Standardizing AI Model Transparency in U.S. Federal Agencies: A Framework for Implementing AI Model Cards and Data Provenance Auditing

Chidinma Queen Adieze
Department of Data Analytics-Data science, Western Governor University

Fabian Emesiani, LL.B BL.
LL.M(Privacy and Cybersecurity Law)
USC. D.B.A (Westcliff University).

Taiwo Paul Onyekwuluje
University of West Georgia, Atlanta

Chisom Elizabeth Alozie
Department. Information Technology,
University of the cumberlands,
Kentucky, United States

ABSTRACT
As artificial intelligence becomes deeply embedded in U.S. federal agency operations spanning healthcare delivery, law enforcement, benefits administration, national security, and regulatory rulemaking the imperative for systematic, standardized transparency has never been more pressing. This article examines the landscape of AI model transparency in U.S. federal agencies, synthesizing recent policy developments including OMB Memoranda M-25-21 and M-25-22 (April 2025), Executive Order 14179, the December 2025 OMB "Unbiased AI" guidance, the NIST AI Risk Management Framework (AI RMF), and the evolving doctrine of data provenance auditing. We analyze the 2024 Federal AI Use Case Inventory which disclosed more than 2,133 use cases including 227 rights- and safety-impacting deployments and evaluate the current patchwork of transparency requirements against emerging best practices for AI model cards. We then propose a comprehensive, five-tier standardization framework for implementing model cards and data provenance auditing across federal agencies, including governance structures, technical standards, enforcement mechanisms, and workforce capacity requirements. Our findings reveal significant inconsistencies in current reporting, dangerous gaps in third-party auditing capacity, and a critical need for interoperable provenance infrastructure. We conclude with policy recommendations for Congress, OMB, NIST, and Chief AI Officers.

Keywords
AI transparency, model cards, data provenance, federal agencies, AI governance, OMB, NIST AI RMF, algorithmic accountability, AI auditing, Chief AI Officers

1. INTRODUCTION
The United States federal government is now one of the largest institutional consumers of artificial intelligence (AI) in the world. From the Department of Veterans Affairs' clinical decision-support tools to the Department of Homeland Security's facial recognition systems, AI is being deployed at scale in contexts where errors, biases, or opaque decision-making can have profound consequences for citizens' rights, safety, and access to government services.[1] Yet the frameworks governing what agencies must disclose about the AI systems they deploy how models are trained, on what data, with what known limitations, and to what end remain fragmented, inconsistently enforced, and insufficiently specific.

The core accountability instruments at stake in this article are AI model cards and data provenance auditing. Model cards, first formalized by Mitchell et al. (2019) at Google, are standardized documentation artifacts that describe a model's intended use, performance metrics across demographic subgroups, training data characteristics, ethical considerations, and operational caveats.[2] Data provenance auditing, as defined by NIST, refers to systematic verification of "the chronology of the origin, development, ownership, location, and changes" to data and systems throughout their lifecycle.[3] Together, these instruments represent the technical backbone of AI accountability the mechanisms by which governments, oversight bodies, and the public can verify whether an AI system is fit for its stated purpose.

As of December 2024, U.S. federal agencies collectively reported over 2,133 AI use cases a 200% increase from the prior year with 227 classified as directly impacting citizens' rights or safety.[4] Of those high-risk deployments, 206 received extension requests from OMB because agencies had not yet implemented the required risk management safeguards as of the December 1, 2024 compliance deadline.[5] This data alone reveals a structural gap between the scale of AI adoption and the maturity of governance frameworks designed to make that adoption transparent and accountable [79].

The political context is complex. The Biden administration laid significant groundwork through OMB M-24-10 (March 2024) and a healthcare AI model card requirement issued by the Office of the National Coordinator for Health Information Technology (ONC). The Trump administration, upon taking office in January 2025, rescinded the prior OMB guidance and replaced it with M-25-21 and M-25-22, which accelerate AI adoption while maintaining some transparency requirements.[6] In December 2025, OMB issued additional guidance on "Unbiased AI Principles" for LLM procurement, requiring agencies to obtain model cards as part of minimum transparency conditions.[7] Simultaneously, the administration's HHS proposed eliminating healthcare AI model card requirements, creating a regulatory contradiction at the heart of federal AI governance.[8]

This article proceeds as follows. Section 2 reviews the existing policy and regulatory landscape. Section 3 analyzes the current

state of AI model card implementation. Section 4 examines data provenance auditing frameworks and gaps. Section 5 proposes a standardization framework. Section 6 addresses implementation challenges. Section 7 offers policy recommendations, and Section 8 concludes.

2. POLICY AND REGULATORY LANDSCAPE

2.1 Executive Orders and OMB Guidance

The federal government's engagement with AI transparency dates to Executive Order 13960, signed by President Trump in December 2020, which required agencies to annually inventory AI use cases and make them public.[9] This requirement was codified and expanded by the bipartisan Advancing American AI Act of 2022, which formalized the inventory mandate into statute.[10]

Under President Biden, the landmark Executive Order 14110 on the Safe, Secure, and Trustworthy Development and Use of AI (October 2023) directed the most comprehensive suite of AI transparency actions the federal government had yet undertaken, including requirements for AI safety evaluations, content provenance mechanisms, and sector-specific disclosure rules.[11] OMB's implementing memorandum M-24-10 (March 2024) established requirements for identifying rights- and safety-impacting AI use cases, conducting AI impact assessments, and publishing risk management information publicly.[12]

The Trump administration's Executive Order 14179 (January 23, 2025) rescinded E.O. 14110, emphasizing instead the removal of barriers to AI adoption and the promotion of American AI competitiveness.[13] OMB's replacement memoranda, M-25-21 and M-25-22 (April 3, 2025), reoriented federal AI governance around innovation acceleration while retaining some transparency infrastructure: agencies must annually publish AI use case inventories, designate Chief AI Officers (CAIOs) within 60 days, convene AI Governance Boards within 90 days, and obtain documentation from vendors that "facilitates transparency and explainability" for procured AI.[14]

Most significantly for AI model transparency, OMB's December 11, 2025 memorandum on "Increasing Public Trust in Artificial Intelligence Through Unbiased AI Principles" established minimum vendor disclosure requirements for LLM procurement. These include: acceptable use policies, model or system cards, mechanisms for users to report policy-violating outputs, and in enhanced cases information on pre- and post-training activities, model-bias evaluations, and third-party modifications.[15] Agencies were required to update procurement policies by March 11, 2026.[16]

2.2 NIST AI Risk Management Framework

The National Institute of Standards and Technology AI Risk Management Framework (NIST AI RMF), released in January 2023, provides the foundational technical structure for federal AI governance. Its four core functions Map, Measure, Manage, and Govern (MMMG) form a repeatable cycle for identifying, assessing, and mitigating AI-related risks.[17] The companion Generative AI Profile (NIST AI 600-1, July 2024) extends this framework to address risks specific to large language models and other generative systems, including content provenance, data lineage, pre-deployment testing, and incident disclosure.[18]

The March 2025 update to NIST AI RMF significantly expanded its scope to address model provenance, data integrity, and third-party model assessment, recognizing that most federal agencies rely on externally developed or open-source AI components.[19] It also introduced new threat categories poisoning attacks, evasion attacks, data extraction, and model manipulation relevant to LLMs deployed in government contexts.[20] NIST's concurrent development of security control overlays for AI systems (adapting SP 800-53 controls for AI-specific concerns including model integrity, data provenance, adversarial robustness, and transparency) further illustrates the convergence of cybersecurity and AI governance frameworks.[21]

In September 2025, NIST released an extended outline for a proposed Zero Draft standard on documentation of AI datasets and AI models effectively a precursor to a formal model card standard signaling movement toward mandatory, interoperable documentation requirements.[22]

2.3 Sector-Specific Regulations

In healthcare, the ONC's 2023 Health IT rulemaking required vendors of certified health IT to submit AI model cards described as "quasi-nutrition labels" covering training data, performance characteristics, and patient risk information for clinical decision support tools.[23] This requirement went into effect in early 2025 but was proposed for elimination by the Trump administration's HHS in late December 2025, citing regulatory burden reduction.[24] The rollback was widely criticized by patient safety advocates and healthcare AI researchers as removing a critical safety mechanism at precisely the moment AI was being most aggressively integrated into clinical workflows.[25]

The December 11, 2025 Executive Order on a National AI Policy Framework directed the FCC to consider a federal reporting and disclosure standard for AI models and tasked the FTC with issuing policy statements on how the FTC Act applies to AI signaling ongoing federal movement toward mandatory disclosure, despite deregulatory pressures in other domains.[26]

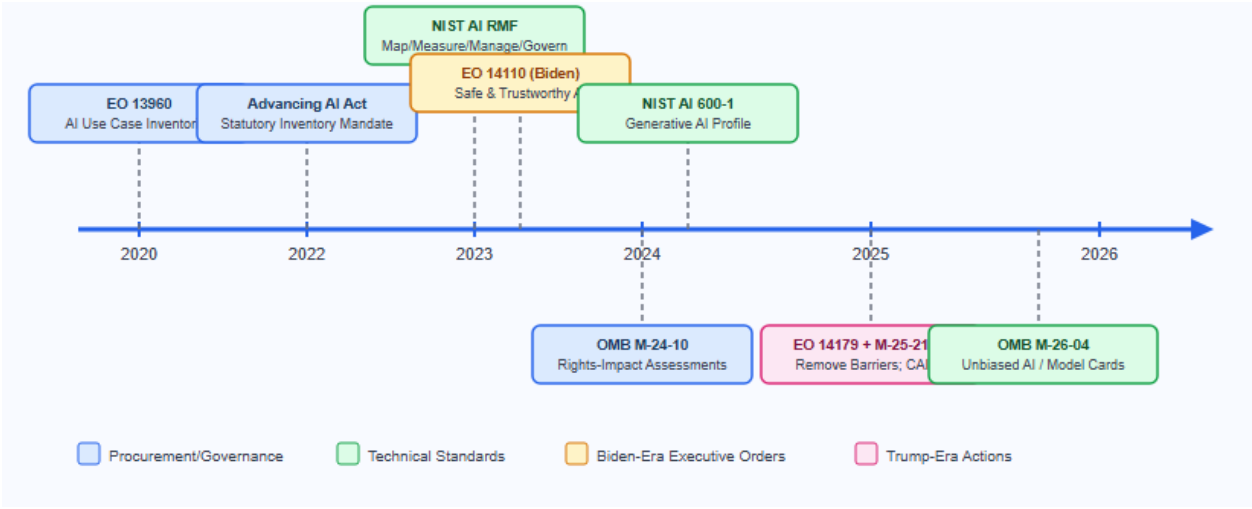

Figure 1: Federal AI Transparency Policy Timeline (2020–2026)

Figure 1: Chronological evolution of federal AI transparency policy from EO 13960 (2020) through OMB M-26-04 (December 2025), showing the layered policy architecture governing model cards and AI accountability in U.S. federal agencies.

3. AI Model Cards: Current State and Standardization Gaps

3.1 What Model Cards Are and Why They Matter

A model card is a concise, standardized report detailing an AI model's intended use, performance across demographic subpopulations, training data provenance, ethical considerations, and operational caveats.[27] First introduced as a formal artifact by Mitchell et al. (2019) and subsequently adopted across industry by Google, Meta, Hugging Face, and others, model cards serve multiple functions: they enable informed procurement decisions, facilitate independent auditing, surface known failure modes before deployment, and create a documented record of accountability should harms arise.[28]

In the federal context, a model card performs additional governance functions. It provides a common language between technical teams and agency leadership, satisfies documentation requirements for rights- and safety-impacting AI systems, enables the Chief AI Officer to maintain situational awareness across an agency's portfolio, and creates an audit trail that Inspector General offices and the GAO can use during oversight reviews.[29] The GAO's AI Accountability Framework (2021) explicitly identifies documentation of training data, model performance, and deployment decisions as key audit procedures for AI systems.[30]

3.2 The Current Patchwork of Model Card Requirements

As of early 2026, no unified federal standard for AI model cards exists. Instead, agencies navigate a patchwork of overlapping and sometimes contradictory requirements. The OMB M-26-04 guidance requires agencies to obtain "model or system cards" from LLM vendors as part of minimum transparency conditions for procurement but provides no standardized template, no specification of required fields, and no validation mechanism.[31] The NIST AI RMF recommends model cards as a governance tool but treats them as voluntary artifacts rather than mandatory deliverables.[32] The now-rescinded ONC healthcare rule provided the most operationally detailed model card specification in federal regulation to date, requiring vendors to document:

(1) training data sources and demographics,

(2) model performance across patient subgroups,

(3) intended clinical use cases,

(4) known limitations and contraindications, and

(5) update and versioning history.[33]

The Center for Democracy and Technology's analysis of the 2024 federal AI inventories found that while agencies reported significantly more detail than in prior years, the information was "reported inconsistently, undermining the usefulness of this greater degree of reporting."[34] Key gaps included: agencies using different definitions of "high-impact AI," varying levels of detail in risk management disclosures, and near-complete absence of model performance data across demographic subgroups.[35] This inconsistency is not merely an administrative inconvenience it makes cross-agency comparisons impossible, frustrates independent oversight, and prevents the public from making meaningful judgments about whether government AI systems pose risks to their rights or safety.

Table 1. Comparison of Model Card Requirements Across U.S. Federal AI Policy Instruments (2024–2026)

Policy Instrument	Year	Model Card Requirement	Required Fields	Enforcement Mechanism	Status (2026)
OMB Memorandum M-24-10	2024	AI impact assessments for rights- and safety-impacting AI	Intended use, risk level, safeguards, AI impact assessment	Deadline-based compliance; extension requests	Rescinded (Jan 2025)
Health Data, Technology, and Interoperability Rule HTI-1	2023	Mandatory AI model cards for certified health IT	Training data, demographics, performance metrics, limitations, update history	Certification denial for non-compliance	Proposed elimination (Dec 2025)
OMB Memorandum M-25-21	2025	Vendor documentation for high-impact AI procurement	Transparency and explainability documentation; performance tracking	Chief AI Officer oversight; procurement contract terms	Active
OMB Memorandum M-26-04	2025	Mandatory model/system cards for all LLM procurement	Acceptable use policy, model card, bias reporting; training details and bias evaluations	Contract eligibility or termination clause	Active; sunsets 2027
NIST AI Risk Management Framework	2023 / 2025	Recommended (voluntary) documentation	Provenance, testing, incident disclosure	None (voluntary)	Active; referenced in federal contracts
GAO Artificial Intelligence Accountability Framework	2021	Audit evidence documentation	Governance, data, performance, monitoring documentation	Audit findings; congressional reporting	Active

3.3 Model Card Components: A Federal Standard Specification

Drawing on the NIST AI 600-1 Generative AI Profile, the GAO AI Accountability Framework, Mitchell et al. (2019), and the ONC HTI-1 requirements, we identify eleven essential components for a federally standardized AI model card:

1. Model Identification and Versioning. Unique identifier, version number, release date, vendor/developer information, and a point-of-contact for technical inquiries. This enables traceability across the model lifecycle and supports CAIO portfolio management.[36]

2. Intended Use and Use Restrictions. Explicit description of the model's designed purpose, appropriate deployment contexts, and out-of-scope uses. Agencies must identify uses that require additional safeguards or are expressly prohibited.[37]

3. Training Data Provenance. Origins of training data, collection methods, data governance controls, consent and licensing arrangements, and geographic/temporal scope. This field must document whether any training occurred outside the United States, per OMB M-26-04 requirements.[38]

4. Data Processing and Filtering. Description of data preprocessing, augmentation, deduplication, and filtering steps. Documentation of red-teaming and bias mitigation applied to training data.[39]

5. Model Architecture and Training Configuration. Model family, parameter scale, training methodology, and hardware infrastructure. For open-source components, documentation of any fine-tuning or adaptation.[40]

6. Evaluation Metrics and Performance Benchmarks. Overall accuracy, fairness metrics disaggregated by race, gender, age, disability status, and other protected characteristics as defined under applicable civil rights law.[41]

7. Known Limitations and Failure Modes. Documented performance degradation scenarios, adversarial vulnerabilities, out-of-distribution behavior, and hallucination rates for generative models.[42]

8. Safety Filters and Content Moderation. Type and scope of content restrictions, red-teaming results, and mechanisms for ongoing safety monitoring. For OMB M-26-04 compliance, documentation of ideological neutrality controls.[43]

9. Third-Party Modifications and Supply Chain. Any fine-tuning, alignment training, or modifications performed by third parties outside the primary developer. Includes RLHF or RLAIF processes.[44]

10. Incident Reporting and Update Protocols. Process for reporting model failures or harms, version update schedule, and re-evaluation triggers. Required to align with NIST AI 600-1 incident disclosure guidance.[45]

11. Governance and Accountability Contacts. Agency CAIO, program manager, legal counsel, and civil rights officer responsible for this deployment. Enables accountability chains required under M-25-21.[46]

4. DATA PROVENANCE AUDITING: FRAMEWORK AND GAPS

4.1 Defining Data Provenance in the Federal AI Context

NIST defines provenance as "the chronology of the origin, development, ownership, location, and changes to a system or system component and associated data."[47] In the context of federal AI deployments, data provenance encompasses:

(1) the origins and collection methods of training data;

(2) data transformation and preprocessing pipelines;

(3) the lineage of any synthetic data used in model development;

(4) post-training modifications including fine-tuning and alignment; and

(5) the provenance of inference-time data used to inform model outputs in operational contexts.[48]

Data provenance is foundational to AI accountability for a simple reason: a model can only be as unbiased, accurate, and fit-for-purpose as the data on which it was trained. Data collection, provenance, and management carry significant risks of encoding and reifying bias deliberately or unintentionally making rigorous provenance documentation not merely a governance formality but a substantive safeguard against discriminatory outcomes.[49]

4.2 Current Federal Provenance Requirements and Their Limitations

The National Telecommunications and Information Administration (NTIA) has identified data provenance as a critical component of AI accountability, noting that "for AI accountability, provenance and authentication help users recognize AI outputs, identify human sources, report incidents of harm, and ultimately hold AI developers, deployers, and users responsible for information integrity."[50] NIST's work on synthetic content provenance (NIST AI 100-4, 2024) has advanced technical methods including digital watermarking, metadata recording, and content authentication for tracking the origin and modification history of AI-generated content.[51]

However, federal requirements for systematic data provenance auditing remain nascent and unevenly applied. OMB M-25-22 requires that agency contracts "permanently prohibit the use of non-public inputted and outputted results to further train publicly or commercially available AI algorithms absent explicit agency consent" a critical protection for government data but does not specify how agencies should verify vendor compliance with this prohibition.[52] The Federal Reserve's AI Use Case Inventory, for example, includes a Section 3 specifically documenting "whether the data used for the AI model is owned, documented, and distributable by the agency" a nascent provenance field that other agencies have not consistently adopted.[53]

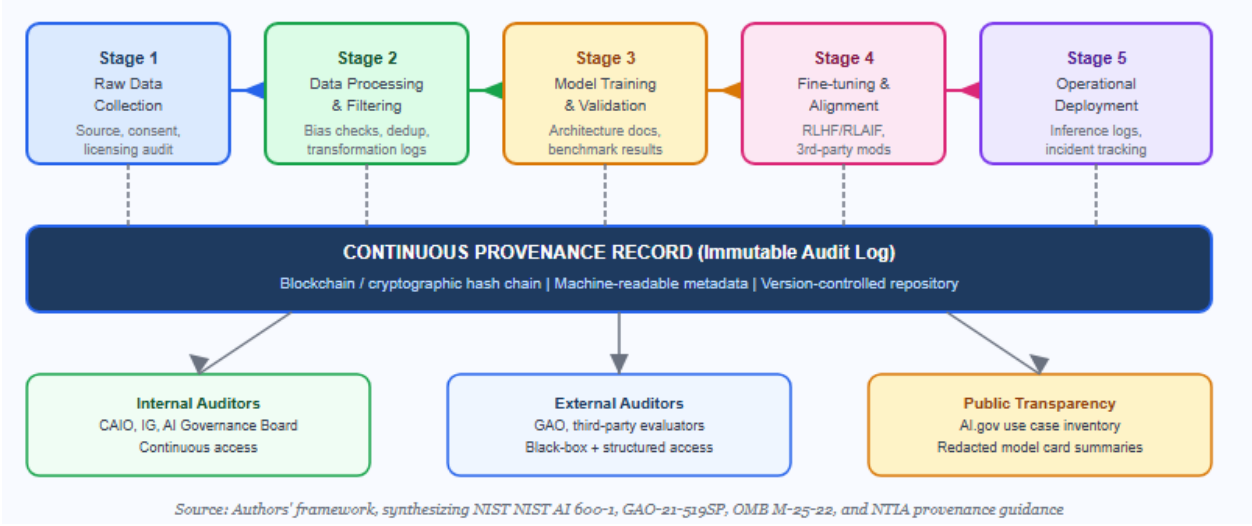

Figure 2: Data Provenance Auditing Chain for Federal AI Systems

Figure 2: Complete data provenance auditing chain for federal AI systems, showing five development stages, a continuous provenance record layer, and tiered audit access for internal, external, and public oversight.

4.3 The Auditing Capacity Gap

A systematic analysis of 435 existing AI audit tools, published by researchers at the 2025 ACM CHI Conference, found that while many tools exist for standards identification (n=206) and performance analysis (n=129), tools for model and data transparency (n=12) and audit communication (n=8) are dramatically underrepresented in the current ecosystem.[54] Among 35 AI audit practitioners interviewed at the same study, a consistent theme emerged: external auditors face significant challenges obtaining data and models needed to conduct genuinely independent audits, often developing their own data collection tools.[55]

For federal AI systems, this auditing capacity gap is especially serious. The Democracy Forward Foundation's October 2025 lawsuit against OPM, GSA, HUD, and OMB alleging failure to respond to FOIA requests about AI use in federal rulemaking illustrates the real-world consequences of opaque AI governance: when agencies deploy AI systems in consequential contexts without adequate documentation, even basic transparency mechanisms like FOIA become ineffective.[56] The complaint specifically highlighted the use of AI to summarize public comments in rulemaking processes

a context where decision-making provenance is directly relevant to the Administrative Procedure Act's requirement for "reasoned decision-making."[57]

Table 2. 2024 Federal AI Use Case Inventory: Key Statistics and Transparency Indicators

Metric	Value	Comparison (2023)	Significance
Total reported AI use cases	2,133	710	~200% increase reflecting improved reporting and expanded adoption
Rights- and safety-impacting use cases	227	Not previously tracked	First systematic identification of high-risk deployments
High-risk use cases requesting compliance extensions	206 of 227 (91%)	N/A	Indicates major compliance gap
Agency with most use cases	HHS (271)	+66% growth	Healthcare is the largest federal AI deployment sector
DHS use case growth	136% increase	Prior baseline	Rapid law-enforcement and border-security AI expansion
GitHub-accessible consolidated inventory	First year (2024)	Previously siloed agency sites	Enables cross-agency comparability and machine-readable access
States adopting similar AI inventory laws	At least 12	Fewer	Federal reporting model influencing state policy
Top AI use categories	Mission-enabling; health and medical; government services	Similar distribution	Internal operational uses still dominate

5. A FIVE-TIER STANDARDIZATION FRAMEWORK

We propose a five-tier framework for standardizing AI model transparency and data provenance auditing across U.S. federal agencies. This framework is designed to be risk-proportionate matching documentation depth and audit intensity to the potential impact of the AI system while establishing universal baseline requirements that all agencies must meet regardless of use case risk level.

Figure 3: Five-Tier Federal AI Transparency & Provenance Standardization Framework

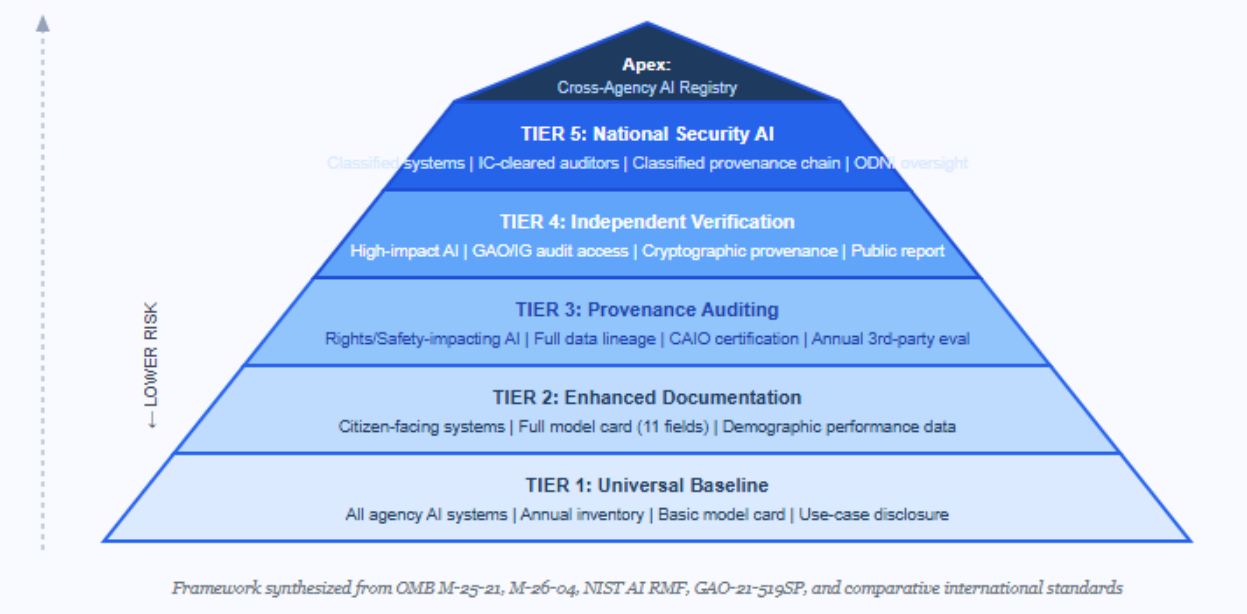

Figure 3: Five-tier risk-proportionate framework for federal AI transparency standardization. Tier 1 applies to all agency AI systems; Tier 5 applies to classified national security AI. Each higher tier inherits all requirements of lower tiers.

5.1 Tier 1 Universal Baseline

Every AI system deployed by any federal agency, regardless of risk level, must meet Tier 1 requirements: annual public disclosure in the Federal AI Use Case Inventory, a minimum model card covering identification, intended use, and vendor information, and designation of an agency point-of-contact. This tier aligns with existing OMB M-25-21 inventory requirements and is achievable with current agency capacity.

5.2 Tier 2 Enhanced Documentation

AI systems deployed in citizen-facing contexts including benefits determination, customer service, information

provision, and service delivery must meet Tier 2 requirements: the complete eleven-field model card specified in Section 3.3, including demographic performance data. Vendors must provide updated documentation when integrating new AI enhancements, features, or model components, per OMB M-25-22 guidance.[58]

5.3 Tier 3 Provenance Auditing

Rights- and safety-impacting AI systems as defined by OMB M-25-21 must meet Tier 3 requirements: complete data lineage documentation across all five provenance stages (Figure 2), CAIO certification of documentation accuracy, and annual third-party evaluation of model performance and fairness metrics. This tier requires agencies to implement the minimum risk management practices mandated by M-25-21 before deployment.

5.4 Tier 4 Independent Verification

High-impact AI systems those serving as a principal basis for decisions affecting civil rights, access to critical government resources, human health and safety, or critical infrastructure must meet Tier 4 requirements: black-box access for GAO and Inspector General auditors; cryptographic provenance documentation using immutable audit logs; and annual public reports on model performance, incident history, and safeguard effectiveness. This tier draws on the ACM research finding that "at minimum, black-box access... should be granted to auditors" as it "effectively balances concerns related to proprietary technology, data privacy, audit standardization, and audit efficiency."[59]

5.5 Tier 5 National Security AI

Classified AI systems operated by intelligence community elements and national security agencies meet modified Tier 5 requirements: IC-cleared auditors from the ODNI and relevant Inspector General offices, classified provenance documentation chains, and classified congressional oversight reporting through the Intelligence Committees. This tier acknowledges the sensitivity of national security systems while preserving meaningful oversight through cleared channels.

Table 3. Five-Tier Federal AI Transparency Framework: Requirements Matrix

Requirement	Tier 1	Tier 2	Tier 3	Tier 4	Tier 5
Annual use-case inventory	✓	✓	✓	✓	Classified
Minimum model card (4 fields)	✓	✓	✓	✓	✓ (classified)
Full 11-field model card	—	✓	✓	✓	✓ (classified)
Demographic performance data	—	✓	✓	✓	✓ (classified)
Complete data lineage documentation	—	—	✓	✓	✓ (classified)
CAIO certification	—	✓	✓	✓	✓
Third-party evaluation (annual)	—	✓	✓	Cleared evaluators	—
GAO / IG black-box audit access	—	—	—	✓	Cleared IGs only
Cryptographic provenance chain	—	—	—	✓	✓ (classified)
Public annual performance report	—	—	✓	✓	—
Congressional oversight reporting	—	—	✓	✓	Intelligence committees

6. IMPLEMENTATION CHALLENGES

6.1 Workforce Capacity and Technical Expertise

Federal agencies face a significant AI workforce capacity gap. Implementing comprehensive model card and provenance auditing requirements demands expertise in machine learning, data governance, algorithmic fairness assessment, and technical documentation skills that are scarce in most agencies' civil service workforce. OMB M-25-21 acknowledges this challenge, directing agencies to prioritize "building an AI-ready federal workforce" through training, resources, talent acquisition, and accountability measures.[60] However, the memo provides little operational guidance on how agencies with limited technical capacity should practically implement transparency requirements, particularly for legacy AI systems already in operation.

The auditing gap is equally severe. A 2025 study of AI audit practitioners found that the available tools "do not currently support the full scope of AI audit practitioners' needs" and that the field needs to "move beyond evaluation" to tools for harms discovery, transparency infrastructure, and audit communication.[61] For government Inspector General offices, which are already resource-constrained and rarely have deep AI expertise, auditing AI model cards and provenance chains represents an entirely new competency that will require substantial investment to develop.

6.2 Vendor Resistance and Proprietary Information Claims

Many AI vendors particularly large commercial providers of foundation models have strong incentives to resist comprehensive model card and provenance disclosure requirements. They may claim that training data composition, model architecture details, or fine-tuning methodologies constitute trade secrets or provide a competitive advantage. The OMB M-25-22 procurement framework anticipates this tension, requiring contracts to include vendor knowledge transfers and transparency documentation, while acknowledging the need to balance transparency with intellectual property protection.[62] The balance is genuinely difficult: overly aggressive disclosure requirements may deter high-quality vendors from competing for federal contracts,

while insufficient requirements leave agencies unable to meaningfully assess the AI systems they deploy.

A practical resolution is the tiered disclosure model described in Section 5: minimum public-facing disclosure in model cards, enhanced documentation available to agency technical staff under appropriate non-disclosure agreements, and full provenance access restricted to cleared auditors for the most sensitive systems. This approach mirrors the "black-box access" model advocated in the academic literature as striking the appropriate balance between proprietary concerns and audit effectiveness.[63]

6.3 Legacy Systems and Technical Debt

A substantial portion of federal AI deployments are built on legacy systems with limited documentation, inadequate data management practices, and fragmented development histories. For these systems, creating comprehensive provenance records retrospectively tracing training data origins, preprocessing pipelines, and model iterations from years prior may be technically infeasible. The DHS 2025 AI Use Case Inventory acknowledged this challenge by refining risk categories and streamlining reporting for deployed systems, while requiring agencies to implement M-25-21 minimum practices for high-impact AI by April 3, 2026, or discontinue use.[64] The compliance deadline approach creates appropriate incentives but risks forcing agencies to discontinue operationally important systems rather than invest in transparency infrastructure.

6.4 Definitional Inconsistencies

A persistent challenge in federal AI transparency is the lack of consistent definitions across agencies and policy instruments. The CDT's analysis of 2024 federal AI inventories found that agencies used materially different definitions of "high-impact AI," resulting in incomparable disclosures that frustrated cross-agency oversight.[65] OMB M-25-21 introduced a more precise definition AI where the output "serves as a principal basis for decisions or actions with legal, material, binding, or significant effect" on enumerated categories including civil rights, housing, health and safety, and critical infrastructure but implementation of this definition remains uneven.[66] NIST's forthcoming standard on documentation of AI datasets and models (the Zero Draft initiative launched September 2025) offers an opportunity to establish interoperable definitional standards that could resolve these inconsistencies.[67]

Figure 4: Implementation Challenge Assessment for Federal AI Transparency Framework

Figure 4: Radar chart depicting current implementation maturity (red polygon) versus required maturity (blue dashed) across six key challenge dimensions for the proposed federal AI transparency framework. Gap areas highlight investment priorities.

7. POLICY RECOMMENDATIONS

7.1 Recommendations for Congress

Congress should enact a Federal AI Transparency Act that codifies the five-tier framework into statute, moving federal AI model card and provenance requirements from the realm of executive memoranda which can be rescinded overnight into durable legal obligations. The Act should: (1) require GAO to develop and maintain government-wide standards for AI model card fields and data provenance documentation, updated on a triennial basis; (2) mandate that the IG community develop shared AI auditing competencies, including dedicated AI audit teams at major agencies; (3) authorize funding for the NIST AI Standards program to accelerate completion of the model card and dataset documentation standards currently in Zero Draft development; and (4) create a statutory right for the public to access redacted model card summaries for all Tier 3 and above federal AI deployments through a centralized portal at AI.gov.[68]

7.2 Recommendations for OMB

OMB should issue a follow-on memorandum to M-25-21 that: (1) adopts the standardized eleven-field model card template as the mandatory documentation format for all high-impact AI procurement; (2) requires CFO Act agencies to publish machine-readable model cards alongside their annual AI use case inventory submissions, enabling automated cross-agency analysis; (3) establishes a GSA-hosted, government-wide model card repository accessible to cleared auditors; and (4) clarifies that the March 11, 2026 procurement policy update

deadline under M-26-04 includes adoption of the standardized model card template.[69] OMB should also issue detailed instructions on AI inventory scope as promised in M-25-21, particularly addressing the inconsistent application of "high-impact" determinations that undermined the utility of the 2024 inventories.[70]

7.3 Recommendations for NIST

NIST should prioritize completion of the Zero Draft standard on documentation of AI datasets and models, with a target for a final standard by December 2026. The standard should specify: mandatory fields for federal agency model cards stratified by risk tier, machine-readable JSON schema for automated compliance checking, interoperability requirements with international standards including ISO/IEC 42001 and EU AI Act documentation requirements, and a provenance metadata standard compatible with the NTIA's watermarking and content authentication framework.[71] NIST should also expand the NIST AI RMF with a specific Federal Government Profile that translates voluntary AI RMF guidance into concrete, implementable requirements calibrated to the five-tier framework.[72]

7.4 Recommendations for Chief AI Officers

Agency CAIOs now mandated under M-25-21 within 60 days of the memorandum's issuance should develop agency-specific AI transparency roadmaps that: (1) inventory all existing AI deployments against the five-tier criteria; (2) prioritize full model card documentation for all Tier 3 and above systems within 12 months; (3) establish relationships with the agency's Inspector General for ongoing AI audit support; (4) negotiate model card and provenance documentation requirements into all new AI procurement contracts; and (5) create an agency-wide model card repository integrated with the GSA shared services infrastructure.[73] CAIOs should also establish clear escalation paths for AI incident reporting, drawing on the incident disclosure guidance in NIST AI 600-1.[74]

Table 4. Governance Architecture for the Proposed Federal AI Transparency Standardization Framework

Institution	Primary Role	Key Responsibilities	Timeline
U.S. Congress	Statutory authority	Federal AI Transparency Act; GAO AI standards mandate; IG funding; AI.gov portal	Legislation within 18 months
Office of Management and Budget	Government-wide policy	Standardized model card template; machine-readable AI inventories; procurement policy updates	Follow-on memo within 6 months
National Institute of Standards and Technology	Technical standards	Model card and dataset documentation standard; federal AI RMF profile; provenance metadata standard	Final standard by Dec 2026
General Services Administration	Procurement and shared services	Government-wide model card repository; AI transparency contract terms; LLM procurement guardrails	Repository within ~200 days
Agency Chief AI Officers	Agency-level implementation	AI transparency roadmap; tier classification; model card production; vendor oversight	Roadmap within 90 days
Inspectors General	Independent oversight	AI audit capability development; Tier-4 audits; annual oversight reports	Framework within 12 months
Government Accountability Office	Congressional accountability	Cross-agency audits; best practices; updates to AI accountability framework	Ongoing
National Telecommunications and Information Administration / Federal Communications Commission	Disclosure standards	Federal AI disclosure rules; watermarking and provenance standards; public registry	FCC proceeding within 12 months

8. INTERNATIONAL COMPARISONS AND LESSONS

The United States is not operating in a vacuum. The European Union's AI Act, which entered into force in August 2024, establishes a risk-tiered mandatory framework requiring high-risk AI systems to maintain technical documentation including training data characteristics, model architecture, performance metrics, and human oversight measures requirements closely analogous to this article's Tier 3 and Tier 4 standards.[75] The EU framework's explicit requirement for demographic subgroup performance data is more rigorous than any current U.S. federal requirement and represents a meaningful benchmark for the standardized model card framework proposed here.[76]

The United Kingdom's approach, while less prescriptive, has pioneered the concept of an AI regulatory sandbox that allows agencies and vendors to test transparency mechanisms in controlled environments before mandatory adoption a model that could inform GSA's development of procurement transparency standards.[77] Canada's Directive on Automated Decision-Making has required algorithmic impact assessments for federal government AI systems since 2019, providing seven years of implementation experience that U.S. agencies could draw on in developing their own impact assessment and model card procedures.[78]

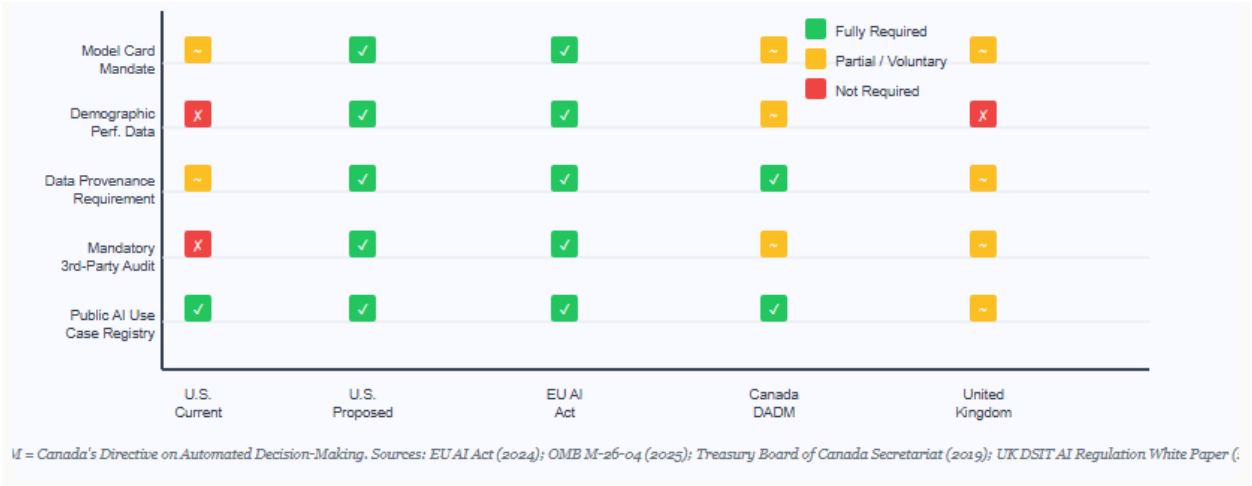

Figure 5: Comparative Federal AI Transparency Requirements U.S. vs. International Peers

Figure 5: Comparative analysis of AI transparency requirements across five jurisdictions. The proposed U.S. framework ("U.S. Proposed") would achieve full parity with EU AI Act requirements while remaining compatible with the current deregulatory policy environment through risk-proportionate implementation.

9. CONCLUSION

The United States federal government is at a critical inflection point in AI governance. With over 2,133 AI use cases now documented across federal agencies including 227 deployments that agencies themselves have identified as directly impacting citizens' rights and safety the current patchwork of transparency requirements is no longer adequate to ensure meaningful accountability for AI-driven government decisions. The December 2025 OMB guidance requiring model cards as a condition of LLM procurement is an important step, but it addresses only one slice of the federal AI portfolio and contains no standardized template, no validation mechanism, and a two-year sunset provision that creates governance uncertainty.

The five-tier standardization framework proposed in this article offers a path toward systematic, risk-proportionate transparency that can withstand changes in administration because it is grounded in enduring principles of democratic accountability: the public's right to know how consequential government decisions are made, the government's obligation to use AI systems that have been validated as fit for their stated purpose, and the oversight community's need for sufficient information to identify and remedy failures before they harm citizens.

Standardized AI model cards and data provenance auditing are not bureaucratic formalities they are the technical foundations upon which meaningful human oversight of government AI is built. In an era when AI systems are influencing determinations about veterans' benefits, refugees' placements, citizens' access to housing, and potentially the content of federal regulations themselves, the transparency infrastructure we build today will shape the trustworthiness of American governance for decades to come.

Table 5. Evidence Base Supporting the Proposed Federal AI Transparency Framework

Finding	Data / Evidence	Source	Framework Implication
Scale of federal AI deployment	2,133 use cases in 2024; ~200% increase from 2023	OMB AI inventory repository	Universal Tier-1 baseline needed for all AI systems
High-risk AI compliance gap	206 of 227 high-risk systems lacked safeguards	Federal reporting analysis	Tier-3 requirements should mandate pre-deployment certification
AI audit tooling gap	Only 12 transparency tools vs. 206 standards tools among 435 surveyed	ACM CHI research	GAO and IG investment in AI audit tooling is required
LLM procurement transparency	Model cards required for all LLM procurement; 43 agencies in federal AI contracts	OMB and federal procurement reports	Procurement policy is fastest route to model-card adoption
Inconsistent high-risk definitions	Agencies use different definitions of "high-impact AI"	Policy research reports	NIST must standardize risk-tier classification
Healthcare model-card rollback risk	Proposal to remove health-AI model card requirement (2025)	Health policy reporting	Statutory codification required
AI market growth context	Global AI market ~$638B in 2024; ~$757B in 2025	Market research	Early governance cheaper than post-deployment reform
NIST documentation standard progress	Draft AI model/dataset documentation standard released 2025	NIST standards process	Final federal documentation standard targeted for 2026

10. REFERENCES

[1] Department of Homeland Security. (2025). 2025 AI Use Case Inventory. DHS.gov. https://www.dhs.gov/ai/use-case-inventory

[2] Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I.D., & Gebru, T. (2019). Model cards for model reporting. Proceedings of the Conference on Fairness, Accountability, and Transparency (FAccT 2019), 220–229.

[3] National Institute of Standards and Technology. (2018). Risk Management Framework for Information Systems and Organizations: A System Life Cycle Approach for Security and Privacy. NIST SP 800-37 Rev. 2. U.S. Department of Commerce.

[4] Center for Democracy and Technology. (2025, April 15). Exploring the 2024 Federal AI Inventories: Key Improvements, Trends, and Continued Inconsistencies. CDT.org.

[5] MeriTalk. (2025). Feds Double AI Use Cases in 2024 to 1,700. MeriTalk.com.

[6] Office of Management and Budget. (2025, April 3). Accelerating Federal Use of AI through Innovation, Governance, and Public Trust. OMB Memorandum M-25-21. Executive Office of the President.

[7] Office of Management and Budget. (2025, December 11). Increasing Public Trust in Artificial Intelligence Through Unbiased AI Principles. OMB Memorandum M-26-04. Executive Office of the President.

[8] Ross, C. (2025, December 22). Trump administration to scrap federal rule requiring transparency into health AI tools. STAT News.

[9] Executive Office of the President. (2020, December 3). Executive Order 13960: Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government. Federal Register.

[10] Congress of the United States. (2022). Advancing American AI Act. Pub. L. 117-263, § 7225.

[11] Executive Office of the President. (2023, October 30). Executive Order 14110: Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. Federal Register.

[12] Office of Management and Budget. (2024, March 28). Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence. OMB Memorandum M-24-10. Executive Office of the President.

[13] Executive Office of the President. (2025, January 23). Executive Order 14179: Removing Barriers to American Leadership in Artificial Intelligence. Federal Register.

[14] Epstein Becker Green. (2025). New Federal Agency Policies and Protocols for Artificial Intelligence Utilization and Procurement. Workforce Bulletin.

[15] Lawfare Media. (2025, December 12). OMB releases guidance on Trump's 'Woke AI' executive order. Lawfare.

[16] FedScoop. (2025, December 11). OMB lays out requirements for agencies to prevent 'woke AI.' FedScoop.

[17] National Institute of Standards and Technology. (2023, January). Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1. U.S. Department of Commerce.

[18] National Institute of Standards and Technology. (2024, July). Generative AI Profile: NIST AI 600-1. U.S. Department of Commerce.

[19] IS Partners LLC. (2025, November 11). NIST AI RMF 2025 Updates: What You Need to Know About the Latest Framework Changes. ispartnersllc.com.

[20] National Institute of Standards and Technology. (2025, March). Adversarial Machine Learning: A Taxonomy and Terminology of Attacks and Mitigations. NIST AI 100-2e2025. U.S. Department of Commerce.

[21] National Institute of Standards and Technology. (2025). NIST Webinar: Security control overlays for AI systems. NIST CSRC Updates. csrc.nist.gov.

[22] National Institute of Standards and Technology. (2025, September). NIST's AI Standards Zero Drafts Pilot Project. NIST.gov.

[23] Office of the National Coordinator for Health Information Technology. (2023). Health Data, Technology, and Interoperability: Certification Program Updates, Algorithm Transparency, and Information Sharing (HTI-1). HHS/ONC Final Rule.

[24] Healthcare Dive. (2025, December 23). Trump administration nixes Biden-era health IT policies, including AI 'model cards.' Healthcare Dive.

[25] Olsen, E. (2025, December 22). HHS seeks input on speeding AI adoption in clinical care. Healthcare Dive.

[26] Wiley Law. (2025). White House issues executive order to promote national AI policy framework. Wiley Law Alert.

[27] Majety, S. (2025, October 12). AI model transparency in action: Model cards & NIST's Map-Measure-Manage-Govern framework. Medium / LinkedIn.

[28] Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., Chaudhary, V., Young, M., Crespo, J-F., & Dennison, D. (2015). Hidden technical debt in machine learning systems. Advances in Neural Information Processing Systems 28. NeurIPS.

[29] Government Accountability Office. (2021, June). Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities. GAO-21-519SP. GAO.

[30] Government Accountability Office. (2021). Standards for Internal Control in the Federal Government. GAO-14-704G. GAO.

[31] Arrieta, J. (2025, December 15). Expert commentary on OMB M-26-04. In Federal News Network: OMB sets procurement guardrails for buying AI tools. FederalNewsNetwork.com.

[32] Clarifai. (2026, January 8). AI Risk Management Frameworks & Strategies for Enterprises. Clarifai.com.

[33] ISACA. (2025, September 19). Collaboration and the new triad of AI governance. ISACA Industry News.

[34] Government Accountability Office. (2021, June). AI Accountability Framework. GAO-21-519SP. See also audit procedures sections, pp. 14–47.

[35] FedScoop. (2024, December 18). Federal government discloses more than 1,700 AI use cases. FedScoop.

[36] Sidhpurwala, H., et al. (2025). Blueprints of trust: AI system cards for enterprise deployment. arXiv preprint arXiv:2509.20394.

[37] Meta AI. (n.d.). System-level transparency of machine learning. Meta AI Research. ai.meta.com.

[38] Office of Management and Budget. (2025, December 11). M-26-04, Section C: Enhanced LLM transparency requirements. Executive Office of the President.

[39] Schellaars, T., Haaker, T., Gordijn, J., & Kartseva, V. (2022). Establishing data provenance for responsible artificial intelligence systems. ACM Transactions on Management Information Systems, 13(2), 1–27. https://doi.org/10.1145/3503488

[40] Office of Management and Budget. (2025, April 3). Driving Efficient Acquisition of Artificial Intelligence in Government. OMB Memorandum M-25-22. Executive Office of the President.

[41] Executive Office of Equal Employment Opportunity Commission. (2023). The Americans with Disabilities Act and the Use of Software, Algorithms, and Artificial Intelligence to Assess Job Applicants and Employees. EEOC Technical Assistance Guidance.

[42] D'Amour, A., Heller, K., Moldovan, D., et al. (2022). Underspecification presents challenges for credibility in modern machine learning. Journal of Machine Learning Research, 23(1), 1–61.

[43] Office of Management and Budget. (2025). M-26-04, Section D: Ideological neutrality controls. Executive Office of the President.

[44] National Institute of Standards and Technology. (2024). NIST AI 600-1: Generative AI Profile, Appendix A: Third-party modifications and supply chain risk. U.S. Department of Commerce.

[45] National Institute of Standards and Technology. (2024). NIST AI 600-1, Section on Incident Disclosure. U.S. Department of Commerce.

[46] Office of Management and Budget. (2025, April 3). M-25-21, Section 4: Chief AI Officer roles and responsibilities. Executive Office of the President.

[47] National Institute of Standards and Technology. (2018). NIST SP 800-37 Rev. 2, Glossary: Provenance, p. 104. U.S. Department of Commerce.

[48] National Telecommunications and Information Administration. (2023). AI Output Disclosures: Use, Provenance, Adverse Incidents. NTIA AI Accountability Policy Report. NTIA.gov.

[49] Farley, R., & Lansang, M. (2025). AI auditing: First steps towards the effective regulation of data collection, provenance, and management. Harvard Journal of Law & Technology. jolt.law.harvard.edu.

[50] National Telecommunications and Information Administration. (2023). AI Accountability Policy Report, Section on Provenance and Authentication. NTIA.gov.

[51] National Institute of Standards and Technology. (2024). Reducing Risks Posed by Synthetic Content. NIST AI 100-4. U.S. Department of Commerce. (See KPMG Regulatory Insights summary, 2024.)

[52] Office of Management and Budget. (2025, April 3). M-25-22, Use of Government Data section. Executive Office of the President.

[53] Federal Reserve Board. (2024). AI Use Case Inventory Section 3: Data Ownership and Documentation. FederalReserve.gov.

[54] Lam, K., Metaxa, D., et al. (2025). Towards AI accountability infrastructure: Gaps and opportunities in AI audit tooling. Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems. ACM. https://doi.org/10.1145/3706598.3713301

[55] Gerchick, M., et al. (2025). Auditing the audits: Lessons for algorithmic accountability from Local Law 144's bias audits. Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency (FAccT 2025).

[56] Mintz. (2025, October 20). AI accountability and governance in focus: Complaint filed in DC District Court for AI transparency. AI: The Washington Report. Mintz.com.

[57] Democracy Forward Foundation. (2025, October 1). Complaint re: FOIA Requests on AI Use in Federal Rulemaking. U.S. District Court for the District of Columbia.

[58] Office of Management and Budget. (2025). M-25-22, Section on vendor documentation and contract terms. Executive Office of the President.

[59] Cen, S.H., & Shah, N.B. (2022). From transparency to accountability and back: A discussion of access and evidence in AI auditing. Proceedings of the 4th ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization. https://doi.org/10.1145/3689904.3694711

[60] Office of Management and Budget. (2025, April 3). M-25-21, AI workforce development section. Executive Office of the President.

[61] Lam, K., et al. (2025). Towards AI accountability infrastructure. ACM CHI 2025, Section 5: Conclusions and future directions.

[62] Hunton Andrews Kurth. (2025). OMB issues revised policies on AI use and procurement by federal agencies. Privacy and Information Security Law Blog. Hunton.com.

[63] Cen, S.H., & Shah, N.B. (2022). From transparency to accountability and back. ACM EAAMO 2022.

[64] Department of Homeland Security. (2025). 2025 DHS AI Use Case Inventory. DHS.gov.

[65] Center for Democracy and Technology. (2025). 2024 Federal AI Inventories. Section: Lack of consistency. CDT.org.

[66] Office of Management and Budget. (2025, April 3). M-25-21, Section 5: Definition of high-impact AI. Executive Office of the President.

[67] National Institute of Standards and Technology. (2025, September). NIST AI Standards Zero Drafts: Extended Outline for Dataset and Model Documentation Standard. NIST.gov.

[68] Office of the CIO.gov. (2024). Executive Order 13960 AI Use Case Inventories Reference. CIO.gov.

[69] Federal News Network. (2025, December 15). OMB sets procurement guardrails for buying AI tools. Federal News Network.

[70] Office of Management and Budget. (2025). M-25-21, Annual inventory and publicly published AI use case requirements. Executive Office of the President.

[71] National Institute of Standards and Technology. (2025). Plan for Global Engagement on AI Standards. NIST AI 100-5. U.S. Department of Commerce.

[72] National Institute of Standards and Technology. (2025, March). AI RMF Update: Federal Government Profile (Draft). NIST AI RMF Playbook. nistairc.nist.gov.

[73] Office of Management and Budget. (2025, April 3). M-25-21, Chief AI Officer designation and AI Governance Board. Executive Office of the President.

[74] National Institute of Standards and Technology. (2024). NIST AI 600-1: Generative AI Profile, Incident Disclosure section. U.S. Department of Commerce.

[75] European Parliament and Council of the European Union. (2024, August). Regulation (EU) 2024/1689: Artificial Intelligence Act. Official Journal of the European Union.

[76] European Union Agency for Fundamental Rights. (2024). Fundamental Rights and the Rule of Law in the EU AI Act. FRA.europa.eu.

[77] UK Department for Science, Innovation and Technology. (2023). A Pro-Innovation Approach to AI Regulation. AI Regulation White Paper. DSIT/DCMS, UK Government.

[78] Treasury Board of Canada Secretariat. (2019, updated 2023). Directive on Automated Decision-Making. Canada.ca.

[79] Agumagu, E. R., Ekine, E. U., Onyekwuluje, T. P., & Uwaezuoke, E. C. (2024). The Role of Automation in Integrating Various Telecommunication Projects. *World Journal of Innovation and Modern Technology*, *8*(6). DOI:10.56201/wjimt.v8.no6.2024.pg217.228

[80] Agumagu, E. R., Ekine, E. U., &Uwaezuoke, E. (2024). The Challenges of Managing International Telecom Projects Across Diverse Cultural Contexts and How to Overcome Them. *International Journal of Social Sciences and Management Research E-ISSN 2545-5303* , *10*(11), 513–524.
DOI:10.56201/ijssmr.v10.no11.2024.pg.513.524

[81] Agumagu, E. R., Ekine, E. U., &Uwazuoke, E. (2024). The Challenges of Managing International Telecom Projects Across Diverse Cultural Contexts and How to Overcome Them. *International Journal of Social Sciences and Management Research*, *10*(11).