Cambridge Forum on AI: Law and Governance (2026), **2**, e14, 1–14
doi:10.1017/cfl.2026.10055

Cambridge Forum

RESEARCH ARTICLE

Between law and code: Operationalizing responsible AI in public administration

Prema Elumalai[1] and Ragul Olakkur Vijayakumar[2]

[1]VIT School of Law, Vellore Institute of Technology, Chennai Campus, Chennai, Tamil Nadu, India and [2]Law, Supreme Court of India, New Delhi, India
Corresponding author: Prema Elumalai; Email: prema.e@vit.ac.in

(Received 4 October 2025 UTC; revised 31 May 2026 UTC; accepted 3 July 2026 UTC)

Abstract
This article investigates how to translate abstract responsible artificial intelligence (AI) principles into practical constraints for public-sector AI systems. Focusing on due process, transparency, fairness, reason-giving and proportionality, the paper uses administrative and constitutional law doctrines to scrutinize AI deployments in government. The purpose of this comparative analysis is to evaluate how the divergent frameworks of the United States (US), the European Union (EU) and Singapore fulfill responsible AI principles. The paper contrasts the EU's unified hard-law model with the fragmented preemption approach of the US and the facilitative soft-law guidelines of Singapore. Furthermore, the analysis explores how legal norms can be integrated throughout the AI lifecycle, addressing the inherent tensions between technical performance and legal compliance. Illustrative case studies – including the Dutch childcare-benefits algorithm, the UK exam-grading fiasco, the US Internal Revenue Service audit controversies and Estonia's successful Kratt AI implementation – highlight the severe consequences of algorithmic opacity and the benefits of proactive governance. Drawing on these insights, the paper argues that multidisciplinary governance is essential. The article's original scholarly contribution lies in systematically mapping administrative and constitutional law doctrines onto AI requirements across the AI lifecycle, proposing a concrete techno-legal architecture that ensures automated public decisions comply with enduring legal norms.

Keywords: responsible artificial intelligence; public administration; administrative law; generative AI; transparency; accountability

1. Introduction

The deployment of artificial intelligence (AI) in government – from automating benefit allocation to predictive policing – raises profound questions about accountability, rights and the rule of law (OECD, 2024). Algorithmic systems can make errors or embed biases in ways that are opaque to affected individuals (National Institute of Standards and Technology, 2023). It is necessary to avoid conflating modern machine-learning AI with broader automated decision-making (ADM) or algorithmic administration, as these concepts raise fundamentally different legal and technical issues regarding explainability and systemic opacity (Wolswinkel, 2022). While traditional ADM often relies on hard-coded, rule-based decision trees that are inherently auditable, modern machine-learning AI evolves based on complex data patterns, creating a mathematical "black box" that confounds traditional legal scrutiny. This opacity collides with long-standing legal norms: public administrations are subject to due process, transparency and non-arbitrariness, and governments must justify and record their decisions (Wolswinkel, 2022).

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial licence (http://creativecommons.org/licenses/by-nc/4.0), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original article is properly cited. The written permission of Cambridge University Press or the rights holder(s) must be obtained prior to any commercial use.

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 09:26:50, subject to the Cambridge Core terms of use.

Public AI systems demand high levels of legal, ethical and technical responsibility to ensure trust and legitimacy. Key principles for governing AI in the public sector include fairness (avoiding bias and non-discrimination), due process (right to notice, hearing and reasoned decisions), proportionality (tailoring AI use to the risk posed), transparency (clear information about how decisions are made), accountability (being answerable for outcomes) and privacy (protecting personal data) (UNESCO, 2021). Bridging law and code remains a central challenge, as law speaks in broad, principle-laden language, whereas technical teams work with statistical data pipelines and models (Nouws & Dobbe, 2024).

This article addresses not only law-as-code and code-as-law mechanisms but also broader issues of institutional design, governance structures and administrative processes related to responsible AI in public administration. The article's original scholarly contribution lies in systematically mapping administrative and constitutional law doctrines onto responsible AI requirements across the AI lifecycle, while undertaking a comparative evaluation of global regulatory regimes.

The selection of the United States (US), the European Union (EU) and Singapore as the jurisdictions examined in this article is driven by their distinct geopolitical and regulatory positions, representing three distinct paradigms of technology governance. The EU represents a regulatory superpower asserting a unified, hard-law approach, despite its heavy reliance on externally developed AI models. The US represents a decentralized, market-driven regime currently leaning toward federal preemption and deregulation, heavily influenced by its status as the dominant hub for frontier AI developers. Singapore represents a smaller, highly agile governance structure prioritizing facilitative, soft-law frameworks that foster rapid technological integration without stifling innovation. Factors such as the dominant role of US companies in AI model development, Singapore's smaller population and agile governance structure, and the EU's regulatory influence despite its reliance on externally developed models significantly affect the implementation of responsible AI principles across the AI lifecycle.

This article explicitly frames the jurisdictional examinations as part of a structured comparative analysis, evaluating these three systems against administrative and constitutional law principles. By exploring thematic challenges – risk classification, human oversight, transparency, explainability, auditability and accountability – the paper outlines a techno-legal architecture that weaves legal rules into technical design and embeds technical checks into legal review.

2. Legal and normative foundations

2.1 Administrative law doctrines

Public administration is traditionally governed by doctrines of procedural fairness and rationality. At a minimum, due process requires that individuals receive notice of a decision affecting them, an opportunity to contest it and a reasoned explanation of the outcome (United Nations Department of Economic and Social Affairs, 2024). In many jurisdictions worldwide, administrative laws impose a duty on public bodies to state reasons for their decisions. This duty serves multiple goals: it promotes accuracy by forcing decision-makers to think more carefully, facilitates judicial or administrative review and fosters public confidence by ensuring decisions are not arbitrary (Donnelly, 2025). Historically, the law has justified reason-giving on instrumental grounds, as a reasoned record makes it easier to detect errors or bias. Beyond mere instrumentalism, reason-giving rests on fundamental human dignity: an individual subject to a public decision has a right to understand why it was made, especially when a powerful state actor is involved (Ullits, 2025).

Introducing an opaque machine-learning model in place of a human decision-maker profoundly challenges these values. An algorithm might take actions – such as disproportionately targeting specific demographics for tax audits – without an obvious logical trail (Elzayn et al., 2025). To satisfy administrative-law duties, an AI system should generate explanations or logs that a human

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 09:26:50, subject to the Cambridge Core terms of use.

auditor or appellant can understand. Without such measures, AI-driven decisions risk flouting the core procedural norms that legitimize state action (Bignami, 2022).

2.2 Fundamental rights and equality

Governments are bound by substantive rights as well as process rules. Antidiscrimination and equality laws prohibit administrative practices that arbitrarily disadvantage protected groups (Wolswinkel, 2022). For example, the EU Charter of Fundamental Rights and the equal protection guarantees under the US Constitution's 5th and 14th Amendments bar government actions that classify based on race, gender or other protected classes without strong justification (Students for Fair Admissions, Inc. v. President and Fellows of Harvard College, 2023). Any public AI system must be designed so as not to encode bias against minorities, necessitating the use of unbiased data, disparate impact auditing and specialized scrutiny for sensitive categories.

Furthermore, privacy and data-protection regimes, notably the EU's General Data Protection Regulation (GDPR), mandate that personal data must be handled lawfully and with informed consent where required. The law provides a right to an effective remedy when rights are violated. If an algorithm errs, affected individuals must have avenues to challenge the decision through an appeal or judicial review, mirroring the contestability of human adjudication (Dun & Bradstreet Austria and Others, 2025). This implies that automated systems must be contestable: they must record inputs and logic sufficiently that an appeals court can audit compliance with legal norms.

2.3 Good governance principles

Beyond individual rights, public administration is guided by broader principles of transparency, accountability and legitimacy. Public agencies operate under a general duty to justify their actions and avoid caprice (OECD, 2025c). Openness is key: laws require that significant administrative decisions be published or communicated in an understandable form. Embedding AI in government must therefore promote these principles of good administration (Wolswinkel, 2022). For instance, documenting how an AI came to a policy-critical output, allowing public audit of algorithmic criteria and ensuring clear lines of responsibility when errors occur are foundational requirements. The core challenge is balancing these norms so that algorithmic governance does not make agencies opaque or unaccountable in the name of efficiency (Government Digital Service, 2025). Good governance demands that AI systems used by governments be designed to reinforce – and never to supplant – the rule-of-law ideals of transparency, reasonableness and accountability (Council of Europe, 2024).

3. Comparative evaluation of regulatory frameworks

This article explicitly frames these sections as a structured comparative analysis, contrasting the EU's unified hard-law model with the fragmented, preemptive approaches of the US and the soft-law guidelines of Singapore. The criteria for this comparative evaluation include the rigidity of the legal mandates, the prioritization of innovation versus fundamental rights and the mechanisms employed to enforce accountability across the AI lifecycle.

3.1 The European Union: the unified hard-law model

The EU has pioneered a hard-law, risk-based regime that prioritizes fundamental rights over unfettered technological innovation. The 2024 AI Act classifies many public-sector AI uses as high-risk by default. Annex III explicitly lists AI in critical government functions – such as systems used to assess eligibility for essential public assistance benefits, healthcare, law enforcement and immigration screening – as high-risk by default, though Article 6 provides specific derogations if the system merely performs a narrow procedural task, does not perform profiling of natural persons and does

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 09:26:50, subject to the Cambridge Core terms of use.

not pose a significant risk to fundamental rights (European Parliament & Council of the European Union, 2024).

Providers must undergo conformity assessments and maintain detailed technical documentation, while deployers are required to perform a Fundamental Rights Impact Assessment (FRIA) before deployment (Regulation (EU) 2024/1689, 2024). Furthermore, Article 14 mandates that high-risk systems be engineered for effective human oversight, ensuring that they can be actively monitored and overridden by natural persons (Regulation (EU) 2024/1689, 2024). Article 13 requires clear transparency information and instructions for use to be provided to deployers, whereas transparency obligations directed at end users are governed by Article 50 (European Parliament & Council of the European Union, 2024).

The EU model successfully establishes a stringent, legally binding framework that forces accountability into the engineering pipeline. However, as a comparative critique, the EU framework's heavy regulatory burden has been criticized for potentially stifling domestic technological agility. Because the EU relies heavily on externally developed AI models – often originating from the US or China – the rigorous compliance requirements may slow down the modernization of public services compared to more permissive jurisdictions. The bureaucratic overhead required to maintain continuous post-market monitoring and audit logs can be prohibitively expensive for smaller municipalities seeking to deploy AI solutions (Jørgensen & Ma, 2025).

3.2 The United States: fragmentation and federal preemption

The US approach stands in stark contrast to the EU, lacking a unified, comprehensive AI statute. Historically, US policy has been shaped by executive guidance and decentralized agency frameworks (Harris, 2025). For example, the AI in Government Act of 2020 mandated that the Office of Management and Budget (OMB) issue guidance to federal agencies regarding AI acquisition and use (Office of Management and Budget, 2024). Consequently, the OMB issued Memorandum M-25-21, which rescinded and replaced the earlier Memorandum M-24-10, directing federal agencies to implement minimum risk-management practices, develop AI use-case inventories and designate Chief AI Officers (Office of Management and Budget, 2025). The National Institute of Standards and Technology (NIST) AI Risk Management Framework provides core concepts of trustworthy AI but remains purely voluntary and non-prescriptive (National Institute of Standards and Technology, 2023).

However, significant changes in US policy during the Trump Administration in 2025 and 2026 have aggressively steered the federal landscape toward deregulation and the preemption of state-level AI laws (The White House, 2026). Issued in December 2025, Executive Order 14365, titled "Ensuring a National Policy Framework for Artificial Intelligence," establishes a minimally burdensome national standard explicitly designed to sustain the US's global AI dominance (Trump, 2025). EO 14365 fundamentally restructures US AI governance by establishing a doctrine of federal preemption. It tasks the Department of Justice with forming an AI Litigation Task Force to actively challenge state AI laws – such as Colorado's regulations addressing algorithmic discrimination – that conflict with federal deregulatory goals. Furthermore, the Order leverages federal purse strings, explicitly mandating a Policy Notice to specify conditions rendering states with onerous AI laws ineligible for non-deployment funds under the Broadband Equity Access and Deployment Program to the maximum extent allowed by Federal law (Exec. Order No. 14,365, 2025).

Comparatively, the US model prioritizes rapid technological advancement and market dominance over stringent rights protections. While this approach fosters unparalleled innovation within the private sector, it creates severe gaps in public administration accountability. The reliance on voluntary adherence at the federal level, combined with the active suppression of state-level civil rights protections via federal preemption, leaves vulnerable populations with limited legal recourse against algorithmic harms (Bipartisan House Task Force on Artificial Intelligence, 2024).

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 09:26:50, subject to the Cambridge Core terms of use.

3.3 Singapore: facilitative soft-law governance

Singapore employs a facilitative, soft-law strategy that seeks a middle ground between the EU's rigidity and the US's deregulatory preemption. In 2018, the Monetary Authority of Singapore introduced the FEAT principles – fairness, ethics, accountability and transparency – as ethical guidelines designed explicitly for the financial sector. This framework translates high-level ethics into practice, emphasizing explainability and a human-centered ethos without binding legislation. In May 2022, Singapore advanced this by launching AI Verify as a Minimum Viable Product for an international pilot, a testing toolkit that assesses systems against 11 AI ethics and governance principles, performing technical tests for specific principles like fairness, explainability and robustness, while generating standardized evaluation reports, which was subsequently open-sourced culminating in the establishment of the AI Verify Foundation in June 2023 (AI Verify Foundation, n.d.).

Singapore's administrative ethos prizes legitimacy and efficiency, encouraging internal governance measures such as ethics boards and impact assessments rather than imposing burdensome legal mandates (World Economic Forum, Infocomm Media Development Authority, & Personal Data Protection Commission, 2020). Singapore's agile regulatory environment allows it to adapt rapidly to new technologies like generative AI, updating its frameworks to address hallucination risks and automated decision ambiguity. However, a comparative critique reveals that the lack of binding, enforceable legal penalties for algorithmic failures may result in accountability deficits. If public sector agencies or private vendors choose to ignore voluntary compliance standards, the soft-law approach provides no statutory basis for holding anyone accountable when algorithms cause public harm (AI Verify Foundation & Infocomm Media Development Authority, 2024).

In summary, a comparative overview reveals distinct paradigms across the three jurisdictions based on their primary governance instruments, regulatory approaches, prioritization goals and accountability mechanisms. The EU relies on the AI Act (2024) and the GDPR to enforce a unified hard-law approach. This regime prioritizes fundamental rights and harm prevention, enforcing accountability through mandatory conformity assessments, fines and algorithmic audits (European Parliament & Council of the European Union, 2024). Conversely, the US relies on Executive Orders, such as EO 14365, and OMB guidance, embodying a fragmented and preemptively deregulatory approach. The US prioritizes global market dominance and innovation speed, relying primarily on voluntary frameworks and fragmented, agency-led enforcement for accountability (Office of Management and Budget, 2025). Finally, Singapore utilizes the FEAT principles and AI Verify as its primary instruments, adopting a facilitative soft-law approach. Singapore prioritizes agile integration and public–private trust, maintaining accountability through internal ethics boards and voluntary standardized testing (Allen et al., 2025).

4. Embedding legal-ethical principles in the AI lifecycle (comparative and prescriptive analysis)

This section systematically evaluates how the principles mapped above must be technically operationalized, distinguishing between the descriptive reality of current laws in the EU and US and the prescriptive standards necessary for true accountability.

4.1 Risk classification and proportionality

European and domestic laws increasingly insist that public agencies classify AI by risk. Descriptively, the EU AI Act explicitly lists automated benefits determination and visa screening as high-risk, mandating pre-deployment impact assessments like the FRIA (European Parliament and Council of the European Union, 2024). US guidance under the OMB defines high-impact AI as any tool

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 09:26:50, subject to the Cambridge Core terms of use.

whose output serves as a principal basis for decisions or actions that have a legal, material, binding or similarly significant effects on individuals or entities (Office of Management and Budget, 2025).

Prescriptively, proportionality dictates graduated oversight: minimally invasive cases require a lighter touch, whereas high-stakes cases trigger full-fledged checks. Concretely, this means embedding legal safeguards directly into risk management protocols via rigorous data governance and built-in technical measures such as bias penalties (European Parliament and the Council, 2024). For example, a welfare screening tool must undergo algorithmic fairness audits to minimize disparate impacts on protected groups. Furthermore, the principle of least encroachment dictates that fully autonomous AI judgments should be rare, reserved only for the most routine administrative tasks, while any significant deprivation of rights requires human intervention (European Union, 2024).

4.2 Human oversight and human-in-the-loop

Both law and policy assert the necessity of keeping a human in the loop for automated public decisions. Descriptively, Article 14 of the EU AI Act strictly requires that high-risk systems be designed for effective human oversight, providing operators with the ability to interrupt, deactivate or override the system (European Union, 2024). Moving beyond the descriptive requirements of EU law, this article advances the prescriptive recommendation that administrators should be trained to question AI flags to actively combat automation bias.

Automation bias – the psychological tendency for humans to over-trust machine outputs and forgo independent judgment – can render human-in-the-loop safeguards entirely performative if overseers operate as mere rubber stamps (Colombi Ciacchi et al., 2025). Prescriptively, agencies must ensure that any AI-driven refusal or penalty is substantively reviewed and ratified by a human authority, integrating oversight into the workflow through mandatory secondary evaluations rather than solely relying on the AI system's output (European Data Protection Supervisor, 2025). Meaningful oversight requires clear processes; agencies must formalize review checkpoints and ensure that procedural rights attach to a human-signed decision. Courts, prescriptively, should demand evidentiary proof that human review was active and substantive.

4.3 Transparency vs. trade secrecy

Governments must reconcile the administrative duty of openness with private vendors' intellectual property claims. Descriptively, under data-protection laws like the EU's GDPR, individuals possess the right to meaningful information about the logic of automated decisions (SCHUFA Holding and Others (Scoring), 2023). However, agencies frequently procure proprietary systems shielded by trade secret exemptions under freedom-of-information laws. This creates a systemic barrier to transparency, allowing vendors to protect source code while public agencies deploy algorithmic systems that affect fundamental rights (Matulionyte, 2024).

This tension severely frustrates accountability. If an individual denied benefits cannot access even a high-level rationale due to corporate secrecy, due process is fundamentally violated. A prescriptive compromise requires legislation that forces vendors to escrow key algorithmic information, such as model documentation and feature data, with the agency (Global Partnership on Artificial Intelligence, 2024). Public accountability demands that agencies retain the right to audit and debug all code; the public-facing side of the algorithm must yield sufficient information for internal audits and adequate notice to citizens (Koene et al., 2019). Agencies must not blindly cede control to black boxes, and policy prescriptions must call for tiered transparency that balances proprietary rights with constitutional duties (Bignami, 2022).

4.4 Explainability, reason-giving and the limits of XAI

Closely tied to transparency is the requirement to give reasons. Descriptively, the EU AI Act obliges providers to give comprehensible information regarding system outputs, and the US White House

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 09:26:50, subject to the Cambridge Core terms of use.

Blueprint for an AI Bill of Rights encourages clear explanations for automated tools (Solaiman, 2025). While describing these frameworks, this analysis prescriptively asserts that agencies must prepare plain-language justifications translating statistical logic into legal rule contexts, rather than relying on abstract scores. If an algorithm flags a visa application as high-risk, the agency must articulate the specific legal reasons reflected in the data (Gugliotta & Elbi, 2024).

However, the technical execution of explainability is fraught with complications. Technical AI safeguard tools, such as explainability mechanisms like Local Interpretable Model-agnostic Explanations (LIME) and SHapley Additive exPlanations (SHAP), are imperfect and can themselves produce erroneous outputs or misleading explanations (Slack et al., 2020). Post hoc explainable AI (XAI) methods often approximate the behavior of a complex model locally, meaning they do not show what the model actually did, but rather what a linear model would do if trained to approximate it (Burkart & Huber, 2021). Recent analyses reveal that while both LIME and SHAP are widely utilized for explainability, SHAP is mathematically designed to provide high local fidelity through cooperative game theory, whereas relying indiscriminately on post hoc tools can create a dangerous false confidence in public-sector settings if the generated explanations are miscalibrated (European Data Protection Supervisor, 2023). Furthermore, tools such as feature importance scores – which quantify the relative contribution of each input variable to the model's final prediction – can surface conflicting results depending on the specific XAI algorithm employed (Krishna et al., 2024).

The legal and governance implications of these erroneous explanations are severe. When public-sector AI systems generate misleading explanations, it triggers "accountability diffusion" (Busuioc, 2021). Citizens denied public benefits may be provided with an inaccurate rationale generated by an unstable XAI wrapper, preventing them from effectively exercising their right to appeal on substantive grounds. If the explanation does not accurately reflect the AI's actual internal processing, the agency violates fundamental administrative due process and is exposed to severe legal liability. Therefore, prescriptively, post hoc XAI tools should be used strictly for internal exploration and debugging, and never as the sole basis for legally binding public justifications. For high-stakes decisions, agencies must default to inherently interpretable models rather than attempting to explain deep neural networks with flawed approximations.

4.5 The fairness-accuracy trade-off in technical implementation

Translating constitutional equality into algorithmic code introduces profound technical paradoxes. Abstract legal principles like fairness cannot be directly or perfectly translated into technical AI standards without confronting mathematical conflicts (Garcia et al., 2024). In the literature of fair machine learning, several prevailing criteria for fairness include demographic parity, equalized odds and equal opportunity (Castelnovo et al., 2022).

Crucially, these technical definitions of fairness are frequently mathematically incompatible. A model optimized to achieve equalized odds (ensuring false-positive rates are identical across racial groups) cannot simultaneously achieve demographic parity (ensuring the base rate of positive predictions is identical across groups) if the underlying base rates in the training data differ (Weerts et al., 2024). A public agency attempting to satisfy one constitutional interpretation of equality will necessarily violate another.

Furthermore, because machine-learning algorithms are designed to optimize for predictive accuracy, enforcing fairness constraints almost universally degrades the model's overall accuracy – a phenomenon known as the fairness-accuracy trade-off (Pessach & Shmueli, 2022). To manage these conflicts within a regulatory framework, public agencies must adopt multi-objective training approaches or post-processing threshold optimizations. This requires explicit, multidisciplinary policy decisions regarding which specific fairness metric aligns with the legal context of the deployment, consciously accepting a scalarized reduction in raw predictive accuracy to fulfill constitutional mandates against discrimination. Transitioning from a descriptive overview of existing law to a

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 09:26:50, subject to the Cambridge Core terms of use.

prescriptive governance framework, this article asserts that any system of ADM must preserve enforceable human accountability (National Institute of Standards and Technology, 2023).

4.6 Auditability, monitoring and accountability

Ensuring responsible AI requires continuous post-market monitoring. Descriptively, the EU AI Act mandates ongoing risk management, logging and documentation (European Parliament and Council of the European Union, 2024). In the US, frameworks from NIST envision Test, Evaluation, Verification and Validation – a structured testing regime whereby domain experts regularly test deployed systems against benchmarks for accuracy, bias and safety (National Institute of Standards and Technology, 2023).

Courts will likely treat a lack of post hoc explanation as a violation of procedural fairness. This observation applies not only to EU courts but broadly to courts in any jurisdiction that formally incorporates administrative and constitutional law principles into its governance frameworks (Williams, 2022). Agencies must explicitly notify individuals of their right to contest algorithmic decisions (Office of Management and Budget, 2024). Data protection authorities, privacy commissioners and specialized bodies must be granted the statutory power to audit algorithms, enforce compliance utilizing technologies like secure multiparty computation – a cryptographic method allowing multiple parties to jointly compute a function over their inputs while keeping those inputs private – and sanction misconduct (OECD, 2025c). Thorough audit trails provide the legal paper trail necessary to demonstrate good faith compliance in judicial proceedings.

5. Case studies in public administration failures and successes

The necessity of the aforementioned legal-technical integration is best illustrated by examining historical deployments of public-sector AI. This section evaluates both high-profile failures and a modern success story, connecting them directly to the governance frameworks discussed.

5.1 The Dutch childcare benefits scandal (Toeslagenaffaire)

In the Toeslagenaffaire, the Dutch tax agency deployed a machine-learning system to flag allegedly fraudulent childcare subsidy claims, utilizing proxies like dual nationality and low income. This system disproportionately targeted minority families, falsely accusing over 20,000 parents of fraud, forcing them into severe debt and violating fundamental rights on an unprecedented scale (Frey et al., 2025).

The algorithm operated as an opaque black box, offering no explanations to the targeted parents, stripping them of administrative due process and effective appeal rights. The Dutch case study explicitly underscores how this catastrophic failure was exacerbated by a complete absence of technical audits and technical review mechanisms within the legal oversight structure. The targeted parents received no notice of how they were flagged and no justification for the debt claims. Had the agency conducted rigorous technical audits for disparate impact or mandated data provenance documentation, the biased proxies would have been flagged. If the EU AI Act's mandatory FRIA and continuous post-market monitoring requirements had been enforced at the time, alongside GDPR fair-processing obligations, this algorithmic injustice could have been mitigated or entirely prevented.

5.2 The UK A-level exam algorithm (2020)

The controversy over canceled A-level grades arose directly from the COVID-19 disruption of school examinations and the subsequent emergency use of algorithmic grading methods. England's exam regulator, Ofqual, deployed a statistical model that adjusted teachers' predicted grades based

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 09:26:50, subject to the Cambridge Core terms of use.

heavily on each school's historical performance data. This process entrenched inequality, systematically downgrading high-achieving students from historically underperforming state schools while inflating grades for students at elite private schools with smaller cohorts (House of Commons Education Committee, 2020).

Although Ofqual published a comprehensive 319-page technical report detailing the standardization model and the statistical formula on the day the results were released, students lacked a clear and accessible appeals process to challenge their algorithmic grades. This episode demonstrates a profound violation of procedural fairness and proportionality. The deployment occurred without algorithmic accountability mechanisms or stakeholder input. Had a framework akin to the US OMB requirements for AI Impact Assessments been strictly applied, the systemic bias against disadvantaged cohorts would have been quantifiable prior to deployment, preventing the subsequent public uproar and the rapid abandonment of the tool (Ofqual, 2020a).

5.3 The US IRS audit algorithm (2022)

In 2023, researchers revealed that an automated tool used by the US Internal Revenue Service (IRS) to select tax returns for audit disproportionately targeted Black taxpayers at a rate 2.9–4.7 times higher than non-Black filers. The main source of the disparity was differing audit rates by race among taxpayers claiming the Earned Income Tax Credit, where algorithmic choices such as minimizing the no-change rate and targeting erroneously claimed refundable credits disproportionately burdened Black taxpayers (Elzayn et al., 2025).

Victims had no insight into the algorithmic selection process, raising profound concerns regarding fundamental procedural fairness and equal protection principles inherent in the 5th Amendment Due Process Clause. The IRS case highlights the severe limitations of the US's fragmented, voluntary governance models (Kemker, 2024). Without mandatory bias testing or hard-law obligations to document algorithmic logic, federal agencies can unknowingly perpetuate severe discriminatory outcomes. The application of mandatory bias mitigation tools and external auditing frameworks is legally essential to translate 5th Amendment equal-protection principles into concrete engineering practices at the federal level. Self-governance is insufficient when civil rights are at stake (Office of Management and Budget, 2024).

5.4 A successful implementation: Estonia's Kratt AI and Bürokratt

In contrast to these failures, this section introduces Estonia's Kratt AI and the Bürokratt network as a successful case study demonstrating proactive, human-centric AI design in the public sector. Legitimized by Estonia's national AI strategy and specific sectoral legal amendments, such as those made to the Taxation Act, Bürokratt functions as an interoperable network of AI virtual assistants allowing citizens to perform public service procedures seamlessly across institutional boundaries, though broad amendments to the Administrative Procedure Act for automated decisions were notably withdrawn (Nordic Council of Ministers, 2024).

Estonia's implementation succeeded because it strictly aligned with good governance principles. The system was designed with a user-centric focus, incorporating continuous citizen feedback, multichannel accessibility and robust data security by design. By integrating legal compliance directly into the modular development of the AI components, Estonia managed to automate over 140 public sector use cases – from the Unemployment Insurance Fund's job matching algorithms to the Rescue Board's fire hazard identification systems – without violating public trust (E-Governance Academy Foundation, 2025). While Estonia maintains strict regulatory oversight, specific sectoral legislation such as Section 46 of the Taxation Act explicitly permits the tax authority to issue administrative acts in an automated manner without the direct intervention of an official (Taxation Act, 2024, § 46^2). The Estonian government mandated transparency, ensured human oversight remained in

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 09:26:50, subject to the Cambridge Core terms of use.

the loop and utilized explainable system architectures, proving that public-sector AI can enhance administrative efficiency while adhering strictly to legal rights.

These case studies collectively illustrate the profound consequences of algorithmic deployments in public administration and highlight the mitigating framework mechanisms required for responsible integration. The Dutch Toeslagenaffaire utilized AI primarily for fraud detection, resulting in a severe legal and ethical failure characterized by disparate impact, opacity and a lack of due process. This failure underscores the necessity of mitigating frameworks like the EU AI Act, which mandates FRIA and continuous auditing (OECD, 2025c). The UK A-Level Algorithm, tasked with exam grade standardization, failed by encoding historical bias and lacking appeal mechanisms; frameworks such as the NIST AI Risk Management Framework and OMB AI Impact Assessments are designed to mitigate such harms (Ofqual, 2020a). Similarly, the US IRS Audit Algorithm for tax audit selection resulted in significant racial disparities, demonstrating the critical need for mandatory disparate impact testing to prevent potential Equal Protection violations (Elzayn et al., 2025). In stark contrast, Estonia's Kratt and Bürokratt systems, which provide interoperable public services, represent a resounding success. By maintaining a transparent, human-centric and auditable architecture supported by comprehensive public sector data governance frameworks and sectoral legislation, Estonia successfully avoided the ethical pitfalls that plagued the other deployments.

6. Integrating law and code: bridging disciplines

6.1 Multidisciplinary governance

Bringing AI into the public sphere requires cross-functional governance teams comprising lawyers, ethicists, domain experts and engineers. Diverse expertise reduces technical blind spots and prevents algorithmic discrimination. Training is paramount: developers must attain basic literacy in privacy and discrimination law, while legal professionals must understand the technical realities of model drift and the mathematical limitations of explainability. Embedding collaborative governance ensures that civil rights and data protection rules are treated as foundational architectural requirements rather than post hoc compliance checkboxes.

6.2 Translating law into code: a step-by-step case study

This section provides an illustrative case study showing how an AI legal principle can be translated into a concrete technical standard by applying the proposed law–code integration guidelines. Consider a municipal government deploying a machine-learning algorithm to allocate emergency housing assistance.

Step 1: Legal Principle Identification. The cross-functional legal team determines that the deployment is subject to the Equal Protection Clause (prohibiting racial discrimination) and administrative due process (requiring actionable, legally sound explanations for denial of housing).

Step 2: Risk Classification. Applying comparative frameworks, such as the EU AI Act Annex III, the allocation of essential public welfare is classified a priori as high-risk. This triggers mandatory conformity assessments and strict data provenance documentation by the provider during development, followed by an FRIA conducted by the deployer prior to deployment.

Step 3: Technical Translation of Fairness. The engineering team must translate the legal mandate of "non-discrimination" into a specific mathematical metric. Recognizing the fairness-accuracy trade-off, the team selects the demographic parity metric, aiming to ensure equal approval rates across demographic groups regardless of historically biased base rates. During model training, developers utilize bias mitigation libraries to penalize the model mathematically if the approval rates for minority applicants statistically deviate from majority applicants.

Step 4: Operationalizing Explainability. Acknowledging the fundamental instability and legal liability flaws of post hoc XAI tools like LIME, the team opts to deploy an inherently interpretable model

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 09:26:50, subject to the Cambridge Core terms of use.

(e.g., a generalized additive model) rather than a deep neural network. This allows the system to output precise, legally valid reasons for denial (e.g., "Applicant income exceeds the statutory threshold of $30,000"), directly satisfying the administrative requirement for reason-giving without risking accountability diffusion.

Step 5: Post-Deployment Auditing and Oversight. The agency establishes a continuous auditing pipeline. If the demographic parity metric drops below an agency-defined acceptable threshold due to data drift, an automated kill switch halts the algorithm. All housing decisions are immediately reverted to human caseworkers until the model is recalibrated, ensuring continuous compliance with constitutional rights.

This translation process demonstrates the practical applicability of techno-legal architecture, proving that legal norms can be hard-coded into software parameters to guarantee civil rights compliance.

6.3 Regulatory sandboxes and institutionalizing accountability

Experimental AI sandboxes serve as vital living laboratories, allowing agencies and legal experts to test new systems under controlled regulatory oversight (European Parliament and Council of the European Union, 2024). Singapore has actively utilized AI pilot programs to jointly explore technical challenges and regulatory requirements before public rollout. These sandboxes foster essential communication between law and code, allowing legal teams to observe how the AI operates on real cases while engineers understand which legal rules are difficult to encode (Wong, 2025). This reduces the risk that broad regulations either stifle innovation or leave dangerous loopholes.

Ultimately, accountability requires building feedback loops into the procurement and operational phases of AI deployment (Organisation for Economic Co-operation and Development & United Nations Educational, Scientific and Cultural Organization, 2024). Contracts must require vendors to submit comprehensive documentation of bias testing, impact assessments and compliance results, effectively making constitutional and administrative norms contractual obligations. Public reporting is crucial; executive orders mandating the publication of AI inventories keep citizens and legislators informed.

7. Conclusion

Public-sector AI holds exceptional promise for improving administrative efficiency but poses acute risks to core democratic values and legal rights. This comparative analysis demonstrates that the classic doctrines of due process, proportionality, non-discrimination and transparency cannot be bypassed when decision-making is automated. Every stage of the AI lifecycle must uphold these principles through explicit techno-legal integration. In design, systems must be risk-classified and mathematically bias-tested; in deployment, they require interpretable documentation and human review points; post-deployment, there must be continuous audits and robust grievance procedures.

To effectively govern AI in public administration, no single approach suffices. Relying purely on the EU's rigid hard-law model may hinder necessary public-sector modernization by imposing overwhelming compliance burdens, while trusting the US's deregulatory preemption risks widespread civil rights violations and corporate opacity. Conversely, Singapore's soft-law guidelines, while agile, lack the enforcement mechanisms necessary to guarantee strict accountability when fundamental rights are breached. Therefore, an optimal strategy must draw upon the strengths of these comparative regimes: combining legally binding risk classifications and human oversight mandates with agile, multidisciplinary technical standards and rigorous auditing pipelines.

Legal safeguards, such as judicial review and privacy laws, must be rendered meaningful in the algorithmic context. This demands that the code itself be engineered for compliance – utilizing interpretable models rather than flawed post hoc approximations, executing mathematically sound

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 09:26:50, subject to the Cambridge Core terms of use.

fairness optimizations and maintaining immutable audit trails. Only by institutionalizing hybrid law-and-code mechanisms, where legal experts continually inform technical practice and engineers strictly operationalize constitutional limits, can governments harness the vast benefits of AI without sacrificing justice, dignity and democratic control.

Acknowledgements. The authors would like to extend their sincere thanks to Associate Professor of Law Jason Grant Allen and Dr. Florian Gamper at Singapore Management University (SMU), Professor Jon Truby at the National University of Singapore (NUS) and Jane Loo at the University of Auckland, New Zealand, for their valuable support and insights. We also thank the organizers of the Cambridge Forum on AI: Law and Governance for providing a robust platform to discuss the operationalization of responsible AI within public administration.

Funding statement. The authors did not receive financial support or funding from any organization for the research, authorship and/or publication of this submitted work.

Competing interests. The authors have no financial or nonfinancial competing interests to declare that are relevant to the content of this article. There are no professional, contractual or personal relationships that could be perceived to exert an undue influence on the content or publication of this work.

References

AI Verify Foundation. (n.d.). *What is AI Verify?*. https://aiverifyfoundation.sg/what-is-ai-verify/

AI Verify Foundation & Infocomm Media Development Authority. (2024). *Model AI governance framework for generative AI*. https://aiverifyfoundation.sg/wp-content/uploads/2024/05/Model-AI-Governance-Framework-for-Generative-AI-May-2024-1-1.pdf

Allen, J. G., Loo, J., & Luna, J. (2025). Governing intelligence: Singapore's evolving AI governance framework. *Cambridge Forum on AI: Law and Governance*, *1*(e12), 1–20. https://doi.org/10.1017/cfl.2024.12

Bignami, F. (2022). Artificial intelligence accountability of public administration. *The American Journal of Comparative Law*, *70*(Suppl._1), i312–i346. https://doi.org/10.1093/ajcl/avac012

Bipartisan House Task Force on Artificial Intelligence. (2024). *Leading AI progress: Policy insights and a U.S. vision for AI adoption, responsible innovation, and governance*. U.S. House of Representatives. https://www.speaker.gov/wp-content/uploads/2024/12/AI-Task-Force-Report-FINAL.pdf

Burkart, N., & Huber, M. F. (2021). A survey on the explainability of supervised machine learning. *Journal of Artificial Intelligence Research*, *70*, 245–317. https://doi.org/10.1613/jair.1.12228.

Busuioc, M. (2021). Accountable artificial intelligence: Holding algorithms to account. *Public Administration Review*, *81*(5), 825–836. https://doi.org/10.1111/puar.13293

Castelnovo, A., Crupi, R., Greco, G., Regoli, D., Penco, I. G., & Cosentini, A. C. (2022). A clarification of the nuances in the fairness metrics landscape. *Scientific Reports*, *12*(1), 4209. https://doi.org/10.1038/s41598-022-07939-1.

Colombi Ciacchi, A., Flórez Rojas, M. L., Lane, L., & Nowak, T. (Eds.). (2025). *AI and public administration: The (legal) limits of algorithmic governance*. Università degli Studi di Trento; Universitat Pompeu Fabra. https://doi.org/10.15168/11572_461731

Council of Europe. (2024). *Council of Europe framework convention on artificial intelligence, human rights, democracy and the rule of law*. CETS No. 225. https://rm.coe.int/1680afae3c

Donnelly, J. (2025). Timely written reasons: Enhancing rationality, objectivity and legal accuracy in decision-making. *Alternative Law Journal*, *50*(1), 27–33. https://doi.org/10.1177/1037969X241305181

Dun & Bradstreet Austria and Others. (2025). (Case C-203/22, ECLI:EU:C:2025:117) (Court of Justice of the European Union, First Chamber, 27 February).

E-Governance Academy Foundation. (2025). February 1. *Unlocking AI's potential: A blueprint for governance and prosperity*.

Elzayn, H., Smith, E., Hertz, T., Guage, C., Ramesh, A., Fisher, R., Ho, D. E., & Goldin, J. (2025). Measuring and mitigating racial disparities in tax audits. *The Quarterly Journal of Economics*, *140*(1), 113–163. https://doi.org/10.1093/qje/qjae027

European Data Protection Supervisor. (2023). EDPS TechDispatch on explainable artificial intelligence. https://doi.org/10.2804/802043.

European Data Protection Supervisor. (2025). *TechDispatch #2/2025 - Human oversight of automated decision-making*. https://www.edps.europa.eu/data-protection/our-work/publications/techdispatch/2025-09-23-techdispatch-22025-human-oversight-automated-making_en

European Parliament & Council of the European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act). *Official Journal of the European Union*, L 2024/1689. https://eur-lex.europa.eu/eli/reg/2024/1689/oj

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 09:26:50, subject to the Cambridge Core terms of use.

Executive Order No. 14,365. (2025). Ensuring a National Policy Framework for Artificial Intelligence, 90 Fed. Reg. 58,499–58,501.

Frey, V., Wijker, D., & Hyee, R. (2025). *AI and the future of social protection in OECD countries*. OECD Publishing.

Garcia, A. C. B., Garcia, M. G. P., & Rigobon, R. (2024). Algorithmic discrimination in the credit domain: What do we know about it? *AI and Society*, *39*(4), 2059–2098. https://doi.org/10.1007/s00146-023-01676-3.

Global Partnership on Artificial Intelligence. (2024). *Algorithmic transparency in the public sector: A state-of-the-art report of algorithmic transparency instruments*.

Government Digital Service. (2025). *Artificial intelligence playbook for the UK government*. https://www.gov.uk/government/publications/ai-playbook-for-the-uk-government

Gugliotta, L., & Elbi, A. (2024). Will AI "subtly" take over decision-making in the EU migration context? Warnings and lessons from ETIAS and VIS. *European Papers*, *9*(3), 1018–1047. https://doi.org/10.15166/2499-8249/797

Harris, L. (2025). *Regulating artificial intelligence: U.S. and international approaches and considerations for Congress* (CRS Report No. R48555). Congressional Research Service.

House of Commons Education Committee. (2020). *Cancellation of 2020 exams and 2021 exams series*. https://committees.parliament.uk/publications/3401/documents/32495/default/

Jørgensen, B. N., & Ma, Z. G. (2025). Impact of EU regulations on AI adoption in smart city solutions: A review of regulatory barriers, technological challenges, and societal benefits. *Information*, *16*(7), 568. https://doi.org/10.2861/59990

Kemker, D. (2024). Do Black taxpayers matter? A critical tax analysis of IRS audit practices. *Stanford Journal of Civil Rights & Civil Liberties*, *20*, 133–178. https://law.stanford.edu/publications/do-black-taxpayers-matter-a-critical-tax-analysis-of-irs-audit-practices/

Koene, A., Clifton, C., Hatada, Y., Webb, H., Patel, M., Machado, C., LaViolette, J., & Richardson, R. (2019). *A governance framework for algorithmic accountability and transparency*. PE 624.262. European Parliament, European Parliamentary Research Service. https://doi.org/10.2861/59990

Krishna, S., Han, T., Gu, A., Wu, S., Jabbari, S., & Lakkaraju, H. (2024). *The disagreement problem in explainable machine learning: A practitioner's perspective*. *Transactions on Machine Learning Research*.

Matulionyte, R. (2024). Government automation, transparency, and trade secrets. *Melbourne University Law Review*, *47*(3), 679–722. https://doi.org/10.2139/ssrn.4771120

National Institute of Standards and Technology. (2023). *Artificial intelligence risk management framework (AI RMF 1.0)* (NIST AI 100-1). https://doi.org/10.6028/NIST.AI.100-1

Nordic Council of Ministers. (2024). *Public digitalisation in a legal perspective: Status, new challenges and perspectives for Nordic-Baltic cooperation*. https://pub.norden.org/temanord2024-503/

Nouws, S. J. J., Dobbe, R. I. J. (2024). The rule of law for artificial intelligence in public administration: A system safety perspective. In K. Prifti, E. Demir, J. Krämer, K. Heine, & E. Stamhuis (Eds.), *Digital governance: Confronting the challenges posed by artificial intelligence* (pp. 183–208). T.M.C. Asser Press. https://doi.org/10.1007/978-94-6265-639-0_9

OECD. (2024). *Governing with artificial intelligence: Are governments ready?* OECD Artificial Intelligence Papers, No. 20. https://doi.org/10.1787/26324bc2-en

OECD. (2025c). *Sharing trustworthy AI models with privacy-enhancing technologies* (OECD Artificial Intelligence Papers, No. 38. OECD Publishing. https://doi.org/10.1787/a266160b-en

Office of Management and Budget. (2024). *Advancing governance, innovation, and risk management for agency use of artificial intelligence* (OMB Memorandum M-24-10). https://www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf

Office of Management and Budget. (2025). *Accelerating federal use of AI through innovation, governance, and public trust* (Memorandum M-25-21). https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-21-Accelerating-Federal-Use-of-AI-through-Innovation-Governance-and-Public-Trust.pdf

Ofqual. (2020a). *Standardisation of grades in general qualifications in summer 2020: Outliers*.

Organisation for Economic Co-operation and Development, & United Nations Educational, Scientific and Cultural Organization. (2024). *G7 toolkit for artificial intelligence in the public sector*. OECD/UNESCO.

Pessach, D., & Shmueli, E. (2022). A review on fairness in machine learning. *ACM Computing Surveys*, *55*(3), Article3494672. https://doi.org/10.1145/3494672

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act). (2024). Official Journal of the European Union. L, 2024/1689, 12 July 2024.

SCHUFA Holding and Others (Scoring) (2023). Case C-634/21, ECLI:EU:C:2023:957 (Court of Justice of the European Union, December 7, 2023).

Slack, D., Hilgard, S., Jia, E., Singh, S., & Lakkaraju, H. (2020). Fooling LIME and SHAP: Adversarial attacks on post hoc explanation methods. In *Proceedings of the 2020 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 180–186). Association for Computing Machinery. https://doi.org/10.1145/3375627.3375830

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 09:26:50, subject to the Cambridge Core terms of use.

Solaiman, B. (2025). The European Union's Artificial Intelligence Act and trust: Towards an AI Bill of Rights in healthcare? *Law, Innovation and Technology*, *17*(1), 318–334. https://doi.org/10.1080/17579961.2025.2469986

Students for Fair Admissions, Inc. v. President and Fellows of Harvard College. (2023). [600 U.S. 181].

Taxation Act. (2024). Riigi Teataja, RT I, 05.08.2024, 3, § 46^2.

The White House. (2026). *National policy framework for artificial intelligence: Legislative recommendations*.

Trump, D. J. (2025). Executive Order 14365—Ensuring a national policy framework for artificial intelligence. *Federal Register*, *90*, 58499–58501. https://www.whitehouse.gov/presidential-actions/2025/12/eliminating-state-law-obstruction-of-national-artificial-intelligence-policy/

Ullits, J. (2025). The justification machine. Automating reason-giving. *European Review of Digital Administration & Law (Erdal)*, *6*(2), 23–40. https://doi.org/10.53136/97912218237763

UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. https://www.unesco.org/en/artificial-intelligence/recommendation-ethics

United Nations Department of Economic and Social Affairs. (2024). *E-government survey 2024: Accelerating digital transformation for sustainable development: With the addendum on artificial intelligence*. United Nations. https://publicadministration.un.org/egovkb/en-us/

Weerts, H., Xenidis, R., Tarissan, F., Olsen, H. P., & Pechenizkiy, M. (2024). The neutrality fallacy: When algorithmic fairness interventions are (not) positive action. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (pp. 2060–2070). ACM. https://doi.org/10.1145/3630106.3659025

Williams, R. (2022). Rethinking administrative law for algorithmic decision making. *Oxford Journal of Legal Studies*, *42*(2), 468–494. https://doi.org/10.1093/ojls/gqab032

Wolswinkel, J. (2022). *Artificial intelligence and administrative law: Comparative study on administrative law and the use of artificial intelligence and other algorithmic systems in administrative decision-making in the member states of the Council of Europe*. (CDCJ(2022)31). Council of Europe.

Wong, D. (2025). *Opening remarks at PECC 2025 Conference on "Asia-Pacific AI Governance Accelerator and Preparing for AI Transformation"* [Speech]. Infocomm Media Development Authority.

World Economic Forum, Infocomm Media Development Authority, & Personal Data Protection Commission. (2020). *Companion to the model AI governance framework: Implementation and self-assessment guide for organizations*. World Economic Forum.

Prema Elumalai is a Professor of Law at the VIT School of Law, Chennai, with over 16 years of rich academic expertise specializing in Constitutional Law and Public International Law. Currently pursuing an LL.D. postdoctoral degree, her research critically examines the interface between emerging digital technologies, international regulatory frameworks and domestic governance systems. Dr. Prema is a highly prolific scholar with extensive global exposure, including institutional research accolades, fully funded United Nations panel selections and paper presentations at prestigious legal forums across Europe and the United States. Her recent scholarship focuses heavily on techno-legal architectures, algorithmic accountability and responsible public-sector AI governance. Having previously addressed the Cambridge Forum on AI and Law on operationalizing responsible AI in public administration, her deep expertise in constitutional due process makes her an asset to contemporary global dialogues on automated state frameworks.

Ragul Olakkur Vijayakumar is an accomplished Advocate at the High Court of Madras and a prominent legal researcher specializing in public law, constitutional jurisprudence and emerging technology regulations. He graduated with top academic honors and consecutive Meritorious Awards from the VIT School of Law. Ragul possesses deep expertise in the systemic intersections of public international law and algorithmic governance, maintaining a highly prolific publication record spanning decentralized blockchain arbitration, digital infrastructure security and AI-driven healthcare liability in platforms like the UIC Law Review. As an expert in techno-legal compliance architectures for public governance, he has actively contributed findings to the Cambridge Forum on AI and Law. Ragul's unique combination of high-court appellate research insight and cutting-edge algorithmic risk analysis enables him to offer practical, cross-disciplinary solutions for embedding legal safeguards directly into technical code environments.

Cite this article: Elumalai, P., & Olakkur Vijayakumar, R. (2026). Between law and code: Operationalizing responsible AI in public administration. *Cambridge Forum on AI: Law and Governance*, *2*, e14, 1–14. https://doi.org/10.1017/cfl.2026.10055

Downloaded from https://www.cambridge.org/core. 12 Nov 2026 at 09:26:50, subject to the Cambridge Core terms of use.