White House, 2025). Afterwards, the AI in Government Act (2020) was introduced. This act mainly aimed to ensure that AI catalogues and fundamental organizational structures were established by appropriate institutions (AIRDIWG, 2019). In addition, President Biden’s government employed a more regulatory and accountable framework to keep America aggressive in leveraging AI in its administration, the Biden Administration's Comprehensive Approach (2021-2025). This framework prioritized privacy, safeguarding, and fiscally prudent growth in terms of AI administration (Biden, 2023).

According to the Federal Register of 2023, Artificial intelligence (AI) holds extraordinary potential for promise and peril. Drawing from this, in October 2023, Executive Order 14110 was put in place (The White House, 2025). The core purpose of this order was to ensure responsible applications of AI. This framework showed potential in solving urgent government challenges while making the administration more prosperous, productive, innovative, and secure. Prior to the execution of this order, irresponsible use was exacerbating societal harms such as fraud, discrimination, bias, and disinformation (Angwin et. al., 2022). This framework did away with displacing and disempowering workers, stifling competition between departments, and risking national security. It was founded under the idea that reducing the significant hazards associated with AI is necessary to use it for

good and reap its many advantages. Government, business, academia, and civil society must all work together to accomplish this goal. This decree was the most extensive government AI policy, from civil rights protections to AI safety research.

Another Federal policy put in place to regulate AI adoption is the Return of Deregulation (2025-Present). According to a publication by The White House in July 2025, by favoring American investments in AI through liberalization and less federal supervision, the current Trump administration has adopted this policy as a drastically different strategy (The White House, 2025). This policy ensures that only AI that seeks justice, honesty, and rigorous neutrality will be dealt with by the US government going forward.

Objectives and Scope of Major Federal AI Policies

Executive Order 13960: Promoting Trustworthy AI (December 2020)

The main objective of this order was to continue to have humans monitor AI systems and specify guidelines for the executive administration's implementation of reliable AI (The White House, 2025). This policy is also intended to safeguard fundamental privileges and confidentiality while fostering creativity and ensuring that AI remains secure, credible, and impartial. This policy has been applied to every government agency that makes decisions that impact the general population, utilizing AI technologies. It requires all agencies to have and maintain AI inventories. It also calls for human oversight of decisions made by AI systems, the exercise of risk evaluation protocols into practice, and guidelines for evaluating and assessing AI.

Executive Order 14110: Safe, Secure, and Trustworthy AI (October 2023)

The 2023 Executive Order 14110 required that systems that use artificial intelligence be freed from prejudice by authorities (The White House, 2025). This safeguarded employees whose employment may be impacted by AI. The order required agencies to eliminate bias from AI systems, protect employees affected by automation, and ensure enterprises developing the systems report directly to the federal government (*Federal Register*, 2023). Its main goals were to boost the use of AI across government agencies, stimulate ethical engagement and creativity, and handle the threats to national defense. Other goals included safeguarding confidentiality and its users, encouraging global collaboration, helping staff impacted by AI, and encouraging human rights and egalitarianism.

OMB Memorandum M-24-10: AI Governance Framework (March 2024)

In March 2024, the Office of Management and Budget produced a memorandum M-24-10 with the inaugural legally obligatory standards requiring institutions to improve oversight, creativity, and risk mitigation for use (White House OMB, 2024). The memo required that there should be considerations for the safety of artificial intelligence and data devices. It also outlines conditions of contracts that guarantee oversight of AI by the government and calls for transparency of the possibilities and restrictions of AI systems by AI companies, as well as criteria for evaluating and assessing acquired AI. IT infrastructure, data, cybersecurity, and generative artificial intelligence are the four main obstacles listed in the memo (US Government Accountability Office, 2025). The compliance plans outline the agencies' strategies for dealing with these problems and the state of key AI-enabling technology inside the organization (Federal AI Compliance Plans for OMB Memorandum M-24-10, 2024). The memo focused on safeguarding public and federal information while purchasing AI. This ensured that federal agencies purchased AI systems responsibly, encouraged equitable competition for governmental tenders involving AI, and defined guidelines for government-partnering AI contractors.

Executive Order: Removing Barriers to American AI Leadership (January 2025)

Although rescinded on January 20, 2025, Executive Order 14110 identified any actions taken in compliance with the directive that clash with or might be inconsistent with the guidelines described in this regulation (White House, 2025). The 2025 Trump Administration proposed to undo several AI policies from the Biden administration. Repealing this order revoked some existing AI regulations and directives impeding American AI innovation (The White House, 2025). By rescinding this order, the Donald Trump administration aimed at

minimizing the need for AI businesses to submit reports, examine, and repeal the AI laws from the Biden administration, put national supremacy ahead of security regulations, and streamline government AI procurement. According to this administration, due to the strength of our free markets, top-notch research institutions, and spirit of entrepreneurship, the United States has long been at the forefront of artificial intelligence (AI) innovation (The White House, 2025). Moreover, the government must create AI systems free from social agendas and ideological bias to keep the administration functional. The president mentioned that doing this guarantees a better future for all Americans and maintains America's position as the world leader in AI with the correct government policies.

Implementation Effectiveness

Federal entities have successfully met the fundamental operational requirements. For instance, by March 2024, all federal agencies met the 13 AI leadership and staffing requirements. To acknowledge this, the government laid a foundation for government-wide AI initiatives (GAO, 2024). Some of the requirements met are the appointment of Chief AI Officers as required by agencies, and staff training, where many corporations are training their staff members on the oversight of AI. Additionally, organizations have developed and maintained detailed registries of AI, and the government implements basic AI frameworks of governance.

Still on the same note, Huergo (2023) notes that there has been significant progress by the National Institute of Standards and Technology (NIST) in terms of developing standards and guidelines as per the government's policies. Some of these developments include bias detection tools and effective methods for assessing the effectiveness of AI systems. Other developments are the prerequisites for safeguarding AI systems against online attacks and thorough recommendations for controlling AI threats (NITS, 2023).

Moreover, following the terms of President Biden's presidential directive on machine intelligence, NITS is seeking data to help it carry out several objectives, such as developing assessment capacities and developing red-teaming test instructions. Additionally, the reported AI use cases nearly doubled among the 11 agencies GAO assessed with its computational intelligence inventory, from 571 in 2023 to 1,110 in 2024 (The White House, 2025). This means that the governmental use of AI is growing rapidly. Some of the growth sectors include solutions for cybersecurity and risk monitoring; interaction with customers, including avatars and autonomous response mechanisms; managerial responsibilities, including time management, handling paperwork, and regular decision-making; and pattern identification and predictive analytics in data analysis (Brantingham et al., 2020; Kroll et al., 2017; Rahwan et al., 2019).

Implementation in DoD, DHS, HHS, and smaller agencies

Implementation Strengths and Challenges in the Department of Defense (DoD)

DOD released a letter in July 2024 to guarantee the department's dynamic artificial intelligence research, implementation, and purchase in an accountable and secure manner**.** This agency has made substantial investments in the study and advancement of AI, procedures, and robust assessment programs, with explicit guidelines for implementing AI. Most of these investments revolve around the adoption of AI in armed forces contexts and a strong AI governance structure with dedicated leadership (GAO, 2024). Nevertheless, it has had its shortcomings. To begin with, the training high-performing AI typically requires precisely labeled material (pictures, text files, videos, etc. that have been tagged with one or more IDs) that the AI system can learn from is one difficulty that GAO noted in a recent evaluation of DoD's AI initiatives (Hoadley & Lucas, 2018). Currently, most of DoD's data is unlabeled. Simultaneously, incorporating AI into current weaponry platforms is difficult. According to the 2023 DoD report, integration necessitates physical space for potentially unavailable computing equipment since AI capabilities integrated into weapon platforms must be able to operate in places without access to digital infrastructure, such as the cloud.

Implementation Strengths and Challenges in the Department of Homeland Security (DHS)

The Department of Homeland Security (DHS) regularly handles large volumes of data. Its mandate covers many areas and is to "keep America safe " (US Homeland Security, 2022). In October 2023, DHS released a regulation

outlining the conditions for authorized commercial generative AI use by DHS employees (GAO, 2025). However, AI for Homeland Security has several disadvantages, and the public is most concerned about two issues with the widespread deployment of AI. They worry that due to the intense competition for government funding and the haste with which AI and ML solutions are implemented to solve pressing issues (US Homeland Security, 2022). Some members of the public worry that the accuracy and security of these systems may not be as high as they should be. There are also concerns regarding the privacy of data as well as vulnerability to malevolent manipulation and hacking.

Resolving these public worries regarding the application of AI in immigration and law enforcement and working on AI standards in coordination with regional and state partners are some of the other challenges the agency faces (Department of Homeland Security, 2023). On the flipside, there has been substantial expenditure on AI supervision skills, robust administration of artificial intelligence guidelines, engaging actively in collaboration among agencies, and a strong emphasis on protecting human rights and detecting prejudice (Department of Homeland Security, 2023).

Implementation Strengths and Challenges in the Department of Health and Human Services (HHS)

The government has put much effort into protecting patients ' private data. HHS has also shown efforts to ensure that AI standards are positively integrated into the medical community, and research about AI is very active and robust within this agency. This agency has also been working to ensure a strong AI governance framework for healthcare applications (Huergo, 2023). However, there has been trouble ensuring that artificial intelligence has met quality standards required in the healthcare sector, and incorporating artificial intelligence in several HHS sub-agencies. Additionally, the agency faces concerns about prejudice and unfairness in the health care docket and difficulties in managing environments for medical artificial intelligence that are extra complex (Corbett-Davies & Goel, 2018)

Common Implementation Challenges in Small and Medium Agencies

Small and medium agencies lack well-trained people to oversee the final decisions that artificial intelligence systems make. These agencies also lack enough funds to advance AI. GAO (2024) mentions that small and medium agencies also tend to depend a lot on other agencies for guidance and AI support, and they have difficulty justifying AI investments given limited use cases. However, they have received some support mechanisms, which include technical and training assistance from more advanced agencies and shared services for AI governance (GAO, 2024).

Implementing federal AI policies in these agencies shows a mixed record of successes and challenges (Text - H.R.7532 - 118th Congress, 2023-2024). While agencies have made significant progress in establishing basic AI governance structures and rapidly adopting AI technologies, substantial gaps remain in ensuring effective oversight, managing risks, and maintaining consistent implementation across government.

Current State of Federal AI Governance

- ✓ **Office of Management and Budget (OMB, 2024)**: Issues government-wide AI policy guidance
- ✓ **National Institute of Standards and Technology (NIST)**: Develops AI standards and guidelines
- ✓ **Office of Science and Technology Policy (OSTP, 2023)**: Coordinates AI policy across agencies
- ✓ **Agency-Level Implementation**: **Chief AI Officers:** Senior officials responsible for AI governance at each agency.
- ✓ **AI Governance Boards**: Multi-stakeholder groups that oversee AI use and policy.
- ✓ **Technical Teams**: Staff responsible for implementing and monitoring AI systems

Challenges and Implementation Gaps

One of the major challenges is the back-and-forth changes in the policies that govern artificial intelligence, as each administration has its own view on using the system. The regulations that will continue to be in force are unknown to departments, as Trump revoked Biden's AI executive order. For example, during Biden's regime, the

2023 executive order required agencies to appoint chief AI officials and establish accountability strategies for using AI. However, Trump's era revoked those policies, leaving much uncertainty about what will replace them (The White House, 2025). This makes it difficult for agencies to allocate resources because it may be difficult to account for long-term costs associated with AI monitoring. Additionally, private companies are not fully aware of the government's AI regulations, and the future of AI governance projects is unclear. This automatically means there will be inconsistencies in applying the set guidelines.

The inconsistent application of AI policies by certain agencies across federal agencies further shows the gap between practical oversight and formal mandates. Although some businesses have set resources and frameworks to manage AI governance, others are substantially behind. According to the Government Accountability Office (2024), twenty out of twenty-three agencies indicated 1,200 AI use cases, mirroring a wide dedication to exploring algorithmic solutions. However, three agencies stated they had no intentions to implement AI (GAO, 2024). This uneven adoption shows disparities in institutional capacities. Larger organizations like the DoD, DoH, and HHS have actively and strongly invested in AI oversight mechanisms and development.

On the other hand, smaller agencies mostly have insufficient qualified personnel and resources to build solid AI governance systems. This difference solidifies the central oversight difficulty: while government policies apply widely, their adoption heavily relies on the priorities and capacity of individual agencies. This leads to inconsistency in accountability.

Another challenge some agencies face while implementing AI policies is the technical difficulties. Some of these systems are hard to understand and evaluate. According to Kroll et al. (2017), some agencies lack people with enough skills to crack through this challenge. Also, there are issues with the data required for AI assessment and education. In addition, Kroll et al. (2017) mention that some entities lack skills in protecting AI systems from cyber threats and data breaches and face difficulties in integrating AI systems with current government technologies. Still on the same note, inadequate funding prevents plenty of organizations from putting extensive oversight of artificial intelligence into practice (GAO, 2024). There are not enough funds to provide thorough staff training on AI supervision. This results in organizations struggling with workforce capacity, implementation, and compliance, hence inconsistent application of AI policies. This imbalance undermines nationwide standardization of AI governance, resulting in varying levels of risk management and fragmented policies.

Auditing and Monitoring Tools for US Government AI Systems

With uses in commerce, transportation, healthcare, agriculture, defense, and many other federal institutions, artificial intelligence is a game-changing technology and has much potential to enhance government functions. Therefore, ensuring AI is accountable, fair, easily identifiable, trustworthy, and governable has been the main goal of federal advice. Third-party audits and assessments are necessary to achieve these goals. However, AI systems pose unique challenges for this type of oversight. This is because their inputs and capabilities do not align with some of these goals. To promote the US government's responsible adoption of the new technology, the White House has mandated that federal agencies evaluate artificial intelligence systems for hazards and assign officials to provide oversight. In order to help assure transparency as well as accountable use of AI by government organizations and other entities participating in the planning, creation, implementation, and periodic evaluation of artificial intelligence (AI) platforms, GAO set out to identify essential approaches. This thorough examination uses federal case studies and implementation examples to examine the best US-based methodologies, instruments, and frameworks for government AI auditing.

During the Biden regime, significant requirements for governmental governance of AI were established, requiring monitoring and auditing systems to be implemented. This marked a turning point in government AI accountability. Chief AI officers and bias testing were needed for US government agencies (Federal Register, 2023). However, a significant change in AI policy was signaled by removing regulations impeding innovation in the paper Removing Barriers to American Leadership in Artificial Intelligence. The future of federal AI auditing regulations is unclear due to this shift, even though bias detection and accountability procedures are still

essential. Currently, there is a platform aiming to retain the necessary control and monitoring skills while offering mission-ready AI advances.

IBM AI Fairness 360 (AIF360)

In order to bring computational study from the lab into the real world of fields as diverse as finance, the administration of human resources, medical care, and education, LF AI incubation project AI Fairness 360 is an extensible open-source toolkit that can assist users in examining, reporting, and mitigating discrimination and bias in machine learning models throughout the AI application lifecycle. Both R and Python versions of the toolbox are available. In July 2020, IBM transferred AI Fairness 360 to LF AI. *Home - AI Fairness 360* (2020) mentions that this data pack has methods for reducing bias in models and datasets. In fields as diverse as finance, human resource management, healthcare, and education, algorithmic research is intended to be converted from theory into real-world applications (AI Fairness 360 (AIF360), 2022). It also aims to ensure a thorough collection of measures for algorithms and statistics to check for prejudices. This will be alongside justifications for the measurements.

Some of the reasons why it is the most effective tool for the government to work with are its comprehensive coverage, as it has over 70 fairness metrics. This tool is specifically designed for government use cases, and they have a proven track record. It has also been successfully applied in several government fields and is compliant with regulations put forth by the government, such as the automatic adherence to NIST AI RMF specifications. Additionally, this tool has complete transparency of auditing procedures and no vendor lock-in. Compared to manual auditing procedures, federal agencies that use AIF360 report 40–60% improvements in bias detection accuracy and notable decreases in incorrect positive percentages for prejudice recognition.

Aequitas Bias Audit Toolkit

Aequitas, which was created by the Center for Data Science and Public Policy at the University of Chicago, has emerged as the industry standard for federal prejudice monitoring. Machine learning developers, analysts, and policymakers can use the freely available Aequitas prejudice assessment toolbox to check models built with machine learning for discrimination and prejudice. This tool can be used to make fair and sound choices about creating and implementing prognostic risk-assessment technologies. According to Corbett-Davies & Goel (2018), this tool was designed for scenarios involving societal consequences and the government, and it has strong support from academic and policy research. This tool is also in accordance with legal safeguards and human rights obligations, as well as detailed documentation and audit trail capabilities, making it an easy-to-use tool.

Microsoft Fair-Learn with Azure Government

When used with Azure Government cloud services, Microsoft's Fairlearn toolset offers an integrated system for identifying prejudices and management tailored to federal agencies' needs. Comparing Fairlearn to human auditing procedures, government departments indicate 45–65% faster bias detection and 30–40% more accurate algorithmic disparity identification. By implementing a Code of Conduct that users should consider while using Azure OpenAI and integrating Microsoft's principles for responsible AI use, the company has made considerable expenditures to prevent abuse and unintended harm (Huergo, 2023). This toolkit has government-favoring features such as smooth interaction with the current Microsoft government tools, committed federal funding and educational materials, enterprise-grade efficiency for huge government databases, and complete adherence to legislative privacy laws. It can survey ethnic demography and has metrics of individual fairness (Corbett-Davies & Goel, 2018). This tool also has frequency equality estimation, evaluation, and threshold modification techniques.

Google's What-If Tool for Government Applications

Through the What-If Tool, practitioners can view model behavior across several government models and subsets of input data, evaluate performance in hypothetical governmental scenarios, and assess the significance of various data aspects (Robinson & Wexler, 2019). Practitioners can also use it to evaluate systems using a variety

of ML fairness metrics. For some federal use cases, Google's What-If Tool has been quite successful, especially in research and development settings, even though it is not government-specific. The toolkit is ideal for presenting to stakeholders alongside reporting accountability. It allows for in-depth analysis of computational judgments and is a powerful teaching tool for government data scientists (Robinson & Wexler, 2019). This tool has been used to look for policy consequences by testing "what-if" scenarios. On the flipside, its limitations to the government include restricted integration with government-specific compliance standards, reliance on Google cloud infrastructures, and security and privacy restrictions in particular administrative contexts. These limitations slow down its use to leverage the full potential of AI applications in government practices.

GSA's Combating Bias in AI Initiative

The General Services Administration (GSA) has been researching to investigate and validate the possibility of racial prejudice in facial recognition technology systems. The agency's investigation includes a review of the equity of some remote identity-proofing technology that the American public may utilize to obtain federal benefits. Sometimes, the federal government's deployment of AI and ML capabilities leads to unfairness and harms, which are addressed by the combating bias in AI project (GAO, 2024; Corbett-Davies & Goel, 2018). The solution created by the government is the most customized way to meet federal AI auditing criteria. Its features include displays for tracking prejudice in the present moment, modernization of regulatory reporting, connectivity to government identification and authorization networks, enhanced capacity to evaluate statistics, and monitoring and exchanging data among agencies.

Case Studies

US Citizenship and Immigration Services (USCIS) Application Processing

The tool used is the Aequitas toolkit with unique federal adaptations. This is because, although USCIS used AI technologies to expedite the processing of immigration applications, it was necessary to guarantee equity for all applicant groups. Some of its functions are tracking approval ratings by demographic and ethnicity. This includes civil rights office oversight protocols, consistent third-party verification of audit findings, and monitoring application algorithms screening for partiality.

Also, adopting Aequitas's government-specific design has made the fast detection of immigration-specific bias patterns possible. This tool’s statistical rigor of its remedial measures helped make them legally defensible. The tool can detect substantial prejudice found in the program's initial implementation and increase openness for the public through consistent bias reporting. This has resulted in algorithmic changes that have resulted in a 60% reduction in differential impact and the establishment of continuous monitoring to end prejudice build-up in the future.

Department of Veterans Affairs (VA) Benefits Processing

To check for possible bias against specific veteran demographics, the VA had to audit its AI-powered disability benefits determination system and, therefore, IBM AIF360 in conjunction with specially created federal surveillance equipment. The tool was supposed to integrate with existing databases and workflows in the agency and test for pre-deployment prejudice among veteran demographics. The tool was also required to regularly update monitoring and leadership entities within VA and ensure continuous tracking of benefit approval rates according to protected attributes. Because of the openness of its source code, AIF360 was able to meet its obligations to transparency while identifying subtle bias patterns that manual assessments had overlooked due to its extensive analytics. The results were perfect since the tool enhanced trust among stakeholders in AI-supported choice-making. Proper steps were taken, reducing the gap to less than 3% and finding a 15% difference between ethnic categories' compensation clearance frequencies. Also, continuous monitoring was established to stop bias from recurring. IBM AIF360 is the most effective for pre-deployment auditing due to its detailed capabilities to analyze and test bias. This tool has a proven track record across several federal agencies and excellent compatibility with government environments for innovation.

Legal and Procedural Mechanisms for Accountability

An increasing body of research indicates that algorithmic systems used in public service delivery can be harmful and often lack openness in their operation, particularly opacity around choices regarding their use. Most nations have not yet committed resources to educating and involving the general public about using algorithms to provide public services. Few systematic, cross-jurisdictional examinations of the application of these policies have been conducted, despite some attempts to assess algorithmic responsibility inside specific organizations or circumstances.

Disparate Impact Legal Framework

Indirect discrimination in employment practices is addressed by the legal notion of "disparate impact theory," rooted in Title VII of the 1964 Civil Rights Act. Griggs v. Duke Power Company. This 1971 Supreme Court decision established that employment practices resulting in disparate results must be based on economic necessity, which helped popularize this approach (Anderson, 2025). This case was built to conclude that even unintentional repercussions that may disproportionately harm members of protected classes based on race, color, religion, sex, and can be legally challenged.

Applied to algorithmic governance, this theory generates a legal foundation for accountability when AI-fueled decisions harm communities or their individuals. For instance, if a federal agency implements an AI hiring model or a real estate organization adopts an automated tenant screening tool that disadvantages minority groups, victims can bring a disparate impact claim. In this case, the organization or agency must illustrate that the algorithm is necessary for legitimate operational purposes (Anderson, 2025). If claimants prove the model disproportionally harms protected communities, courts can require compensation for losses, impose remedial measures, or mandate policy changes.

This is a perfect example of how current civil rights laws serve as a tangible accountability model in the US. As the government agencies go big in using algorithmic tools in fields such as law enforcement, hiring, and benefits allocation, disparate effect litigation is one of the clearest legal avenues to deal with any harm caused, even when the bias is unintentional. Therefore, it narrows the gap between functional accountability in practice and formal AI policy dedications.

Algorithmic Accountability Act

According to this bill, certain business enterprises must evaluate extremely dangerous technologies that handle personal data or make programmed judgments, like those that employ data mining or intelligent technology. In particular, computerized assessment technologies that pose substantial hazards include those that assist in making decisions on delicate areas of customers' lives by analyzing their behavior. Additionally, an automated decision-making system or information system that uses personal data is deemed high-risk if it has security or privacy issues. It is also deemed high-risk if it uses many people's personal data, may contribute to inaccuracy, bias, or discrimination, or conducts routine monitoring of a sizable, open space. High-risk automated decision systems must be thoroughly described, the corresponding costs and advantages evaluated, potential threats to the security and privacy of personal data identified, and the measures taken to reduce those risks, if any, explained in the evaluation process. The degree to which the system safeguards the security and privacy of personal data must be considered in assessments of high-risk information systems.

Section 1983 Civil Rights Lawsuits

When federal constitutional or statutory rights are violated, Section 1983 allows anyone to sue state and municipal officials and workers, including police officers, jailers, prosecutors, mayors, and even cities. People can sue government entities under Section 1983 primarily when algorithmic systems violate their constitutional rights, such as when AI systems discriminate against people based on their race, deny them due process, or violate their equal protection obligations. However, as an attorney's argument against a Section 1983 lawsuit, exemption from liability shields a police officer or other official from responsibility if the alleged violation of a constitutional right was not "clearly established under federal law."

Federal Civil Rights Enforcement: Department of Justice Civil Rights Division

Top legislators and federal agency civil rights office principals met on October 9, 2024, in the Justice Department's Civil Rights Division to promote synergy between AI and civil rights. The Civil Rights Division had its fourth such meeting. Government organizations that participate in patterns of algorithmic discrimination are investigated and prosecuted by the Civil Rights Division. Some of its recent enforcement actions include: State benefit distribution schemes, which are examined for potential adverse implications, monitoring of risk assessment instruments in criminal justice for compliance with equal protection, racial prejudice in the analysis of police predictive law enforcement systems, and Congressional recruitment computations are evaluated for equitable hiring methodologies.

Administrative Procedure Act (APA)

Noting the potential for AI to "dehumanize" parts of the law, leading commentators in law enforcement have warned against a greater dependence on AI tools. Given these concerns, legal quandaries will likely arise when federal agencies attempt to include AI/ML tools in the notice-and-comment rulemaking process. Section 553 of the Administrative Procedure Act (APA), which mandates notice, disclosure, and commenting options for all review-and-comment regulations, is particularly where complaints can be raised. Under APA section 706, a court of appeals may declare a governmental activity unconstitutional and set it aside as arbitrary and capricious if the agency fails to adhere to these procedural criteria. When government algorithmic choices lack sufficient examination of pertinent factors or rational justification, they may be challenged under the APA's arbitrary and capricious standard.

Freedom of Information Act (FOIA) Transparency

FOIA promotes accountability through openness by offering a way to learn more about government algorithmic systems. Congress presented the Federal AI Governance and Transparency Act of 2024 (H.R. 7532) to close the transparency gaps in AI, which established governance criteria for federal agencies' AI systems. The proposed law will target AI system supervision directly by amending current disclosure rules (Olsen et al., 2024). However, the high potential costs of disclosure result from FOIA's requirement that public information be made available to anyone upon request, which encourages businesses to try to avoid it. Furthermore, evaluating the benefits of public disclosure is frequently challenging due to the complexity of algorithmic systems and the circumstances in which they are employed. To prevent public exposure of AI algorithms, government organizations and their commercial sector partners commonly assert trade secret rights (Olsen et al., 2024).

Federal Trade Commission (FTC) Authority

In an effort to combat deceptive and unfair corporate behavior that involves the use of artificial intelligence (AI), the Federal Trade Commission (FTC) declared the official start of "Operation AI Comply" on September 25, 2024. The campaign intends to address misleading or exaggerated representations regarding prospective Services available from AI products. The agency has been highly active in deciding what criteria it deems appropriate for AI enterprises in the past few years through its guidelines and enforcement actions. It has used this power when companies deploy AI tools in potentially discriminatory ways. The FTC uses its client safeguarding authority under Section 5 of the FTC Act to regulate business practices connected to artificial intelligence, focusing on deceptive and fraudulent actions. The FTC has issued recommendations and taken enforcement action to counteract deceptive AI actions, including overstating AI capabilities, gathering data in an unreliable manner, and making biased conclusions. The FTC advises companies to remain transparent, validate AI models, and avoid discriminatory results to achieve its standards. Consent orders, fines, and eliminating algorithms trained on improperly obtained data are possible outcomes of violations.

Comparative Analysis: International Perspectives on AI Oversight and Accountability

United Kingdom: Sector-Specific Governance

Five fundamental principles are the foundation for the UK federal government's cross-sector, outcome-oriented approach for overseeing AI. These include contestability and redress, accountability and governance, safety,

security, robustness, and proper transparency, explainability, and justice. As anticipated, the UK has no intention of establishing a new AI body to supervise the framework's application. Using current rules and restrictions, administrators are supposed to employ a balanced, context-dependent approach. As they oversee and regulate AI in their respective fields, current regulators, including the FCA, Ofcom, and the Information Commissioner's Office (ICO), have been requested to implement the five principles. This strategy is based on voluntary safety and transparency standards for AI system creators.

Canada: Federal Leadership with Provincial Coordination

Attard-Frost et al. (2024) state that to address the various possible effects of AI systems, Canada's federal and provincial governments have created and carried out several programs. For instance, risk is approached horizontally and graduated in the EU and Canada, with impact levels (Canada) and risk thresholds (EU) established (Attard-Frost et al., 2024). Canada shares the UK's emphasis on "impact," which stresses the real effects of AI technology rather than potential threats, and the USA's Blueprint for an AI Bill of Rights, which focuses on the potential effects of AI systems on democratic principles and human rights. Meanwhile, the USA and UK perceive AI threats as domain-specific and lack a comprehensive, legally binding risk framework. They emphasize the sectoral impact of these technologies more than the EU and Canada, which take a horizontal approach.

Singapore: Smart Nation Innovation

In the international conversation on AI ethics and governance, Singapore thinks its well-rounded strategy can promote innovation, protect consumer interests, and act as a universal benchmark (Singapore’s Model AI Governance Framework - Securiti, 2024). Singapore has established a practical strategy for striking a balance between innovation and governance through its model, AI Governance Framework, and other sector-specific initiatives. This model was created to handle generative AI challenges methodically and equitably while promoting innovation (GAO, 2025). It calls for the combined efforts of all major stakeholders, including the public, industry, government, and the research community.

Japan: Society 5.0 Integration

In contrast to the Fourth Industrial Revolution, Japan's new design for a super-smart society, Society 5.0, aims to radically revolutionize Japanese culture by erasing the boundaries between the internet and physical space (Deguchi et al., 2020). This model has been described as a human-centered society in which community problem settlement and revenue generation are harmonious through an intricately connected network of cyberspace and physical space (Cabinet Office, 2019).

Recommendations for Improvement

First, the US government should strengthen its legal and regulatory framework by enacting detailed legislation for Congress's AI governance. The US Congress should enact federal AI governance regulations that fill in the deficiencies in the present regulatory framework for computational decision-making processes. The Congress should also enact governance policies that generate obligatory minimum criteria for AI oversight across all organizations, fix non-compliance with the law, and provide clear enforcement mechanisms. Additionally, it should increase the FOIA requirements for AI systems to implement mandatory transparency and accountability measures by addressing AI transparency limitations through FOIA overhauls. Simultaneously, it should restrict the exclusion of confidential information for federal AI uses and establish clear transparency guidelines for government AI systems that impact citizens. The government arms should also demand prompt transparency of AI systems' fundamental operations and supervision protocols, and provide consistent public information on biased tests and AI system effectiveness.

Additionally, they should consider the creation of a publicly accessible AI Software Archive by compiling an extensive library of government-funded AI systems. This may include competency indicators and findings from prejudice assessment, technology operation, managerial authority, grades of individual engagement, oversight methods, and public complaint and appeal procedures. By doing this, the US government will strengthen

algorithmic transparency and accountability. This is because the public can review how the government-funded AI is developed and adopted. It will also enable independent auditing with the available information and support policy evaluation, building public trust.

Kroll et al. (2017) suggest that the federal government organize a National AI Oversight Authority to provide a Uniform Federal AI Governance Structure. The United States should create a specialized federal AI oversight body within the President's Executive Office, modeled after the EU's centralized AI Office. This authority should establish uniform technical specifications and supervision processes and guarantee that the criteria for monitoring are applied consistently. It should also organize the governance of AI for all federal agencies and assist organizations without AI capabilities by providing them with specific knowledge and resources. Lastly, enforcing meaningful human oversight will strengthen evaluation mechanisms and human control by establishing clear standards for human evaluation based on international best practices.

Results Synthesis and Policy Implications for US Government AI Oversight

The findings of this study claim that there is a fundamental misalignment between formal AI regulations and practical monitoring practices across US government organizations. Agencies vary widely in how they execute broad frameworks like Executive Order 14110 and OMB Memorandum M-24-10, which establish high requirements for AI governance. To begin with, the many policy reversals between governments create ambiguity in execution. The contrast between Biden's all-encompassing regulatory strategy and Trump's deregulatory stance demonstrates how political changes could endanger the long-term development of AI governance (The White House, 2025). This volatility affects private sector compliance strategies, resource allocation, and staff training investments. The study also found that even agencies with ample resources encounter substantial technical obstacles. Implementing policies can be hampered by basic technical constraints, as demonstrated by the DoD's problem with unlabeled information.

Parallel to this, operational obstacles that cannot be handled by legislation alone are caused by cybersecurity flaws and difficulties in integrating legacy systems. This disparity is most evident in the supervisory monitoring requirements. Although laws require people to analyze AI decisions meaningfully, practice reveals "rubber-stamping" techniques, an over-reliance on AI recommendations, and undertrained employees who cannot evaluate algorithmic outputs. The discrepancy between desired human-AI collaboration and actual implementation compromises the fundamental tenet of human-centered AI governance. Finally, research indicates that formal policies necessitate human monitoring, bias testing, and transparency measures; nonetheless, implementation flaws persist due to various systemic issues. Larger agencies like DoD, DHS, and HHS have made great progress in creating AI governance systems, whereas smaller agencies struggle with a lack of funding and technical expertise. This creates an unequal environment where agency capabilities, rather than legal requirements, determine AI responsibility.

Implications for Future Policy

A thorough disclosure of AI system operations, including algorithmic logic, training data properties, bias test outcomes, and human oversight procedures, should be required by future policy. Trade secret protections should be interpreted carefully for government AI systems that impact citizens' rights. The protection of individual rights would be strengthened by extending Section 1983 protections, elucidating disparate effect applications to AI systems, and establishing specialized administrative procedures for complaints about AI. The current legal framework offers minimal remedies for algorithmic injury. Closing the skills gap across agencies requires systematic investment in interdisciplinary collaborative capacities, technical knowledge growth, and AI literacy training.

Future rules should establish training requirements for the entire government and mandate a certain degree of technical proficiency from AI oversight personnel. The budgetary constraints of smaller organizations also demonstrate the need for funding sources for AI governance. Future rules should establish common services for technical competency, centralized funding for AI supervision skills, and economies of scale for bias detection systems. Last but not least, the paper claims that executive directives cannot provide the stability needed for effective AI governance. The basis would be strengthened by legislation establishing clear enforcement

procedures, mandatory minimum requirements for AI supervision, and legal remedies for algorithmic misuse, rather than by executive action alone.

CONCLUSION

This research article examines the gaps between formal AI oversight policies in the American government and their functional deployment across its agencies. Although governmental directives continuously call for accountability in AI adoption, the findings of this study show uneven compliance due to disparities in governance structures, funding, technical capacity, and staffing. This research also highlights that despite the abundance of legal and procedural frameworks for accountability; they cannot adequately address the unique challenges of algorithmic governance. Traditional administrative law frameworks, such as the APA and civil rights enforcement procedures, must be modified to handle algorithmic decision-making processes affecting millions daily effectively. Additionally, there are still significant gaps in ensuring efficient human oversight, managing algorithmic risks, and upholding consistent implementation standards. This is despite agencies effectively establishing the fundamental AI governance structures, designating Chief AI Officers, keeping AI inventories, and creating compliance frameworks.

Although AIF360 and Aequitas are promising bias detection technologies, their use is still uneven throughout the government. Special emphasis should be placed on creating legal frameworks tailored to AI. Federal courts have yet to explore how much current statutory and constitutional protections extend to algorithmic decision-making. Both policymaking and legal practice would benefit from research on successful legal challenges and their precedent-setting impacts. Studies on the patterns of human-AI interaction in governmental contexts may influence better oversight and training procedures. The selection of government tools would also be optimized with comparative effectiveness evaluations of various bias detection approaches. Therefore, several important areas need more research. A sustainable AI governance model might be developed with the help of longitudinal research looking at the long-term efficacy of the current oversight mechanisms. Lastly, intellectual efficacy, administrative stability, and significant responsibility must be prioritized in any future AI supervision framework.

To do this, comprehensive legislative frameworks resilient to political shifts and adaptable to technological advancements must replace executive action. The only way the US government can fulfill the potential of AI while upholding democratic accountability and safeguarding citizen rights is by acknowledging that ongoing investment in institutional capacity, technical infrastructure, and human capital is necessary for effective AI governance, in addition to policy declarations.

REFERENCE

1. Anderson, A. J. (2025). What Is Disparate-Impact Discrimination? (IF13057) https://www.congress.gov/crs-product/IF13057
2. Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2022). Machine bias*. Ethics of Data and Analytics, 254–264. https://doi.org/10.1201/9781003278290-37
3. Artificial Intelligence Research & Development Interagency Working Group. (2019). 2016–2019 PROGRESS REPORT: ADVANCING ARTIFICIAL INTELLIGENCE R&D. https://trumpwhitehouse.archives.gov/wp-content/uploads/2019/11/AI-Research-and-Development-Progress-Report-2016-2019.pdf
4. Attard-Frost, B., Brandusescu, A., & Lyons, K. (2024). The governance of artificial intelligence in Canada: Findings and opportunities from a review of 84 AI governance initiatives. Government Information Quarterly, 41(2), 101929. https://doi.org/10.1016/j.giq.2024.101929
5. Biden White House Archives. (2024). FACT SHEET: OMB issues guidance to advance the responsible acquisition of AI in government. Office of Management and Budget. https://bidenwhitehouse.archives.gov/omb/briefing-room/2024/10/03/fact-sheet-omb-issues-guidance-to-advance-the-responsible-acquisition-of-ai-in-government/
6. Brantingham, P. J., Valasik, M., & Mohler, G. O. (2022). Does predictive policing lead to biased arrests? Results from a randomized controlled trial. Statistics and Public Policy, 5(1), 1–6. https://doi.org/10.1080/2330443x.2018.1438940

7. Brookings Institution. (2025). Trump's executive orders politicize AI. https://www.brookings.edu/articles/trumps-executive-orders-politicize-ai/
8. Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. Big Data & Society, 3(1). https://doi.org/10.1177/2053951715622512
9. Cabinet Office. (2019). Society 5.0. Cao.go.jp. https://www8.cao.go.jp/cstp/english/society5_0/index.html
10. Corbett-Davies, S., & Goel, S. (2018). The measure and mismeasure of fairness: A critical review of fair machine learning. arXiv. https://doi.org/10.48550/arXiv.1808.00023
11. Deguchi, A., Hirai, C., Matsuoka, H., Nakano, T., Oshima, K., Tai, M., & Tani, S. (2020). What Is Society 5.0? Society, 1–23. https://doi.org/10.1007/978-981-15-2989-4_1
12. Deloitte. (2020). AI in Government: A Guide to Implementing AI Solutions.
13. Department of Homeland Security. (2023). Artificial intelligence use case inventory. https://www.dhs.gov/ai-use-cases
14. Federal AI Compliance Plans for OMB Memorandum M-24-10. (2024). https://static.carahsoft.com/concrete/files/2317/3888/1325/Compliance_Federal_Agency_AI_OMB_M-24-10_Compliance_Plans.pdf
15. Federal Register. (2023, October 30). Federal Register:: Request Access. Unblock.federalregister.gov.https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence
16. Garvie, C., Bedoya, A., & Frankle, J. (2016). The Perpetual Line-Up: Unregulated Police Face Recognition in America. Upturn.
17. Government Accountability Office. (2024). Artificial intelligence: Agencies are implementing management and personnel requirements (GAO-24-107332). https://www.gao.gov/products/gao-24-107332
18. Government Accountability Office. (2024). Artificial intelligence: Agencies have begun implementation but must complete key requirements (GAO-24-105980). https://www.gao.gov/products/gao-24-105980
19. Government Accountability Office. (2025). Artificial intelligence: Generative AI use and management at federal agencies (GAO-25-107653). https://www.gao.gov/products/gao-25-107653
20. Hoadley, D. S., & Lucas, N. J. (2018). Artificial intelligence and national security. Congress. https://www.congress.gov/crs_external_products/R/PDF/R45178/R45178.3.pdf
21. Home - AI Fairness 360. (2020, October 2). AI Fairness 360. http://ai-fairness-360.org/
22. Huergo, J. (2023, February 7). NIST calls for information to support safe, secure, and trustworthy development and use of artificial intelligence. https://www.nist.gov/news-events/news/2023/12/nist-calls-information-support-safe-secure-and-trustworthy-development-and
23. Kroll, J. A., Barocas, S., Kleinberg, J., & Levy, K. (2017). Accountable Algorithms. University of Pennsylvania Law Review.
24. Kroll, J. A., Huey, J., Barocas, S., Felten, E. W., Reidenberg, J. R., Robinson, D. G., & Yu, H. (2017). Accountable algorithms. University of Pennsylvania Law Review, 165(3), 633–705.
25. Office of Management and Budget. (2024). Memorandum M-24-10: Advancing governance, innovation, and risk management for agency use of artificial intelligence. https://www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf
26. Office of Science and Technology Policy. (2023, November 1). Executive Order 14110: Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence (Presidential Document). Federal Register, Vol. 88, No. 210, 75191–... Federal RegisterGovInfo
27. Olsen, H. P., Hildebrandt, T. T., Wiesener, C., Larsen, M. S., & Flügge, A. W. (2024). The right to transparency in public governance: Freedom of information and the use of artificial intelligence by public agencies. Digital Government: Research and Practice, 5(1), 1-15. https://doi.org/10.1145/3632753
28. Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J., Breazeal, C., … & Wellman, M. (2019). Machine behaviour. Nature, 568(7753), 477–486

29. Robinson, S., & Wexler, J. (2019, July 19). Introducing the what-if tool for cloud AI platform models. Google Cloud Blog. https://cloud.google.com/blog/products/ai-machine-learning/introducing-the-what-if-tool-for-cloud-ai-platform-models
30. Text - H.R.7532 - 118th Congress (2023-2024): Federal AI Governance and Transparency Act. (2023). Congress.gov. https://www.congress.gov/bill/118th-congress/house-bill/7532/text
31. The White House. (2024). Memorandum on Advancing the United States' Leadership in Artificial Intelligence; Harnessing Artificial Intelligence to Fulfill National Security Objectives; and Fostering Artificial Intelligence's Safety, Security, and Trustworthiness. 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/
32. The White House. (2025). America's Action Plan. EOP. https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf
33. The White House. (2025, January 23). Removing barriers to American leadership in artificial intelligence. https://www.whitehouse.gov/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence/
34. Trump White House Archives. (2019). Artificial intelligence for the American people. https://trumpwhitehouse.archives.gov/ai/
35. US Government Accountability Office. (2025). Generative AI Use and Management at Federal Agencies (GAO-25-107653). https://www.gao.gov/products/gao-25-107653
36. US Homeland Security. (2022). US Department of Homeland Security Annual Performance Report. Homeland Security. https://www.dhs.gov/sites/default/files/2022-04/DHS%20FY21-23%20APR.pdf
37. White House. (2023). Executive Order 14110: Safe, secure, and trustworthy artificial intelligence development and use. Federal Register, 88(219), 75191–75236.
38. White House. (2025). Executive Order: Removing barriers to American leadership in artificial intelligence. https://www.whitehouse.gov/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence/