Journal of Public Policy and Local Government (JPPLG)
https://journal.ashapublishing.co.id/index.php/jpplg/
E-ISSN: 3047-6984
Volume 3, Issue 1, 2026
Pages: 12-22

Governing Algorithmic Fairness in Climate-Health Systems: A Policy Framework for Bias Mitigation in Public Sector Decision-Making

Prince Arthur
University of Utah
Corresponding Author's Email: p.arthur@utah.edu

ABSTRACT

Artificial intelligence (AI) systems are increasingly integrated into climate-health decision-making processes across the United States public sector, offering considerable capacity to model complex environmental and public health interactions while supporting resource allocation and policy planning. Despite these technological advances, the deployment of AI in this domain raises significant concerns regarding algorithmic bias, which can systematically disadvantage vulnerable populations and undermine the equity objectives of climate-health governance. The existing literature reveals a critical gap: no comprehensive policy framework has been specifically designed to govern algorithmic fairness within the climate-health nexus of American public sector decision-making. This article addresses that gap by developing an original, structured policy framework designated the Algorithmic Fairness Climate-Health Governance (AFCHG) Framework. Drawing on a systematic synthesis of peer-reviewed scholarship across AI governance, algorithmic fairness, public administration, and climate-health systems, this article employs a conceptual research design grounded in thematic analysis and policy analysis methodology. Conceptual frameworks play a critical role in emerging policy domains where empirical evidence remains fragmented, providing structured guidance that can subsequently be tested through applied research. The AFCHG Framework comprises four interdependent pillars: Policy and Legal Foundations, Governance and Accountability, Technical Bias Mitigation, and Ethics, Equity, and Inclusion. Each pillar is theorized in relation to existing scholarly evidence and grounded in the governance realities of United States public sector institutions. The article further proposes an eight-step governance flow for AI bias mitigation, supported by two summary tables that translate the framework into actionable policy guidance. The findings suggest that addressing algorithmic bias in climate-health systems requires not merely technical solutions but coordinated, multi-level governance architectures embedded within broader equity and human rights frameworks. This work contributes a theoretically rigorous and practically applicable framework for American policymakers, public administrators, and AI governance scholars.

Keywords: algorithmic fairness, AI governance, climate-health systems, public sector decision-making, bias mitigation, policy framework, United States

INTRODUCTION

The intersection of artificial intelligence (AI), climate change, and public health represents one of the most consequential governance frontiers of the twenty-first century. Across the United States, federal and state agencies are increasingly deploying AI-driven systems to enhance decision-making in climate-health management, including tools used by the Centers for Disease Control and Prevention (CDC), the Environmental Protection Agency (EPA), the

DOI: https://doi.org/10.70188/tmar1n73

Federal Emergency Management Agency (FEMA), and the Department of Health and Human Services (HHS) to allocate climate adaptation resources across vulnerable communities. While these innovations carry immense promise, significant risks remain that have yet to receive adequate policy attention. Chief among these risks is algorithmic bias, the systematic tendency of AI systems to produce discriminatory outputs that disadvantage particular social groups, frequently along lines of race, geography, income, and gender.

Algorithmic bias refers to the systematic deviation of an AI system's outputs from equitable standards due to flawed data inputs, biased training sets, or structurally embedded assumptions within model design (Chen et al., 2023; Green, 2022). In the context of public sector decision-making, the consequences are especially profound because algorithmic outputs carry the authority of institutional decisions. Alon-Barkat and Busuioc (2023) demonstrated that public sector decision-makers exhibit patterns of selective adherence to algorithmic advice, meaning that biased recommendations can be reinforced through the very human-AI interaction dynamics intended to serve as corrective safeguards. This compounding effect is particularly dangerous in high-stakes, resource-constrained governance environments, where algorithmic errors translate directly into inequitable distributions of public health resources and climate adaptation investments.

Climate change is increasingly recognized as a fundamental driver of public health outcomes across the United States, amplifying vulnerabilities among marginalized populations through heat-related illness, vector-borne disease spread, food insecurity, and displacement (Ansah et al., 2024; Martins et al., 2024). When AI systems are deployed to manage these climate-health interactions, algorithmic bias can translate directly into inequitable health outcomes. Mosadeghrad et al. (2023) underscored the need for climate-resilient health systems responsive to structural vulnerabilities, yet the role of algorithmic governance in ensuring such responsiveness remains theoretically underdeveloped within the American policy context. Birkstedt et al. (2023), Roberts et al. (2024), and Papagiannidis et al. (2023) each identified governance gaps surrounding AI in public institutions, including fragmented regulatory frameworks, unclear accountability structures, and the absence of binding equity-focused instruments. These gaps are especially pronounced within the United States federal system, where AI adoption in climate-health governance has proceeded rapidly without commensurate regulatory development.

Existing frameworks for algorithmic fairness tend to be technical in orientation, focusing on mathematical definitions of fairness that, as Green (2022) demonstrated, are often mutually incompatible and insufficient to address substantive equity concerns in public sector contexts. Wilson and van der Velden (2022) proposed a sustainable AI model for public sector decision-making, yet the specific domain of United States climate-health governance has not been addressed in a dedicated policy framework. This article addresses that gap by developing the Algorithmic Fairness Climate-Health Governance (AFCHG) Framework, an original multi-pillar architecture that is simultaneously theoretically grounded and practically applicable. The article makes three primary contributions: it establishes conceptual linkages between algorithmic bias, climate-health governance, and public sector equity; it proposes an original, named policy framework; and it provides an eight-step governance flow supported by tables that translate the framework into operational guidance for American policymakers and public administrators.

DOI: https://doi.org/10.70188/tmar1n73

METHODS

This article employs a conceptual research design, appropriate when the objective is to develop theoretical frameworks, synthesize existing knowledge, and propose policy architectures rather than to collect primary empirical data (Di Vaio et al., 2021; Wilson and van der Velden, 2022). Conceptual frameworks play a critical role in emerging policy domains where empirical evidence remains fragmented, providing structured guidance that can subsequently be tested through applied research. This is particularly salient in algorithmic governance within climate-health systems, where the pace of AI deployment in United States federal agencies has substantially outstripped the development of evidence-based governance standards. The methodology is grounded in three complementary processes: a systematic literature review and synthesis, a policy analysis approach, and a thematic framework construction process.

The literature review covered five primary domains: algorithmic bias and fairness in AI systems; AI governance in public sector contexts; climate-health systems and resilience; public sector decision-making and human-AI interaction; and equity, ethics, and inclusion in computational systems. Sources were selected on the basis of relevance, methodological quality, and disciplinary diversity, drawing from leading journals in public administration, philosophy and technology, health sciences, computer science, and governance studies, with priority given to peer-reviewed articles published within the last five years.

The framework construction process followed a structured thematic synthesis methodology operationalized through four distinct analytical stages. In the first stage, all reviewed sources were subjected to thematic coding, with recurring governance-relevant themes extracted and recorded, including regulatory mandates for AI equity, institutional accountability mechanisms, technical fairness interventions, and normative equity principles. In the second stage, coded themes were systematically grouped into clusters based on conceptual affinity and governance function, guided by the criterion that each cluster must address a distinct and non-overlapping governance dimension. In the third stage, each cluster was evaluated against the policy literature on public sector AI governance to assess operational feasibility within the institutional architecture of the United States federal system. In the fourth stage, validated clusters were elevated into the four pillars of the AFCHG Framework, with each pillar assigned specific policy components derived from the scholarly evidence and arranged in a logical governance sequence. The framework was then subjected to internal consistency review to ensure conceptual coherence and comprehensive coverage without significant overlap.

The thematic synthesis produced four discrete cluster groupings that correspond directly to the AFCHG Framework's four pillars. The first cluster, aggregating themes of regulatory mandate, impact assessment, and anti-discrimination law, constituted Pillar I. The second cluster, grouping accountability structures, oversight mechanisms, and transparency requirements, formed Pillar II. The third cluster, consolidating pre-processing, in-processing, and post-processing technical interventions alongside monitoring protocols, produced Pillar III. The fourth cluster, synthesizing intersectionality, vulnerable population protection, and participatory governance norms, yielded Pillar IV. During the Stage 3 feasibility evaluation, several emergent themes, including blockchain-based audit trails and automated penalty enforcement, were assessed as operationally premature within current United States federal institutional architecture and were consequently excluded from the framework's core pillars, though acknowledged as directions for future development. This exclusion process ensured

DOI: https://doi.org/10.70188/tmar1n73

that the AFCHG Framework remains practically actionable within existing governance constraints rather than aspirationally over-specified.

The policy analysis approach examined the governance landscape for AI in public sector decision-making across federal and state levels, reviewing frameworks from NIST, the Office of Management and Budget (OMB), and the White House Office of Science and Technology Policy (OSTP). This analysis informed the governance flow and the United States contextual layer embedded within the AFCHG Framework. The methodology produces no empirical results in the conventional sense. The contribution is conceptual: an original, evidence-grounded framework that organizes existing knowledge into a governance architecture with direct policy applicability, and a research agenda pointing toward empirical work that must follow in future scholarship.

RESULT AND DISCUSSION

Algorithmic Bias in the United States Climate-Health Context

Algorithmic bias operates across the full AI lifecycle. Drawing on the taxonomy established by Chen et al. (2023), bias can be classified as pre-processing, in-processing, or post-processing, each presenting distinct governance challenges. Pre-processing bias is particularly consequential in American climate-health contexts because historical health and environmental data reflect structural inequality, including uneven monitoring infrastructure and the underrepresentation of racial and ethnic minorities in health databases (Xu et al., 2022). This is illustrated by the EPA's EJScreen mapping tool: when AI systems allocating environmental health resources are trained on monitoring data that disproportionately represents affluent communities with denser sensor networks, the resulting algorithms systematically underestimate pollution exposure in underserved areas, directing fewer resources where resources are most needed (Fabris et al., 2022; Chen et al., 2023).

In-processing bias reflects normative assumptions embedded in model objectives. When systems optimize for aggregate efficiency, populations whose health outcomes are structurally more costly to improve may be systematically deprioritized, reflecting what Green (2022) described as the tension between formal and substantive algorithmic fairness. The CDC's disease surveillance analytics illustrate this pattern: models optimized for population-level accuracy can perform poorly for indigenous communities and recent immigrant populations whose health profiles are underrepresented in training data, precisely the accuracy-fairness trade-off documented by Dehghani et al. (2025). Post-processing bias is mediated by selective adherence dynamics (Alon-Barkat and Busuioc, 2023): FEMA's disaster relief allocation systems illustrate how officials may consistently adopt algorithmic recommendations for high-property-value areas while applying less consistent adoption for low-income communities, laundering institutional bias through an algorithmic interface. The climate-health literature underscores these equity stakes: Ansah et al. (2024), Smith et al. (2025), and Martins et al. (2024) collectively demonstrate that the double burden of climate change falls disproportionately on communities already facing the greatest health inequities in the United States, establishing the equity imperatives embedded in the AFCHG Framework.

The AFCHG Framework: Architecture and Interdependencies

The AFCHG Framework is an original, multi-pillar policy framework for governing algorithmic bias in AI-driven climate-health systems used in United States public sector decision-making. Figure 1 presents its full visual architecture. The framework's design reflects a deliberate governance layering logic: the four pillars are not independent interventions but

DOI: https://doi.org/10.70188/tmar1n73

a sequential and mutually reinforcing architecture in which each pillar creates the enabling conditions for the next. Pillar I establishes the mandatory regulatory environment without which the other pillars lack institutional authority. Pillar II builds on that legal foundation by providing the oversight mechanisms through which legal mandates are operationalized and enforced. Pillar III operates within those governance structures, providing the technical tools that oversight mechanisms need to perform meaningful scrutiny. Pillar IV provides the normative compass that determines the objectives of all three preceding pillars, specifying what counts as a fair outcome and whose interests must be centered. The integration layer coordinates pillar interactions, aligns the framework with NIST AI Risk Management Framework standards and federal equity executive orders, and provides the capacity-building infrastructure that sustains implementation. Table 1 summarizes the pillars, components, policy objectives, and governance flow step linkages.

The integration layer warrants particular analytical attention as it performs three distinct governance functions that Figure 1 encodes visually but that require explicit elaboration. First, it provides cross-pillar coordination by establishing shared definitions, common metrics, and joint reporting protocols that prevent the four pillars from operating as parallel silos within federal agency structures. In practice, Algorithmic Impact Assessments mandated by Pillar I must use the fairness metrics defined by Pillar IV, evaluated through the oversight bodies established by Pillar II, using the technical tools specified by Pillar III. This coordination demand is institutionalized by the integration layer as a mandatory governance requirement rather than an aspirational norm. Second, the integration layer operationalizes alignment with the NIST AI Risk Management Framework's four core functions, specifically Govern, Map, Measure, and Manage, providing agencies already invested in NIST-aligned programs with a direct adoption pathway. Third, the integration layer provides the capacity-building infrastructure that makes the framework sustainable across electoral cycles and budget fluctuations, encompassing training programs, inter-agency knowledge transfer, and technical assistance for state and local agencies whose resources differ significantly from those of federal agencies.

The interdependencies are further clarified by considering what happens when any single pillar is absent. Without Pillar I, Pillar II's mechanisms lack legal authority and become voluntary, a condition the literature consistently identifies as insufficient to produce equitable AI outcomes at scale (Papagiannidis et al., 2023; Roberts et al., 2024). Without Pillar II, Pillar III's technical interventions are conducted without independent scrutiny, creating the risk of agency self-certification of fairness, analogous to allowing financial institutions to conduct internal regulatory audits. Without Pillar III, Pillar II's oversight bodies possess no technical tools with which to evaluate AI systems, rendering oversight substantively empty. Without Pillar IV, the entire framework risks optimizing for narrow formal definitions of fairness that satisfy regulatory requirements while leaving substantive equity harms intact, precisely the failure mode that Green (2022) identified as the central limitation of existing algorithmic fairness approaches.

Table 1: Summary of the AFCHG Framework: Pillars, Components, Policy Objectives, and Governance Flow Linkages

Pillar	Components	Policy Objective	Key Citations	Governance Flow Steps
I. Policy and Legal Foundations	Algorithmic Impact Assessment; Anti-Discrimination Legislation; Mandatory Bias Audits; AI Deployment Standards	Establish a regulatory environment mandating equity assessments prior to AI deployment in public health and climate governance	Alrawahna et al. (2025); Green (2022); Maes et al. (2020)	Steps 1 and 2
II. Governance and Accountability	Multi-Stakeholder Oversight; Transparent Decision Logging; Independent Review Panels; Community Feedback Mechanisms	Ensure institutional accountability through structured oversight, independent auditing, and meaningful community participation	Papagiannidis et al. (2023); Roberts et al. (2024); Alon-Barkat and Busuioc (2023)	Steps 3 and 5
III. Technical Bias Mitigation	Pre-processing Detection; In-processing Fairness Constraints; Post-processing Calibration; Continuous Monitoring	Operationalize bias detection and correction across the full AI lifecycle, from training data through output and deployment monitoring	Chen et al. (2023); Dehghani et al. (2025); Wang and Singh (2025); Ashokan and Haas (2021)	Step 4
IV. Ethics, Equity, and Inclusion	Intersectional Equity Analysis; Vulnerable Population Protections; Inclusive Data Collection; Community-Led Validation	Embed substantive equity imperatives into all governance decisions, ensuring AI systems serve populations most vulnerable to climate-health risks	Green (2022); Xu et al. (2022); Erman et al. (2025); Yilma (2025)	Steps 6 to 8

Source: Author (2025). Synthesized from Chen et al. (2023); Green (2022); Papagiannidis et al. (2023); Roberts et al. (2024); Dehghani et al. (2025).

a. Pillar I: Policy and Legal Foundations

Pillar I establishes the policy and legal foundations for governing algorithmic fairness in United States climate-health systems. Alrawahna et al. (2025) documented that the benefits of AI in public sector institutions are most reliably realized when deployment is governed by clear policy frameworks rather than agency discretion. A concrete illustration of the current Pillar I deficit is the absence of a mandatory Algorithmic Impact Assessment (AIA) requirement at the federal level: the EPA's Office of Environmental Justice has developed voluntary equity screening tools, but without mandatory AIA requirements, agencies can deploy AI systems that have not been evaluated for disparate impact on communities of color or low-income populations. The AIA concept adapts established practices from environmental impact assessment to algorithmic systems, requiring agencies to document training dataset demographics, model outcome distributions across social groups, identify vulnerability hotspots, and consult affected communities before deployment. Maes et al. (2020) demonstrated that systematic evaluation processes improve both the quality and equity of institutional choices. Anti-discrimination legislation alignment, mandatory bias audit protocols, and AI deployment standards complete the Pillar I architecture, creating the enabling environment within which the other three pillars function.

To illustrate Pillar I in operation: under the AFCHG Framework, the EPA's deployment of AI-driven environmental justice screening tools would require a mandatory AIA to be filed before deployment, documenting the demographic composition of training datasets, auditing historical monitoring coverage across census tracts stratified by race and income, and publishing outcome distributions across social groups. This assessment would be subject to public comment and independent review under Pillar II before the system proceeds to bias mitigation intervention at Step 4. The current voluntary architecture of EJScreen, in which equity evaluation is discretionary rather than mandatory, represents precisely the Pillar I gap the framework addresses. The contrast between existing practice and the AIA requirement illustrates the governance distance the framework asks agencies to close.

b. Pillar II: Governance and Accountability

Pillar II addresses the governance and accountability structures necessary to operationalize Pillar I mandates. Di Vaio et al. (2021) demonstrated that human-AI decision-making effectiveness in public organizations is strongly conditioned by governance structure quality, and within the American federal system, the multiplicity of agencies involved in climate-health governance creates coordination challenges that can easily produce accountability gaps. HHS's predictive risk scoring systems, used in several states to allocate child welfare interventions, illustrate this: independent analyses have identified racial disparities in predicted risk scores, but in the absence of mandatory multi-stakeholder oversight and independent review panels, these disparities have persisted across multiple deployment cycles without triggering corrective action (Alon-Barkat and Busuioc, 2023; Ochmann et al., 2024). Multi-stakeholder oversight bodies, transparent decision logging, independent algorithmic review panels, and community feedback mechanisms complete the Pillar II architecture. These mechanisms are mutually reinforcing: oversight provides strategic direction, logging provides evidentiary accountability, review panels provide technical scrutiny, and community feedback ensures responsive governance connected to lived experience.

Applying Pillar II to the HHS child welfare scoring scenario: the framework would require, first, that a multi-stakeholder oversight body comprising representatives of affected communities, civil rights organizations, independent data scientists, and HHS administrators formally review the system's equity metrics on a defined schedule. Second, transparent decision logging would require that every algorithmic recommendation and the human decision recorded following that recommendation be made available for audit, creating an evidentiary trail that enables independent analysts to detect patterns of selective adherence of the type documented by Alon-Barkat and Busuioc (2023). Third, an independent Algorithmic Review Panel convened at Step 5 would audit these records, assess whether racial disparities in predicted risk scores persist, and issue mandatory corrective recommendations. This institutional architecture transforms accountability from a retrospective complaint mechanism into a prospective governance requirement.

c. Pillar III: Technical Bias Mitigation

Pillar III addresses the technical interventions necessary to detect and mitigate algorithmic bias at each stage of the AI lifecycle. Pre-processing methods identify and correct bias in training data before model development; Ashokan and Haas (2021) demonstrated that resampling and synthetic data augmentation can significantly reduce disparate impact in outputs. In CDC disease surveillance AI, this would involve auditing historical datasets for demographic coverage gaps, partnering with community health organizations to supplement official data sources, and applying reweighting techniques to ensure proportional minority representation in model training. These steps are governed by Pillar I's AIA requirements and subject to Pillar II's oversight mechanisms, illustrating the framework's layered architecture in practice. In-processing fairness constraints, incorporated directly into model training, impose mathematical requirements on outputs; Yang and Lee (2024) demonstrated effectiveness in improving equity without unacceptable accuracy losses, while Erman et al. (2025) argued that such constraints must be grounded in a coherent theory of algorithmic justice guided by Pillar IV's substantive equity goals. Post-processing calibration addresses residual output bias, while continuous monitoring recognizes that bias is not a static property. Al-Zawqari et al. (2024) demonstrated that latent bias can emerge and intensify over time, underscoring ongoing monitoring as a governance requirement rather than an optional enhancement.

d. Pillar IV: Ethics, Equity, and Inclusion

Pillar IV establishes the normative foundation of the AFCHG Framework. Green (2022) argued that formal algorithmic fairness consistently fails to address the substantive inequities arising from the social hierarchies in which AI systems are embedded. The disproportionate health burden borne by Black and Latino communities in urban heat islands exemplifies this: AI tools designed to optimize city-wide heat mitigation investments may formally treat all census tracts equally while systematically underinvesting in communities where housing quality, outdoor labor, and limited green space create the greatest heat vulnerability (Xu et al., 2022; Martins et al., 2024). Intersectional equity analysis, inclusive data collection protocols, vulnerable population protection clauses, and community-led validation processes operationalize the pillar's commitments. Mosadeghrad et al. (2024) demonstrated that governance frameworks must explicitly account for differential population vulnerability to climate-health risks. Community-led validation, implemented through participatory governance models where community representatives assess whether algorithmic recommendations align with experienced needs and priorities, receives empirical support from Ochmann et al. (2024), who found that transparency and community engagement significantly improve both perceived fairness and governance quality. Pillar IV provides both the normative substance that gives the other pillars a defined purpose and the evaluative criteria by which the framework's success must ultimately be assessed.

The Governance Flow: Operational Pathway and Figure Interpretation

The AFCHG Framework's four pillars are operationalized through an eight-step governance flow illustrated in Figure 2 and summarized in Table 2. Understanding this flow requires attending not only to its sequential structure but to the governance logic that makes that sequencing necessary, and to the explicit connections between each step and the pillar architecture of Figure 1.

DOI: https://doi.org/10.70188/tmar1n73

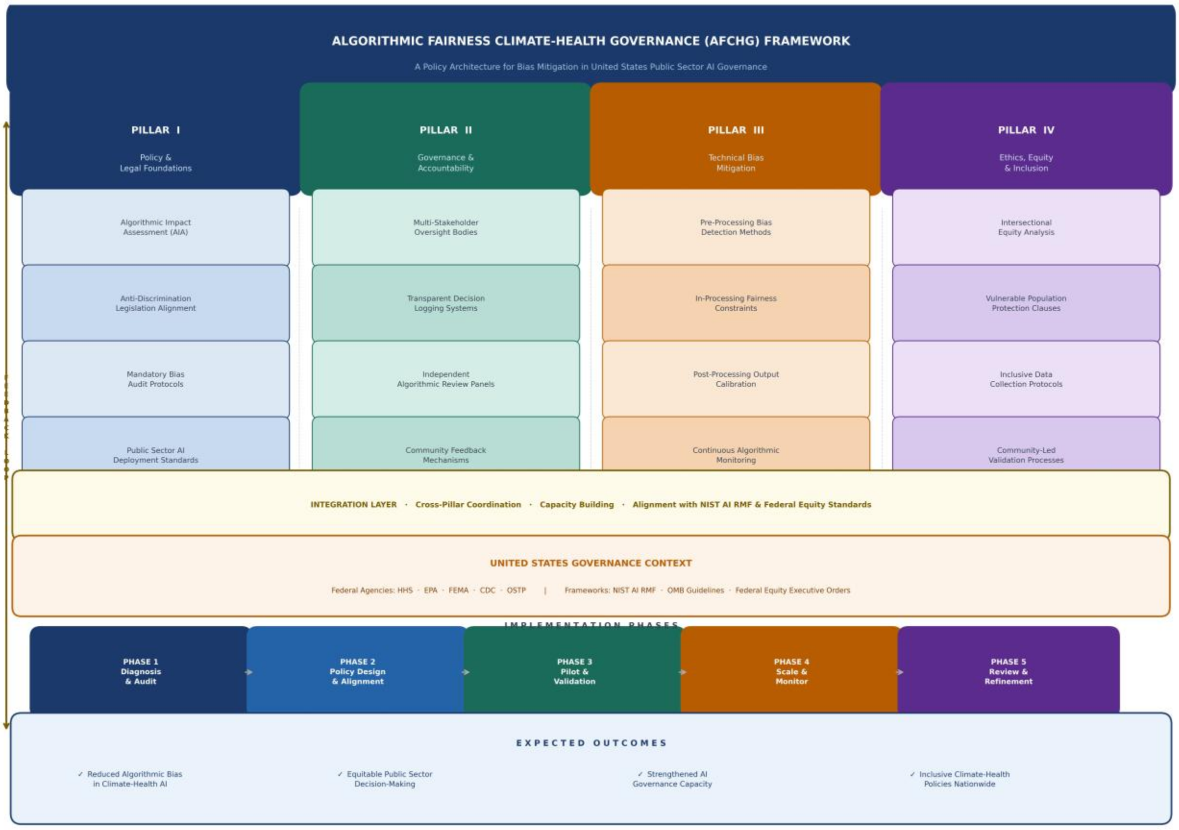

Figure 1. The Algorithmic Fairness Climate-Health Governance (AFCHG) Framework
Source: Author (2025), synthesized from Alrawahna et al. (2025); Green (2022); Papagiannidis et al. (2023); Roberts et al. (2024); Chen et al. (2023); Dehghani et al. (2025); Wang and Singh (2025); Xu et al. (2022); Erman et al. (2025); Alon-Barkat and Busuioc (2023); Wilson and van der Velden (2022).

Step 1, AI System Identification, establishes the governance baseline by requiring agencies to formally register all AI deployments in climate-health decision-making, activating the AIA mandate of Pillar I. Step 2, Bias Risk Assessment, implements that AIA mandate in full, requiring structured equity evaluation using the substantive fairness criteria of Pillar IV before any system proceeds. The deliberate placement of stakeholder consultation at Step 3, prior to bias mitigation, reflects the framework's commitment to community participation as a prerequisite for legitimate governance rather than a post-hoc exercise, operationalizing Pillar II's community feedback mechanisms and Pillar IV's inclusive validation principles. Step 4, Bias Mitigation Intervention, draws on the full technical toolkit of Pillar III, with the specific intervention selected based on the bias type and governance context identified in Steps 2 and 3. Step 5, Independent Validation, implements Pillar II's independent review panel mechanism, providing external scrutiny necessary to prevent agency self-certification. Steps 6 and 7, Policy Integration and Deployment with Monitoring, translate validated standards into regulatory practice and establish the continuous monitoring infrastructure that Pillar III requires. Step 8, Review and Iteration, closes the governance loop.

The feedback loop connecting Step 8 back to Step 2 is the most structurally significant feature of Figure 2 and merits explicit analytical attention. It transforms the governance flow from a linear compliance exercise, one completed once and filed, into an adaptive governance system that responds to the dynamic character of both AI technology and climate-health risk. The feedback loop operates through three specific mechanisms. First, community feedback gathered at Step 8 is formally re-entered into the Bias Risk Assessment at Step 2 in subsequent

DOI: https://doi.org/10.70188/tmar1n73

cycles, ensuring that lived-experience evidence shapes technical evaluation rather than remaining confined to consultation exercises. Second, updated equity evidence produced by the continuous monitoring of Step 7 is incorporated into the AIA criteria governing Step 2, so that emerging bias patterns detected post-deployment trigger reassessment before compounding. Third, model recalibration findings from Step 8 inform the selection of bias mitigation strategies at Step 4 in subsequent cycles, creating a technical learning loop governed by the equity imperatives of Pillar IV. Without this feedback architecture, the governance flow would address bias only at the point of initial deployment, leaving systems vulnerable to performance drift as climate conditions, demographic compositions, and data environments change over time, a vulnerability that Al-Zawqari et al. (2024) identified as a persistent failure mode in algorithmic governance.

The relationship between Figure 1 and Figure 2 is one of architectural specification and operational enactment. Figure 1 specifies the governance architecture, encompassing the four pillars, the associated components, and the interdependencies among pillars, while Figure 2 enacts that architecture as an operational sequence of steps distributed across time and institutional actors. The "Governance Flow Steps" column in Table 1, added in this revision, makes the pillar-to-step mapping explicit: Pillar I governs Steps 1 and 2; Pillar II's mechanisms are engaged at Steps 3 and 5 and embedded in the monitoring of Step 7; Pillar III's technical toolkit is deployed at Step 4; and Pillar IV's normative criteria inform Steps 2, 3, 6, and 8, with the evaluative function of Pillar IV active throughout. Neither figure is interpretively complete without the other. Figure 1 without Figure 2 specifies what governance architecture is required but not how that architecture is sequenced in practice. Figure 2 without Figure 1 specifies operational steps without identifying the governance logic that determines the order, content, and accountability requirements of each step. Read together, the two figures constitute a unified framework in which theoretical architecture and operational pathway are mutually specified and analytically inseparable.

DOI: https://doi.org/10.70188/tmar1n73

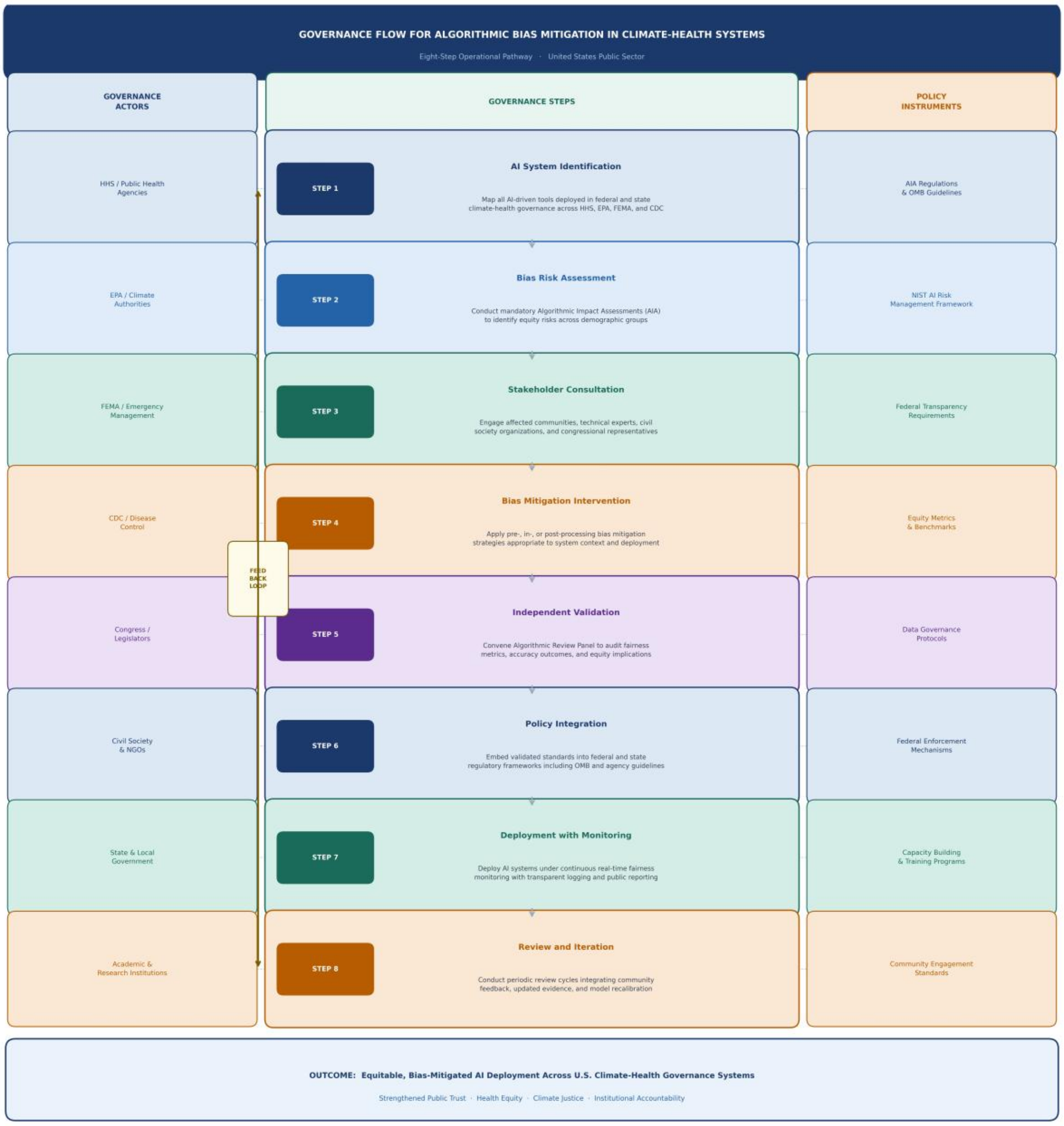

Figure 2. Governance Flow for Algorithmic Bias Mitigation in Climate-Health Systems
Source: Author (2025), synthesized from Birkstedt et al. (2023); Papagiannidis et al. (2023); Wilson and van der Velden (2022); Roberts et al. (2024); Ochmann et al. (2024); Chen et al. (2023); Mosadeghrad et al. (2023); Al-Zawqari et al. (2024); Smith et al. (2025); Alrawahna et al. (2025).

DOI: https://doi.org/10.70188/tmar1n73

Table 2. Eight-Step Governance Flow: Actions, Descriptions, Supporting Evidence, and Pillar Linkages

Step	Action	Description	Supporting Evidence	Pillar Linkage
1	AI System Identification	Map all AI-driven tools deployed in U.S. federal and state climate-health governance, including systems used by HHS, EPA, FEMA, and CDC	Wilson and van der Velden (2022); Alrawahna et al. (2025)	Pillar I: AIA mandate activation
2	Bias Risk Assessment	Conduct mandatory Algorithmic Impact Assessments to identify equity risks across demographic groups before and during deployment	Maes et al. (2020); Green (2022); Alon-Barkat and Busuioc (2023)	Pillar I and Pillar IV: equity criteria
3	Stakeholder Consultation	Engage affected communities, technical experts, civil society organizations, and congressional representatives in governance planning	Roberts et al. (2024); Papagiannidis et al. (2023)	Pillar II: community mechanisms and Pillar IV
4	Bias Mitigation Intervention	Apply pre-processing, in-processing, or post-processing bias mitigation strategies appropriate to the system and deployment context	Chen et al. (2023); Dehghani et al. (2025); Wang and Singh (2025)	Pillar III: full technical toolkit
5	Independent Validation	Convene an Algorithmic Review Panel of technical experts, ethicists, and health equity scholars to audit fairness metrics and accuracy outcomes	Birkstedt et al. (2023); Fabris et al. (2022)	Pillar II: independent review panels
6	Policy Integration	Embed validated algorithmic standards into federal and state regulatory frameworks, including OMB guidelines and agency procurement standards	Papagiannidis et al. (2023); Xu et al. (2024)	Pillar I and Pillar II: regulatory embedding
7	Deployment with Monitoring	Deploy AI systems under continuous real-time fairness monitoring with transparent decision logging and public reporting requirements	Ochmann et al. (2024); Di Vaio et al. (2021)	Pillar II: logging and Pillar III: monitoring
8	Review and Iteration	Conduct periodic review cycles integrating community feedback, updated equity evidence, and model recalibration; feeds back to Step 2	Mosadeghrad et al. (2023); Smith et al. (2025)	All pillars: adaptive feedback loop

Source: Author (2025). Synthesized from Alon-Barkat and Busuioc (2023); Birkstedt et al. (2023); Wilson and van der Velden (2022); Mosadeghrad et al. (2023); Smith et al. (2025).

DOI: https://doi.org/10.70188/tmar1n73

Integration with Broader U.S. AI Governance Architecture

The AFCHG Framework's integration layer, as detailed in Section 3.2, provides a direct bridge to existing federal AI governance infrastructure through alignment with the NIST AI Risk Management Framework, facilitating adoption by agencies already invested in NIST-aligned risk management programs. Wilson and van der Velden's (2022) integrated sustainable AI model emphasized systemic interdependencies between AI governance, environmental sustainability, and social equity, a perspective directly embedded in the framework's integration layer. Roberts et al. (2024) documented significant structural barriers to effective global AI governance, including fragmented regulatory approaches and absent enforcement mechanisms. The AFCHG Framework responds by providing a nationally grounded architecture that is simultaneously compatible with emerging international governance norms, a dynamic that Xu et al. (2024) demonstrated is achievable when national frameworks are designed with global alignment explicitly in mind, while remaining anchored in the specific institutional realities of the United States federal system, including the multi-agency jurisdictional complexity that distinguishes American climate-health governance from more centralized national contexts.

CONCLUSION

This article has developed the Algorithmic Fairness Climate-Health Governance (AFCHG) Framework as an original conceptual contribution to the governance of algorithmic bias in AI-driven climate-health systems within the United States public sector. Through systematic synthesis of peer-reviewed scholarship spanning AI governance, algorithmic fairness, public administration, and climate-health systems, the article has constructed a multi-pillar framework that addresses the governance challenge at its legal, institutional, technical, and ethical dimensions simultaneously. The framework's four pillars are theorized with analytical precision and grounded in current evidence, providing both a conceptual architecture and an operational pathway for policy reform. The eight-step governance flow and two summary tables translate this architecture into actionable guidance for federal and state policymakers and public administrators.

The article's major intellectual insights are threefold. First, algorithmic bias in climate-health systems is a multi-dimensional governance challenge requiring a comprehensive architecture spanning legal, institutional, technical, and normative dimensions. The compounding effects of pre-processing, in-processing, and post-processing bias, combined with the selective adherence dynamics identified by Alon-Barkat and Busuioc (2023), make technical solutions alone insufficient. Second, formal algorithmic fairness approaches are structurally inadequate to the substantive equity challenges of American climate-health governance: drawing on Green's (2022) analysis, the AFCHG Framework adopts a substantive approach focused on the actual distribution of climate-health outcomes across social groups. Third, the explicit pillar-to-step mapping, formalized in this article through Table 1 and the analytical discussion of Figure 2, represents a methodological contribution to AI governance framework design extending beyond the climate-health domain, demonstrating how theoretical architecture and operational sequencing can be unified within a single coherent governance framework.

The policy implications extend across multiple governance levels. At the federal level, the framework recommends mandatory Algorithmic Impact Assessment procedures, independent algorithmic review panels, and community-led validation processes as minimum governance requirements for AI deployment by HHS, EPA, FEMA, and CDC. At the state and local level, the governance flow provides a practical implementation pathway adaptable to diverse institutional contexts. Several directions for future research emerge from this work: empirical testing of the AFCHG Framework within specific federal agency contexts; comparative analysis of AI governance approaches across state-level implementations; and further investigation of the accuracy-fairness trade-off documented by Dehghani et al. (2025) in the climate-health governance context. The deployment of AI in climate-health governance is proceeding rapidly. The governance challenge is to ensure this deployment is guided by a commitment to equity as rigorous and persistent as the technical ambition driving AI adoption. The AFCHG Framework represents a scholarly contribution toward that governance challenge, grounded in the best available evidence, attentive to the realities of American institutional governance, and oriented toward the substantive equity outcomes that must be the ultimate measure of AI governance success.

REFERENCES

Al-Zawqari, A., Peumans, D., and Vandersteen, G. (2024). Latent space bias mitigation for predicting at-risk students. Computers and Education: Artificial Intelligence, 7, 100300. https://doi.org/10.1016/j.caeai.2024.100300

Alon-Barkat, S. and Busuioc, M. (2023). Human-AI interactions in public sector decision making: "Automation bias" and "selective adherence" to algorithmic advice. Journal of Public Administration Research and Theory, 33(1), 153-169. https://doi.org/10.1093/jopart/muac007

Alrawahna, A. S., Alzghoul, A., and Awad, H. (2025). The impact of artificial intelligence on public sector decision-making: Benefits, challenges, and policy implications. International Review of Management and Marketing, 15(5), 125-138. https://doi.org/10.32479/irmm.19419

Ansah, E. W., Amoadu, M., Obeng, P., et al. (2024). Health systems response to climate change adaptation: A scoping review of global evidence. BMC Public Health, 24, 2015. https://doi.org/10.1186/s12889-024-19459-w

Arthur, P., Adams, D. P., Fisk, J. M., et al. (2025). Exploring the role of multi-agency coordination in California's climate investment program 2015-2024. Public Organization Review, 25, 1861-1878. https://doi.org/10.1007/s11115-025-00958-2

Ashokan, A. and Haas, C. (2021). Fairness metrics and bias mitigation strategies for rating predictions. Information Processing and Management, 58(5), 102646. https://doi.org/10.1016/j.ipm.2021.102646

Birkstedt, T., Minkkinen, M., Tandon, A., and Mantymaki, M. (2023). AI governance: Themes, knowledge gaps and future agendas. Internet Research, 33(7), 133-167. https://doi.org/10.1108/INTR-01-2022-0042

Chen, Z., Zhang, J. M., Sarro, F., and Harman, M. (2023). A comprehensive empirical study of bias mitigation methods for machine learning classifiers. ACM Transactions on Software Engineering and Methodology, 32(4), 106. https://doi.org/10.1145/3583561

Dehghani, F., Paiva, P., Malik, N., Lin, J., Bayat, S., and Bento, M. (2025). Accuracy-fairness trade-off in machine learning for healthcare: A quantitative evaluation of bias mitigation strategies. Information and Software Technology, 188, 107896. https://doi.org/10.1016/j.infsof.2025.107896

Di Vaio, A., Hassan, R., and Alavoine, C. (2021). Data intelligence and analytics: A bibliometric analysis of human-artificial intelligence in public sector decision-making effectiveness. Technological Forecasting and Social Change, 173, 121201. https://doi.org/10.1016/j.techfore.2021.121201

Erman, E., Furendal, M., and Moller, N. (2025). Algorithmic fairness and feasibility. Philosophy and Technology, 38, 3. https://doi.org/10.1007/s13347-024-00835-8

Fabris, A., Messina, S., Silvello, G., et al. (2022). Algorithmic fairness datasets: The story so far. Data Mining and Knowledge Discovery, 36, 2074-2152. https://doi.org/10.1007/s10618-022-00854-z

Green, B. (2022). Escaping the impossibility of fairness: From formal to substantive algorithmic fairness. Philosophy and Technology, 35, 90. https://doi.org/10.1007/s13347-022-00584-6

Maes, M. J. A., Jones, K. E., Toledano, M. B., and Milligan, B. (2020). Accounting for natural capital has cross-cutting relevance for UK public sector decision-making. Ecosystem Services, 45, 101127. https://doi.org/10.1016/j.ecoser.2020.101127

Martins, F. P., Paschoalotto, M. A. C., Closs, J., Bukowski, M., and Veras, M. M. (2024). The double burden: Climate change challenges for health systems. Environmental Health Insights, 18. https://doi.org/10.1177/11786302241298789

Mosadeghrad, A. M., Isfahani, P., Eslambolchi, L., et al. (2023). Strategies to strengthen a climate-resilient health system: A scoping review. Global Health, 19, 62. https://doi.org/10.1186/s12992-023-00965-2

Mosadeghrad, A. M., Afshari, M., Dehnavi, H., et al. (2024). Strategic analysis of a climate-resilient health system. BMC Health Services Research, 24, 1149. https://doi.org/10.1186/s12913-024-11621-9

Myhre, S. L., Fronsdal, K. B., Ames, H. M. R., and Papadopoulou, E. (2025). A scoping review of climate-resilient health system strategies in low-resource settings. Public Health. https://doi.org/10.1016/j.puhe.2025.106026

Ochmann, J., Michels, L., Tiefenbeck, V., Maier, C., and Laumer, S. (2024). Perceived algorithmic fairness: An empirical study of transparency and anthropomorphism in algorithmic recruiting. Information Systems Journal, 34(2), 384-414. https://doi.org/10.1111/isj.12482

Papagiannidis, E., Enholm, I. M., Dremel, C., et al. (2023). Toward AI governance: Identifying best practices and potential barriers and outcomes. Information Systems Frontiers, 25, 123-141. https://doi.org/10.1007/s10796-022-10251-y

Roberts, H., Hine, E., Taddeo, M., and Floridi, L. (2024). Global AI governance: Barriers and pathways forward. International Affairs, 100(3), 1275-1286. https://doi.org/10.1093/ia/iiae073

Smith, C. L., Rojas, C., and Zurynski, Y., et al. (2025). From policy to practice: Building a resilient, climate-aware health system from the ground up. Climatic Change, 178, 69. https://doi.org/10.1007/s10584-025-03914-7

Wang, Y. and Singh, L. (2025). Impact of bias mitigation algorithms on variations in inferred sensitive attribute uncertainty. Frontiers in Artificial Intelligence, 8, 1520330. https://doi.org/10.3389/frai.2025.1520330

Wilson, C. and van der Velden, M. (2022). Sustainable AI: An integrated model to guide public sector decision-making. Technology in Society, 70, 101926. https://doi.org/10.1016/j.techsoc.2022.101926

Xu, J., Xiao, Y., Wang, W. H., Ning, Y., Shenkman, E. A., and Bian, J., et al. (2022). Algorithmic fairness in computational medicine. EBioMedicine. https://doi.org/10.1016/j.ebiom.2022.104250

Xu, J., Lee, T., and Goggin, G. (2024). AI governance in Asia: Policies, praxis and approaches. Communication Research and Practice, 10(3), 275-287. https://doi.org/10.1080/22041451.2024.2391204

Yang, Q. and Lee, Y.-C. (2024). Ethical AI in financial inclusion: The role of algorithmic fairness on user satisfaction and recommendation. Big Data and Cognitive Computing, 8(9), 105. https://doi.org/10.3390/bdcc8090105

Yilma, K. (2025). Ethics of AI in Africa: Interrogating the role of Ubuntu and AI governance initiatives. Ethics and Information Technology, 27, 24. https://doi.org/10.1007/s10676-025-09834-5