PHIG PARTNERS

Public Health Infrastructure Grant

Policy Analysis: State-Level Artificial Intelligence Policies and Provisions

Supplement to *Developing Artificial Intelligence Policies for Public Health Organizations: A Template and Guidance*

Health and Human Services Region 7

January 2025

PHIG PARTNERS
Public Health Infrastructure Grant

About This Initiative

Developing Artificial Intelligence Policies for Public Health Organizations: A Template and Guidance is a collaboration between the Kansas Health Institute (KHI), Health Resources in Action (HRiA) and the Wichita State University Community Engagement Institute (WSU CEI). The project scope included an environmental scan that informed the template's development, comprising a literature review and policy analysis at both state and city levels. The role of each organization varied by project component. This document, *Policy Analysis: State-Level Artificial Intelligence Policies and Provisions,* was developed by the Kansas Health Institute.

KANSAS HEALTH INSTITUTE
Informing Policy. Improving Health.

HEALTH RESOURCES IN ACTION

WSU WICHITA STATE UNIVERSITY COMMUNITY ENGAGEMENT INSTITUTE

Authors

- **Shelby C. Rowell, M.P.A.**, Analyst, KHI
- **Emma Uridge, M.P.H.**, Analyst, KHI

Contributors

- **Tatiana Y. Lin, M.A.**, Director of Business Strategy and Innovation, KHI
- **Linda J. Sheppard, J.D.**, Team Leader, KHI

The authors and contributors also would like to thank Health Resources in Action and the Wichita State University Community Engagement Institute for providing feedback throughout the process.

Acknowledgement

This work is supported by funds made available from the Centers for Disease Control and Prevention (CDC) of the U.S. Department of Health and Human Services (HHS), National Center for STLT Public Health Infrastructure and Workforce, through OE22-2203: Strengthening U.S. Public Health Infrastructure, Workforce, and Data Systems grant. The contents are those of the author(s) and do not necessarily represent the official views of, nor an endorsement by, CDC/HHS or the U.S. Government.

Disclosure Statement

During the policy review process, the research team utilized AI tools, specifically Petal and ChatGPT, to identify search terms, locate relevant policies and support the development of policy summaries and cross-policy analyses. These tools were used to generate initial drafts of the summaries. All content was subsequently reviewed and refined by the authors to ensure accuracy and quality. The authors take full responsibility for the final content presented in this document.

PHIG PARTNERS
Public Health Infrastructure Grant

Table of Contents

(This page intentionally left blank.)

Introduction

In 2024, the Kansas Health Institute (KHI), Health Resources in Action (HRiA) and Wichita State University Community Engagement Institute (WSU CEI) collaborated on a project which resulted in the document titled *Developing Artificial Intelligence Policies for Public Health Organizations: A Template and Guidance*. This document is designed to assist public health organizations, including nonprofits and government agencies at all levels, in creating policies or guidelines that facilitate ethical experimentation with artificial intelligence (AI) systems while addressing potential risks and promoting health equity and innovation.

To inform the development of the template, the research team conducted an environmental scan focused on considerations surrounding the use of AI, specifically identifying what should be included in the policies for public health organizations. This scan included a review of relevant literature, state policies that were introduced or passed, and policies or guidelines passed by local-level jurisdictions, specifically cities.

The primary goal of this document, *Policy Analysis: State-Level Artificial Intelligence Policies and Provisions,* is to identify key themes and common provisions that are found within current state AI policies to enhance provision language within the template. To achieve this, the analysis utilized data from the *Artificial Intelligence Legislation Tracker* from the National Conference of State Legislatures (NCSL) and the *Public Health Infrastructure Legal Map* from the Association of State and Territorial Health Officials (ASTHO), focusing on AI-specific legislation introduced in 2023 and 2024 state legislative sessions. The information discussed within the summaries is derived from the 152 legislative documents included within the analysis. Individual provisions were not extracted for direct discussion within the document, rather the provisions were aggregated to create comprehensive summaries.

During the initial phase of template development, the list of included provisions was based upon findings from the review of policies. To ensure that the provisions in the template aligned with current research, the literature review findings were used to verify, modify, remove or propose new provisions. Additionally, the literature review findings were used to develop sections of the template that explain why specific issues are important to include in the policy and provide rationale for their inclusion.

Structure of the Document

The document begins with an overview of the policy analysis methodology, providing readers with a clear understanding of the approach taken. Following this, Summary of Provisions section organizes the content into 10 topics, including Acquisition of AI Tools, Bias Mitigation, Community/Public Engagement, and Data Privacy, among others. For the full list of topics, please refer to the Methodology section of the report.

Within each topic, the provisions from relevant policies are summarized under subtopics, offering detailed descriptions of the types of information they encompass. The document does not include a conclusion, as its primary purpose is to present comprehensive information to support the development of *Developing Artificial Intelligence Policies for Public Health Organizations: A Template and Guidance (referred to hereafter as AI Template and Guidance*. The document concludes with an overview of the Petal and ChatGPT protocols used to conduct the policy analysis (*Appendix A*, page A-1), along with a bibliography (*Appendix B*, page B-1) and a list of the legislative documents reviewed (*Appendix C*, page C-1).

Methodology

The purpose of the analysis of state policies was to identify common policies related to AI areas such as bias mitigation, human oversight, data privacy, transparency, among others. The full list of examined topics is below.

The focus areas were determined based on issues typically raised by the public health community and cited in both media and research.

The findings from the review also were essential in establishing a foundational basis for the document *Policy Analysis: State-Level Artificial Intelligence Policies and Provisions*. The review focused on policies introduced during the 2023 and 2024 legislative sessions across various states. These policies ranged in status, from those newly introduced to those passed by state legislatures and enacted into law with the governor's signature. The team also utilized direct online searches to ensure comprehensive coverage. The selected legislation primarily focused on the regulation, oversight and responsible use of AI in key areas such as government, public services, public assistance and health care.

List of Examined AI Topical Areas

1. Acquisition of AI Tools
2. Bias Mitigation
3. Community/Public Engagement
4. Data Privacy
5. Environmental Impact
6. Evaluation and Quality Improvement
7. Human Oversight
8. Impacted Populations
9. Transparency
10. Unauthorized Use and Prohibitions

The policies were identified from multiple sources, including tracking tools such as the *Artificial Intelligence Legislation Tracker* (for the 2023 and 2024 legislative sessions) from the National Conference of State Legislatures (NCSL) and the *Public Health Infrastructure Legal Map* from the Association of State and Territorial Health Officials (ASTHO).

Analysis

The policies were downloaded, reviewed and detailed in a table referred to as the evidence table, which captured information such as state, legislation number, status, type of document, date enacted and detailed provisions related to the examined topical areas, which are listed above. These areas of focus included data privacy, human oversight other relevant topics. Three evidence tables were developed, one for each of the AI Policy Trackers sourced from ASTHO and NCSL (2023 and 2024). An example of the columns included in the evidence table can be found in *Appendix A*, page A-1.

The research team utilized Petal and ChatGPT 4.0 for policy review and analysis. The purpose was to evaluate the feasibility of using these tools for policy analysis and facilitate the examination of all the policies across 21 dimensions. The tools were used interchangeably and used for several purposes: to identify policies based on set search parameters, to summarize policies for inclusion in the evidence table and to support the creation of final summaries across all policies by topic.

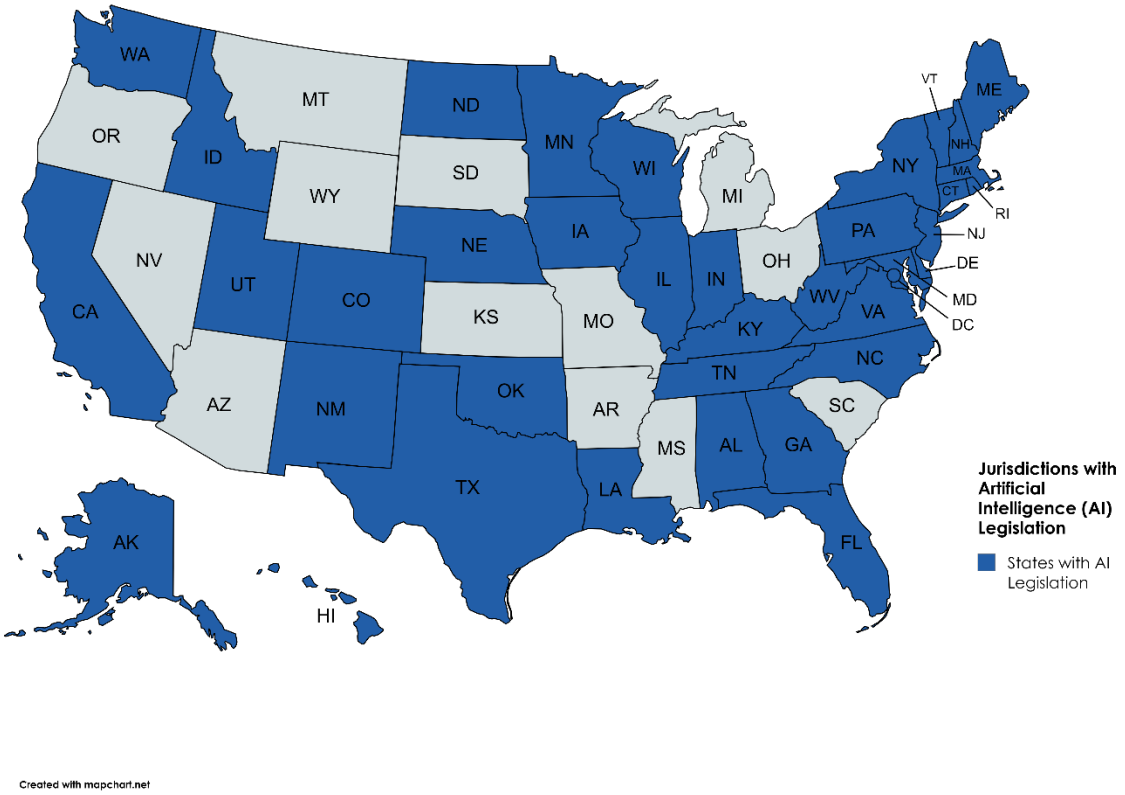
The research team prioritized a human-in-the-loop approach in the quality assurance process, which was implemented throughout the review. Outputs were examined and validated against

the policies and provisions to ensure accuracy and reliability. For more information about the process of using Petal and ChatGPT for policy analysis, see *Appendix A*, page A-1.

Overview of Reviewed Policies

Overall, 152 state legislative documents on artificial intelligence (AI) governance spanning from 2023 to 2024 were identified, representing senate, house and assembly bills, as well as resolutions. These legislative documents address the responsible use and regulation of AI technologies across 36 states and the District of Columbia. *Figure 1* shows the locations of jurisdictions with legislation that was included in this review.

Figure 1. **Locations of Jurisdictions with AI Legislation (2023-2024)**



Sources: National Conference of State Legislatures, 2023 Legislation, available at https://www.ncsl.org/technology-and-communication/artificial-intelligence-2023-legislation; Association of State and Territorial Health Officials (ASTHO), Public Health Infrastructure, available at https://www.astho.org/advocacy/state-health-policy/public-health-legal-mapping-center/public-health-infrastructure.

Scope of Use

State policy provisions across the U.S. define a range of governmental entities responsible for enacting and enforcing artificial intelligence (AI) policies. These provisions outline the roles of legislative bodies, state departments, specialized councils and professional boards in overseeing the responsible use of AI. The entities involved in AI policy enactment are diverse, reflecting the complexities of AI governance. This summary categorizes the roles of these enacting entities, focusing on their responsibilities in managing, enforcing and overseeing AI policy implementation across various sectors.

Legislative and Judicial Authorities as Key Enacting Bodies

A central theme in AI policy enactment is the role of legislative bodies, which are responsible for drafting, proposing and enacting AI-related policies. In several instances, legislative bodies play the lead role in establishing frameworks that govern the use of AI in various sectors, including health care, education and government services. Legislative involvement is often focused on addressing the ethical use of AI, ensuring transparency in AI systems and developing civil rights protections.

Judicial authorities also play a role in enforcing AI-related provisions, particularly when AI-generated evidence or algorithmic systems are involved in legal proceedings. Judicial authorities are tasked with ensuring that AI-generated evidence is handled according to legal standards, especially in verifying the authenticity and accuracy of synthetic or AI-manipulated media. In addition, attorneys general and other legal offices are frequently charged with enforcing AI policy provisions, investigating violations and initiating legal actions to ensure compliance with AI regulations.

State Departments as Enforcing and Implementing Bodies

State departments are key entities responsible for the enforcement and implementation of AI-related policies. Various departments, such as those overseeing banking, insurance, commerce and technology, are tasked with ensuring that AI systems are integrated responsibly into state operations. These departments may be responsible for developing procurement standards for AI technologies, managing the implementation of AI in government services and overseeing compliance with AI-related policies.

In sectors like health care, state health departments play a crucial role in regulating the use of AI. These departments are responsible for certifying AI tools used in patient care, ensuring the accuracy of diagnostic algorithms and enforcing health care related policies that involve AI-driven decision-making. Similarly, state departments of human services often manage the implementation of AI policies in public health care systems, ensuring that AI tools adhere to ethical standards and regulatory requirements.

The Role of the Health and Education Sectors in Policy Enactment

The health and education sectors are heavily involved in AI policy enactment. State health departments and agencies are frequently tasked with overseeing the implementation of AI systems in health care delivery. These provisions often relate to patient care, diagnostics and health care staffing, ensuring that AI is used in a manner that benefits patient outcomes while maintaining transparency and safety.

In education, state departments of education and related commissions are responsible for enacting AI policies that impact school operations and student welfare. These entities often manage programs related to AI-driven educational tools, school safety and resource officer training. The goal of these provisions is to integrate AI responsibly into educational systems, ensuring that AI is used ethically to support both teaching and administration.

Specialized Councils and Commissions as Policy Enactors

Specialized councils and commissions play a significant role in overseeing the enactment and enforcement of AI policies. These bodies are often created to provide sector-specific oversight, ensuring that AI is developed and deployed in compliance with legal and ethical standards. For example, technology modernization councils and health care innovation councils are established to guide the implementation of AI technologies within specific sectors.

These councils often consist of members from different governmental departments, academic institutions and private industries, and their roles may include providing recommendations, conducting impact assessments and ensuring that AI policies align with state objectives. By focusing on specialized sectors, such as health care or education, these councils help manage the complex ethical, technical and legal challenges that arise with AI adoption.

Local Government and State Agencies in Policy Enactment

Local government and state agencies are frequently charged with implementing AI-related policies, particularly in sectors like education and health care. Local authorities are often tasked with enacting AI policies that address community-specific needs, such as school safety programs or public health initiatives. These policies may involve the use of AI in educational tools, public mental health programs or other community services, ensuring that AI is applied effectively at the local level.

State agencies also are responsible for conducting AI system inventories, performing impact assessments and reporting on AI usage within their jurisdiction. These agencies often collaborate with other governmental bodies to ensure that AI systems are integrated responsibly across different sectors and comply with relevant state regulations.

Interagency Coordination and Cross-Departmental Responsibility

AI policy implementation often requires coordination across multiple state agencies. Various departments may be required to work together to ensure that AI systems are developed and deployed according to policy provisions. This includes the integration of AI into public services, the establishment of procurement standards for AI systems and the development of policies that address the ethical use of AI technologies.

Interagency collaboration also is critical for addressing the multifaceted challenges posed by AI. For instance, departments related to technology, health care, human services and education may need to work together to ensure that AI is used responsibly across these different sectors. These collaborations help ensure that AI policies are consistently enforced and that AI tools are applied in ways that protect public interests.

Professional Licensing Boards and Industry-Specific Oversight

Professional licensing boards play an important role in the enactment and enforcement of AI policies, particularly in areas where AI intersects with licensed professions. These boards are responsible for overseeing the ethical use of AI by professionals in fields such as health care, law and mental health services. They ensure that AI tools comply with industry standards and that professionals using AI adhere to established ethical guidelines.

For example, licensing boards may regulate the use of AI in patient care by physicians or oversee how AI is integrated into legal or mental health services. This industry-specific oversight

ensures that AI is deployed in ways that protect consumers and adhere to the ethical standards of each profession.

Enforcement Mechanisms and Investigative Powers

AI policy provisions often grant specific entities the authority to enforce AI-related regulations. These enforcement bodies, which may include state insurance departments, labor standards offices and civil rights departments, are tasked with investigating violations, enforcing compliance and ensuring that AI systems operate within the legal framework.

State attorneys general are frequently granted enforcement powers to pursue legal actions against entities that violate AI-related policies. This includes investigating complaints, conducting inquiries into AI-related practices and imposing penalties when necessary. Enforcement mechanisms are essential for ensuring that AI technologies are used responsibly and in compliance with state regulations.

Summary of Provisions

Acquisition of AI Tools

The policy provisions across various U.S. states regarding the acquisition and deployment of artificial intelligence (AI) tools outline a detailed approach aimed at ensuring the ethical, transparent and accountable use of these technologies. These provisions, which focus on the acquisition phase, cover multiple key areas such as documentation, transparency, ethical standards, risk management, sector-specific applications and regulatory oversight. The growing role of AI in public and private sectors necessitates strong policies that safeguard public interests while promoting responsible AI use.

Defining AI and Framework for Acquisition

A critical starting point in the policy provisions for AI acquisition is the establishment of a clear definition of AI. Several states mandate the formation of working groups or task forces to develop standardized definitions for AI within legislative contexts. These definitions play a pivotal role in determining how AI systems are classified, regulated and ultimately acquired by state agencies or other entities. By ensuring consistency in how AI is understood across various sectors, these provisions create a solid foundation for the legal governance of AI tools.

In addition to definitions, policy provisions emphasize the need for structured frameworks to guide the procurement and deployment of AI tools. These frameworks typically require agencies to evaluate AI systems comprehensively before acquisition. The evaluation process includes assessing the AI tool's potential impacts, especially concerning discrimination, data misuse or other ethical risks. The goal is to ensure that AI systems do not result in unlawful discrimination or adverse effects on historically marginalized populations. Acquiring AI tools is thus framed within the context of ongoing oversight to maintain compliance with ethical and legal standards throughout the tool's operational lifecycle.

Documentation and Transparency Requirements

Transparency in the acquisition and deployment of AI tools is a major theme in these policy provisions. States require that AI developers and vendors provide detailed documentation regarding their AI systems. This documentation must include information on the AI tool's intended purpose, limitations, benefits, risks and technical specifications, such as algorithmic processes and data sources. This ensures that the stakeholders acquiring these tools — whether public agencies or private entities — have a clear understanding of the technology they are implementing.

Additionally, developers are mandated to make AI systems and the data used in their creation available for auditing and independent testing. This accessibility allows third-party evaluators to assess whether the systems operate fairly, accurately and without bias, particularly in cases where the AI tool could impact historically marginalized groups. Furthermore, regular inventories of AI systems, listing their data inputs and outputs, must be compiled and submitted to state authorities. These inventories promote transparency and provide a mechanism for continuous scrutiny and accountability in the acquisition and use of AI tools.

Ethical and Safety Guidelines

The ethical use of AI tools, particularly during and after their acquisition, is a recurring focus within state policy provisions. The acquisition of AI tools classified as high-risk — those that could significantly impact individuals' rights or welfare — is subject to heightened ethical scrutiny. Before acquisition, these AI systems must undergo a detailed safety and impact assessment, which evaluates the tool's intended use, foreseeable risks and any mitigation strategies to address potential harm. Such assessments are not limited to pre-acquisition stages

but are required to be updated regularly to ensure the AI system remains compliant with ethical standards over time.

Moreover, the policy provisions demand that AI developers address algorithmic discrimination, especially concerning race, gender and disability. To mitigate potential biases, AI systems must be tested rigorously before they are acquired. Continuous monitoring after deployment also is mandated to ensure that any discriminatory effects are detected and corrected promptly. These ethical guidelines ensure that the acquisition of AI tools does not result in unintended harm or exacerbate existing inequalities.

Human Oversight and Accountability

Policy provisions also emphasize the necessity of human oversight in decisions involving AI systems, particularly during the acquisition and deployment phases. One key requirement is that decisions made or influenced by AI tools must be subject to human review, ensuring that individuals retain the ability to challenge or appeal decisions made by AI systems. This human oversight is crucial in preventing over-reliance on automated systems for decisions that could have significant legal, social or personal impacts.

In health care, for instance, policy provisions stipulate that AI tools used for diagnostic purposes must have their recommendations reviewed and approved by licensed medical professionals before any clinical decisions are made. Similarly, AI tools in public assistance programs cannot be used exclusively to identify fraud; human personnel must investigate any issues flagged by the AI system to prevent unwarranted denials of service. These provisions underscore the importance of human judgment in maintaining fairness and accountability when AI tools are integrated into decision-making processes.

Sector-Specific Applications of AI

The policy provisions governing the acquisition of AI tools recognize that different sectors have varying needs and challenges when it comes to AI integration. Tailored regulations ensure that AI tools are used appropriately in sectors such as health care, education and public assistance.

In health care, AI systems used for diagnostic purposes or insurance determinations must comply with clinical guidelines, undergo regular reviews and be approved by qualified physicians before deployment. For example, AI algorithms used in diagnostic imaging, such as breast tissue analysis, must meet stringent ethical and safety standards to ensure that patient

care is not compromised. In mental health services, policy provisions require continuous human supervision of AI-assisted treatment, ensuring that licensed practitioners maintain control over patient care decisions.

In public assistance programs, AI tools are prohibited from making determinations on their own. Instead, they may only flag potential issues for human review, preventing the undue reliance on automated systems for critical decisions that could affect access to essential services. In the education sector, AI tools are used to support specialized services, such as providing interpreter services for individuals with hearing impairments or enhancing science, technology, engineering and mathematics (STEM) education through AI-driven platforms. The acquisition of these AI tools is governed by provisions that ensure their continued effectiveness and compliance with ethical standards.

Impact Assessments and Risk Management

Impact assessments are a critical component of the policy provisions related to AI acquisition. Before an AI tool is acquired, its developers and deployers are required to conduct comprehensive impact assessments to evaluate the system's fairness, reliability and potential risks. This is especially important for high-risk AI systems, where the consequences of failure or bias could be severe. The assessments often cover multiple dimensions, such as the system's data inputs, outputs and the potential for discrimination or other adverse effects.

Additionally, risk management programs are required to accompany the acquisition of AI tools, particularly high-risk ones. These programs involve identifying, mitigating and documenting risks throughout the AI system's deployment. Developers must provide the necessary technical capacity for agencies to monitor these risks, ensuring that the AI tools remain safe, effective and aligned with ethical guidelines. These risk management protocols, along with regular impact assessments, are crucial in preventing AI systems from causing harm or legal liability.

Bias Audits and Non-Discrimination

Bias audits are another significant policy requirement for the acquisition and deployment of AI tools, particularly in sectors like employment and health care. Policy provisions stipulate that AI systems must undergo regular bias audits to detect and correct any discriminatory outcomes that may disproportionately affect historically marginalized groups. These audits, typically conducted by independent third parties, are intended to ensure that AI systems operate fairly and equitably.

In the employment sector, for example, AI systems used for hiring decisions are subject to bias audits to prevent discrimination based on race, gender or other characteristics. Employers are also required to notify job candidates when AI tools are used in the hiring process and provide transparent information about the tool's data inputs and decision-making criteria. The results of these bias audits must often be made public or reported to relevant state authorities to ensure accountability.

In health care, policy provisions explicitly prohibit the use of AI tools in ways that could perpetuate biases in clinical algorithms. AI systems used to inform medical decisions, such as diagnosis or treatment recommendations, must avoid discrimination based on race, gender or disability. These provisions aim to ensure that AI tools contribute to equitable outcomes in health care and do not exacerbate existing disparities.

Data Privacy, Security and Ethical Considerations

The policy provisions surrounding the acquisition of AI tools also place a strong emphasis on data privacy and security. AI systems, particularly those processing sensitive personal data, must comply with stringent state and federal privacy laws. Agencies acquiring AI tools are required to develop comprehensive data protection policies that cover the collection, retention and sharing of personal data. These policies must ensure that data is handled securely, particularly in high-risk AI systems where breaches could have serious consequences.

In addition to privacy protections, ethical considerations play a prominent role in these provisions. AI systems must be designed with safeguards against foreseeable harm and the use of AI for malicious purposes — such as predictive policing or invasive surveillance — is explicitly prohibited. The policy provisions also reflect concerns about the social and economic impacts of AI, particularly its potential to displace workers. Provisions often call for measures to prevent AI tools from being used in ways that unfairly impact employment or job security.

Governance and Regulatory Oversight

Governance and regulatory oversight are crucial components of the policy provisions for acquiring AI tools. Many states have established specialized councils or task forces to oversee the acquisition and deployment of AI technologies. These bodies are tasked with evaluating the impact of AI systems, recommending policy reforms and ensuring that AI tools are acquired and used in compliance with legal and ethical standards.

Interagency coordination is often required to maintain consistency in AI governance. Agencies are required to submit detailed reports on the AI systems they acquire, including information about the tool's capabilities, data inputs, outputs and any bias audits or impact assessments conducted. These inventories provide a mechanism for transparency and enable continuous monitoring of AI systems to ensure they remain aligned with legal and ethical requirements.

Bias Mitigation

Various legislative provisions across U.S. states aim to mitigate bias in AI systems, focusing on ensuring fairness, transparency and accountability. These provisions, both enacted and proposed, target the risks of algorithmic discrimination in sectors like employment, health care and public services. Central to these efforts are task forces, mandatory impact assessments, bias audits and regulations around automated decision tools. This summary outlines key provisions aimed at preventing discriminatory outcomes in AI systems.

Task Forces and Working Groups on Bias Mitigation

A key feature in AI bias mitigation efforts is the establishment of task forces and working groups. These bodies are responsible for studying the potential risks of AI systems, particularly in high-risk areas, and making recommendations to prevent biases. They are tasked with developing standardized metrics to evaluate AI system performance and trustworthiness. These groups ensure AI systems are tested for bias, especially concerning historically marginalized populations, and aim to prevent unintended outcomes. They also focus on addressing privacy concerns and protecting civil liberties, offering policy suggestions to guide future AI regulations.

Impact Assessments and Risk Mitigation Strategies

Several policy provisions require comprehensive impact assessments before AI systems can be deployed. These assessments are designed to evaluate potential biases that could affect individuals based on characteristics such as race, gender and disability. Developers are expected to assess the risks posed by their AI systems and document the measures they take to mitigate these risks.

In some provisions, high-risk AI systems — particularly those used in critical areas like health care, housing and employment — must undergo detailed evaluations that explore the potential for bias and provide strategies for addressing it. These assessments also are intended to be conducted on an ongoing basis to ensure that AI systems remain fair throughout their use.

Documenting and reporting these steps help ensure that AI systems align with ethical standards and do not result in discriminatory outcomes.

Bias Audits and Transparency Measures

Bias audits form a significant part of the legislative effort to reduce discrimination in AI systems. Many provisions require that AI tools, particularly automated decision-making tools used in employment, undergo regular audits. These audits are meant to evaluate the system's compliance with anti-discrimination laws and detect any potential biases that may lead to disparate impacts on historically marginalized groups.

The results of these bias audits must often be made publicly available, ensuring transparency and accountability. Employers and developers are required to notify individuals when AI systems are used to make consequential decisions, such as hiring, and to provide accessible summaries of these audits. This level of transparency helps to foster public trust in AI systems by ensuring that biases are identified and corrected before the systems can cause harm.

Transparency and Accountability in AI Deployment

Transparency is central to many legislative provisions on AI deployment. Developers and deployers of AI systems are required to disclose information about how the AI makes decisions and the potential biases inherent in the data used. These provisions ensure that AI systems do not rely on proxy variables for characteristics that could lead to unintended discrimination.

Additionally, some provisions require equity assessments that ensure AI systems use diverse and representative data sets, minimizing the risk of biased outcomes. Deployers of AI systems also must establish governance programs to continuously manage and mitigate risks related to bias. By publicly disclosing bias audits and impact assessments, these provisions promote accountability, allowing affected individuals to understand how AI systems operate and take action if they suspect discrimination.

Protection of Individual Rights

A primary concern in AI bias mitigation legislation is the protection of individual rights. Provisions focus on ensuring that AI systems do not lead to discriminatory or unfair treatment of marginalized groups. Legislative frameworks include the requirement for pre-deployment and ongoing testing of AI systems to detect biases and prevent unlawful discrimination.

Moreover, some provisions mandate that independent evaluations of AI systems be conducted with the goal of ensuring fairness and preventing bias throughout the system's lifecycle. These measures help safeguard civil liberties, ensuring that AI technologies do not exacerbate existing inequalities or disproportionately impact historically marginalize groups in areas such as housing, health care and employment.

Regulation of Automated Decision Tools

Regulations targeting automated employment decision tools (AEDTs) form a critical aspect of bias mitigation efforts, particularly in employment contexts. Legislative provisions often require these tools to undergo disparate impact analyses, which test whether the systems disproportionately affects certain groups, such as those defined by race, gender or ethnicity.

Before these tools can be deployed, they must be audited for bias, and these audits must be updated regularly. Employers are typically required to notify individuals when AEDTs are used in hiring, promotion or other employment-related decisions. Furthermore, some provisions prohibit the sale or use of AEDTs unless they have recently undergone bias audits, ensuring that discriminatory impacts are identified and addressed before the tools are implemented.

In addition to employment, automated decision tools used in other sectors, such as health care or public services, are subject to similar regulatory scrutiny. Provisions ensure that AI systems do not make decisions solely based on automated processes without meaningful human oversight, particularly when such decisions could impact individuals' rights.

Enforcement and Legal Recourse

Enforcement mechanisms provide individuals with avenues for challenging AI-driven decisions that may result in discriminatory outcomes. Provisions in this area often give individuals the right to contest decisions made by AI systems, particularly if they believe these decisions were influenced by bias.

State agencies also are tasked with developing policies and conducting continuous assessments to ensure that AI systems do not result in unlawful discrimination. Additionally, in some legislative frameworks, AI developers and deployers are required to report any detected bias or discriminatory outcomes to regulatory authorities, which have the power to investigate and enforce corrective actions.

These provisions establish clear legal pathways for individuals to seek recourse if their rights are violated by AI systems, ensuring that developers and deployers are held accountable for any discriminatory impacts their systems may cause.

Community/Public Engagement

Various U.S. state-level legislative provisions address aspects of community and public engagement in the governance AI. These provisions outline the involvement of diverse stakeholders, public notification about AI use and mechanisms for transparency and accountability in AI-related decision-making. This summary highlights key themes from these legislative efforts, focusing on the specific ways public participation is incorporated into AI governance.

Stakeholder Inclusion and Collaboration

Many AI-related legislative provisions mandate the inclusion of diverse stakeholders in the policy-making process. Academic experts, industry representatives, consumer advocacy groups and businesses of varying sizes are brought together to ensure that AI systems are designed with public welfare in mind. Some states establish formal task forces or working groups, which include representatives from academia, advocacy organizations and the business community to guide AI governance. This broad inclusion ensures that diverse perspectives, particularly from historically marginalized or affected communities, are considered throughout the development and deployment of AI technologies.

Another key aspect of collaboration emphasized in these provisions is the partnership between government agencies, academic institutions and the private sector. These collaborations are crucial for advancing AI research while safeguarding public interests. Legislative provisions often call for centralized hubs or committees to foster innovation while addressing critical issues like privacy, security and societal risks associated with AI. These collaborations are designed to ensure that AI development benefits society as a whole, rather than prioritizing private or sectoral interests.

Transparency and Public Notification

Transparency is a fundamental theme in AI legislation, particularly in contexts where AI systems are involved in decision-making processes that significantly affect individuals. Many provisions require entities deploying AI tools to notify the public about their use, ensuring that individuals understand when and how AI systems are impacting decisions that affect their lives. For

example, in employment, job candidates must be informed if AI tools are used during hiring processes. Transparency extends to providing clear explanations of how the AI system functions, what data it uses and the potential implications for those affected by it.

Legislative provisions also often mandate regular bias audits of AI systems to ensure fairness and prevent discriminatory impacts. These audits, particularly in areas such as employment or financial services, assess whether AI systems disproportionately affect historically marginalized groups. The results of these audits must be publicly available, ensuring that the public can see how AI tools are being monitored for fairness and what steps are being taken to address any detected biases. This type of transparency is key to maintaining public trust in AI systems and ensuring accountability for their use.

Public Engagement in Decision-Making

Public input plays a crucial role in AI governance, with many legislative provisions establishing mechanisms to ensure that communities are actively involved in shaping AI policy. Public hearings, advisory committees and task forces are common structures for soliciting input from community members, particularly those most likely to be impacted by AI technologies. Policymakers are required to consult with these groups, including historically marginalized communities, before making decisions regarding AI systems.

Public hearings provide a formal platform for individuals and organizations to voice concerns or offer support for AI policies. This ensures that the voices of those most affected by AI systems, especially in areas such as health care, employment and public services, are heard and considered in the decision-making process. Some provisions further require that public hearings be held before any significant regulatory decisions are made regarding AI technologies, reinforcing the importance of community involvement in AI governance.

Educational Outreach and Public Resources

Public education on AI and its ethical implications is another critical theme in these legislative provisions. Governments are tasked with providing resources that help demystify AI technologies, making them more accessible to the general public. By educating the public about how AI systems work, their potential benefits and the ethical challenges they present, these resources empower individuals to engage meaningfully in AI-related discussions and advocate for policies that align with their values.

Additionally, many provisions require regular public reporting on the use of AI systems by public agencies and private entities. These reports often include findings from bias audits, updates on the deployment of AI technologies and recommendations for ethical AI usage. By making these reports publicly available, the legislation promotes ongoing transparency and allows for continuous community oversight of AI systems, ensuring that public feedback is incorporated into the refinement and regulation of AI technologies.

Accountability and Public Oversight

Ensuring accountability in AI governance is a key objective of the legislative provisions, and many of them include formal mechanisms for public feedback. These mechanisms range from public comment periods to advisory boards that gather input on the use of AI in various sectors. Provisions often enable the temporary suspension of AI systems if significant public concerns about their fairness or safety are raised, allowing for a thorough review and remediation of any issues before systems are re-deployed.

Public hearings and open meetings also serve as crucial platforms for public oversight. In many cases, legislation requires that these meetings be made accessible online, ensuring broader participation from individuals who may not be able to attend in person. These hearings allow for direct engagement between community members and policymakers, fostering a transparent decision-making process where public concerns are addressed in real-time. By integrating public oversight into AI governance, these provisions help ensure that AI systems remain accountable to the communities they serve.

Reporting Mechanisms and Public Participation

Several legislative provisions emphasize the importance of accessible reporting mechanisms to facilitate public participation in monitoring AI systems. In many cases, state agencies are required to create formal reporting channels — such as dedicated webpages or email addresses — where individuals can report violations of AI-related regulations. These reporting mechanisms are designed to empower the public to hold organizations accountable for non-compliance with AI laws, including cases of identity theft, fraud or discriminatory practices. Some provisions require annual notifications to the public about the existence of these channels, ensuring that individuals remain aware of their rights and how to report violations.

In certain contexts, public engagement is encouraged in environmental monitoring, with AI systems playing a role in tracking environmental risks. For instance, individuals can submit

evidence, such as photos or data, to state authorities regarding harmful atmospheric activities, such as geoengineering. This approach highlights the importance of public participation not only in AI governance but also in broader issues of public and environmental safety.

Prioritizing Historically Marginalized and Underserved Populations

Many provisions prioritize engagement with historically marginalized and underserved populations, particularly in the areas of public health and safety. Task forces or working groups often are required to consult with these populations — such as underserved cultural and linguistic groups and individuals with limited English proficiency (LEP) — to ensure that their perspectives are taken into account when developing AI systems that affect their communities. These provisions help prevent the deepening of existing inequalities by ensuring that AI systems are designed and implemented in ways that address the specific needs and concerns of marginalized groups.

In some cases, legislation mandates that AI-driven public health initiatives specifically engage with historically marginalized populations, ensuring that these groups have access to the services and protections provided by AI technologies. By focusing on marginalized communities, the legislation seeks to create AI systems that promote equity and do not exacerbate disparities in access to services or resources.

Data Privacy

Data privacy in the realm of AI has become a major focal point in policy discussions as automated decision systems (ADS) increasingly shape interactions in both public and private sectors. Policy provisions emphasize the need to safeguard personal data, ensure transparency and prevent algorithmic discrimination. These provisions aim to balance the advantages of AI technology with robust privacy protections and ethical practices. The following summary synthesizes key themes from data privacy provisions in AI-related policy, including transparency, accountability, data security, bias mitigation and regulatory oversight.

Establishment of AI Working Groups and Regulatory Bodies

A common theme across policy provisions is the creation of AI working groups or advisory councils tasked with ensuring ethical and responsible AI use. These bodies are responsible for investigating potential risks associated with AI technologies, particularly concerning data privacy, and developing recommendations for future privacy regulations. These groups consist

of a wide range of stakeholders, including experts from academia, industry and consumer advocacy groups, who work collaboratively to shape policies that protect individuals from data misuse or unauthorized access.

In addition to shaping policy, these working groups oversee compliance with current data protection laws, ensuring that AI systems operate in ways that prioritize user privacy. They also play a role in monitoring how AI systems interact with sensitive personal data and make recommendations for improving safeguards.

Risk Assessments and Impact Evaluations

A key requirement in the provisions is the performance of risk assessments and impact evaluations before deploying AI systems that could impact individual privacy or civil rights. These assessments evaluate the types of data collected, the potential for biases in AI algorithms and any risks of unlawful discrimination.

The provisions often require organizations to maintain detailed documentation of the data processing methods, the known limitations of the AI systems and the strategies used to mitigate risks. High-risk AI systems, in particular, must be subject to frequent review, especially when changes occur in the system's purpose or the type of data processed. The goal of these assessments is to ensure transparency and accountability, providing a clear pathway for identifying and addressing any risks that may arise from the deployment of AI systems.

Transparency and Accountability in AI Operations

Transparency is central to the policy provisions related to data privacy in AI systems. Organizations deploying AI tools are required to notify individuals when AI systems are used in ways that affect them. This includes providing individuals with clear and accessible explanations about how their data is being used and how AI-driven decisions are made.

The provisions stipulate that organizations must make publicly available the results of any bias audits or impact assessments conducted on their AI systems. This transparency fosters trust and ensures that organizations remain accountable for the fairness and accuracy of their AI systems. Individuals affected by AI-driven decisions should have access to actionable information about how these systems operate and be able to challenge or inquire about those decisions.

Data Minimization and Retention Policies

Another core theme is the restriction of data collection to what is strictly necessary for the functioning of AI systems. The provisions mandate that AI systems collect only the minimum amount of data needed to perform their intended tasks. Once the purpose of data collection is fulfilled, or the relationship between the individual and the organization ends (e.g., employment), the collected data must be securely deleted.

Certain provisions go further by requiring organizations to obtain informed consent from individuals before retaining their data for extended periods. This aligns with broader data privacy objectives that emphasize user control over personal information and limit opportunities for misuse or unauthorized access to sensitive data.

Protections Against AI-Driven Discrimination

To prevent AI systems from exacerbating discrimination, the policy provisions require continuous bias audits and impact assessments. These mechanisms are in place to evaluate whether AI systems disproportionately affect certain groups based on race, gender, age, disability or other characteristics.

Governance programs established by these provisions include safeguards to document, assess and manage algorithmic bias. The provisions also require that the results of these bias audits be made publicly available, ensuring transparency in how organizations address issues of discrimination. In employment, housing and financial services, these safeguards are essential for preventing unfair outcomes and promoting fairness across diverse populations.

Employee Data Privacy and Surveillance Restrictions

The provisions also focus on protecting employee data privacy, particularly concerning the use of AI-driven surveillance tools in the workplace. Employers are required to notify employees about the use of AI systems for monitoring purposes and ensure that such systems collect only necessary data for legitimate safety or security reasons.

Additionally, the provisions emphasize that employees must have access to their data and the ability to correct inaccuracies. The use of invasive AI technologies, such as facial recognition or emotion detection, is generally discouraged unless absolutely necessary, with restrictions ensuring the least invasive methods are used to monitor employees. This promotes a balanced approach to workplace surveillance while upholding employee privacy rights.

Consent and Individual Data Rights

Consent and control over personal data are essential elements of the data privacy provisions. Individuals must provide informed consent for their data to be collected and processed by AI systems, and they must be informed of how their data will be used, shared or retained.

The provisions allow individuals to opt out of AI systems if they do not wish to have their data processed. Additionally, individuals are granted the right to access the data collected about them and request corrections or deletions if their data is inaccurate or no longer needed. This ensures that individuals maintain a high degree of control over their personal information and can take action if they believe their data is being misused.

Public Accountability and Reporting

The provisions emphasize public accountability by requiring organizations to publish regular reports on their AI systems and the associated privacy risks. These reports must detail the types of data collected, the methods used to process and store data and the outcomes of any impact assessments or audits.

Regulatory bodies also are granted oversight authority, ensuring that organizations comply with privacy standards and implement necessary safeguards to protect individuals' data. These reports are typically made available to the public, offering transparency into how AI systems function and providing stakeholders with the information needed to hold organizations accountable.

Enforcement Mechanisms and Civil Penalties

To ensure compliance with data privacy provisions, many policies include enforcement mechanisms that allow for civil penalties in cases of noncompliance. Organizations found in violation of these provisions may face fines, legal action or other penalties, especially if they have compromised individuals' data privacy or allowed discriminatory outcomes to occur.

In some instances, individuals who have been harmed by AI systems can pursue legal remedies through private rights of action, further ensuring that organizations are held accountable for their use of AI technologies. Regulatory bodies also are empowered to subpoena relevant documentation, such as bias audits or impact assessments, to ensure compliance during investigations.

Environmental Impact

The review of legislative provisions on AI across various U.S. state policies reveals a gap in addressing the environmental impacts of AI technologies. Although AI is increasingly being integrated into state-level governance and public sector operations, specific attention to the environmental dimensions of these technologies remains limited. The legislative documents frequently point out this absence, reflecting a broader trend in AI policy development that focuses on ethical, societal and legal concerns while overlooking environmental consequences. Despite this, some provisions offer a framework for potential future environmental assessments.

Absence of Direct Environmental Provisions

One of the most notable themes across the legislative texts is the explicit acknowledgment that environmental impacts are not directly addressed. Multiple sections emphasize that the legislative documents do not contain specific guidelines or mandates regarding the environmental consequences of AI. The documents clarify that “the document does not contain any provisions related to the environmental impact of AI,” a phrase reiterated throughout. This recurring absence highlights a significant gap in the approach to AI regulation, where environmental concerns such as energy consumption and resource depletion are not yet integrated into state policy frameworks.

AI in Public Sector and Environmental Risk Assessments

Several legislative provisions emphasize the responsible use of AI in government operations, aligning AI systems with ethical and sustainability goals. In some cases, these provisions mandate risk assessments that consider both societal and environmental factors, though the primary focus is often on fairness, privacy and security. The inclusion of environmental considerations, even if secondary, demonstrates a growing awareness of AI’s broader impacts.

For example, public sector AI systems are subject to algorithmic impact assessments (AIAs), which include evaluations of the long-term consequences of using AI. While these assessments primarily focus on human welfare, they offer a procedural framework that could be adapted to include environmental risk assessments. By requiring government AI systems to undergo these evaluations, states are positioning themselves to monitor AI’s environmental impacts, even though current provisions tend to emphasize social and ethical concerns.

Additionally, some legislative provisions introduce the establishment of AI advisory councils tasked with overseeing AI systems used by state agencies. These councils are responsible for

monitoring the impacts of AI technologies, including their environmental implications. Although the text regarding environmental assessments is brief, it presents an opportunity for more detailed evaluations in the future. The advisory council's duties include assessing the need for a state code of ethics for AI systems and reviewing automated decision systems to understand their broader implications. While the level of rigor required for environmental assessments is not clearly defined, the inclusion of environmental impact as part of the council's responsibilities indicates a step toward integrating sustainability into AI governance within government operations.

Generative AI Systems and Environmental Risks

Another provision that touches on environmental concerns relates to the potential risks posed by generative AI systems. Generative AI systems, which are capable of producing content or making autonomous decisions, require large volumes of data and extensive computational resources to function. This operational complexity translates into significant energy consumption, which in turn raises concerns about the environmental footprint of such systems.

To address this, the provision mandates a joint risk analysis to evaluate the dangers posed by AI, including potential environmental emergencies. The provision explicitly mentions the need to assess risks that could lead to "environmental emergencies," suggesting a recognition of the broader impacts AI technologies could have on natural systems. However, the focus is primarily on catastrophic scenarios, such as the role AI might play in large-scale environmental disasters, rather than on the everyday environmental impacts, such as the high energy demands of generative AI systems.

Evaluation and Quality Improvement

The legislative provisions surrounding the evaluation and quality improvement (QI) of AI systems reflect the growing need for oversight and continuous refinement in AI deployment. These policies aim to ensure that AI systems operate safely, ethically and effectively across various sectors, with a strong emphasis on transparency, accountability and risk mitigation. This summary explores the key themes in AI legislation focused on evaluation and quality improvement, organizing the narrative around critical topics such as impact assessments, bias audits, ongoing monitoring, public transparency and governance structures.

Evaluation and Impact Assessments of AI Systems

A foundational element in the quality improvement of AI systems is the requirement for comprehensive evaluations and impact assessments before and after deployment. Legislative provisions mandate that developers and deployers of AI systems conduct detailed evaluations to assess performance, risks and potential biases.

These assessments focus on various aspects of AI system functionality, ensuring that AI tools meet performance metrics and legal standards. For example, developers are required to conduct design evaluations that focus on potential biases and risks of unlawful discrimination. These evaluations analyze how AI systems interact with human users and other systems, identifying risks such as algorithmic discrimination based on race, gender or socioeconomic status. The evaluations are designed to ensure that AI systems meet trustworthiness standards and do not lead to unintended negative outcomes.

Legislative mandates also call for annual impact assessments, with a requirement for reassessment if there are material changes to the AI system's usage or data processing methods. These impact assessments ensure continuous monitoring and document the purpose, data inputs and outputs of the AI system, ensuring transparency and accountability at every stage of its deployment.

Bias Audits and Mitigation of Discriminatory Outcomes

Bias mitigation plays a critical role in ensuring that AI systems do not disproportionately affect certain demographic groups, particularly in high-impact areas such as employment and public services. Provisions across different states require regular bias audits to detect and address any discriminatory outcomes resulting from AI systems.

The legislative provisions mandate that bias audits be conducted annually and that the results be made publicly available to ensure transparency. These audits focus on critical metrics such as selection rates and impact ratios for different demographic groups, especially in AI-driven decision-making tools like automated hiring systems. In employment settings, bias audits are essential for evaluating how AI tools influence hiring, promotions and other decisions, ensuring compliance with anti-discrimination laws. Deployers of AI systems must maintain records of these audits, which can be requested by regulators to ensure ongoing oversight.

By mandating bias audits, the legislation underscores the importance of mitigating algorithmic discrimination and promoting fairness in AI decision-making processes. These audits help identify any unintentional biases embedded in AI systems and provide mechanisms to address them proactively.

Ongoing Monitoring and Post-Deployment Evaluation

In addition to pre-deployment evaluations, continuous monitoring and post-deployment assessments are critical to ensuring that AI systems remain effective and unbiased as they operate. Legislative provisions emphasize the importance of ongoing evaluation to track system performance, identify errors or biases and ensure compliance with ethical standards over time.

Developers and deployers are required to conduct ongoing monitoring of AI systems, regularly assessing their accuracy, fairness and alignment with legal regulations. In some cases, provisions call for independent third-party evaluations of AI systems to verify their performance, particularly regarding safety and unintended harmful outcomes. This external validation ensures that AI systems are subject to unbiased scrutiny and remain trustworthy.

For government agencies, regular evaluations are particularly important. State agencies using AI tools are required to conduct biennial assessments to ensure that AI systems do not result in unlawful discrimination or disparate impacts on historically marginalized groups. These assessments help ensure that AI technologies used in public services are aligned with legal and ethical standards throughout their lifecycle.

Transparency, Accountability and Public Reporting

Transparency is a key theme in AI quality improvement legislation, with provisions mandating detailed reporting and public disclosure of AI system evaluations. Public reporting requirements are designed to ensure that AI systems, especially those used in employment and public services, are transparent in their operations and impact.

Legislative provisions often require employers and agencies to publish the results of bias audits and impact assessments on their websites, making the findings accessible to all stakeholders. This level of transparency allows individuals affected by AI systems to understand how decisions are made and provides a means for challenging outcomes that may seem unfair or biased.

To further support transparency, provisions also emphasize algorithmic transparency, requiring developers to disclose the methodologies used in AI systems. This includes details about data sources, how data is processed and the safeguards in place to mitigate foreseeable risks of bias or error. These transparency measures help build public trust in AI technologies by ensuring that stakeholders are informed about how AI systems function and are aware of any potential risks.

Establishing AI Governance and Ethics Guidelines

Effective governance structures are essential for the responsible deployment and management of AI systems. Legislative provisions focus on the creation of AI governance programs and the establishment of ethical guidelines to oversee the development and deployment of AI systems.

Governance programs are tasked with ensuring that AI systems adhere to ethical standards, including fairness, transparency and non-discrimination. These programs are often responsible for overseeing the continuous evaluation of AI systems, addressing any risks or issues that arise during their operation and ensuring that AI technologies are aligned with ethical principles.

Many states also have established AI advisory councils or commissions to monitor AI usage and provide recommendations for legislative reforms. These bodies conduct periodic audits and evaluations to ensure that AI systems comply with ethical and legal standards, particularly in public sectors such as health care, education and public assistance programs.

Human Oversight and Safeguards

Human oversight is a crucial safeguard in mitigating the risks associated with AI systems. Legislative provisions emphasize the need for independent human oversight in AI-assisted decision-making processes, particularly in sectors like employment, health care and public services.

In employment settings, provisions require that human reviewers have the authority to dispute or correct AI-generated outputs if they appear inaccurate, biased or discriminatory. This ensures that human judgment remains a critical element in decision-making processes, preventing over-reliance on AI systems. In addition to human oversight, provisions mandate that risk mitigation strategies be put in place to address any potential unintended consequences or algorithmic discrimination that may arise from AI tools.

By incorporating human oversight, these legislative provisions ensure that AI systems do not operate in isolation and that human professionals retain the ability to review and intervene in AI-driven decisions.

Recommendations and Legislative Actions for AI Improvement

Continuous improvement in AI systems requires ongoing feedback and recommendations for policy changes. Many legislative provisions call for regular reports from AI oversight bodies, which are submitted to legislative entities outlining the progress, challenges and proposed improvements for AI systems.

These reports often include recommendations for enhancing AI systems' effectiveness, fairness and safety. Legislative bodies use these findings to refine existing regulations, address emerging challenges and support the responsible use of AI technologies in both public and private sectors.

Risk Mitigation and Safeguards Against Algorithmic Discrimination

Provisions focused on risk mitigation ensure that AI systems are designed and deployed with safeguards to prevent biases, discrimination and unintended harmful outcomes. These safeguards are particularly important in high-risk AI systems, where the consequences of failure or bias can be severe.

In sectors like health care, policy provisions require that diagnostic algorithms used in clinical settings undergo regular evaluations to ensure they do not introduce discriminatory biases into medical decision-making. Insurers and health care providers must also conduct utilization reviews to ensure that AI algorithms adhere to clinical guidelines and minimize bias when determining care eligibility.

By establishing rigorous risk mitigation protocols, these legislative provisions help ensure that AI systems operate fairly and do not contribute to discriminatory outcomes, particularly in critical sectors like health care and employment.

Addressing Privacy and Security Concerns

Data privacy and security are critical considerations in AI quality improvement legislation. Provisions often require developers and deployers of AI systems to conduct data protection

assessments to ensure that personal data is handled securely and in compliance with privacy laws.

In sectors like health care, where AI systems process sensitive personal data, privacy safeguards are essential to protect patient information. Legislative provisions mandate that AI systems used in health care comply with privacy regulations such as the Health Insurance Portability and Accountability Act (HIPAA) and that appropriate cybersecurity measures are in place to prevent data breaches.

Integration of AI into Government Services

In addition to sector-specific provisions, some legislative frameworks include measures for the integration of AI into government services. These provisions require detailed feasibility studies to assess the potential benefits and risks of implementing AI technologies in public administration. Research plans and cost-benefit analyses are used to guide government agencies in determining how AI can improve service delivery while ensuring that privacy, equity and cybersecurity concerns are properly managed.

Provisions related to recommendations for AI implementation ensure that government agencies follow clear guidelines and risk management strategies when adopting AI technologies, ensuring that they enhance efficiency without compromising ethical standards or individual rights.

Human Oversight

Human oversight in AI policy provisions is essential for ensuring ethical standards, accountability and fairness in AI deployment. These provisions emphasize the need for human involvement in monitoring, reviewing and regulating AI systems to prevent biases, discrimination and other risks. The following summary outlines key themes surrounding human oversight in AI policy provisions, focusing on risk management, accountability, governance, transparency and the ethical use of AI technologies.

Risk Management and Continuous Monitoring

Risk management and ongoing monitoring are central to ensuring AI systems operate safely and fairly. Many provisions require human oversight in high-risk AI systems to manage potential issues after deployment. This includes continuous post-deployment assessments to detect and correct biases or errors that may arise during the AI system's operational phase.

For high-risk AI applications, such as those used in public services or employment, human reviewers are responsible for actively monitoring the AI's outputs to ensure that any adverse effects are identified and addressed in real-time. This ensures that AI systems are not left unchecked and that potential harms are mitigated promptly.

Furthermore, some provisions mandate that companies deploying AI tools in critical areas like health care, law enforcement and employment establish comprehensive risk management programs. These programs are structured to include human oversight mechanisms, ensuring that decisions made by the AI systems can be reviewed and corrected as necessary. In the employment context, for example, AI systems used for recruitment and hiring decisions must undergo regular bias audits to prevent discriminatory outcomes. Human reviewers are required to intervene when AI outputs are flagged as biased or unfair, ensuring a continuous review process that upholds fairness in decision-making.

Accountability in Employment Decision Tools

AI systems in employment decision-making contexts are subject to provisions requiring significant human oversight. Employers are often mandated to notify candidates when AI tools are used in assessing their qualifications, providing transparency in the process. Beyond notification, human intervention is required when decisions are based on AI outputs.

The policy provisions suggest that employers must regularly conduct bias audits and impact assessments to ensure AI systems do not disproportionately affect historically marginalized groups, such as those based on race or gender. These audits must be conducted annually, with human oversight ensuring that any adverse effects identified are addressed through corrective measures. In cases where AI systems lead to decisions with significant consequences, such as hiring or promotions, human reviewers must have the authority to alter or override AI-generated outcomes.

In addition, provisions often require that employers document their use of AI in employment decisions and ensure that internal reviewers possess the expertise to evaluate the fairness and accuracy of AI-generated outcomes. This dual approach — combining regular audits with active human oversight — aims to maintain transparency, accountability and fairness in employment-related AI applications.

Governance and Safeguards for Algorithmic Decision Systems

Governance frameworks within the provisions emphasize the need for structured oversight programs to manage the use of automated decision systems (ADS). These governance programs are designed to map, measure and mitigate the risks associated with AI-driven decision-making, particularly in contexts where AI systems could lead to biased or discriminatory outcomes.

Many provisions advocate for the establishment of dedicated oversight bodies within organizations, such as an AI governance board or a designated officer responsible for overseeing the ethical use of AI systems. These bodies or individuals are tasked with conducting regular risk assessments and ensuring compliance with legal and ethical standards. Such governance frameworks are crucial for maintaining the integrity of AI systems in sectors where decisions can significantly impact individuals' lives, such as public services, law enforcement and social welfare programs.

Safeguards in these governance programs also include regular audits, bias detection mechanisms and detailed documentation of AI processes. In cases where automated systems may disproportionately affect marginalized communities, human oversight is required to adjust or discontinue the use of these tools until the risks are adequately addressed.

Transparency, User Rights and Accountability

Transparency and protecting user rights are critical components of human oversight in AI provisions. Many of these provisions ensure that individuals impacted by AI systems are informed of the decisions made by these systems and provided with recourse to contest or appeal these decisions.

Provisions often require that individuals be notified when AI tools are used to make significant decisions, such as those affecting employment, health care or public services. These notices must clearly explain how the AI system operates, they data they use and how individuals can request a human review of any AI-generated decisions. The aim is to ensure that AI systems do not operate in a "black box" manner, where their inner workings are opaque and inaccessible to those affected.

Further, many provisions mandate that AI systems provide users with a clear path to request human intervention. For example, if an AI system produces a decision that negatively impacts

an individual's legal rights or employment status, the affected individual has the right to appeal the decision and request a review by a qualified human agent. This ensures that AI systems do not act as the final arbiters in decisions that carry significant personal or societal consequences.

Ethical Use of Generative AI

As generative AI systems become more prevalent, provisions call for heightened human oversight to ensure these systems are used ethically and responsibly. Provisions related to generative AI stress the importance of maintaining human agency, particularly in high-risk areas where AI-generated content or decisions could be manipulated for disinformation or fraud.

Generative AI systems must undergo rigorous testing and security evaluations before being deployed for public use. Human oversight is necessary to ensure that these systems do not produce harmful or misleading content, especially in political or commercial contexts. Moreover, the provisions stress that generative AI must always be supplemented by human decision-making to prevent the spread of disinformation or unethical practices.

Impact Assessments and Bias Audits

Impact assessments and bias audits are key to ensuring fairness and non-discrimination in AI systems. Provisions often require that AI systems undergo regular audits to detect and mitigate potential biases, particularly in employment, health care and public service applications. Independent third-party auditors may be employed to evaluate the extent to which AI systems produce fair and equitable outcomes.

For instance, in the employment sector, AI systems used for recruitment or performance evaluations must be audited to ensure that they do not disproportionately disadvantage specific demographic groups. These audits are conducted regularly and are supplemented by internal human oversight mechanisms to ensure that any biased outcomes are corrected swiftly.

Impact assessments focus on analyzing the broader societal effects of AI systems, ensuring that they comply with anti-discrimination laws and promote equitable outcomes. These assessments are particularly emphasized in provisions related to public service delivery and health care, where biased AI outputs could have severe consequences for historically marginalized populations.

Training and Expertise for Oversight Personnel

Provisions underscore the importance of training and expertise for personnel tasked with overseeing AI systems. Human reviewers must have a thorough understanding of how AI systems function and be capable of interpreting AI-generated outcomes critically. In employment contexts, for example, internal reviewers must have the authority to override AI decisions when discrepancies or biases are detected.

The policy provisions also emphasize that individuals responsible for AI oversight receive ongoing training to keep up with technological developments and regulatory requirements. This training is vital in ensuring that human oversight personnel are equipped to manage AI systems responsibly and maintain the integrity of decision-making processes.

Additionally, provisions often require that organizations deploy oversight personnel with the authority and resources to intervene when necessary. This ensures that AI systems remain subject to human judgment and that their outputs are continuously evaluated for fairness and accuracy.

Impacted Populations

State legislation on artificial intelligence (AI) increasingly emphasizes protecting historically marginalized populations from the risks posed by AI systems. These laws focus on preventing discrimination, ensuring oversight, safeguarding data privacy and promoting inclusivity in AI governance. This summary highlights how various provisions address the concerns of impacted populations such as racial minorities, children, LGBTQ+ communities and individuals with disabilities, drawing solely from the detailed legislative summaries provided.

Protecting Historically Marginalized Populations

Many AI policy provisions explicitly prioritize the protection of historically marginalized groups. A major focus is placed on safeguarding children, with specific working groups dedicated to examining the potential harms AI could pose to young people, including the spread of misinformation and the exacerbation of social inequalities. These working groups collaborate with various stakeholders, such as consumer advocacy organizations and academic institutions, to develop mitigation strategies tailored to protect children from AI-assisted risks. This proactive approach reflects a legislative commitment to shielding younger populations from the unintended consequences of AI technologies.

Similarly, individuals with disabilities are recognized as a key group requiring protection in the AI landscape. AI systems deployed in areas such as employment, education and health care must be assessed for accessibility to ensure that they do not create barriers for people with disabilities. For example, AI used in employment screening or educational tools is required to accommodate the needs of disabled individuals, ensuring that these populations are not excluded from essential services or opportunities. Where potential accessibility issues are identified, developers and employers are mandated to take corrective actions. These measures ensure that people with disabilities are not inadvertently left behind by AI technologies and that AI systems foster inclusivity rather than exclusion.

Preventing Discrimination and Bias in AI Systems

Preventing unlawful discrimination is one of the primary objectives of state-level AI legislation. AI systems, especially those involved in decision-making processes across sectors like employment, housing, health care and credit, are at risk of perpetuating societal biases if left unchecked. Legislators have introduced provisions that require bias audits and impact assessments to mitigate these risks, ensuring that AI outputs do not disproportionately affect individuals based on characteristics such as race, gender, ethnicity, disability and sexual orientation.

For example, in the context of employment, AI tools used to screen job applicants must be regularly audited to ensure they do not unfairly disadvantage certain groups. Employers are required to conduct these audits to ensure that their AI systems do not disproportionately affect individuals based on their race, gender or disability. Similarly, in the housing and credit sectors, AI decision-making tools are subject to stringent evaluations to prevent discriminatory outcomes. These audits ensure that marginalized populations are not unfairly denied access to essential services or opportunities, and developers must take immediate corrective actions if biases are detected.

Public services are another area of concern. Legislative provisions require AI systems used in government services, such as public benefits allocation, to be continually monitored and subject to human review. This oversight is critical in preventing discriminatory outcomes for populations reliant on government assistance, such as individuals who are low-income or from racial and ethnic minority groups. By mandating that these AI systems be regularly assessed for fairness, legislators aim to prevent the widening of existing inequalities in access to public resources.

Oversight and Accountability for High-Risk AI Systems

High-risk AI systems — are typically referred as high risks because they are involved in highly consequential decision-making — play a critical role in determining access to fundamental resources and opportunities, such as employment, health care, housing and education. These domains significantly influence individuals' quality of life, economic stability and social mobility, making fairness and accountability in decision-making processes essential. For instance, AI systems used in hiring processes may screen resumes or rank candidates, potentially introducing biases that disadvantage certain groups. In health care, AI-driven diagnostic tools or treatment recommendations could inadvertently overlook or misdiagnose conditions in underserved populations. Housing-related AI systems, such as those used for mortgage approvals or tenant screenings, can perpetuate discriminatory practices by relying on historical data reflecting systemic biases. Similarly, in education, AI tools for admissions, resource allocation or performance evaluation may unintentionally reinforce existing inequities, further disadvantaging certain populations of students.

Some legislative provisions require thorough evaluations of high-risk AI systems to identify potential harms across these key areas, particularly for historically marginalized populations. For example, AI systems deployed in health care settings must be assessed to ensure they do not deny or limit care based on an individual’s race or socioeconomic background. Similarly, AI systems used in housing must be evaluated for their impact on minority groups to prevent discriminatory denials of access to housing services.

Developers and deployers of high-risk AI systems are required to certify compliance with anti-discrimination regulations and regularly report on their systems’ impacts. These public certification requirements foster transparency and ensure that AI systems in critical sectors operate in ways that promote equity rather than perpetuate biases. Continuous oversight, including bias audits and impact assessments, is mandated to ensure that any discriminatory impacts are identified and addressed promptly.

This focus on accountability reflects a growing recognition of the risks AI poses to historically marginalized populations when used in decision-making processes that can have lasting, life-altering consequences. Legislative efforts aim to ensure that AI systems, particularly those classified as high-risk, are continually monitored and held accountable for their impacts on impacted populations.

Data Privacy and Protection for Sensitive Populations

As AI systems rely heavily on personal data to function, there is significant concern about the potential misuse of sensitive data, particularly data belonging to historically marginalized populations. Legislative provisions emphasize the protection of personal data, especially in areas such as health, education and criminal justice, where the inappropriate use of data can disproportionately harm disadvantaged groups.

For instance, provisions require health insurers to disclose their use of AI in processing claims and to provide transparency around how claims are denied or approved. This level of transparency is intended to protect individuals, particularly those from low-income or underserved communities, from being unfairly denied access to essential health services due to opaque AI decision-making processes.

In public services, where automated systems may determine eligibility for benefits or assistance, provisions mandate regular human oversight to ensure that data processing does not result in biased outcomes. For example, databases used to verify individuals' eligibility for reemployment benefits or public assistance must be carefully monitored to prevent false exclusions based on discriminatory criteria. These safeguards ensure that sensitive data is not used to unfairly disadvantage individuals, particularly those relying on government support.

Additionally, protections are in place for populations at risk of data exploitation, such as children and minors. Legislative provisions shield the personal data of students and minors participating in public or recreational programs from public disclosure. This reflects a broader recognition of the importance of data privacy, particularly for young people who are increasingly subjected to data collection in educational and social environments.

Inclusive AI Governance and Representation

A recurring theme in AI legislation is the inclusion of historically marginalized communities in the decision-making processes that govern AI development and deployment. To ensure that AI policies reflect the needs of impacted populations, many legislative provisions require the establishment of advisory councils or working groups that include representatives from historically marginalized groups, such as racial and ethnic minorities, LGBTQ+ communities and people with disabilities.

These advisory bodies are tasked with monitoring the impact of AI on specific communities and making recommendations to prevent discriminatory outcomes. By ensuring that individuals from impacted populations are part of the governance structure, legislators aim to create a more equitable and inclusive approach to AI regulation. These councils provide a platform for marginalized voices to influence how AI technologies are developed and implemented, particularly in areas where algorithmic bias is likely to arise.

In addition to including impacted populations in governance, these provisions also emphasize the importance of transparency and public accountability in AI systems. Policies require that the use of AI in critical areas, such as public services and health care, be clearly communicated to affected individuals. This ensures that populations impacted by AI-driven decisions are aware of the technology's role and can hold decision-makers accountable.

Transparency

As AI becomes increasingly integrated into various sectors, state legislation has focused on ensuring transparency to promote ethical AI use and accountability. These provisions address the impact of AI technologies, particularly automated decision systems, on areas such as health care, employment, government services and consumer protection. This summary examines the key themes of transparency in AI-related policy provisions, covering topics like automated decision-making systems, data privacy, government accountability, public engagement and the governance of high-risk AI systems.

Transparency in Automated Decision Systems and AI-Driven Services

One major focus of AI policy provisions is ensuring transparency in the use of automated decision systems (ADS) and AI-driven services, especially in health care, employment and public services. These provisions ensure that institutions using AI notify affected individuals, provide clear information about how decisions are made and offer options to challenge or seek human intervention in those decisions.

In health care, provisions require providers to inform patients when diagnostic algorithms are used in their care. Patients also are given the option to opt out of the use of AI-driven diagnostic tools, allowing them to maintain control over their treatment decisions. Additionally, health care facilities are required to disclose information about staffing methodologies, including the use of algorithms to determine nurse-to-patient ratios, ensuring that both staff and patients are aware of how technology influences health care services.

In employment settings, transparency provisions focus on ensuring that employers notify employees and job applicants when AI tools are used in recruitment, promotion or other employment decisions. These tools must undergo bias audits to ensure compliance with anti-discrimination laws. The provisions are designed to protect workers from potential bias or unfair outcomes produced by AI systems and to promote fair employment practices.

Similarly, provisions for government agencies require transparency in the use of AI in public services. Agencies must notify individuals when AI is used in communications or decision-making processes and provide options for citizens to bypass AI systems in favor of interacting with a human representative. These provisions are intended to ensure that citizens remain fully informed about how AI affects their interactions with government services.

Data Privacy, Algorithmic Transparency and Consumer Protection

Another central theme in these policy provisions is transparency in data privacy and algorithmic decision-making, particularly given the extensive use of personal data in AI systems. These provisions require organizations to disclose how personal data is collected, processed and used, especially when this data informs AI models that impact high-stakes decisions such as health care or employment outcomes.

For example, health care providers must disclose their data collection practices, particularly when AI systems are involved in processing claims or determining treatment options. This ensures that patients are fully aware of how their sensitive health data is managed and what safeguards are in place to protect their privacy.

Consumer protection also is emphasized in provisions related to AI systems. Companies developing AI systems, particularly those involved in creating synthetic media or AI-generated content, are required to inform consumers about the nature of AI-generated material. These provisions also mandate that developers disclose information about the datasets used to train AI models, especially in cases where personal or copyrighted information is included. These measures help consumers understand the ethical and legal implications of AI-generated content.

Furthermore, provisions require that organizations using algorithms for decision-making, whether in employment, housing or credit evaluations, conduct regular audits to prevent discriminatory outcomes. The results of these audits must be made publicly available, ensuring that consumers can hold organizations accountable for how AI systems affect them.

Government Accountability and Public Reporting on AI Use

Provisions also focus on ensuring government accountability and transparency in the use of AI technologies. Government agencies are required to maintain inventories of all AI systems in use, detailing the systems' capabilities, vendors, the decisions they support and the extent of human oversight involved. These inventories are made publicly accessible, allowing for external scrutiny and ensuring transparency in governmental AI usage.

For instance, some provisions require that state agencies conduct annual inventories of AI systems and report these findings to the public. This allows citizens to understand how AI is being implemented in government operations and ensures public oversight.

In addition to these inventories, provisions mandate regular reporting by agencies to legislative bodies on the use of AI systems. These reports detail the impact of AI on critical areas like equity, cybersecurity and legal rights. By mandating public reporting and audits, these provisions aim to foster accountability in how government agencies use AI.

Public Engagement and Ethical AI Governance

Public engagement and ethical governance also are emphasized in AI policies. Provisions call for involving diverse stakeholders in the development and implementation of AI policies. The establishment of advisory councils, task forces or working groups to oversee AI use is a key component of these provisions.

These advisory bodies are tasked with engaging a broad spectrum of stakeholders, including academia, businesses and consumer advocacy groups, to ensure that AI policies reflect a wide range of perspectives. They also are responsible for making policy recommendations on how AI can be used responsibly and ethically in both public and private sectors.

Additionally, these advisory bodies issue regular reports on the societal impacts of AI technologies, including risks such as algorithmic bias or misinformation. By mandating open meetings and public access to records of discussions, these provisions ensure that the public remains informed and involved in ongoing policy development. Public engagement is seen as essential to shaping AI governance that is inclusive, transparent and aligned with societal values.

Transparency in High-Risk and Generative AI Systems

High-risk AI systems — those that significantly impact individuals' lives, such as AI used in hiring, credit decisions or health care — are subject to particularly stringent transparency provisions. Organizations deploying these systems must notify individuals when AI is used in decision-making processes, explain the system's function and disclose any associated biases or risks.

In employment settings, for example, AI systems used for hiring or promotion decisions must be regularly audited for bias, with the results shared with both employees and applicants. These transparency measures are intended to prevent discriminatory outcomes and ensure that AI systems are used fairly and ethically.

Generative AI systems, which create synthetic media or content, also are addressed in these policy provisions. Developers of such systems are required to provide clear disclosures to users, informing them when content is generated by AI. This is particularly important in preventing misuse or deception associated with AI-generated media. Provisions also require developers to publicly disclose information about the datasets used to train generative AI models, ensuring that these systems operate within ethical and legal boundaries.

Unauthorized Use and Prohibitions

This summary examines the themes related to unauthorized use and prohibitions. It is organized under distinct areas such as prohibitions on granting AI legal personhood, limitations on AI in critical decision-making, protections against algorithmic discrimination and restrictions on AI use in surveillance, health care and employment.

Prohibitions on AI Legal Personhood

A central provision in many state policy frameworks prohibits the recognition of legal personhood for AI systems. These provisions ensure that AI, as well as other nonhuman entities like inanimate objects or natural elements, cannot be conferred legal status or rights. By clearly limiting legal personhood to human beings, these provisions prevent AI systems from functioning as independent legal actors. The effect of this restriction is to ensure that AI remains a tool under human control, with no ability to hold rights or responsibilities akin to human beings.

Restrictions on AI in Critical Decision-Making Areas

Provisions frequently restrict AI from making independent decisions in critical areas where human oversight is necessary. In sectors such as health care, public services and insurance, AI is prohibited from being the sole decision-making entity, particularly in situations where decisions affect individuals' access to essential services or rights.

For instance, in health care, AI systems are prohibited from making autonomous clinical decisions without review by a human professional. Medical personnel must have the authority to override AI-generated recommendations to ensure patient care is not compromised by automated decision-making. In public assistance programs, AI systems may assist in detecting fraud or managing resources, but human review is required for any final decisions regarding eligibility or the provision of benefits. This ensures that critical decisions remain subject to human judgment, protecting individuals from potentially harmful or unfair outcomes caused by automated systems.

Prohibitions on Algorithmic Discrimination

One of the core themes in state AI policy provisions is the prevention of algorithmic discrimination. These provisions prohibit AI from making decisions based on characteristics such as race, gender, national origin and disability. Whether in employment, health care, housing or financial services, the use of AI to perpetuate or introduce discriminatory practices is expressly forbidden.

In employment, AI systems used for hiring or promotion decisions must undergo regular bias audits to ensure that they do not unfairly disadvantage certain groups. Employers are often required to provide transparency regarding how these systems operate and must ensure that any AI tools used comply with anti-discrimination laws. Similar prohibitions apply in health care, where AI algorithms used to determine treatment plans or insurance coverage cannot factor in demographic characteristics in ways that could result in unequal access to services. These provisions help ensure that AI systems do not reinforce or create new forms of bias and discrimination in critical decision-making processes.

Restrictions on AI in Health Care

AI's role in health care is tightly regulated by policy provisions that prohibit the unauthorized use of AI to make independent clinical decisions or replace health care workers. These provisions

mandate that human professionals retain the final authority in determining patient care. For instance, registered nurses and other health care professionals have the right to override AI-generated clinical recommendations if they believe it is in the patient's best interest. This ensures that AI is used as a supportive tool rather than a replacement for human judgment in patient care.

Additionally, hospitals and health care facilities are prohibited from reducing staff levels or replacing essential health care workers with AI systems in an attempt to meet cost-cutting objectives. This restriction ensures that AI-driven efficiency does not come at the expense of patient care quality or workforce stability. AI tools used for diagnostics must be certified for accuracy, and patients must be informed and provide consent before such tools are applied.

Restrictions on AI in Employment Decisions

Several policy provisions restrict the use of AI in employment, especially when it comes to hiring, promotion and workplace monitoring. Employers are prohibited from using AI tools in employment decisions unless these tools have been rigorously tested for bias and transparency. Candidates must be informed when AI is used in hiring processes, and their consent is required before such tools are employed.

Moreover, employers must ensure that AI systems used for employment purposes are regularly audited to identify any potential biases. If AI systems are found to be discriminatory, they may be restricted or banned from use. Employers who violate these provisions may face penalties, including fines and legal action, underscoring the importance of ethical and fair AI deployment in the workplace.

Prohibitions on AI in Public Benefits and Fraud Detection

Provisions governing the use of AI in public benefits programs place strong limitations on the role AI can play in fraud detection or resource management. AI systems are prohibited from being the sole determinants in decisions that impact individuals' access to public benefits, such as Social Security or unemployment assistance. Instead, human oversight is required to review and confirm any AI-generated findings before benefits can be denied, reduced or terminated.

These restrictions are designed to prevent errors or biases in AI systems from unfairly harming individuals who rely on public assistance. While AI may be used to flag potentially fraudulent activities or streamline resource distribution, the final decision must always involve a human review process to ensure fairness and accuracy.

Prohibitions on AI-Driven Surveillance and Data Collection

AI systems used for surveillance are subject to strict prohibitions, especially when it comes to unauthorized monitoring or biometric data collection. Policy provisions often prohibit AI from conducting surveillance in public spaces or using biometric technologies, such as facial recognition, without proper legal authorization. These restrictions are meant to safeguard individual privacy and prevent the misuse of AI by government agencies or private entities.

Furthermore, employers are prohibited from using AI systems to conduct unauthorized surveillance of employees, including monitoring their personal spaces, such as homes or private vehicles. Provisions also restrict the use of AI tools, such as emotion detection or gait analysis, to monitor workers in ways that could infringe on their privacy or autonomy. These policies ensure that AI-driven surveillance remains within ethical and legal bounds, protecting individual privacy in both public and private settings.

Prohibitions on AI for Deceptive Practices

The unauthorized use of AI for fraudulent or deceptive purposes is strictly prohibited by various state policy provisions. AI systems are barred from creating or disseminating false information, such as deepfakes[1] or misleading media, designed to deceive the public. Provisions mandate that any AI-generated content must include clear indicators, such as watermarks, to signal that it was produced by an AI system.

Additionally, AI tools cannot be used to impersonate individuals without their consent, whether in commercial, political or personal contexts. These prohibitions aim to prevent the manipulation of public perception and the spread of disinformation using AI-generated content, protecting individuals and the broader public from deceptive practices.

Prohibitions on AI in Algorithmic Rent Setting

A specific prohibition that has emerged in the housing sector involves the use of AI systems to set or recommend rental prices. Landlords are prohibited from using AI-driven algorithms to

[1] A specific kind of synthetic media where a person in an image or video is swapped with another person's likeness. Somers, M. (2020, July 21). *Deepfakes explained*. MIT Sloan School of Management. Retrieved January 6, 2025, from https://mitsloan.mit.edu/ideas-made-to-matter/deepfakes-explained

determine rental amounts, as these practices are viewed as potentially exploitative and unfair. The use of AI in rent setting could result in opaque pricing models that lack accountability, disproportionately affecting tenants. These provisions are designed to protect renters from being subject to inflated or biased pricing systems controlled by AI algorithms.

Appendix A. Methodology Protocol for Completing Steps in Using Petal and ChatGPT for Policy Analysis

This protocol was used to analyze AI policies and guidelines at the state level, specifically introduced and enacted state legislation. The goal was to identify common components and provisions, with an emphasis on equity, ethical considerations, transparency, human oversight and data privacy.

The analysis aimed to answer the question: *What are the common components and provisions found in AI policies across different states, with a specific emphasis on equity, ethical considerations, transparency, human oversight and data privacy?*

Step-by-Step Approach

Step 1: Upload and Prompt

After downloading the policies, researchers attached the policy files from the legislative trackers found on the websites of the Association of State and Territorial Health Officials (ASTHO) and the National Conference of State Legislatures (NCSL) into Petal using the upload/import feature. Then, three separate tables were created, one for each set of legislative policies included within our analysis — the NCSL 2023 AI Policies, NCSL 2024 AI Policies and ASTHO AI Policies. Within each table, columns (listed below) were added via the “Add Question” button. When adding each column, the user input the desired title of their column along with the specialized prompt that was developed by the users to obtain accurate, specific summaries from each uploaded policy. Once the columns were added, the researchers began running their analysis through Petal by clicking on the half-circle arrow button at the top right corner of each column, which instructs the system to generate responses for each document in the table within the selected column. The prompts asked for direct quotes, which in many cases included section and, if applicable, subsection where the policy provision was found within the document.

The columns in the evidence table included:

1. Document
2. State
3. Legislation Number
4. Status
5. Type of Document
6. Enacted By
7. Collaborators
8. Relevant Policies
9. General Provisions
10. Bias Mitigation
11. Data Privacy
12. Equity and Ethics
13. Community/Public Engagement
14. Acquisition of AI Tools
15. Quality Improvement (QI)
16. Transparency
17. Human Oversight
18. Environmental Impact
19. Impacted Populations
20. Outline
21. Unauthorized Use/Prohibitions

Step 2: Review Results

Researchers read the results included in the table and compared them to the policy by reading the provision.

1. If any information was incorrect, a correction was made and it was noted that AI had produced an error.
2. If any information was unclear, a follow-up prompt was used and Petal was instructed to re-run the column for additional analysis.

Step 3: Complete Evidence Tables

Researchers completed the evidence tables.

- Researchers downloaded each evidence table from Petal into Excel.
- By using direct quotations as well as section and subsection numbers referenced by Petal in summaries, researchers spot-checked the documents to ensure that the information was accurate.

Step 4: Summarize Information Using ChatGPT

Researchers summarized the information using ChatGPT.

- Researchers copied the information from each relevant column, one at a time, for all three evidence tables.
- Researchers pasted all the copied information into ChatGPT and instructed the system to provide a summary for each column.
- Once all summaries are developed, the researchers developed comprehensive summaries that integrated information from all three evidence tables by copying and pasting each column category from all three tables into ChatGPT. The researchers then asked ChatGPT to develop the summaries utilizing the following prompt:

I am going to provide you with a set of state policy provisions (legislation) relating to (insert column name) in AI policy. Your task is to create a narrative summary based only on the content I provide, adhering strictly to the following guidelines:

The summary should be approximately 2,000 words, but no longer. Slight flexibility in word count is acceptable, but do not exceed 2,000 words.

Organize the summary using subtitles based on themes that emerge from the provisions. Ensure that the sub-titles are clear and accurately represent the key sections or ideas within the provisions. Do not use any external information, opinions or interpretations. Every detail in the summary must come exclusively from the provisions I provide. Avoid adding any information that is not explicitly stated or implied in the provisions. Do not make anything up — stay entirely faithful to the original content of the provisions. After I provide the provisions, please proceed to create the summary accordingly. Ask me anything before you begin.

Figure A-1. **Sample Prompts for Summary Creation and Quality Assurance Strategies**

Category	ChatGPT/Petal Prompt	Quality Assurance
A-1. State	"What state does this policy involve? Write the state only."	Cross reference listed state with uploaded legislative document.
B-1. Status	"Using the title of the resource, note the status of the state legislation. Example: Enacted, Pending - Carryover, Pending, Failed."	Verify status utilizing ASTHO, NCSL and online bill tracking platforms.
C-1. Type of Document	"Note the type of legislation. Example: Bill, resolution."	Cross reference listed document type with document title.
D-1. Enacted By	"Specifically, what entity, agency, organization or office will be responsible for enacting the policy?"	Check that the entities charged with enacting each policy are thoroughly summarized.
E-1. Collaborators	"What entities or who specifically is also involved or will be impacted by this state policy?"	Check that the organizations, agencies and individuals impacted by the policy are thoroughly summarized.
F-1. Relevant Policies	"What current state statutes are being amended for this state legislation?"	Verify the completeness and accuracy of the relevant policies.
G-1. General Provisions	"Provide a summary of provisions included in the policy."	Verify the completeness and accuracy of the general provisions.
H-1. Bias Mitigation	"Provide a description of the provisions specific to bias mitigation and provide the provisions verbatim."	Verify the completeness and accuracy of the bias mitigation provisions.
I-1. Data Privacy	"Provide a summary description of all data privacy provisions and provide provisions verbatim."	Verify the completeness and accuracy of the data privacy provisions.
J-1. Equity and Ethics	"Provide a description of all provisions related to equity or ethics and provide provisions verbatim"	Verify the completeness and accuracy of the equity and ethics provisions.
K-1. Community/Public Engagement	"Provide a description of community/public engagement and provide these relevant provisions verbatim."	Verify the completeness and accuracy of the community/public engagement provisions.
L-1. Acquisition of AI Tools	"Provide the requirements for acquisition of AI tools/models and note specific provisions verbatim."	Verify the completeness and accuracy of the acquisition of AI tools provisions.
M-1. Quality Improvement	"Provide a description of provisions related to evaluation and quality improvement and provide specific provisions verbatim."	Verify the completeness and accuracy of the quality improvement provisions.

Figure A-1 (continued). **Sample Prompts for Summary Creation and Quality Assurance Strategies**

Category	ChatGPT/Petal Prompt	Quality Assurance
N-1. Transparency	"Provide a description of provisions related to transparency and provide specific provisions verbatim."	Verify the completeness and accuracy of the transparency provisions.
O-1. Human Oversight	"Provide a description of provisions related to human oversight and provide specific provisions verbatim."	Verify the completeness and accuracy of the human oversight provisions.
P-1. Environmental Impact	"Provide a description of provisions related to environmental impact of AI and provide specific provisions verbatim."	Verify the completeness and accuracy of the environmental impact provisions.
Q-1. Impacted Populations	"Provide a description of provisions related to specific populations/demographics and provide specific provisions verbatim"	Verify the completeness and accuracy of the impacted populations provisions.
R-1. Unauthorized Use/Prohibitions	"Provide a description of provisions related to unauthorized use/prohibitions and provide specific provisions verbatim."	Verify the completeness and accuracy of the unauthorized use/prohibitions provisions.
S-1. Legislation Number	"List bill/resolution number with prefix (example: SB, HB, AB, S, H, etc.)"	Cross-reference the listed legislation number with the uploaded legislative document.

Lessons Learned from Using AI Tools

The information provided below is not intended to endorse any specific AI tool but is primarily meant to provide a high-level description of the lessons learned related to their usage in the conducted literature review.

In general, both tools — ChatGPT and Petal — demonstrated their utility in the policy review process. The team utilized both free and paid subscription versions of ChatGPT. Advantages included the identification of articles and the development of summaries organized by themes based on prompts with reasonable accuracy. However, in some cases, the summaries were too general and required additional editing through further prompting. An additional issue occasionally observed was that some summaries included information not sourced from the policy provided. To mitigate this, the research team implemented measures, including disabling

ChatGPT's memory and internet browsing capabilities in the settings, to ensure outputs were based on the policy content.

The second AI tool used was Petal. Given that the functionality of the Petal tool differs from ChatGPT — it does not browse the web and works solely with the documents provided — the research team did not need to implement any extra measures beyond standard review and quality assurance to ensure that the summaries accurately reflected the content of the provided policies. However, some limitations noted included challenges in recognizing specialized terminology and phrases used in the policies which sometimes resulted in outputs that did not fully capture the intended meaning. The team addressed this issue by conducting manual reviews to correct AI-generated outputs, ensuring they aligned with the intended meaning and by adding additional prompts to guide the tool toward producing more accurate and relevant results.

Further testing of these and other AI tools is essential to identify the most effective ways to leverage their capabilities for policy reviews in the future, while also addressing potential ethical considerations.

Appendix B: Bibliography

Association of State and Territorial Health Officials. (2024) *Public Health Infrastructure Legal Map*. Retrieved from https://www.astho.org/advocacy/state-health-policy/public-health-legal-mapping-center/public-health-infrastructure/

National Conference of State Legislatures. (2024, January 12). *Artificial Intelligence 2023 Legislation*. Retrieved from https://www.ncsl.org/technology-and-communication/artificial-intelligence-2023-legislation

National Conference of State Legislatures. (2024, September 9). *Artificial Intelligence 2024 Legislation*. Retrieved from https://www.ncsl.org/technology-and-communication/artificial-intelligence-2024-legislation

Appendix C: Reviewed Legislative Documents

Alabama H 168 (2024)

Alaska H 107 (2023)

California AB 1282 (2023)

California AB 1502 (2023)

California AB 2013 (2023)

California AB 2602 (2023)

California AB 2885 (2023)

California AB 2930 (2024)

California AB 302 (2023)

California AB 3050 (2023)

California AB 331 (2023)

California AB 892 (2023)

California AB 2930 (2024)

California AJR 6 (2023)

California ACR 96 (2023)

California SB 294 (2023)

California SB 313 (2023)

California SB 398 (2023)

California SB 721 (2023)

California SB 893 (2024)

California SB 896 (2023)

California SB 942 (2023)

California SB 970 (2023)

California SB 1120 (2023)

California SCR 17 (2023)

Colorado H 1057 (2024)

Connecticut SB 1103 (2023)

Connecticut Substitute Bill 2 (2024)

Delaware HB 333 (2024)

District of Columbia B 114 (2023)

Florida HB 1289 (2024)

Florida HB 1459 (2024)

Florida S 1680 (2024)

Florida S 7018 (2024)

Florida SB 972 (2024)

Georgia H 887 (2023)

Georgia HB 890 (2023)

Georgia HB 988 (2023)

Hawaii H 2152 (2023)

Hawaii HB 2245 (2024)

Hawaii HCR 71 (2023)

Hawaii S 2284 (2023)

Hawaii S 2285 (2023)

Hawaii SB 2879 (2024)

Idaho HB 568 (2024)

Illinois H 1002 (2023)

Illinois H 2472 (2023)

Illinois H 3338 (2023)

Illinois H 3563 (2023)

Illinois H 3773 (2023)

Illinois H 4705 (2023)

Illinois H 5099 (2023)

Illinois H 5115 (2023)

Illinois H 5116 (2023)

Illinois H 5395 (2023)

Illinois H 5649 (2023)

Illinois HB 4836 (2024)

Illinois HB 5322 (2024)

Illinois S 2314 (2023)

Illinois S 2795 (2023)

Indiana SB 150 (2024)

Iowa H 2090 (2023)

Kentucky H 732 (2024)

Louisiana SCR 49 (2023)

Maine H 1270 (2023)

Maine LD 949 (2023)

Maine S 656 (2023)

Maryland H 582 (2024)

Maryland H 1141 (2024)

Maryland H 1174 (2024)

Maryland H 1255 (2024)

Maryland H 1271 (2024)

Maryland H 1297 (2024)

Maryland H 883 (2024)

Maryland SB 818 (2024)

Maryland SB 955 (2024)

Maryland SB 957 (2024)

Maryland SB 979 (2024)

Maryland SB 1068 (2024)

Massachusetts H 64 (2023)

Massachusetts H 1873 (2023)

Massachusetts H 1974 (2023)

Massachusetts H 4024 (2023)

Massachusetts S 33 (2023)

Minnesota H 1814 (2023)

Nebraska LB 1253 (2023)

Nebraska LB 1284 (2023)

New Hampshire H 1688 (2023)

New Hampshire HB 556 (2023)

New Jersey A 3854 (2024)

New Jersey A 3855 (2024)

New Jersey A 3858 (2024)

New Jersey A 4292 (2024)

New Jersey S 1588 (2024)

New Jersey S 2964 (2024)

New Jersey S 3298 (2024)

New Jersey S 3876 (2023)

New Mexico HB 184 (2024)

New Mexico SB 130 (2024)

New York A 5309 (2023)

New York A 567 (2023)

New York A 7838 (2023)

New York A 7859 (2023)

New York A 8129 (2023)

New York A 8328 (2023)

New York A 9149 (2023)

New York A 9314 (2024)

New York A 9315-A (2024)

New York A 9430-B (2024)

New York S 5641-A (2023)

New York S 7543 (2023)

New York S 7623 (2023)

New York S 8755 (2024)

North Carolina S 508 (2023)

North Dakota HB 1003 (2023)

Oklahoma H 3577 (2023)

Oklahoma HB 3828 (2024)

Pennsylvania H 1304 (2023)

Pennsylvania H 1663 (2023)

Pennsylvania HB 1729 (2023)

Pennsylvania HR 170 (2023)

Pennsylvania S 913 (2023)

Rhode Island H 5866 (2023)

Rhode Island H 6285 (2023)

Rhode Island H 6423 (2023)

Rhode Island H 7521 (2024)

Rhode Island H 7786 (2024)

Rhode Island H 8073 (2023)

Rhode Island S 117 (2023)

Rhode Island S 143 (2023)

Rhode Island S 2888 (2024)

Tennessee HB 2325 (2024)

Tennessee HB 2341 (2024)

Tennessee HB 2747 (2024)

Tennessee SB 2461 (2024)

Tennessee SB 2530 (2024)

Texas H 2060 (2023)

Texas H 4695 (2023)

Utah H 249 (2024)

Vermont H 114 (2023)

Vermont H 710 (2023)

Vermont H 711 (2023)

Virginia H 747 (2024)

Washington HB 1616 (2023)

Washington HB 1934 (2023)

Washington HB 1951 (2023)

Washington SB 5356 (2023)

Washington SB 5838 (2023)

Washington Sub HB 1934 (2024)

West Virginia HB 5690 (2024)

Wisconsin AB 1068 (2023)

Wisconsin SB 1010 (2023)