DRAFT LEGISLATION

Government Algorithmic Accountability

and AI Fiduciary Act

GAAFA *Companion statute to the Verifiable Identity and Digital Autonomy Act (VIDA) and the Personal Data Trusteeship Act (PDTA)* FOR the purpose of governing government use of artificial intelligence and automated decision systems; requiring algorithmic deployment assessments before rights-affecting AI deployment; establishing the right to algorithmic explanation; classifying autonomous AI agents as statutory secondary fiduciaries; mandating interpretable models for high-stakes government decisions; requiring algorithmic impact assessments with binding remediation; establishing technical AI oversight capacity; and generally relating to government accountability in artificial intelligence.

Prepared by Michael G. Leahy

Initial Draft | March 2026

Be it enacted by the [Legislative Body] as follows:

Section 1. Short Title

(a) This Act shall be known and may be cited as the “Government Algorithmic Accountability and AI Fiduciary Act” or the “GAAFA.”

Section 2. Legislative Findings and Purpose

(a) Findings

The [Legislative Body] finds that:

(1) Artificial intelligence systems, including large language models, automated decision systems, and autonomous AI agents, are being deployed by government agencies in contexts that directly affect citizens’ rights, benefits, liberties, and opportunities, without adequate frameworks governing the accountability, interpretability, or fiduciary obligations of such systems;

(2) Existing statutes, including the Verifiable Identity and Digital Autonomy Act (VIDA) and the Personal Data Trusteeship Act (PDTA), establish the architectural and legal foundations for government’s fiduciary relationship with citizen data and identity, but do not fully address the specific risks arising when government uses AI systems to make or substantially influence decisions affecting citizens;

(3) Artificial intelligence systems used in government decision-making are frequently opaque, in that neither the citizens affected by their outputs nor the government officials nominally responsible for their operation can readily ascertain how any specific outcome was reached, what inputs most influenced it, or whether it reflects bias, error, or unlawful discrimination;

(4) The opacity of government AI systems constitutes a structural threat to due process, equal protection, and the transparency duties derived from the Fourth Amendment's particularity requirement, as developed in the companion scholarship underlying VIDA and PDTA;

(5) Autonomous AI agents deployed by government agencies to query, process, or act upon citizen data exercise fiduciary power over citizens' fundamental interests and must therefore be governed by the same fiduciary obligations that PDTA imposes on human secondary fiduciaries;

(6) The European Union's AI Act, the world's most ambitious attempt at comprehensive AI governance, has been materially delayed and weakened because it relied on harmonized technical standards that were not developed in time and on enforcement authorities that were not constituted by statutory deadlines, demonstrating that durable AI governance must be grounded in constitutional principles rather than technical specifications;

(7) Citizens have a due process right to understand the basis of adverse government decisions, including decisions made or substantially influenced by automated systems, and to have those decisions reviewed by accountable human officials who possess the ability to exercise genuine, informed oversight;

(8) The deployment of black-box AI models in rights-affecting government contexts, where the model's decisional logic cannot be explained to the citizens it affects or the officials responsible for it, constitutes a breach of the fiduciary duties of loyalty, care, and transparency already imposed on government agencies by PDTA;

(9) Government AI systems that process citizen data within architecturally compliant infrastructure under VIDA can still generate unlawful surveillance profiles, discriminatory outcomes, and unchallengeable adverse decisions unless statutory requirements govern the decisional layer specifically; and

(10) A complete framework for constitutional digital governance requires three complementary statutes: VIDA governing the technical architecture of government data systems; PDTA governing the legal fiduciary relationship between government and citizens regarding data; and this Act governing the use of AI within that architecture and under those obligations.

(b) Purpose

The purpose of this Act is to:

(1) Require government agencies to demonstrate algorithmic understanding before deploying AI or automated decision systems in rights-affecting contexts, through mandatory Algorithmic Deployment Assessments filed with the Data Fiduciary Oversight Office established under VIDA Section 10(c);

(2) Establish the right of every citizen to receive a specific, intelligible explanation of any adverse government decision made or substantially influenced by an automated system, enforceable through the private right of action provided by PDTA Section 8;

(3) Classify autonomous AI agents deployed by government agencies as statutory secondary fiduciaries subject to the full fiduciary duty regime of PDTA, with the deploying agency jointly and severally liable for fiduciary breaches by agents it deploys;

(4) Establish an interpretable model preference rule requiring government agencies to use transparent, auditable decision logic for rights-affecting determinations, with black-box alternatives permitted only upon demonstrated necessity and subject to enhanced oversight and remedies;

(5) Give algorithmic impact assessments binding legal force, with mandatory remediation timelines, suspension of non-compliant functions, and reopener rights for citizens adversely affected during periods of identified bias or systematic error;

(6) Mandate technical AI competency within the oversight bodies established by VIDA and PDTA, ensuring those bodies can audit algorithmic systems and not merely data handling practices; and

(7) Establish a framework that governs government AI regardless of the specific technology deployed, grounded in constitutional fiduciary principles that remain applicable as technology evolves.

Section 3. Definitions

In this Act, the following terms have the meanings indicated. Capitalized terms defined in VIDA or PDTA have the same meanings as in those Acts unless this Act provides otherwise.

(a) "Adverse Automated Decision" means any decision, determination, denial, suspension, termination, reduction, assignment, classification, or other action taken by or on behalf of a government agency that (i) negatively affects a citizen's rights, benefits, entitlements, licenses, eligibility, or legal status; (ii) was made or substantially influenced by an Automated Decision System; and (iii) was not reviewed and affirmatively approved by an Accountable Human Official prior to communication to the affected citizen, except where this Act expressly permits automated issuance at Tier 2 or Tier 3 levels.

(b) "Agentic AI System" means an Automated Decision System that operates with a degree of autonomy to pursue goals, execute multi-step tasks, query or process data, take actions with external effect, or generate outputs that the system itself determines, without requiring explicit human direction for each individual step, including systems commonly described as AI agents, autonomous agents, or agentic AI.

(c) "Algorithmic Deployment Assessment" or "ADA" means the pre-deployment documentation required by Section 5 of this Act, in which a government agency demonstrates its understanding of an AI or Automated Decision System's decision logic, training data, known failure modes, and compliance with this Act before deploying that system in a rights-affecting context.

(d) ”Algorithmic Impact Assessment” or “AIA” means the ongoing evaluation required by Section 9 of this Act, in which a government agency systematically tests an Automated Decision System for bias, disparate impact, accuracy, and fiduciary compliance throughout its operational lifecycle.

(e) ”Automated Decision System” or “ADS” means any system, software, algorithm, model, or process, including systems derived from machine learning, statistical inference, neural networks, large language models, or other artificial intelligence techniques, that uses computation to make, substantially influence, support, or produce outputs that serve as the basis for government decisions or determinations affecting citizens’ rights, benefits, or legal status. The term includes systems in which a human nominally reviews an automated output but lacks the information or practical ability to exercise genuine independent judgment about that output.

(f) ”Accountable Human Official” means a specific, identified government employee or officer who (i) possesses the authority under applicable law to make or approve the decision in question; (ii) has reviewed the specific case and the automated output pertaining to it; (iii) understands the principal factors that produced the automated output; and (iv) has affirmatively approved the decision based on that understanding. A human official who approves an automated output without access to the system’s decision rationale, or who approves outputs in batches without individual case review, does not satisfy this definition.

(g) ”Black-Box Model” means an Automated Decision System whose decision logic, including the relative weight of inputs, the rules or patterns used to generate outputs, and the basis for any specific output, cannot be expressed in terms that are intelligible to a person with reasonable intelligence and without specialized technical expertise in machine learning. A model is presumptively a Black-Box Model if the deploying agency cannot generate an explanation of any specific output that satisfies the requirements of Section 6(b) of this Act.

(h) ”Disparate Impact” means a statistically significant difference in the rate of adverse outcomes produced by an Automated Decision System across groups defined by race, color, national origin, sex, disability, age, religion, sexual orientation, gender identity, or any other characteristic protected under applicable state or federal civil rights law, where that difference is not justified by a legitimate, non-discriminatory, and proportionate operational necessity.

(i) ”Interpretable Model” means an Automated Decision System whose decision logic can be expressed in terms that a person with reasonable intelligence and without specialized technical expertise can understand, including decision trees, rule-based systems, linear or logistic regression with documented features, scorecard models, and other approaches in which the factors contributing to any specific output can be identified and explained in plain language. Whether a model satisfies this definition is assessed at the level of individual decision explanation, not aggregate model description.

(j) ”Rights-Affecting Decision” means any government decision or determination that affects a citizen’s access to benefits, services, licenses, permits, eligibility determinations, enforcement actions, law enforcement encounters, housing, employment with a government entity, immigration status, voting eligibility, educational access, healthcare access, child welfare determinations, or any other matter in which the government exercises discretionary authority over a citizen's fundamental interests or legal rights.

(k) "Tier 1 AI Use" means the deployment of an Automated Decision System to make or substantially influence a Rights-Affecting Decision, as further defined in Section 7(a) of this Act.

(l) "Tier 2 AI Use" means the deployment of an Automated Decision System for internal administrative or operational functions that do not directly produce Rights-Affecting Decisions, as further defined in Section 7(b) of this Act.

(m) "Tier 3 AI Use" means the deployment of an Automated Decision System to provide citizens with information, navigation assistance, or general guidance in connection with government services, as further defined in Section 7(c) of this Act.

(n) "Training Data" means the dataset or datasets used to train, fine-tune, or adapt an Automated Decision System, including any data used to calibrate, validate, or test the system prior to deployment.

(o) "Vendor" means any private entity, contractor, subcontractor, or licensee that develops, provides, operates, maintains, or hosts an Automated Decision System used by a government agency, including entities that provide AI capabilities as a service, platform, or application programming interface.

(p) "VIDA" means the Verifiable Identity and Digital Autonomy Act, and references to specific sections of VIDA refer to the sections of that Act as enacted and subsequently amended.

(q) "PDTA" means the Personal Data Trusteeship Act, and references to specific sections of PDTA refer to the sections of that Act as enacted and subsequently amended.

Section 4. Constitutional and Fiduciary Foundations

Cross-references: PDTA Sec. 4 (Establishment of Personal Data Trusteeship); PDTA Sec. 4A (Bilateral Trust Structure); PDTA Sec. 7 (Government Duties); VIDA Sec. 2 (Legislative Findings and Purpose); VIDA Sec. 3A (Prohibition on Surveillance Architectures); VIDA Sec. 4A (Protections Against AI-Enabled Threats)

(a) General Principle. The fiduciary duties imposed on government agencies by PDTA, including the duties of loyalty, care, confidentiality, and transparency, apply with full force to all Automated Decision Systems deployed by those agencies, regardless of whether those systems are operated by agency personnel or by Vendors, and regardless of whether the system's outputs are characterized as recommendations, scores, risk assessments, flags, eligibility determinations, or any other term. No classification of an AI system's output as advisory, supplementary, or non-binding relieves the deploying agency of its fiduciary obligations under PDTA and this Act with respect to citizens affected by that output.

(b) Duty of Loyalty in AI Deployment. Consistent with PDTA Section 6(a)'s requirement that secondary fiduciaries act in the best interests of the data subject trustee without opportunism, a government agency deploying an Automated Decision System fulfills its duty of loyalty only if:

(1) The system is deployed solely for the specific, authorized purpose for which citizen data was collected, consistent with PDTA Section 4A(c)'s prohibition on repurposing beyond the original justified purpose;

(2) The system does not use citizen data to generate inferences, profiles, or risk scores that exceed the scope of the original data collection purpose, consistent with VIDA Section 8's prohibition on centralized multipurpose data uses; and

(3) The agency has not permitted any Vendor to retain, use, or derive value from citizen data processed through the system beyond the scope of the agency's authorized purpose.

(c) Duty of Care in AI Deployment. Consistent with PDTA Section 6(a)'s requirement that secondary fiduciaries exercise reasonable care to protect and manage data accurately and securely, a government agency deploying an Automated Decision System fulfills its duty of care only if:

(1) The agency has completed an Algorithmic Deployment Assessment under Section 5 of this Act before deployment;

(2) The agency conducts ongoing Algorithmic Impact Assessments under Section 9 of this Act throughout the system's operational lifecycle;

(3) Accountable Human Officials exercise genuine, informed oversight of Tier 1 AI Uses as required by Section 7(a) of this Act; and

(4) The agency maintains the technical competency required by Section 10 of this Act to monitor, audit, and override the system's operations.

(d) Transparency Duty in AI Deployment. Consistent with the transparency duty derived from the Fourth Amendment's warrant requirement as developed in the scholarship underlying VIDA and PDTA, and with VIDA Section 10(a)'s enumeration of individual rights, a government agency deploying an Automated Decision System fulfills its transparency duty only if:

(1) The agency provides the right to algorithmic explanation required by Section 6 of this Act to any citizen who receives an Adverse Automated Decision;

(2) The agency publicly registers the system in the transparency registry maintained by the Data Fiduciary Oversight Office under VIDA Section 10(c)(2); and

(3) The agency maintains the immutable audit trail required by PDTA Section 6A(c), extended by this Act to include a complete record of all Automated Decision System outputs and the specific inputs that produced them.

(e) Confidentiality Duty in AI Deployment. Consistent with PDTA Section 6(a)'s prohibition on unauthorized disclosure or secondary use of personal data, a government agency deploying an Automated Decision System fulfills its confidentiality duty only if the system does not enable Vendor access to citizen data beyond what is strictly necessary for the contracted function, and the agency's contract with any Vendor expressly prohibits the Vendor from using citizen data to train, fine-tune, or improve any model beyond the scope of the specific contracted deployment.

Section 4A. Algorithmic inference as identity construction

Cross-references: GAAFA Sec. 4 (Constitutional and Fiduciary Foundations); VIDA Sec. 4B (Digital likeness and synthetic identity protections); PDTA Sec. 4 (Establishment of Personal Data Trusteeship)

(a) Algorithmic inference, meaning the use of machine learning, statistical modeling, pattern recognition, or any other automated method to derive personal attributes, characteristics, conditions, or classifications about an identifiable individual from data that does not explicitly contain those attributes, is hereby recognized as a form of digital likeness construction within the meaning of VIDA Section 4B.

(b) For purposes of this Act, inferred attributes include but are not limited to:

(1) Health conditions, diagnoses, or medical risks;

(2) Sexual orientation or gender identity;

(3) Political affiliation, beliefs, or persuadability;

(4) Emotional or psychological states;

(5) Religious practice or belief;

(6) Pregnancy or reproductive status;

(7) Addiction or substance use;

(8) Financial vulnerability or creditworthiness; and

(9) Racial or ethnic classification.

(c) No government entity shall deploy or use an Automated Decision System that generates inferred attributes about identifiable individuals unless:

(1) The specific categories of inference are disclosed in the Algorithmic Deployment Assessment required by Section 5 of this Act;

(2) The inference is strictly necessary for the specific, authorized purpose for which the underlying data was collected, consistent with PDTA Section 4A(c);

(3) The individual has been notified that inferences will be drawn and has been provided the right to inspect, correct, and contest inferred attributes through the mechanisms established by Section 6 of this Act; and

(4) The deploying agency can demonstrate, through the audit trail required by PDTA Section 6A(c), the specific data inputs that produced each inference and the statistical confidence of the inference.

(d) Inferred attributes shall not be treated as established facts for any purpose under this Act, PDTA, or VIDA. No Rights-Affecting Decision shall be based solely or predominantly on inferred attributes without independent verification of the inferred characteristic through means other than algorithmic inference.

Section 4B. Government use of algorithmically processed commercial data

Cross-references: GAAFA Sec. 4 (Constitutional and Fiduciary Foundations); PDTA Sec. 6B (Prohibition on government acquisition of commercially compiled surveillance data); VIDA Sec. 4B (Digital likeness and synthetic identity protections)

(a) Where a government entity uses, or proposes to use, algorithmic outputs, scores, classifications, predictions, or other products derived from commercially compiled personal data, both the constitutional constraints on government data acquisition established by PDTA Section 6B and the algorithmic accountability requirements of this Act apply in full.

(b) The commercial origin of the underlying data does not diminish, modify, or excuse compliance with:

(1) The Algorithmic Deployment Assessment requirements of Section 5;

(2) The right to algorithmic explanation established by Section 6;

(3) The tiered oversight requirements of Section 7;

(4) The audit trail requirements of PDTA Section 6A(c), extended by this Act; and

(5) The individual's right to inspect, correct, and contest under Section 6 of this Act and PDTA Section 5(b).

(c) A government entity may not use commercially derived algorithmic outputs as a substitute for the entity's own independent assessment when making Rights-Affecting Decisions. A risk score, creditworthiness assessment, fraud prediction, or behavioral classification produced by a commercial algorithmic system and purchased or licensed by a government entity is subject to the same validation, audit, and accountability requirements as an Automated Decision System deployed directly by the entity.

Section 5. Algorithmic Deployment Assessment

Cross-references: VIDA Sec. 4A(b) (Government shall not use AI for automated surveillance or profiling without probable cause and judicial oversight); VIDA Sec. 4A(c) (Post-enactment regulations shall address emerging AI applications); PDTA Sec. 6A (Architectural Requirements *for Fiduciary Compliance); PDTA Sec. 7(b)(4) (Government shall require impact assessments for any data initiatives)*

(a) Requirement. No government agency may deploy an Automated Decision System in a Tier 1 AI Use, as defined in Section 7(a), unless the agency has completed and filed an Algorithmic Deployment Assessment with the Data Fiduciary Oversight Office at least [60] days prior to deployment. Deployment of an ADS without a completed and approved ADA, or deployment that materially deviates from a filed ADA, constitutes a per se breach of the duty of care under PDTA Section 6(a) and triggers the enforcement mechanisms of Section 11 of this Act.

(b) Required Contents. An Algorithmic Deployment Assessment shall document, at a minimum:

(1) System Description. A plain-language description of the system, its intended function, the decisions it will make or influence, the population of citizens it will affect, and the legal authority under which it is deployed;

(2) Decision Logic. An explanation, accessible to a person of reasonable intelligence without technical expertise, of how the system reaches its outputs, including the principal input factors, the relative weight or influence of those factors, and the range of outputs the system can produce;

(3) Training Data Documentation. A description of the Training Data used to develop the system, including the sources of that data, the time period it covers, known gaps or underrepresentation of populations that will be affected by the system, any known biases in the Training Data, and the steps taken to identify and mitigate those biases, consistent with PDTA Section 6A's requirement for auditable data provenance;

(4) Accuracy and Error Rate Analysis. Documented testing results showing the system's accuracy, false positive rate, false negative rate, and error rates disaggregated by race, sex, national origin, disability, age, and any other characteristic protected under applicable civil rights law;

(5) Disparate Impact Analysis. An analysis of whether the system produces Disparate Impact on any protected class, including the methodology used to test for disparate impact and the agency's assessment of whether any identified disparate impact is justified by legitimate operational necessity;

(6) Red-Team Testing Results. Documentation of adversarial testing conducted before deployment, including testing with inputs designed to probe failure modes, edge cases, and manipulated inputs representative of the citizen population the system will serve;

(7) Human Oversight Architecture. A description of how Accountable Human Officials will exercise oversight of the system's outputs, including the specific review process for Adverse Automated Decisions, the information that will be available to reviewing officials, and the criteria under which they are expected to override the system;

(8) Interpretability Assessment. A determination of whether the system is an Interpretable Model or a Black-Box Model under Section 3(g) and 3(i), with supporting justification, and if the system is a Black-Box Model, a demonstration that no Interpretable Model alternative achieves the required function at acceptable accuracy;

(9) VIDA Compliance Certification. Certification that the system operates within VIDA-compliant technical architecture, including Zero Trust principles under VIDA Section 7(a)(5), purpose-sequestered databases under VIDA Section 8, and prohibition on phone-home surveillance under VIDA Section 3(6); and

(10) PDTA Compliance Certification. Certification that the system's deployment complies with the fiduciary duties of loyalty, care, confidentiality, and transparency under PDTA Section 6(a) and Section 4 of this Act, including the purpose-limitation and audit-trail requirements of PDTA Section 6A.

(c) Review and Approval. The Data Fiduciary Oversight Office shall review each filed ADA within [45] days of filing and shall:

(1) Approve the ADA if it satisfies the requirements of subsection (b) and the agency's proposed deployment is consistent with this Act, VIDA, and PDTA;

(2) Return the ADA with written deficiency findings if it is materially incomplete or if the proposed deployment would violate this Act, VIDA, or PDTA, specifying the deficiencies to be remediated before resubmission; or

(3) Refer the ADA to the joint AI Audit and Review Panel established under Section 10(d) for independent technical review if the system is novel, high-risk, or raises questions outside the routine competency of the Oversight Office.

(d) Public Transparency. All filed ADAs shall be published in the transparency registry maintained by the Data Fiduciary Oversight Office under VIDA Section 10(c)(2), with redactions limited to information the agency demonstrates would create a specific, documented security vulnerability if disclosed. The existence of an ADA, the system it covers, and the nature of its deployment may not be redacted.

(e) Material Changes. An agency that makes a material change to a deployed Automated Decision System, including retraining on new data, modifying the model architecture, changing the decision threshold, or expanding the use to new populations or decision types, must file an updated ADA and obtain approval before the modified system is placed into operation.

(f) Tier 2 and Tier 3 Systems. For Tier 2 and Tier 3 AI Uses as defined in Section 7, agencies shall complete an abbreviated Deployment Notice, filed with the Oversight Office, documenting the system's function, its VIDA and PDTA compliance, and the safeguards in place to prevent drift toward Tier 1 functions. Abbreviated Deployment Notices are subject to the public transparency requirement of subsection (d).

Section 6. Right to Algorithmic Explanation

Cross-references: VIDA Sec. 10(a) (Individual rights including right to know what data government holds, to access and correct it, and to use digital services without surveillance); PDTA Sec. 5(b) (Trustee right to access comprehensive logs of all data interactions and to veto or revoke access); PDTA Sec. 8 (Private Rights of Action); PDTA Sec. 9 (Remedies)

(a) Right Established. Any individual who receives an Adverse Automated Decision has the right to request and receive a complete algorithmic explanation of that decision. This right is an extension of the individual rights established by VIDA Section 10(a) and the trustee rights established by PDTA Section 5(b), applied specifically to the context of AI-mediated government decisions.

(b) Contents of Explanation. An agency that has made an Adverse Automated Decision must provide, within [30] days of a written request:

(1) A plain-language account of the principal factors that contributed to the adverse outcome, including the specific data points, scores, classifications, or attributes that most significantly influenced the system's output in the requester's individual case, expressed in terms that a person of reasonable intelligence can understand without specialized technical expertise;

(2) Identification of what information or circumstances, if different, would have produced a materially different outcome, including any specific threshold, score, or factor that the requester could address to change the determination;

(3) Identification by name and title of the Accountable Human Official who reviewed and approved the automated determination before it was communicated to the requester;

(4) A copy of the relevant portion of the immutable audit trail required by PDTA Section 6A(c) and this Act, showing the specific inputs used to process the requester's case, the intermediate outputs of the system, and the final determination;

(5) A clear statement of the administrative appeal pathway, including the right to have the decision reviewed de novo by a different Accountable Human Official who has not been involved in the automated processing of the requester's case; and

(6) A reference to the public ADA filed for the system, identifying where the requester may review the complete system description, Training Data documentation, and accuracy statistics.

(c) Inability to Explain as Presumptive Breach. If an agency cannot provide the explanation required by subsection (b) because the system's decisional logic is not accessible to the agency's own personnel, or because the system does not generate case-specific rationale, that inability is presumptive evidence that the agency has breached its duty of care under PDTA Section 6(a) and Section 4(c) of this Act. In any enforcement proceeding, the burden shifts to the agency to demonstrate that the inability to explain did not result from a failure to select an appropriately interpretable model or to maintain adequate oversight of the deployed system.

(d) Enforcement. The right to algorithmic explanation is enforceable through the private right of action established by PDTA Section 8. An agency's failure to provide a complete explanation within the timeframe required by subsection (b), or provision of an explanation that does not satisfy the requirements of subsection (b), constitutes a violation of this Act subject to the remedies provided by PDTA Section 9, including injunctive relief, compensatory damages, statutory damages of not less than [$1,000] per violation, and attorney's fees.

(e) Proactive Notice. An agency need not wait for a request before providing algorithmic explanation for Adverse Automated Decisions. Best practice, and the standard toward which agencies shall work within the implementation timeline established by Section 13, is the proactive provision of a plain-language explanation at the time any Adverse Automated Decision is communicated to the affected citizen.

Section 7. Tiered AI Use Requirements and the Interpretable Model Preference Rule

Cross-references: VIDA Sec. 7(a)(2) (Purpose-limited use: identity verification shall rely on cryptographic yes/no queries rather than raw data transmission where feasible); VIDA Sec. 8 (Data Minimization and Sequestered Databases); PDTA Sec. 6A(b) (Where verification suffices, collection of underlying data is prohibited); PDTA Sec. 6(a) (Secondary fiduciary duties of loyalty and care)

(a) Tier 1: Rights-Affecting AI Uses.

(1) Definition. Tier 1 AI Use means the deployment of any Automated Decision System to make, substantially influence, or generate the primary basis for a Rights-Affecting Decision, including decisions on benefits eligibility, license or permit grants or denials, enforcement actions, risk scoring used in law enforcement contexts, child welfare determinations, immigration-related data processing, employment decisions by government entities, healthcare eligibility, educational access, housing assistance, and any other decision in which government exercises discretionary authority over a citizen's fundamental interests.

(2) Interpretable Model Requirement. A government agency may deploy only an Interpretable Model for a Tier 1 AI Use, unless the agency demonstrates in its ADA, and the Data Fiduciary Oversight Office confirms, that no Interpretable Model alternative achieves the required function at accuracy within [10] percentage points of the best available Black-Box Model for the specific use case, and that the public interest in the function's accuracy justifies the reduction in transparency.

(3) Accountable Human Official Requirement. Every Adverse Automated Decision produced by a Tier 1 system must be reviewed and affirmatively approved by an Accountable Human Official before it is communicated to the affected citizen. Batch approval of automated outputs without individual case review does not satisfy this requirement. Agencies shall document the identity of the reviewing official and the date and nature of their review in the audit trail required by PDTA Section 6A(c).

(4) Black-Box Exception Conditions. Where the Data Fiduciary Oversight Office approves a Black-Box Model exception under paragraph (2), the following additional requirements apply:

(i) The agency must provide enhanced human review, meaning the Accountable Human Official must receive a complete case file and must document specific reasons for approving or modifying the automated output, not merely indicate approval;

(ii) The agency must conduct quarterly Disparate Impact analyses under Section 9 and provide quarterly reports to the Oversight Office;

(iii) Citizens subject to adverse decisions from Black-Box exception systems are entitled to the full algorithmic explanation under Section 6, and an agency's inability to fully explain a Black-

Box system's output in any individual case triggers de novo human review of that case as a matter of right; and

(iv) Black-Box exception approvals expire after [24] months and must be renewed through a new ADA demonstrating continued necessity.

(b) Tier 2: Administrative and Operational AI Uses.

(1) Definition. Tier 2 AI Use means the deployment of an Automated Decision System for internal administrative or operational functions that do not directly produce Rights-Affecting Decisions, including document processing and routing, scheduling, internal workflow management, resource allocation modeling, and preliminary data analysis where outputs are reviewed by personnel before influencing any citizen-affecting decision.

(2) Requirements. Tier 2 systems are subject to: the Abbreviated Deployment Notice requirement of Section 5(f); the VIDA compliance requirements of VIDA Section 8 regarding Zero Trust architecture and purpose-sequestered databases; the audit trail requirements of PDTA Section 6A(c); and the Algorithmic Impact Assessment requirements of Section 9 of this Act, conducted annually.

(3) Drift Prevention. Agencies shall implement technical and administrative controls to prevent Tier 2 systems from drifting into Tier 1 functions. Any expansion of a Tier 2 system to influence Rights-Affecting Decisions requires a full ADA under Section 5 before that expansion is implemented.

(c) Tier 3: Citizen-Facing Information AI Uses.

(1) Definition. Tier 3 AI Use means the deployment of an Automated Decision System, including large language models and AI chatbots, to provide citizens with information, navigation assistance, or general guidance in connection with government services, where the system's outputs do not directly constitute or substantially determine a Rights-Affecting Decision.

(2) Disclosure Requirement. Any government interface using a Tier 3 AI system must clearly disclose to the citizen, at the beginning of every interaction, that they are communicating with an automated AI system and not with a human government employee. This disclosure must be prominent, in plain language, and in any language in which the interface operates.

(3) Escalation Pathway. Every Tier 3 interface must provide a clearly accessible pathway for a citizen to request communication with a human government employee at any point during the interaction. Agencies may not condition receipt of government services on a citizen's willingness to interact exclusively through a Tier 3 AI interface.

(4) Accuracy Liability. An agency is liable for material errors in Tier 3 AI outputs that cause a citizen to take or forego action to their demonstrable detriment in reliance on the erroneous information. Liability under this paragraph is enforceable through the private right of action under PDTA Section 8, with a good-faith defense available to agencies that can demonstrate regular accuracy auditing and prompt correction of identified errors.

(5) LLM Reliability Safeguards. Agencies deploying large language models in Tier 3 uses shall implement safeguards against hallucination and confabulation, including retrieval-augmented generation grounded in official government sources, output confidence scoring, and automatic escalation to human personnel when query complexity exceeds the system's validated accuracy range.

Section 8. Agentic AI Systems as Statutory Secondary Fiduciaries

Cross-references: PDTA Sec. 4(c) (Secondary fiduciaries must obtain explicit, revocable authorization for any collection, processing, or sharing); PDTA Sec. 4A (Bilateral Trust Structure; government as limited secondary fiduciary); PDTA Sec. 6 (Obligations of Secondary Fiduciaries); VIDA Sec. 6(a)(3) (No mandatory use of single private platform); VIDA Sec. 7(a)(12) (Preference for endorsement over issuance); VIDA Sec. 8 (Data Minimization and Sequestered Databases)

(a) Classification. Every Agentic AI System deployed by a government agency to query, process, access, or act upon citizen personal data is hereby classified as a Secondary Fiduciary under PDTA Section 3(c) for all purposes of PDTA and this Act. The deploying agency is jointly and severally liable for fiduciary breaches committed by Agentic AI Systems it deploys. No defense based on the automated nature of the action, the absence of human direction for the specific act, or the involvement of a Vendor in the system's operation is available to a deploying agency in an enforcement proceeding under this Act or PDTA.

(b) Authorization Requirements. An Agentic AI System deployed by a government agency may access or process citizen personal data only if:

(1) The specific type of data access or processing was disclosed to the citizen in the agency's commitment under VIDA Section 4(e)(1)(v) at the time of original data collection, or the citizen has provided explicit, revocable authorization for the specific agentic action consistent with PDTA Section 4(c);

(2) The access or processing is strictly necessary for the specific, authorized purpose for which the data was collected, consistent with PDTA Section 4A(c)'s prohibition on repurposing beyond the original justified purpose;

(3) The agentic action is logged in the immutable audit trail required by PDTA Section 6A(c) and Section 8(c) of this Act before it is executed; and

(4) The system operates within the VIDA-compliant technical architecture, including the Zero Trust requirements of VIDA Section 7(a)(5) and the purpose-sequestered database requirements of VIDA Section 8.

(c) Agentic Audit Trail Requirements. In addition to the audit trail requirements of PDTA Section 6A(c), agencies deploying Agentic AI Systems shall maintain a complete, tamper-evident log of:

(1) Every data query or access made by the agent, including the specific data elements accessed, the timestamp of access, and the stated purpose for the access;

(2) Every action taken by the agent with external effect, including any credential verification, determination, notification, or communication generated by the agent;

(3) The authorization basis for each action, traceable to the specific citizen authorization or agency legal authority that justified it; and

(4) The identity of the Accountable Human Official responsible for the agent's operation during the period in which each logged action occurred.

(d) Citizen Access to Agentic Audit Trails. Citizens have the right to access the complete agentic audit trail generated in connection with their own data, consistent with PDTA Section 5(b)(1)'s trustee right to access comprehensive logs of all data interactions. Access shall be provided through the secure individual-controlled portal required by PDTA Section 6A(c). An agency may redact specific entries only where disclosure would create a specific, documented law enforcement necessity, and shall provide the citizen with notice of any redaction and the general category of information withheld.

(e) Vendor Agentic AI. Where an Agentic AI System is developed or operated by a Vendor, the agency's contract with the Vendor shall:

(1) Explicitly designate the Vendor as a Secondary Fiduciary under PDTA for purposes of its operation of the agentic system;

(2) Require the Vendor to comply with all requirements of this Act and PDTA applicable to Secondary Fiduciaries;

(3) Prohibit the Vendor from retaining, using, or deriving commercial value from citizen data accessed through the agentic system beyond the scope of the contracted function;

(4) Require the Vendor to provide the agency with complete access to the system's audit trails, training data documentation, and decision logic at any time upon agency request; and

(5) Consistent with VIDA Section 6(a)(3)'s prohibition on mandatory use of a single private platform, be structured so that the agency can transition to an alternative provider without loss of citizen data, audit trail records, or operational continuity.

(f) Open Source Preference. Consistent with the principles of monopoly prevention and open participation in VIDA Section 7(a)(15), agencies procuring Agentic AI Systems shall prefer open-source models over proprietary black-box alternatives where the open-source model achieves the required function at comparable accuracy, in order to ensure the agency's ability to audit the system's decision logic, retrain the model on corrected data, and avoid vendor lock-in.

Section 9. Algorithmic Impact Assessments with Binding Consequences

Cross-references: VIDA Sec. 4A(b) (Post-enactment regulations shall mandate privacy impact assessments for AI deployments); VIDA Sec. 9 (Public Participation and Accessibility); PDTA Sec. 7(b)(4) (Government shall require impact assessments for any data initiatives); PDTA Sec. 7(c) (Government breaches trigger heightened scrutiny with remedies favoring restoration of individual control)

(a) Requirement. Every government agency operating a Tier 1 or Tier 2 Automated Decision System shall conduct Algorithmic Impact Assessments at the frequency specified in subsections (b) and (c). Algorithmic Impact Assessments are not aspirational reviews; they are legally required evaluations whose findings trigger mandatory consequences under subsection (e). Failure to conduct a required AIA, or material falsification of AIA findings, constitutes a per se breach of the duty of care under PDTA Section 6(a) and a knowing violation of this Act subject to the willful violation penalties of PDTA Section 9(a)(4).

(b) Frequency of Assessments.

(1) Tier 1 systems using Interpretable Models: annual assessment.

(2) Tier 1 systems operating under Black-Box exception approval: quarterly assessment.

(3) Tier 2 systems: annual assessment.

(4) Any system after a material change under Section 5(e): immediate assessment before the modified system is placed into operation.

(5) Any system following a credible complaint of bias, systematic error, or disparate impact filed with the Data Fiduciary Oversight Office: assessment within [30] days of the complaint.

(c) Required Contents. Each Algorithmic Impact Assessment shall document:

(1) System Performance. Current accuracy metrics, false positive and false negative rates, and outcome distributions, disaggregated by race, sex, national origin, disability, age, and other applicable protected characteristics;

(2) Disparate Impact Analysis. Whether the system currently produces Disparate Impact on any protected class, using a methodology consistent with applicable civil rights law, including the four-fifths rule and other recognized analytical frameworks, and identifying any change in disparate impact rates since the prior assessment;

(3) Data Drift Analysis. Whether the Training Data on which the system is based continues to adequately represent the current population being affected by the system, and whether any data drift has introduced new sources of bias or error;

(4) Human Override Analysis. The rate at which Accountable Human Officials have overridden, modified, or upheld automated outputs during the assessment period, disaggregated by decision type and outcome, as evidence of whether genuine human oversight is functioning;

(5) Complaint and Appeal Analysis. The number and nature of citizen complaints, appeals, and requests for algorithmic explanation received during the assessment period, the outcomes of those processes, and any patterns suggesting systematic error or bias; and

(6) Remediation Status. The status of any remediation required by a prior AIA, including documentation that required corrections have been implemented and their effectiveness verified.

(d) Public Filing. Every completed AIA shall be filed with the Data Fiduciary Oversight Office and published in the transparency registry, subject to the same limited redaction standards applicable to ADAs under Section 5(d). Filing and publication shall occur within [30] days of completion.

(e) Mandatory Consequences of AIA Findings. AIA findings are not advisory. The following findings trigger mandatory agency action:

(1) Finding of Disparate Impact. A finding of Disparate Impact that is not justified by documented legitimate operational necessity triggers:

(i) Mandatory remediation of the identified source of disparate impact within [90] days;

(ii) Suspension of the affected Automated Decision System function if remediation cannot be demonstrated within [90] days, with all pending decisions processed through human review during the suspension;

(iii) Citizen notification: written notice to all citizens who received Adverse Automated Decisions from the affected system during the period the disparate impact was present, with a statement of their right to request de novo human review; and

(iv) Reopener right: every citizen who received an Adverse Automated Decision during the period of identified disparate impact has the right to de novo human review of their case, upon request within [one year] of the notification required by subparagraph (iii).

(2) Finding of Systematic Error. A finding of systematic error exceeding [10] percent in any decision category, or of systematic errors concentrated in any identifiable population, triggers the same mandatory consequences as a finding of Disparate Impact under paragraph (1).

(3) Finding of Human Override Rate Anomaly. A finding that Accountable Human Officials approved automated outputs at a rate of [95] percent or higher without documented case-specific reasoning, across any quarter of the assessment period, triggers a mandatory review of the human oversight architecture under Section 7(a)(3) and an enhanced oversight requirement for the following [12] months.

(f) Enforcement. Citizens who were subject to Adverse Automated Decisions during a period of identified disparate impact or systematic error, and who were not notified as required by subsection (e)(1)(iii), may bring a private right of action under PDTA Section 8. In any such action, the agency's failure to conduct required AIAs, or its failure to notify and provide reopener rights following adverse AIA findings, is presumptive evidence of a knowing violation subject to the punitive damages provision of PDTA Section 9(a)(4).

Section 10. Technical AI Oversight Capacity

Cross-references: VIDA Sec. 10(c) (Data Fiduciary Oversight Office); VIDA Sec. 7(a)(16) (Multi-stakeholder governance of trust registries); VIDA Sec. 18 (Oversight and Reporting); PDTA Sec. 7(b)(3) (Data Trusteeship Authority with audit, penalty, and remediation authority); PDTA Sec. 10 (Enforcement)

(a) Amendment of VIDA Section 10(c). The Data Fiduciary Oversight Office established by VIDA Section 10(c) is hereby supplemented to include the following AI-specific mandates:

(1) Technical Staffing. The Oversight Office shall employ or retain, on a full-time basis, a sufficient number of personnel with demonstrated expertise in machine learning systems, algorithmic auditing, AI safety evaluation, and statistical analysis of automated decision outcomes to carry out the oversight functions required by this Act. Expertise qualifications shall be defined by regulation within [180] days of this Act's effective date and shall be reviewed and updated at least every [3] years;

(2) Audit Capability. The Oversight Office shall maintain the technical capability to conduct or commission independent audits of Automated Decision Systems deployed by government agencies, including access to training data documentation, model architecture specifications, decision logs, outcome data, and the ability to conduct adversarial testing of deployed systems;

(3) Transparency Registry. The Oversight Office shall maintain a publicly accessible transparency registry containing all filed ADAs, Abbreviated Deployment Notices, and AIAs, searchable by agency, system type, and deployment date; and

(4) Annual AI Deployment Report. The Oversight Office shall publish an annual report on government AI deployment, including the number and type of Automated Decision Systems in operation across all state agencies, findings from ADA reviews and AIA filings, patterns of disparate impact or systematic error identified, remediation actions taken, and recommendations for legislative or regulatory updates.

(b) Amendment of PDTA Section 7(b)(3). The Data Trusteeship Authority established by PDTA Section 7(b)(3) is hereby supplemented to include the following AI-specific mandates:

(1) Citizen AI Rights Enforcement. The DTA shall have primary jurisdiction to investigate complaints by citizens arising from Adverse Automated Decisions, failures to provide algorithmic explanation, failures to conduct AIAs, and violations of agentic AI fiduciary duties under Section 8 of this Act;

(2) AI Fiduciary Compliance Audits. The DTA shall conduct or commission independent audits of no fewer than [10] percent of all active Tier 1 Automated Decision Systems annually, selected based on risk, complaint volume, and population impact;

(3) Technical Capacity. The DTA shall employ or retain technical AI expertise on the same terms and standards required of the Oversight Office under subsection (a)(1); and

(4) Coordination. The Oversight Office and the DTA shall coordinate their AI-related functions through a joint annual work plan to prevent duplication and ensure comprehensive coverage of agency AI deployments.

(c) Joint AI Audit and Review Panel. There is hereby established within the Data Fiduciary Oversight Office a Joint AI Audit and Review Panel (the Panel), which shall:

(1) Composition. Include at minimum: two technical AI experts with expertise in machine learning, algorithmic fairness, and AI safety; two civil liberties or digital rights advocates; one representative of affected citizen communities with demonstrated experience in algorithmic discrimination; one legal expert in administrative law and civil rights; and the directors of the Oversight Office and the DTA or their designees;

(2) Functions. Conduct independent technical review of ADAs referred by the Oversight Office under Section 5(c)(3); conduct de novo review of any government AI deployment upon petition by [500] or more citizens or by [5] or more civil rights organizations asserting systemic harm; and issue public advisory opinions on novel AI governance questions arising under this Act;

(3) Independence. Panel members other than the Oversight Office and DTA directors shall be appointed for staggered [3]-year terms, may not be employed by or have material financial relationships with AI Vendors during their terms, and shall not be subject to removal except for cause; and

(4) Reporting. The Panel shall publish a public report on its activities, findings, and recommendations no less than annually, consistent with the multi-stakeholder governance requirements of VIDA Section 7(a)(16).

(d) Penalties from Agency Operating Budgets. Consistent with VIDA Section 4(d)(3)'s requirement that penalties be paid from the operating budgets of agencies found in violation, all civil penalties, fines, and court-ordered payments arising from enforcement of this Act shall be paid from the operating budget of the responsible agency and not from the general fund. No officer or employee found to have committed a willful violation of this Act shall be indemnified with public funds, consistent with VIDA Section 4(d)(4).

Section 11. Enforcement

Cross-references: PDTA Sec. 8 (Private Rights of Action); PDTA Sec. 9 (Remedies); PDTA Sec. 10 (Enforcement); VIDA Sec. 10(a)(5) (Private right of action for violations of VIDA); VIDA Sec. 10(d) (Attorney General enforcement authority)

(a) Private Right of Action. Any individual aggrieved by a violation of this Act may bring a civil action in [appropriate court] against the responsible government agency or, where the violation arises from a Vendor's conduct, against the Vendor. The private right of action under this Act supplements, and does not supersede or limit, the private rights of action established by PDTA Section 8 and VIDA Section 10(a)(5). Where a violation of this Act also constitutes a violation of PDTA or VIDA, the plaintiff may assert all available claims in a single proceeding.

(b) No Administrative Exhaustion Required. No individual is required to exhaust administrative remedies before bringing a private right of action under this Act, consistent with PDTA Section 8(c). Individuals may simultaneously pursue administrative remedies through the Oversight Office or the DTA and civil remedies through the courts.

(c) Remedies. In a successful action under this Act, courts may award:

(1) Injunctive relief, including suspension of the Automated Decision System, mandatory de novo human review of the plaintiff's case, and mandamus requiring compliance with ADA, AIA, or explanation requirements;

(2) Compensatory damages for actual harms, including harm from denial of benefits, services, licenses, or opportunities to which the plaintiff was entitled;

(3) Statutory damages of not less than [$1,000] per violation and [$100] per day for each day a continuing violation remains uncorrected following notice;

(4) Punitive damages for willful or reckless violations, including deployment of a Black-Box Model in a Tier 1 use without ADA approval, failure to conduct required AIAs, and failure to notify citizens of identified disparate impact;

(5) Attorney's fees and costs to prevailing plaintiffs; and

(6) Reopener relief: an order requiring the agency to conduct de novo human review of the plaintiff's original adverse determination and all similar adverse determinations issued during the period of the violation.

(d) Class Actions. Actions under this Act may be maintained as class actions where the automated decision system's violations affected a sufficiently numerous class. Courts shall apply a presumption in favor of class certification where the plaintiff establishes that an Automated Decision System produced the same category of violation across multiple citizens, consistent with the equitable access provisions of PDTA Section 9(b).

(e) Attorney General Enforcement. The Attorney General has enforcement authority under this Act concurrent with the Oversight Office and the DTA, consistent with VIDA Section 10(d). The Attorney General may bring enforcement actions on behalf of the state or on behalf of a class of aggrieved citizens.

(f) Sovereign Immunity. Consistent with PDTA Section 9(c), sovereign immunity is waived for violations of this Act, subject to the monetary limits specified in that section. An individual's right to bring a private right of action against a government agency for violation of this Act is not conditioned on any administrative claim filing requirement.

(g) Statute of Limitations. Actions under this Act shall be brought within [five years] from the later of: (1) the date of the Adverse Automated Decision giving rise to the claim; (2) the date the plaintiff discovered or reasonably should have discovered the violation; or (3) for claims arising from an agency's failure to notify under Section 9(e)(1)(iii), the date notification was due. The discovery rule applies to claims involving AI systems whose bias or error was not disclosed to the affected citizen.

Section 12. Vendor Obligations and Government Procurement Standards

Cross-references: VIDA Sec. 7(a)(15) (Monopoly prevention; FRAND terms for publicly funded identity systems); VIDA Sec. 6(a)(3) (No mandatory use of single private platform); PDTA Sec. 6 (Obligations of Secondary Fiduciaries); VIDA Sec. 4(d)(1) (Grant program for compliant systems)

(a) Vendor Fiduciary Obligations. A Vendor whose Automated Decision System is used by a government agency in a Tier 1 or Tier 2 AI Use is a Secondary Fiduciary under PDTA for purposes of its handling of citizen personal data processed through that system. Vendor fiduciary obligations include:

(1) Providing the deploying agency, upon request and at any time, with complete documentation of the system's Training Data, model architecture, decision logic, accuracy statistics, and known failure modes;

(2) Promptly notifying the deploying agency of any change to the system that materially affects its accuracy, decision logic, or disparate impact characteristics;

(3) Cooperating fully with audits conducted by the Oversight Office, the DTA, or the Joint AI Audit and Review Panel; and

(4) Indemnifying the deploying agency for costs, penalties, and damages arising from violations of this Act caused by the Vendor's failure to comply with the requirements of this section or of its contract with the agency.

(b) Procurement Requirements. All government agency procurements of Automated Decision Systems shall, as conditions of contract award:

(1) Require the Vendor to provide complete Training Data documentation, decision logic documentation, and accuracy statistics prior to contract award;

(2) Require the Vendor to warrant that the system has been tested for Disparate Impact and that any identified disparate impact has been disclosed and assessed;

(3) Require the Vendor to comply with VIDA Section 7(a)(15)'s FRAND terms requirement for any publicly funded AI infrastructure;

(4) Prohibit exclusive licensing arrangements that would prevent the agency from procuring alternative compliant systems;

(5) Require that all citizen data processed through the system remains under the agency's custody and control and is not used by the Vendor for any purpose beyond the contracted function; and

(6) Include a data portability and transition clause ensuring the agency can migrate to an alternative system without loss of citizen data, audit trail records, or operational continuity, consistent with VIDA Section 4(b)(4)'s duty of portability.

(c) Procurement Incentives. Agencies that procure Automated Decision Systems from Vendors who provide open-source models, open Training Data documentation, and open audit log access are eligible for priority in government funding allocations for digital initiatives, consistent with the compliance incentive provisions of VIDA Section 4(d)(2).

Section 13. Implementation and Timelines

Cross-references: VIDA Sec. 11 (Implementation and Timelines; 24-month compliance for new services, 48-month for existing); PDTA Sec. 12 (Effective Date; 180 days after enactment); VIDA Sec. 4(d)(1) (Grant program for compliance)

(a) Effective Date. This Act takes effect [180] days after enactment, consistent with the effective date provisions of VIDA and PDTA.

(b) Technical Standards. Within [7] months of enactment, the Data Fiduciary Oversight Office, in consultation with the DTA and the Joint AI Audit and Review Panel, shall issue technical standards and guidance for:

(1) Algorithmic Deployment Assessment format, required contents, and review procedures;

(2) Abbreviated Deployment Notice format for Tier 2 and Tier 3 systems;

(3) Algorithmic Impact Assessment methodology, including approved frameworks for Disparate Impact analysis;

(4) The criteria for distinguishing Interpretable Models from Black-Box Models under Section 3;

(5) Agentic AI audit trail format and individual access portal requirements;

(6) Right to algorithmic explanation delivery standards, including plain-language requirements and multilingual access consistent with VIDA Section 12(b)(3); and

(7) Vendor documentation requirements and procurement contract templates.

(c) Compliance Timeline for Existing Systems.

(1) Within [12] months of this Act's effective date: all agencies shall complete an inventory of all Automated Decision Systems currently deployed in Tier 1 contexts and file Abbreviated Deployment Notices for all Tier 2 and Tier 3 systems;

(2) Within [24] months: all newly deployed or re-platformed Automated Decision Systems shall comply fully with this Act;

(3) Within [36] months: all agencies shall have completed ADAs for all existing Tier 1 systems and received Oversight Office review; and

(4) Within [48] months: all existing Tier 1 systems shall be in full compliance with this Act, including the interpretable model preference rule of Section 7(a)(2), unless granted a waiver under subsection (e) of this section.

(d) Immediate Suspension for High-Risk Systems. Notwithstanding the compliance timeline in subsection (c), any Automated Decision System currently deployed in a Tier 1 context that has received a credible formal complaint of Disparate Impact, has been the subject of adverse judicial findings, or has been flagged as high-risk by the Oversight Office shall be subject to immediate ADA requirement, with a [90]-day completion deadline. During the period before ADA approval, all Adverse Automated Decisions from flagged systems shall receive de novo human review.

(e) Waiver. An agency may apply to the Oversight Office for a waiver of the [48]-month full compliance deadline for a specific system, demonstrating good cause including technical complexity, resource constraints, or lack of available compliant alternatives. Waivers shall not extend compliance beyond [60] months from the Act's effective date. During any waiver period, the agency shall implement interim safeguards including enhanced human oversight and quarterly AIAs.

(f) Funding. The [state agency] shall allocate funds from [source] to support implementation, consistent with the funding provisions of VIDA Section 17 and PDTA Section 15. The grant program established under VIDA Section 4(d)(1) shall be extended to assist agencies in achieving compliance with this Act, with priority for agencies serving high-need populations and for agencies procuring open-source, interpretable systems.

Section 14. Cross-Referencing and Statutory Construction

(a) This Act shall be construed in harmony with the Verifiable Identity and Digital Autonomy Act (VIDA) and the Personal Data Trusteeship Act (PDTA) to create a comprehensive and integrated framework for constitutional digital governance. Where this Act and VIDA or PDTA address the same subject matter, the provisions shall be read together to provide the most complete protection for citizens' constitutional rights and fiduciary interests. In the event of irreconcilable conflict between this Act and VIDA or PDTA on any specific point, the provision that provides greater protection to the citizen's rights shall govern.

(b) This Act does not limit any right or remedy available to a citizen under VIDA, PDTA, the Fourth Amendment to the United States Constitution, the equal protection guarantees of the state and federal constitutions, applicable civil rights statutes, or any other provision of law. The rights and remedies provided by this Act are cumulative and in addition to all other rights and remedies.

(c) This Act shall not be construed to prohibit government use of Automated Decision Systems. It establishes the conditions under which government may lawfully use such systems in ways that honor its constitutional obligations to the citizens it serves.

Section 15. Severability

(a) If any provision of this Act or its application to any person or circumstance is held invalid, the remainder of the Act and the application of its provisions to other persons or circumstances shall not be affected and shall continue in full force and effect.

Section 16. Sunset Review

(a) The [relevant oversight body], in coordination with the Data Fiduciary Oversight Office and the Data Trusteeship Authority, shall review this Act and report to [the Legislative Body] on its implementation, effectiveness, and recommended amendments every [5] years following the effective date, consistent with the sunset review provisions of VIDA Section 16 and PDTA Section 14. The review shall specifically assess whether the tiered AI use framework in Section 7 remains appropriate given developments in AI technology, whether the Interpretable Model Preference Rule in Section 7(a)(2) requires updating, and whether new categories of Agentic AI deployment require additional statutory treatment.

About this Draft

This is an initial draft of the Government Algorithmic Accountability and AI Fiduciary Act (GAAFA), prepared as a companion statute to the Verifiable Identity and Digital Autonomy Act (VIDA) and the Personal Data Trusteeship Act (PDTA). All bracketed provisions indicate items requiring legislative specification. Cross-references to VIDA and PDTA reflect the section numbering of the March 2026 drafts of those Acts. This draft is prepared for discussion and does not represent the position of any government body.

Prepared by Michael G. Leahy | March 2026