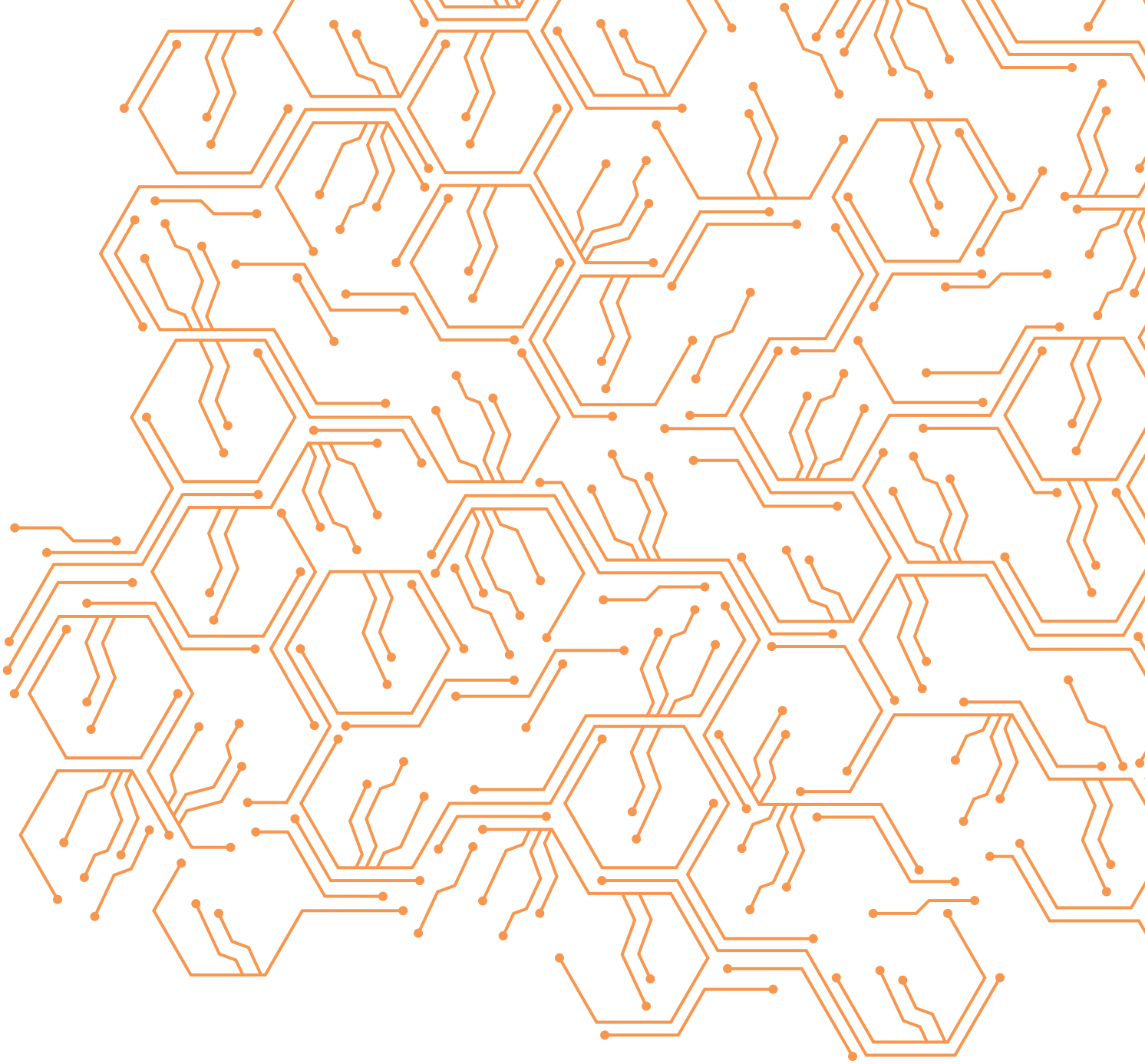

DECEMBER 2025

Artificial Intelligence: Risks and Opportunities

Report Produced for the Washington State Attorney General's Office by The Raben Group

Contents

Introduction

Section 2 (3) (a) of the Artificial Intelligence Task Force's authorizing statute, ESSB 5838, requires the Task Force to provide a "literature review of public policy issues with artificial intelligence, including benefits and risks to the public broadly, historically excluded communities, and other identifiable groups, racial equity considerations, workforce impacts, and ethical concerns." The following report was prepared in June 2025 under contract by The Raben Group and edited by AGO staff. This memo examines the current literature on AI to explain how AI and algorithms work, help policymakers understand the promise and challenge these new technologies present, and make recommendations on how to harness potential while minimizing risk, particularly to our most vulnerable communities.

What is Artificial Intelligence?

This section is not meant to be an exhaustive review of how AI works, but rather a primer for policymakers interested in understanding the basic terminology, design, and functions of AI. For more information, please see technical guidance on AI from McKinsey & Company.[1]

AI operates through the simulation of human intelligence processes in machines, primarily by analyzing data to identify patterns and generate predictions. IBM explains AI as "technology that enables computers and machines to simulate human learning, comprehension, problem solving, decision making, creativity and autonomy."[2]

AI is "only as good as the algorithms and machine learning techniques that guide its actions"[3] and works by "ingesting large amounts of labeled training data, analyzing that data for correlations and patterns, and using these patterns to make predictions about future states."[4]

At its core, AI relies on machine learning (ML), a subset of AI that enables systems to autonomously learn from historical data without explicit programming line by line. ML algorithms process labeled datasets to detect correlations, which inform decision-making or predictive models.[5]

AI is widely used in different sectors. For example:

- **Healthcare:** AI is being integrated into healthcare in multiple ways, including the use of virtual assistants and chatbots to help patients identify symptoms, natural language processing technology to "automate administrative tasks such as documenting patient visits in electronic health records, [which allows] clinicians more time to focus on caring for patients," and the use of large-scale data sets and AI to improve "image-based diagnoses from several medical specialties" such as radiology, dermatology, pathology, and cardiology.[6]
- **Transportation:** The transportation sector is widely using AI, most notably in self-driving cars, but also through intelligent transportation systems that, for example, help with ticketing or provide real-time location and timing updates on public transportation methods.[7]
- **Automated Screening Tools:** AI systems are used to automate many decision-making processes that screen individuals to determine access to goods and services—including government benefits. For example, facial recognition tools use AI software to verify an individual's identification via image, which can then be used to automate the granting of access into a business or other setting.[8] Another form of automated screening is the use of algorithm-based decision making to grant or deny access to services, such as in the case of creditors using algorithmic scoring or predictive models to award or deny consumers' applications for credit.

AI Risks, Challenges, and Limitations

While the advent of AI integrated into multiple sectors can increase efficiency, it can also supplant human decision-making and create meaningful risk in light of the inherent limits of the technology. This section will provide an overview of some of the most salient challenges, including a lack of transparency that creates opportunities for bias and real-world policy consequences in core areas of civil rights concern, e.g. housing, employment, and the criminal justice system.

LACK OF TRANSPARENCY

The concept of trustworthy AI refers to artificial intelligence systems that meet a set of standards that experts have developed that both "create trust and confidence in AI systems among stakeholders and end users" and "mitigate the potential risks associated with the deployment of AI models" such as "harm to people, organizations and ecosystems."[9]

Trust is important in AI because "many AI and machine learning (ML) systems, such as deep learning models, operate as veritable black boxes; they ingest data and create outputs, with little to no transparency into how they arrive at those outputs." This is because, "as AI becomes more advanced, humans are challenged to comprehend and retrace how the algorithm came to a result."[10]

For AI to be successful and functional, it needs high-quality, unbiased data for training, transparent algorithms so people can trust its decisions, and enough computing power to handle large models.[11] Many AI systems function in an opaque manner, which is a key factor eroding public confidence in their reliability. Developing transparent AI systems remains a significant technical hurdle.

There are three distinct layers of transparency needed to be able to fully understand (and therefore assess) an AI model, in increasing order of difficulty. First, there is transparency as to the data being ingested by the system. Second, there is transparency as to the algorithm's design, e.g., how the data is weighted and what the model is encouraged to prioritize. And third, there is transparency as to what happens once machine learning takes over.

The first two aspects of transparency can be regulated, meaning companies creating AI can be required to disclose the data sets they used and the parameters they set. The third aspect is more challenging.

Due to opaque decision-making processes, AI systems are often described as "black boxes." In ML algorithms, "black box models are created directly from data by an algorithm, meaning that humans, even those who design them, cannot understand how variables are being combined to make predictions.[12] Even if one has a list of the input variables, black box predictive models can be such complicated functions of the variables that no human can understand how the variables are jointly related to each other to reach a final prediction."

Modern AI systems, particularly those using deep neural networks, involve layers of interconnected neurons processing data, sometimes with hundreds or thousands of layers, creating intricate patterns that even their developers struggle to interpret.[13]

"Many of the most advanced AI technologies, including generative AI tools, are what one might call 'organic black boxes' where obscuring is not intentional but rather due to complex [deep learning] systems that make it hard to 'understand exactly what happens inside them.'" [14] For example, "large language models, the type of A.I. systems that power ChatGPT and other popular chatbots, are not programmed line by line by human engineers, as conventional computer programs are.[15] Instead, these systems essentially learn on their own, by ingesting vast amounts of data and identifying patterns and relationships in language, then using that knowledge to predict the next words in a sequence."

Although AI models sometimes "become black boxes as a byproduct of their training," in some instances, they can be intentionally created by developers to protect intellectual property by obfuscating "the inner workings of AI tools before releasing them to the public" and "keep[ing] the source code and decision-making process a secret."[16]

This type of secrecy, intentional or otherwise, safeguards competitive advantage, but it also limits transparency and can lead to issues around accountability and trust of the algorithms. The complex AI systems that underpin tools like ChatGPT "essentially learn on their own, by ingesting vast amounts of data and identifying patterns and relationships in language, then using that knowledge to predict the next words in a sequence."[17] This makes it "difficult to reverse-engineer them or to fix problems by identifying specific bugs in the code." So, if an algorithm is yielding incorrect or biased results, it can be challenging to both explain why the result was rendered and go into the system to make corrections.

Societal and Ethical Challenges

The advent of AI and its continual integration into society necessitates a consideration of the societal and ethical challenges that have the potential to arise. These challenges can manifest in real-world policy consequences, largely due to bias.

In part because of the lack of transparency, the use of AI has been controversial. Key concerns include privacy risks tied to data harvesting and surveillance practices, safety, environmental consequences, the future of work, and the potential deepening of societal inequities through social bias (often unwittingly) perpetuated by AI.

01 COLLECTION OF PERSONAL DATA WITHOUT CONSENT

The issue of privacy is a key concern.[18] As AI is widely adopted, as AI systems often "rely heavily on personal data."[19] These challenges are amplified by the scope and nature of the data collected and the fact that individuals often don't know that their information has been garnered or that they may have the right to challenge specific instances of collection.

One example of data misuse is the Cambridge Analytica scandal, which involved the unauthorized harvesting of personal data from up to 87 million Facebook users via a third-party app.[20] The data was fed into an AI model and used to create psychological profiles for microtargeted political advertising, notably in the 2016 U.S. presidential election and Brexit campaigns. Cambridge Analytica deceived users by claiming the app was for academic research, while Facebook allowed data collection from users' friends without explicit consent.

The breach led to widespread backlash, regulatory fines (including a $5 billion FTC penalty for Facebook), and Cambridge Analytica's bankruptcy in 2018.[21] It exposed systemic privacy risks in data monetization and surveillance capitalism of personal data from up to 87 million Facebook users via a third-party app.[22]

02 DEEPFAKES

Generative AI can also "be misused to create fake profiles or manipulate images," also known as deepfakes.[23] Generative AI's ability to create realistic deepfakes, using scraped biometric data and machine learning, poses significant ethical, privacy, and security risks. Deepfakes are created by using ML to "capture and encode…unique biometric characteristics" of targets "to train an AI neural network to combine the subject's unique characteristics with the acquired knowledge of general human expression in order to then synthesize the target's facial features, voice, mannerisms, etc. to generate deepfake material at will."[24]

This data can be used to synthesize convincing audiovisual content, enabling fraud, non-consensual pornography, and misinformation campaigns.[25] Detection remains challenging due to evolving AI techniques and the need for extensive labeled datasets, while misuse threatens privacy, trust, and financial systems by enabling identity theft, extortion, and manipulated narratives. For example, in 2024, a Hong Kong employee transferred $25 million after being deceived by a scam that "saw the worker duped into attending a video call with what they thought were several other members of staff, but all of whom were in fact deepfake recreations,"[26] and believed they were acting on the instructions of their CFO.[27]

03 SURVEILLANCE AND NATIONAL SECURITY CONCERNS

AI systems in the U.S. raise three separate but inter-related concerns around surveillance and national security.

First, there is the civil liberties threat (and accompanying litigation exposure) posed by government collection of private data to build AI used for surveillance of private civilians. AI-driven surveillance tools in law enforcement or public spaces heighten concerns about unwarranted monitoring and erosion of individual autonomy.[28]

Second, there is the threat of private for-profit companies collecting data to surveil private citizens, whether under government contract or for their own motives. Facial recognition tools in particular have been the subject of litigation over bias and privacy violations.[29]

Third, AI can now be used to infiltrate critical systems, as exemplified by the Salt Typhoon breach—an "extensive cyberattack that compromised U.S. telecommunications firms, including AT&T, Verizon, and T-Mobile, and secure government communications networks."[30]

04 ENVIRONMENTAL IMPACT

Another societal challenge of AI is the environmental risk. AI's environmental impact is marked by exponential energy demands, resource depletion and localized inequities. Training LLMs like ChatGPT "consumes[31] thousands of megawatt hours of electricity and emits hundreds of tons of carbon," and "impose[s] a substantial burden on energy and natural resources, often fueled by nonrenewable sources."[32]

Data centers, which power AI, strain grids and exacerbate water scarcity through cooling processes that evaporate freshwater, which can be a threat to drought-prone regions like the American Southwest.[33] Globally, AI-driven data center power demand is projected to grow 160% by 2030, contributing $125-$140 billion in social costs from carbon emissions.[34] Additionally, AI infrastructure relies on the unsustainable mining of critical minerals and generates electronic waste, while also creating localized pollution.[35] "By perpetuating high-carbon-emitting behaviors, AI systems play a significant role in exacerbating the climate crisis."[36] For example, AI data centers that consume water for cooling have increased local water demands and diverted water resources needed to support growing populations in emerging countries such as Lagos, Nigeria.[37]

05 FUTURE OF WORK

AI poses significant risks to the future of work through accelerated job displacement and structural shifts in employment for both blue collar and white-collar workers. McKinsey estimates that "activities that account for up to 30% of hours currently worked across the US economy could be automated—a trend accelerated by generative AI."[38]

In white-collar sectors, AI-driven efficiency improvements compound rapidly and will displace roles in finance, law, and tech to increase corporate profitability.[39] Research from the Pew Research Center[40] finds that "one-fifth of all workers have high exposure jobs" and "women, Asian, college-educated and higher-paid workers are more exposed."[41]

For blue-collar workers, the data shows that women are 21% more exposed to AI automation than men, and lower-wage roles will be impacted by industrial robots performing routine and repetitive tasks, with outsized impact on sectors such as office support and customer service.[42]

AI and Bias

Research from the Brookings Institute defines bias generally as "a term that we define broadly as it relates to outcomes which are systematically less favorable to individuals within a particular group and where there is no relevant difference between groups that justifies such harm."[43] UC Berkeley researchers frame biased AI as "AI systems that result in inaccurate and/or discriminatory predictions or outputs for certain subsets of the population."[44] As AI becomes increasingly adopted across sectors, AI models can reflect bias and the associated consequences on people of color, women and vulnerable communities. These concerns are compounded by the perception that algorithms are inherently less biased than humans and that decisions made by algorithms are objective because they are data driven.[45]

SOURCES OF BIAS

Bias in AI arises from three primary sources: data bias, algorithmic design, and human decision-making processes. This report uses the term "bias" to refer not to intentional discrimination, but rather to results that are factually inaccurate and therefore flawed. As the ACLU notes, "bias is in the data used to train the AI — data that is often discriminatory or unrepresentative for people of color, women, or other marginalized groups — and can rear its head throughout the AI's design, development, implementation, and use."

These three dimensions of bias are interconnected and can work together to yield inaccurate results as the combined factors of biased inputs coupled with biased design and selection decisions, result in a biased algorithm rendering biased results. This feedback loop can then lead to inequitable decisions in areas like hiring, healthcare, and law enforcement—areas that have historically been plagued by bias.

Three Dimensions of Bias

Data Bias

> Data bias occurs when training datasets reflect historical or societal inequities,[46] such as underrepresentation of marginalized groups in facial recognition datasets.[47] If "the raw data already reflects social prejudices, and the algorithm also incorporates biased relationships, [it can lead] to the 'bias in and bias out' phenomenon," which can mean that historical bias can be further replicated in the future by the results of the AI models.[48] For example, historic bias is seen in a 2019 MIT study on AI-based facial recognition systems, which were found to have a far lower error rate when identifying light-skinned men and a far higher error rate when identifying dark-skinned women due primarily to disproportionate representation of white men in the training data.[49] Flawed training data can introduce systemic bias, triggering "model collapse," a scenario wherein AI systems perform worse over time by relying on self-generated or low-quality inputs, resulting in narrowed output diversity and diminishing accuracy.[50]

Algorithmic Design Bias

> Algorithmic design bias results when "systematic errors in machine learning algorithms produce unfair or discriminatory outcomes," often reflecting existing societal biases through skewed training data or design choices.[51] These biases emerge from factors like limited datasets, implicit developer prejudices, socio-technical influences, or from design decisions like outcome metrics, where the designer makes a choice on what outcome the algorithm is optimized to work toward, which allows for bias occur due to "subjective value judgments about how to define amorphous concepts like productivity or creditworthiness in measurable and quantifiable ways."[52]

Human Bias

> Human bias, also referred to as "designer bias, can also be introduced during model design, where engineers' unconscious assumptions or societal prejudices influence feature selection or weighting.[53] These biases stem from developers' lack of awareness of social contexts and the replication of historical inequalities in training data, underscoring the need for diverse teams and ethical oversight. The lack of representation "of people who understand and can work to address the potential harms of these technologies"[54] and pressure placed on those who speak up[55] in the tech industry only exacerbates this problem."[56]

REAL-WORLD AND POLICY CONSEQUENCES OF AI BIAS

Biased AI systems amplify existing inequalities by embedding historical discrimination into automated decision-making. As AI is increasingly adopted and incorporated into multiple facets of society, corporate entities, government, and service delivery, bias can manifest in critical domains like healthcare, criminal justice, education and employment—all areas heavily regulated by civil rights laws because they involve essential goods and services, and government action. Below are just a few illustrative examples of bias in some relevant settings.

Criminal Justice and Policing

Bias can manifest in criminal justice and policing due to discriminatory outcomes that are often the result of biased historical data and algorithmic design. For example, the use of AI algorithms in predictive policing, which are tools "to tailor law enforcement use of assets for efficiency and objectivity," have "increased racial biases" and resulted in "many disparate outcomes, including disproportionate surveillance and policing of Black communities."[57]

One investigation focused on the tool Geolitica (previously known as PredPol), which disproportionately targeted Black and Latino neighborhoods despite similar crime rates in white areas to "produce daily predictions on where and when crimes are most likely to occur,"[58] that "rarely lined up with reported crimes."[59] When researchers examining PredPol "applied the algorithm to Oakland, [CA,] they found that the algorithm targeted Black neighborhoods at twice the rate of white ones…because the algorithm relied on Oakland's racially biased arrest data to make its predictions."[60] This resulted in the tool identifying crime hotspots that "were the same neighborhoods that were already disproportion[ately] targeted by the Oakland police for drug arrests."[61] Increased patrol in overpoliced communities erodes trust in public institutions and upholds biased treatment of Black and Latino individuals in the criminal justice system.

Similarly, the COMPAS risk assessment tool exhibits significant bias, particularly against Black defendants, in predicting recidivism risk. ProPublica's analysis revealed that "Black defendants were far more likely than white defendants to be incorrectly judged to be at a higher risk of recidivism, while white defendants were more likely than Black defendants to be incorrectly flagged as low risk."[62] This racial disparity persists even when controlling for criminal history and other factors, leading to disproportionate pretrial detention and harsher sentences for Black individuals.[63] The tool's opacity exacerbates the issue, because its proprietary nature prevents thorough scrutiny of its methodology.[64] COMPAS also demonstrated gender bias, disadvantaging female offenders.[65]

A number of states are concerned that AI systems can produce biased criminal justice outcomes. Multiple cases have also been brought forth[66] regarding wrongful arrest and due process violations, largely due to misidentification by AI-driven facial recognition systems (see cases in Michigan[67], Georgia[68] and Texas[69]).

Education

In education, AI bias can be rendered through discriminatory outcomes and the further perpetuation of systemic inequality especially if "developers input historical data into the technology that replicate pre-existing biases that the model is trained to believe are accurate."[70] For example, an AI tool called the Dropout Early Warning System deployed in Wisconsin generated "raise[d] false alarms about Black and Hispanic students [being unlikely to graduate high school] at a significantly greater rate than it does White students",[71] despite "state records show[ing] the model is wrong nearly three quarters of the time it predicts a student won't graduate."[72]

Generative AI also presents a host of bias-related risks in the education setting, often related to historic inequities the models learn through training data. A study from Arizona State university and New Mexico State University "found that when presented with identical writing examples along with student demographics, ChatGPT gave differing scores to students whose racial indicators were implied rather than explicitly stated, compared to scores that were given when student descriptors were explicitly stated."[73] When the algorithm was fed the same passage with varying descriptors of the student, including race, socioeconomic status and school type, results showed statistically significant rates of subtle forms of prejudice."[74]

To mitigate these biases, educators and developers must implement inclusive design practices, increase transparency in data sourcing and algorithm development, and foster critical thinking skills among students to question AI recommendations.[75]

Potential litigation exposure regarding education and AI anchors largely around the perpetuation of discrimination. A student is suing "Yale University alleging that he has been falsely accused of using artificial intelligence on a final exam" by the university's AI-powered cheating detection software that unfairly targeted him as a non-native English speaker, which eventually led to disciplinary action against him.[76] This lawsuit showcases that "higher education institutions will have to continually evaluate the potential legal risks related to AI use, as well as methods to detect and deter AI use in academic settings."

Guidance for schools from the U.S. Department of Education's Office for Civil Rights (OCR) released in 2024 highlights multiple discriminatory uses of artificial intelligence, featuring detailed examples of how improper school use of AI and handling of AI-related issues could violate the civil rights of students."[77] The guidance, now removed from the OCR website at the start of the 2025 Trump administration, provided "hypothetical scenarios to illustrate how school AI use could violate this law, including an AI cheating detector that inaccurately flags English learners and AI-enabled facial recognition technology that consistently misidentifies students of color as known criminals."

Government Programs

AI bias can impact access to services and the distribution of government resources by yielding flawed results. In addition to discrepancies in resource delivery that may arise through the adoption of AI in the education, healthcare, and criminal justice policy areas, AI can also have an impact on additional public services, including public benefits programs and social services. In the case of public benefits programs, AI tools can be deployed to help determine eligibility and to detect fraudulent activity; however, if AI systems yield errors due to biased inputs or outputs, those who need access can suffer.

For example, the Electronic Privacy Information Center (EPIC) filed a complaint against Thomson Reuters Corporation for backing an AI-driven software that was sold to agencies used for automated fraud detection to identify public benefits fraud.[78] EPIC claimed that when the California Employment Development Department used the tool to draw upon "personal data like social media information, credit reports, and housing records to predict if public benefits applications are fraudulent."[79] The tool reviewed "10 million unemployment insurance claims paid out since the beginning of the COVID-19 pandemic," and 1.1 million claims were flagged as "suspicious" and those claimants' benefits were suspended.[80] Eventually, "further investigation showed that more than 600,000 (54%) of the claims flagged by the tool as fraudulent, meaning public benefits were revoked, were actually legitimate."[81]

In another case of AI bias impacting public services, the Allegheny Family Screening Tool (AFST), used for child and family services in Pittsburgh, Pennsylvania to assess a rise in child welfare investigations, is under scrutiny for possible discrimination against families with disabilities and for further compounding racial disparities. An AP investigation[82] "revealed potential bias and transparency issues about the opaque algorithm that is designed to assess a family's risk level when they are reported for child welfare concerts in Allegheny County" as the tool was using[83] "data points tied to disabilities in children, parents, and other members of local households because they can help predict the risk that a child will be removed from their home after a maltreatment report." Eventually the investigation led to the agency "updat[ing] its algorithm several times" and "sometimes remov[ing] disabilities-related data points."[84]

The ACLU of Idaho took on the issue of AI bias in government service provision when it represented adults with developmental disabilities in a case where the State of Idaho "drastically cut their assistance to dangerously low levels" and the state "Department of Health and Welfare claimed that the reasons for the cuts were "trade secrets" and refused to disclose the formula it used to calculate the reductions."[85] The ACLU managed to obtain an injunction to stop the cuts and require Idaho to make the Medicaid formulas publicly available, which eventually led to Idaho's federal district court striking down the formulas, ordering the state to bring the system up to order, and a class action lawsuit in play that, if approved, "will require the Department of Health and Welfare to develop a new system, with input and oversight by the ACLU of Idaho and participants and their families throughout the state."[86]

Labor

AI systems are now central to hiring, promotion, and workforce management. However, their increasing use has exposed significant biases, often perpetuating or amplifying existing inequities.

Research from the University of Washington that examined LLMs used for resume screening found significant levels of racial, gender, and intersectional bias. The research indicated that the models "significantly favor[ed] White-associated names in 85.1% of cases and female-associated names in only 11.1% of cases" and "that Black males are disadvantaged in up to 100% of cases, replicating real-world patterns of bias in employment settings."[87]

A study on AI-backed applicant tracking systems (ATS) also demonstrated that these systems are likely to screen out qualified candidates with marginalized identities, such as immigrants, women, people with disabilities and people with non-Anglo names.[88] From keyword filtering to analyzing video interviews, ATS have been deployed in hiring to create efficiency, but can end up replicating bias by filtering out qualified candidates for issues such as non-standard language (impacting immigrants or English language learners), employment gaps (impacting mothers) or other characteristics such as interpreting visual cues on video interviews that don't take into consideration that a candidate may have disabilities preventing them from engaging in typical behaviors.[89]

A significant example of bias in AI-driving workforce tools also surfaced in Amazon's admission that the company developed an experimental AI-based hiring tool to automate resume screening and rank applicants on a five-star scale.[90] The tool's algorithm was trained on biased data because it used resumes submitted to the company over a ten-year period--which largely consisted of male applicants due to their overrepresentation in the technology field.[91] The model learned to favor male candidates and systematically downgraded resumes containing the word "women's" or references to all-women colleges, discriminating against female applicants for technical roles.[92] When Amazon discovered the bias, it attempted to make the algorithm gender-neutral but ultimately abandoned the project in 2015, unable to guarantee the tool would not discriminate in other ways.[93]

Policy Consequences in Washington State

In Washington state, the rapid adoption of AI has raised concerns for residents about job security and government services, while also illustrating the need for transparency and accountability. A 2025 survey on the impact of generative AI on the Washington state workforce indicated that over 37% of surveyed employees believe that AI will likely fully automate some current jobs and over 60% believe AI will fundamentally change how people work.[94]

Labor representatives have expressed worries about roles in education, engineering, customer service, and law enforcement, where generative AI's role in automation may reshape job responsibilities. They have also raised concerns such as "impact on job security, especially regarding job automation, job displacement, skill erosion, reduction in opportunities for career growth, increased workloads, and work intensification that could lead to burnout."[95] Notably, 46.5% of surveyed state and local government employees cited labor rights concerns related to automation as their biggest apprehension, with frontline employees expressing the greatest anxiety.[96] One response seen to these concerns in 2025 is the introduction of a state bill[97] that "gives public workers bargaining power of how AI is used in state agencies," in an effort to protect workers and ensure AI is deployed with their input in mind.[98]

AI is already being deployed in multiple arenas of government service, including education and healthcare. The Washington Office of Superintendent of Public Instruction released guidance on an ethical framework for AI use in K-12 education that highlighted some educators' concerns, including "increasing and/or creating inequitable learning environments; unauthorized access to protected user information and unauthorized data collection; perpetuating institutional and systemic biases; plagiarism and academic dishonesty; and over-relying on technology and undermining the importance of human intelligence in education."[99] These concerns stand to impact over 1.07 million students enrolled in k-12 public education in the state.[100]

In Washington, approximately 1.85 million people are enrolled in Medicaid in Washington State, representing about 21% of the state's population.[101] National studies have indicated that AI is being used in administrative processes surrounding Medicaid to address administrative challenges by automating processes, streamlining eligibility and reducing the incidence of improper payments, which could save billions in healthcare costs.[102] While promising on some fronts, the adoption of AI in government services raises[103] concerns, namely around the risks associated with bias. Some researchers warn that the lack of robust AI regulation can lead to misuse, which could threaten patients' safety and privacy.[104]

Ultimately these concerns indicate the need for transparent policies that proactively address the risk of bias associated with AI, in order to ensure that public services are enhanced by AI, rather than pose risk to civil rights, safety, or equity.

Regulation: Existing Efforts and Best Practices

As AI continues to be adopted in multiple industries and by government, the need for thorough and comprehensive regulation is evident. The integration of AI into key sectors such as education, healthcare, law enforcement, government, and business necessitate a deep examination of how AI can be responsibly adopted while protecting society from the concerning issues such as bias and harm, privacy violations, and lack of accountability and transparency.

Federal efforts have largely stalled, with a change in administration[105] and influential Big Tech actors whose business interests lie in deregulation.[106] But there is potential in the patchwork of state level regulation and useful lessons to be gleaned from the EU and best practices recommendations.[107]

FEDERAL AI POLICY CHALLENGES

Federal AI regulation is fragmented and insufficient. There are no comprehensive federal laws to regulate AI. In place of a new AI law, federal agencies such as the Department of Justice[108] and the Federal Trade Commission[109] have been applying existing law and policy to AI systems—but it is unclear what policies will and will not remain under the new administration.

The Biden administration attempted to advance federal-level AI regulation through several actions, including the Blueprint for an AI Bill of Rights,[110] the 2023 AI Executive Order on the "Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence"[111] and related guidance[112] from the Office of Management and Budget (OMB) that charged the government to address the regulatory and ethical challenges surrounding AI. These efforts were built upon voluntary commitments to AI safety that the administration secured from 15 leading U.S. AI companies. President Biden's approach to AI policy attempted to balance the promotion of competition and innovation by U.S. companies with the need to protect Americans from the risks associated with AI.[113]

President Trump has revoked the Biden administration's large-scale efforts on AI policy in favor of a deregulatory approach.[114] In January 2025, President Trump issued an Executive Order that rescinded Biden's regulatory efforts in favor of a new U.S. AI policy that removed perceived barriers to competition and is framed to be oriented toward economic growth and American AI leadership in the competitive AI field.[115]

Key stipulations in President Biden's Executive Order, including the requirement to implement safeguards for equity and civil rights were removed,[116] and the Trump administration's accompanying White House-issued fact sheet flags that "American development of AI systems must be free from ideological bias or engineered social agendas."[117]

In place of comprehensive federal legislation, multiple narrower bills such as the AI Advancement and Reliability Act[118] and the Creating Resources for Every American to Experiment with Artificial Intelligence (CREATE AI) Act,[119] focus on specific AI policy issues such as safety and fostering fair competition, but lack broad bipartisan support.

PROMISING STATE LAW

As the country awaits the new development of a comprehensive legislative view on AI by the Trump administration, states will continue to be the focus area for regulation. Multiple states released their own state-level AI legislation, including California, Colorado, and Virginia which have emerged as pioneering models for regulating AI through the lifecycle from design to deployment, that will likely have influence on further state legislation developed across the country. Utah and Illinois have also developed AI legislation that is noteworthy.[120]

Across the multiple state-led regulation efforts, some common areas of focus include data privacy, transparency and accountability, algorithmic bias, and safety and effectiveness. Some states have also enacted sector-specific regulations, in areas such as education, healthcare, legal services and employment, to mitigate specific risk of AI harms in critical policy arenas.[121]

An overview of the recent state-level laws developed on AI regulation:

- **CA: California AI Transparency Act (CATA)[122] and California AB 2013[123]**
  Both bills target AI developers. The CATA focusing primarily on transparency and detection of AI-generated content and targets generative AI tools with more than one million monthly users in the state. AB 2013 focuses on transparency of training data in AI systems.
- **CO: Colorado Artificial Intelligence Act (CAIA)[124]**
  Focuses primarily on regulation of high-risk AI systems and algorithmic discrimination, with sector-specific areas of stipulations in healthcare, employment, and education and a call for the state to engage in regular assessments and provide consumers with notifications when they are engaging with AI content.

- **VA: High-Risk Artificial Intelligence Developer and Deployer Act (HB 2094)[125](Vetoed)**
  This bill builds "important safety standards to ensure the responsible, ethical, and transparent use of AI by state government" and, if signed into law, would make Virginia "the second state after Colorado to enact comprehensive AI legislation."[126] It focuses on algorithmic transparency, and requires enforcements of civil penalties of up to $10,000, and offers sector specific rules for high-risk AI systems as those intended to autonomously make or substantially influence consequential decisions.[127]
- **UT: Utah Artificial Intelligence Policy Act (UAIP)[128]**
  Focuses on transparency in generative AI and consumer protection, by setting forth disclosure requirements for entities using AI with customers and has also taken a great step to create a regulatory sandbox model where businesses receive regulatory reprieve for a timeframe in order to continue to innovate without hurdles.

While these state laws are largely lauded for taking on requirements to address critical issues like privacy and bias, some see a gap in that "as more states implement distinct AI regulations, companies will face the challenge of complying with a growing patchwork of laws. The variations in requirements between states, such as disclosure obligations and risk assessments, could complicate operations for developers and deployers working across multiple regions."[129]

Best Practices for Strong AI Policy

Best practice for AI state policy is based on an integration of existing and emerging frameworks on AI regulation, that is principled, addresses risk, and balances ethical safeguarding with innovation potential. It should include the following considerations and guidelines:

Policy Best Practice #1

Transparency

AI systems should provide clear insights into their decision-making processes, data collection methods, and algorithmic mechanisms. This transparency generates stakeholder trust, facilitates accountability, and facilitates back-tracing to address bias.[130]

California, Colorado and Utah are all "prioritiz[ing] disclosure and transparency" in their bills, "whether by focusing on customer interactions (as in Utah), comprehensive documentation requirements (as in Colorado), or mandates for content labeling and dataset disclosures (as in California).[131]

Policy Best Practice #2

Risk-Based Approach

A risk-based approach to AI regulation is vital to striking a balance between the need for strong oversight to keep harms at bay while fostering innovation.[132] This can allow policy areas, such as healthcare and criminal justice, that are high-risk for bias and consequences, to have stricter oversight while lower risk uses face fewer restrictions.[133]

For example, Colorado's CAIA[134] adopts this strategy by implementing "more stringent requirements on 'high-risk' AI systems" that make consequential decisions in critical areas.[135] The EU AI Act also uses a similar framing, with "a proportionate risk-based approach to AI regulation, which imposes a gradual scheme of requirements and obligations depending on the level of risk posed to health, safety and fundamental rights."[136]

The U.S. National Institute of Standards and Technology (NIST) has also developed the AI Risk Management Framework,[137] an approach for recognizing, evaluating, and controlling associated risks of AI systems in their lifecycle, allowing for proportionate governance that does not stymie the ability to further develop AI technology.[138]

Policy Best Practice #3

Oversight and Accountability

Oversight and accountability can have three components: an external regulatory body charged with providing oversight; requirements for organizations that create and use AI systems to hold them accountable; and rules that apportion liability.

Organizations that develop and deploy AI should be required to demonstrate accountability practices, and exemplify internal accountability by strategies that "designate specific roles for overseeing AI compliance and maintaining risk assessments" and create "accountability structures within the organization, with a designated team or individual responsible for AI ethics and regulatory adherence."[139]

Experts note that demonstratable organizational accountability should be "a central element of AI regulations" and include "adoption of accountable AI governance practices."[140] Some scholars suggest that effective AI governance should be conducted by a single regulatory body that has "establish[ed] a baseline for the scope of authority needed to credibly and effectively regulate high-risk technological systems."[141] However, given the reality that AI is cross-sectoral in nature, it is also essential to have mechanisms in place to promote cross-coordination between regulatory bodies "to set high-level AI policies and goals applicable across all sectors and industries, and facilitate alignment, regulatory coordination, and joint action between different regulatory bodies, where necessary and appropriate."[142]

Another element to oversight is the apportionment of liability, "with a focus on the party most closely associated with generating harm."[143]

Policy Best Practice #4

Addressing Algorithmic Bias & Discrimination

A crucial focus of state AI regulation should be to address risk and harm associated with AI systems to prevent bias and discrimination. This can be achieved through stipulations such as mandating regular audits and impact assessments or prohibiting discriminatory AI use.[144]

Prohibition of algorithmic discrimination would "prohibit deployers from using an automated decision tool and prohibit developers from making available an automated decision tool if an impact assessment identifies a reasonable risk of algorithmic discrimination" and is an approach that California regulation has adopted.[145]

The "Duty of Care" approach is where "both developers and deployers are subject to a duty to use 'reasonable care' to protect consumers from "any known or reasonably foreseeable risks of algorithmic discrimination from the intended and contracted uses of the high-risk AI system."[146]

In addition to prohibitions, transparency requirements and risk-based approaches, the use of impact assessments and audits can also help address bias and discrimination. Several states[147] require organizations to engage in regular audits and impact assessments of their AI systems.[148] For example, Colorado state law requires annual impact assessments.[149]

Policy Best Practice #5

Leveraging Existing Frameworks

States should consider adapting or reinterpreting applications of existing law on consumer protections, civil rights, privacy, and data protection to advance AI regulation.[150] Massachusetts has taken this approach and recognizes that AI is subject to already existing legal standards.[151] "Relying on existing hard law frameworks to the extent possible reduces the risk of creating overlapping or conflicting rules that could lead to legal uncertainty and inconsistent protections," and experts believe that existing law can be "relevant to address many of the most important risks associated with AI."[152]

Policy Best Practice #6

Formal Verification and Safety Guarantees

AI safety can be promoted through the requirement of formal verification that "provides mathematical guarantees" and is more reliable than the current approach of extensive testing protocols because "dangerous capabilities can arise unpredictably and it is inherently difficult to exhaustively test models for all possible capabilities."[153]

To avoid danger, experts call for AI systems to be "designed in such a way to be able to demonstrate compliance with common-sense regulations" with "'termination obligations that would function like a circuit breaker in electrical power systems: if an AI system is not under human control, then it must be terminated."[154]

Policy Best Practice #7

Flexibility and Adaptability

AI regulation must be designed with flexibility in mind given the rapid pace of AI development from an industry that has historically wanted to "move fast and break things."[155] While AI "regulation is urgently needed and unpredictable," it can "be counterproductive if not done well.[156] However, governments cannot wait until they have perfect and complete information before they act, because doing so may be too late to ensure that the trajectory of technological development does not lead to existential or unacceptable risks."

In the face of uncertainty and fast-paced growth, regulatory frameworks should incorporate mechanisms for ongoing assessment and adaptation, such as iterative policy development processes.[157] Another option is to use regulatory sandboxes that create "a space where participating [actors] won't be subject to onerous regulations—usually for a limited amount of time," allowing businesses to develop their ideas before "lawmakers will evaluate what's working and what isn't—namely what regulations the business needs to follow once it transitions out of the sandbox."[158]

Conclusion

As the AI industry continues to rapidly grow and AI systems are increasingly embedded into critical decision-making processes, states have a vital role to ensure transparency and protect constituents against privacy risks, and faulty outcomes that are often rooted in bias.

Because our society is imperfect, AI is imperfect: it is unclear whether AI can ever be appropriately designed to produce results that are not flawed when it comes to women, people of color, or people with disabilities. And any use of AI for critically important systems will result in potential litigation exposure. Until states have a better understanding of the data being ingested by algorithms, algorithmic design, and machine learning, AI should likely not be used to determine access to public services or benefits—and only with extreme caution to determine access to private goods and services. States are in a position to adopt these best practices to help mitigate the worst consequences of AI without compromising innovation.

Endnotes

[1]What is AI (artificial intelligence), McKinsey & Company, (Apr. 3, 2024), https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-ai.

[2]AI and Privacy: The privacy concerns surrounding AI, its potential impact on personal data, The Economic Times, (Apr. 25, 2023), https://economictimes.indiatimes.com/news/how-to/ai-and-privacy-the-privacy-concerns-surrounding-ai-its-potential-impact-on-personal-data/articleshow/99738234.cms.

[3]What is artificial intelligence (AI)?, ISO, https://www.iso.org/artificial-intelligence/what-is-ai.

[4]Lev Craig, Nicole Laskowski, Linda Tucci, What is AI? Artificial Intelligence explained, TechTarget Network, (Oct 2024), https://www.techtarget.com/searchenterpriseai/definition/AI-Artificial-Intelligence.

[5]Lev Craig, Nicole Laskowski, Linda Tucci, What is AI? Artificial Intelligence explained, TechTarget Network, (Oct 2024), https://www.techtarget.com/searchenterpriseai/definition/AI-Artificial-Intelligence

[6]Junaid Bajwa, Usman Munir, Aditya Nori, Bryan Williams, Artificial intelligence in healthcare: transforming the practice of medicine, National Library of Medicine, (July 2021), https://pmc.ncbi.nlm.nih.gov/articles/PMC8285156/.

[7]Shahzadi Parveen, Ramneet singh Chadha, Pradeep Kumar, Jasmehar Singh, Artificial Intelligence in Transportation Industry, ResearchGate, (Aug. 2022), https://www.researchgate.net/publication/380000645_Artificial_Intelligence_in_Transportation_Industry.

[8]James Andrew Lewis, William Crumpler, How Does Facial Recognition Work?, CSIS, (June10, 2021), https://www.csis.org/analysis/how-does-facial-recognition-work.

[9]Alice Gomstyn, Amanda McGrath, Alexandra Jonker, What is trustworthy AI?, IBM, (Oct. 5, 2024), https://www.ibm.com/think/topics/trustworthy-ai.

[10]What is explainable AI?, IBM, https://www.ibm.com/think/topics/explainable-ai.

[11]Benjamin van Giffen, Dennis Herhausen, Tobias Fahse, Overcoming the pitfalls and perils of algorithms: A classification of machine learning biases and mitigation methods, Science Direct, Journal of Business Research, (May 2022), https://www.sciencedirect.com/science/article/pii/S0148296322000881.

[12]Cynthia Rudin, Joanna Radin, Why Are We Using Black Box Models in AI When We Don't Need To? A Lesson From an Explainable AI Competition, HDSR, (Nov. 22, 2019), https://hdsr.mitpress.mit.edu/pub/f9kuryi8/release/8.

[13]Matthew Kosinski, What is black box artificial intelligence (AI)?, IBM, (Oct. 29, 2024), https://www.ibm.com/think/topics/black-box-ai.

[14]Matthew Kosinski, What is black box artificial intelligence (AI)?, IBM, (Oct. 29, 2024), https://www.ibm.com/think/topics/black-box-ai.

[15]A.I.'s Black Boxes Just Got a Little Less Mysterious, https://www.nytimes.com/2024/05/21/technology/ai-language-models-anthropic.html.

[16]Matthew Kosinski, What is black box artificial intelligence (AI)?, IBM, (Oct. 29, 2024), https://www.ibm.com/think/topics/black-box-ai.

[17]A.I.'s Black Boxes Just Got a Little Less Mysterious, https://www.nytimes.com/2024/05/21/technology/ai-language-models-anthropic.html.

[18]Alice Gomstyn, Alexandra Jonker, Exploring privacy issues in the age of AI, IBM, (Sept. 30, 2024), https://www.ibm.com/think/insights/ai-privacy.

[19]The Growing Data Privacy Concerns with AI: What You Need To Know, Dataguard (Jan 10, 2025), https://www.dataguard.com/blog/growing-data-privacy-concerns-ai/

[20]How Trump Consultants Exploited the Facebook Data of Millions, The New York Times, https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html.

[21]Eli Watkins, Joe Sutton, Cambridge Analytica files for bankruptcy, CNN, (May 18, 2018), https://www.cnn.com/2018/05/18/politics/cambridge-analytica-bankruptcy/index.html.

[22]The Great Hack': Cambridge Analytica is just the tip of the iceberg, Amnesty International, (July 24, 2019), https://www.amnesty.org/en/latest/news/2019/07/the-great-hack-facebook-cambridge-analytica/

[23]AI and Privacy: The privacy concerns surrounding AI, its potential impact on personal data, The Economic Times, (Apr. 25, 2023), https://economictimes.indiatimes.com/news/how-to/ai-and-privacy-the-privacy-concerns-surrounding-ai-its-potential-impact-on-personal-data/articleshow/99738234.cms.

[24]Rolling in the deepfakes: Generative AI, privacy and regulation, LexisNexis, (Nov. 6, 2024), https://www.lexisnexis.com/blogs/au/b/insights/posts/rolling-deepfakes-generative-artificial-intelligence-privacy-regulation.

[25]Increasing Threat of Deepfake Identities, DHS, https://www.lexisnexis.com/blogs/au/b/insights/posts/rolling-deepfakes-generative-artificial-intelligence-privacy-regulation.

[26]Satish Lalchand, Val Srinivas, Brendan Maggiore, Joshua Henderson, Generative AI is expected to magnify the risk of deepfakes and other fraud in banking, Deloitte, (May 29, 2024), https://www2.deloitte.com/us/en/insights/industry/financial-services/financial-services-industry-predictions/2024/deepfake-banking-fraud-risk-on-the-rise.html.

[27]Heather Chen, Kathleen Magramo, Finance worker pays out $25 million after video call with deepfake 'chief financial officer', CNN, (Feb. 4, 2024), https://www.cnn.com/2024/02/04/asia/deepfake-cfo-scam-hong-kong-intl-hnk/index.html.

[28]The growing data privacy concerns with AI: What you need to know, DataGuard, https://www.dataguard.com/blog/growing-data-privacy-concerns-ai/.

[29]Woodruff v. Oliver, ACLU, https://www.aclu.org/court-cases?issue=face-recognition-technology.

[30]David Kertai, Salt Typhoon Exposes US Cyber Vulnerabilities, ITIF, (Feb. 13, 2025), https://itif.org/publications/2025/02/13/salt-typhoon-exposes-us-cyber-vulnerabilities/.

[31]Shaolei Ren, Adam Wierman, The Uneven Distribution of AI's Environmental Impacts, Harvard Business Review, (July 15, 2024), https://hbr.org/2024/07/the-uneven-distribution-of-ais-environmental-impacts.

[32]Joseph B. Keller, Manann Donoghoe, Andre M. Perry, The US must balance climate justice challenges in the era of artificial intelligence, Brookings, (Jan. 29, 2024), https://www.brookings.edu/articles/the-us-must-balance-climate-justice-challenges-in-the-era-of-artificial-intelligence/.

[33]Joseph B. Keller, Manann Donoghoe, Andre M. Perry, The US must balance climate justice challenges in the era of artificial intelligence, Brookings, (Jan. 29, 2024), https://www.brookings.edu/articles/the-us-must-balance-climate-justice-challenges-in-the-era-of-artificial-intelligence/.

[34]AI is poised to drive 160% increase in data center power demand, Goldman Sachs, (May 14, 2024), https://www.goldmansachs.com/insights/articles/AI-poised-to-drive-160-increase-in-power-demand.

[35]AI has an environmental problem. Here's what the world can do about that, UN Environment Programme, (Sept. 21, 2024), https://www.unep.org/news-and-stories/story/ai-has-environmental-problem-heres-what-world-can-do-about.

[36]Joseph B. Keller, Manann Donoghoe, Andre M. Perry, The US must balance climate justice challenges in the era of artificial intelligence, Brookings, (Jan. 29, 2024), https://www.brookings.edu/articles/the-us-must-balance-climate-justice-challenges-in-the-era-of-artificial-intelligence/.

[37]Data centres 'straining water resources' as AI swells, SciDevNet, (Nov. 15, 2023), https://www.scidev.net/global/scidev-net-investigates/data-centres-straining-water-resources-as-ai-swells/.

[38]Generative AI and the future of work in America, McKinsey Global Institute, (July 26, 2023), https://www.mckinsey.com/mgi/our-research/generative-ai-and-the-future-of-work-in-america.

[39]Greg Iacurci, A.I. is on a collision course with white-collar, high-paid jobs — and with unknown impact, CNBC, (July 31, 2023), https://www.cnbc.com/2023/07/31/ai-could-affect-many-white-collar-high-paid-jobs.html.

[40]Rakesh Kochhar, Which U.S. Workers Are More Exposed to AI on Their Jobs?, Pew Research Center, (July 26, 2023), https://www.pewresearch.org/social-trends/2023/07/26/which-u-s-workers-are-more-exposed-to-ai-on-their-jobs/.

[41]Rakesh Kochhar, Which U.S. Workers Are More Exposed to AI on Their Jobs?, Pew Research Center, (July 26, 2023), https://www.pewresearch.org/social-trends/2023/07/26/which-u-s-workers-are-more-exposed-to-ai-on-their-jobs/.

[42]Rakesh Kochhar, Which U.S. Workers Are More Exposed to AI on Their Jobs?, Pew Research Center, (July 26, 2023), https://www.pewresearch.org/social-trends/2023/07/26/which-u-s-workers-are-more-exposed-to-ai-on-their-jobs/.

[43]Nicol Turner Lee, Paul Resnick, Genie Barton, Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms, Brookings, (May 22, 2019), https://www.brookings.edu/articles/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/.

[44]Genevieve Smith, Ishita Rustagi, Mitigating Bias in Artificial Intelligence: An Equity Fluent Leadership Playbook, Berkeley Haas Center for Equity, Gender and Leadership, (July 2020), https://haas.berkeley.edu/wp-content/uploads/UCB_Playbook_R10_V2_spreads2.pdf.

[45]Hyesun Choung, John S. Seberger, Prabu David, When AI is Perceived to Be Fairer than a Human: Understanding Perceptions of Algorithmic Decisions in a Job Application Context, Taylor & Francis Online, (Mar. 7, 2023), https://www.tandfonline.com/doi/full/10.1080/10447318.2023.2266244.

[46]Shedding light on AI bias with real world examples, IBM, (Oct. 16, 2023), https://www.ibm.com/think/topics/shedding-light-on-ai-bias-with-real-world-examples.

[47]James Manyika, Jake Silberg and Brittany Presten, What Do We Do About the Biases in AI?, Harvard Business Review, (Oct. 25, 2019), https://hbr.org/2019/10/what-do-we-do-about-the-biases-in-ai.

[48]Zhisheng Chen, Ethics and discrimination in artificial intelligence-enabled recruitment practices, Nature Humanities and Social Sciences Communications, (Sept. 13, 2023), https://www.nature.com/articles/s41599-023-02079-x.

[49]Project Gender Shades, Algorithmic Bias Persists, MIT Media Lab, https://www.media.mit.edu/projects/gender-shades/overview/.

[50]Taylor Karl, Understanding AI: Key Concepts and Technologies Explained, New Horizons, (Sept. 30, 2024), https://www.newhorizons.com/resources/blog/understanding-ai-technology.

[51]Alexandra Jonker, Julie Rogers, What is algorithmic bias?, IBM, (Sept. 20, 2024), https://www.ibm.com/think/topics/algorithmic-bias.

[52]Gissela Moya, Vinhcent Le, Algorithmic Bias Explained, How Automated Decision Making Becomes Automated Discrimination, The Greenlining Institute, https://greenlining.org/wp-content/uploads/2021/04/Greenlining-Institute-Algorithmic-Bias-Explained-Report-Feb-2021.pdf.

[53]Bias in AI Design, Queen's University Library, https://guides.library.queensu.ca/ai/ethics.

[54]Kaitlyn Schwanemann, Experts call for more diversity to combat bias in artificial intelligence, CNN, (Dec. 15, 2023), https://www.cnn.com/2023/12/15/us/diversity-artificial-intelligence-bias-reaj/index.html.

[55]Tom Simonite, What Really Happened When Google Ousted Timnit Gebru, WIRED, (June 8, 2021), https://www.wired.com/story/google-timnit-gebru-ai-what-really-happened/.

[56]Olga Akselrod, How Artificial Intelligence Can Deepen Racial and Economic Inequities, ACLU, (July 13, 2021), https://www.aclu.org/news/privacy-technology/how-artificial-intelligence-can-deepen-racial-and-economic-inequities.

[57]Artificial Intelligence in Predictive Policing Issue Brief, NAACP, https://naacp.org/resources/artificial-intelligence-predictive-policing-issue-brief.

[58]Aaron Sankin, Dhruv Mehrotra, Surya Mattu, Annie Gilbertson, Crime Prediction Software Promised to Be Free of Biases. New Data Shows It Perpetuates Them, The Markup, (Dec. 2, 2021), https://themarkup.org/prediction-bias/2021/12/02/crime-prediction-software-promised-to-be-free-of-biases-new-data-shows-it-perpetuates-them.

[59]Aaron Sankin, Surya Mattu, Predictive Policing Software Terrible At Predicting Crimes, The Markup, (Oct. 2, 2023), https://themarkup.org/prediction-bias/2023/10/02/predictive-policing-software-terrible-at-predicting-crimes.

[60]Aaron Sankin, Surya Mattu, Predictive Policing Software Terrible At Predicting Crimes, The Markup, (Oct. 2, 2023), https://themarkup.org/prediction-bias/2023/10/02/predictive-policing-software-terrible-at-predicting-crimes.

[61]Aaron Sankin, Surya Mattu, Predictive Policing Software Terrible At Predicting Crimes, The Markup, (Oct. 2, 2023), https://themarkup.org/prediction-bias/2023/10/02/predictive-policing-software-terrible-at-predicting-crimes.

[62]Jeff Larson, Surya Mattu, Lauren Kirchner, Julia Angwin, How We Analyzed the COMPAS Recidivism Algorithm, ProPublica, (May 23, 2016), https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm.

[63]Jeff Larson, Surya Mattu, Lauren Kirchner, Julia Angwin, How We Analyzed the COMPAS Recidivism Algorithm, ProPublica, (May 23, 2016), https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm.

[64]Cynthia Rudin, Caroline Wang, Beau Coker, The Age of Secrecy and Unfairness in Recidivism Prediction, HDSR, (Mar. 31, 2020), https://hdsr.mitpress.mit.edu/pub/7z10o269/release/7.

[65]Justice served? Discrimination in algorithmic risk assessment, Research Outreach, (Sept. 19, 2019), https://researchoutreach.org/articles/justice-served-discrimination-in-algorithmic-risk-assessment/.

[66]Bruce Barcott, AI Lawsuits Worth Watching: A Curated Guide, Tech Policy Press, (July 1, 2024), https://www.techpolicy.press/ai-lawsuits-worth-watching-a-curated-guide/.

[67]Woodruff v. Detroit, City of (5:23-cv-11886), Court Listener, https://www.courtlistener.com/docket/67661093/woodruff-v-detroit-city-of/. Woodruff v. Detroit, City of (5:23-cv-11886), Court Listener, https://www.courtlistener.com/docket/67661093/woodruff-v-detroit-city-of/.

[68]Reid v. Bartholomew (1:23-cv-04035), Court Listener, https://www.courtlistener.com/docket/67800075/reid-v-bartholomew/.

[69]Murphy v. Essilorluxottica, ETI, https://blogs.gwu.edu/law-eti/ai-litigation-database/case-detail-page/?pid=127.

[70]Hoang Pham, Tanvi Kohli, Emily Olick Llano, Imani Nouri, Anya Weinstock, How will AI Impact Racial Disparities in Education?, SLS, (June 29, 2024), https://law.stanford.edu/2024/06/29/how-will-ai-impact-racial-disparities-in-education/.

[71]Hoang Pham, Tanvi Kohli, Emily Olick Llano, Imani Nouri, Anya Weinstock, How will AI Impact Racial Disparities in Education?, SLS, (June 29, 2024), https://law.stanford.edu/2024/06/29/how-will-ai-impact-racial-disparities-in-education/.

[72]Todd Feathers, False Alarm: How Wisconsin Uses Race and Income to Label Students "High Risk", The Markup, (Apr. 27, 2023), https://themarkup.org/machine-learning/2023/04/27/false-alarm-how-wisconsin-uses-race-and-income-to-label-students-high-risk.

[73]Melissa Warr, Nicole Jakubczyk Oster, Roger Isaac, Implicit bias in large language models: Experimental proof and implications for education, Taylor & Francis Online, (Aug. 28, 2024), https://www.tandfonline.com/doi/full/10.1080/15391523.2024.2395295.

[74]Publications, Mellisa Warr, https://melissa-warr.com/publications/.

[75]Ken Shelton, Dee Lanier, Thinking About Equity and Bias in AI, Edutopia, (Aug. 30, 2024), https://www.edutopia.org/article/equity-bias-ai-what-educators-should-know/.

[76]Ivy League Lawsuit Centers on Alleged Impermissible Use of AI in Academia, Crowell, (Mar. 6, 2025), https://www.crowell.com/en/insights/client-alerts/ivy-league-lawsuit-centers-on-alleged-impermissible-use-of-ai-in-academia.

[77]ED Issues Guidance to Avoid Discriminatory Use of AI, Government Technology, (Nov. 26, 2024), https://www.govtech.com/education/k-12/ed-issues-guidance-to-avoid-discriminatory-use-of-ai.

[78]In Re Thomas Reuters Corporation, FTC, https://epic.org/wp-content/uploads/2024/01/EPIC-FTC-Thomson-Reuters-Complaint.pdf.

[79]Keely Quinlan, Nonprofit behind FTC complaint about automated fraud-detection software hopes for more responsible AI use, StateScoop, (Jan. 10, 2024), https://statescoop.com/nonprofit-ftc-complaint-automated-benefits-responsible-ai/.

[80]Keely Quinlan, Nonprofit behind FTC complaint about automated fraud-detection software hopes for more responsible AI use, StateScoop, (Jan. 10, 2024), https://statescoop.com/nonprofit-ftc-complaint-automated-benefits-responsible-ai/.

[81]In Re Thomas Reuters Corporation, FTC, https://epic.org/wp-content/uploads/2024/01/EPIC-FTC-Thomson-Reuters-Complaint.pdf.

[82]Sally Ho, Garance Burke, An algorithm that screens for child neglect raises concerns, AP, (Apr. 29, 2022), https://apnews.com/article/child-welfare-algorithm-investigation-9497ee937e0053ad4144a86c68241ef1.

[83]Child welfare algorithm faces Justice Department scrutiny, CBS News, (Jan. 31, 2023), https://www.cbsnews.com/pittsburgh/news/child-welfare-algorithm-faces-justice-department-scrutiny/.

[84]Child welfare algorithm faces Justice Department scrutiny, CBS News, (Jan. 31, 2023), https://www.cbsnews.com/pittsburgh/news/child-welfare-algorithm-faces-justice-department-scrutiny/.

[85]K.W., et al. v. Armstrong, Justicia U.S. Law, https://law.justia.com/cases/federal/appellate-courts/ca9/14-35296/14-35296-2015-06-05.html

[86]K.W. VS. ARMSTRONG, ACLU, https://www.acluidaho.org/en/cases/kw-v-armstrong.

[87]Kyra Wilson et al., Gender, Race, and Intersectional Bias in Resume Screening via Language Model Retrieval, Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society 7, (Nov. 1, 2024) https://ojs.aaai.org/index.php/AIES/article/view/31748/33915

[88]Eesha Bayana, Bias in AI Hiring Tools: Impacted Groups, Legal Risks, Historical Foundations, and Next Steps, Research Archive of Rising Scholars, (Jan. 15, 2025), https://research-archive.org/index.php/rars/preprint/view/2177/3055

[89]Eesha Bayana, Bias in AI Hiring Tools: Impacted Groups, Legal Risks, Historical Foundations, and Next Steps, Research Archive of Rising Scholars, (Jan. 15, 2025), https://research-archive.org/index.php/rars/preprint/view/2177/3055

[90]Jeffrey Dastin, Insight - Amazon scraps secret AI recruiting tool that showed bias against women, Reuters, (Oct. 10, 2018), https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G/

[91]Erin Winick, Amazon ditched AI recruitment software because it was biased against women, MIT Technology Review, (Oct. 10, 2028), https://www.technologyreview.com/2018/10/10/139858/amazon-ditched-ai-recruitment-software-because-it-was-biased-against-women/

[92]Rachel Goodman, Why Amazon's Automated Hiring Tool Discriminated Against Women, ACLU, (Oct. 12, 2018), https://www.aclu.org/news/womens-rights/why-amazons-automated-hiring-tool-discriminated-against

[93]Jeffrey Dastin, Insight - Amazon scraps secret AI recruiting tool that showed bias against women, Reuters, (Oct. 10, 2018), https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G/

[94]Impact of generative artificial intelligence on the Washington state workforce, Washington State Office of Financial Management, (Dec. 2024), https://ofm.wa.gov/sites/default/files/public/publications/Impact_generative_AI_on_state_workforce.pdf

[95]Impact of generative artificial intelligence on the Washington state workforce, Washington State Office of Financial Management, (Dec. 2024), https://ofm.wa.gov/sites/default/files/public/publications/Impact_generative_AI_on_state_workforce.pdf

[96]Impact of generative artificial intelligence on the Washington state workforce, Washington State Office of Financial Management, (Dec. 2024), https://ofm.wa.gov/sites/default/files/public/publications/Impact_generative_AI_on_state_workforce.pdf

[97]HB 1622 - 2025-26, Washington State Legislature, (May. 13, 2025), https://app.leg.wa.gov/BillSummary/?BillNumber=1622&Year=2025&Initiative=false

[98]Sophia Fox-Sowell, New Washington bill would let state workers influence how agencies use AI, StateScoop, (Feb. 13, 2025), https://statescoop.com/washington-bill-labor-union-ai-2025/

[99]Human-Centered AI: Guidance for K-12 Public Schools, Washington Office of Superintendent of Public Instruction, (Jul. 1, 2024), https://ospi.k12.wa.us/sites/default/files/2024-07/comprehensive-ai-guidance-accessible-format_0.pdf

[100]Kindergarten through grade 12 (K-12) enrollment, Washington State Office of Financial Management, (Dec. 12, 2024), https://ofm.wa.gov/washington-data-research/statewide-data/washington-trends/budget-drivers/kindergarten-through-grade-12-k-12-enrollment

[101]Medicaid eligibility and enrollment in Washington, Health Insurance, (May. 13, 2024), https://www.healthinsurance.org/medicaid/washington/

[102]Ted Cho et al., Using artificial intelligence to improve administrative process in Medicaid, Health Affairs Scholar, (Jan. 29, 2024), https://pmc.ncbi.nlm.nih.gov/articles/PMC10986276/

[103]Washington State AI Task Force raises AI transparency, accountability as priorities for 2025 session, Transparency Coalition, (Jan. 2, 2025), https://www.transparencycoalition.ai/news/washington-ai-task-force-raises-ai-transparency-accountability-as-top-priorities-for-2025-session

[104]Natalie Shen, AI Regulation in Health Care: How Washington State Can Conquer the New Territory of AI Regulation, Seattle Journal of Technology, Environment & Innovation Law, (2023), https://digitalcommons.law.seattleu.edu/cgi/viewcontent.cgi?article=1043&context=sjteil

[105]Shaleen Khanal, Hongzhou Zhang, Araz Taeihagh, Why and how is the power of Big Tech increasing in the policy process? The case of generative AI, Oxford Academic Policy and Society, (Mar. 27, 2024), https://academic.oup.com/policyandsociety/advance-article/doi/10.1093/polsoc/puae012/7636223#445724760.

[106]Camille Tuutti, What Trump 2.0 means for tech and AI regulation, NextGov, (Nov. 20, 2024), https://www.nextgov.com/policy/2024/11/what-trump-20-means-tech-and-ai-regulation/401161/.

[107]Emerging Trends in AI Governance: Insights from State-Level Regulations Enacted in 2024, Varnum, (Jan 6, 2025), https://www.varnumlaw.com/insights/state-level-ai-regulations-enacted-in-2024/.

[108]Artificial Intelligence and Civil Rights, U.S. Department of Justice Archives, https://www.justice.gov/archives/crt/ai.

[109]AI and the Risk of Consumer Harm, FTC, (Jan. 3, 2025), https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2025/01/ai-risk-consumer-harm.

[110]Blueprint for an AI Bill of Rights, White House Archives, https://bidenwhitehouse.archives.gov/ostp/ai-bill-of-rights/.

[111]Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, Federal Register, https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence.

[112]Fact Sheet: OMB Issues Guidance to Advance the Responsible Acquisition of AI in Government, White House Archives, (Oct. 3, 2024), https://bidenwhitehouse.archives.gov/omb/briefing-room/2024/10/03/fact-sheet-omb-issues-guidance-to-advance-the-responsible-acquisition-of-ai-in-government/.

[113]FACT SHEET: Biden-Harris Administration Announces New AI Actions and Receives Additional Major Voluntary Commitment on AI, The White House Archives, (Jul. 26, 2024), https://bidenwhitehouse.archives.gov/briefing-room/statements-releases/2024/07/26/fact-sheet-biden-harris-administration-announces-new-ai-actions-and-receives-additional-major-voluntary-commitment-on-ai/.

[114]David Shepardson, Trump revokes Biden executive order on addressing AI risks, Reuters, (Jan. 21, 2025), https://www.reuters.com/technology/artificial-intelligence/trump-revokes-biden-executive-order-addressing-ai-risks-2025-01-21/.

[115]Removing Barriers to American Leadership in Artificial Intelligence, The White House, (Jan. 23, 2025), https://www.whitehouse.gov/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence/.

[116]Kathleen D. Parker, Erinn L. Rigney, Isabella F. Sparhawk, The Changing Landscape of AI: Federal Guidance for Employers Reverses Course With New Administration, The National Law Review, (Jan. 31, 2025), https://natlawreview.com/article/changing-landscape-ai-federal-guidance-employers-reverses-course-new-administration#google_vignette.

[117]Fact Sheet: President Donald J. Trump Takes Action to Enhance America's AI Leadership, The White House, (Jan. 23, 2025), https://www.whitehouse.gov/fact-sheets/2025/01/fact-sheet-president-donald-j-trump-takes-action-to-enhance-americas-ai-leadership/.

[118]H.R.9497 - AI Advancement and Reliability Act of 2024, https://www.congress.gov/bill/118th-congress/house-bill/9497/text.

[119]S.2714 - CREATE AI Act of 2024, https://www.congress.gov/bill/118th-congress/senate-bill/2714.

[120]Rachel Wright, Artificial Intelligence in the States: Emerging Legislation, The Council of State Governments, (Dec. 6, 2023), https://csg.org/2023/12/06/artificial-intelligence-in-the-states-emerging-legislation/.

[121]Katrina Zhu, The State of State AI Laws: 2023, Epic, (Aug. 3, 2023), https://epic.org/the-state-of-state-ai-laws-2023/.

[122]SB-942 California AI Transparency Act, https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=202320240SB942.

[123]AB-2013 Generative artificial intelligence: training data transparency, https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=202320240AB2013.

[124]SB24-205, Consumer Protections for Artificial Intelligence, https://leg.colorado.gov/bills/sb24-205.

[125]High-Risk Artificial Intelligence Developer and Deployer Act (HB 2094), https://www.williamsmullen.com/sites/default/files/2025-03/House%20Bill%20No.%202094.pdf.

[126]Beth Burgin Waller, Patrick Sweeping AI Legislation Under Consideration in Virginia, Progam on Coporate Compliance and Enforcement at NYU Law, https://wp.nyu.edu/compliance_enforcement/2025/01/09/sweeping-ai-legislation-under-consideration-in-virginia/.

[127]William A. Wright , Summer L. Elliot , Joseph C. O'Keefe , Scott M. Kosnoff, Virginia Legislature Passes High-Risk AI Regulation Bill, Faegre Drinker, (Feb. 28, 2025), https://www.faegredrinker.com/en/insights/publications/2025/2/virginia-legislature-passes-high-risk-ai-regulation-bill.

[128]S.B. 149 Artificial Intelligence Amendments, https://le.utah.gov/~2024/bills/static/SB0149.html.

[129]John J. Rolecki, Amy L. Baddley, Emerging Trends in AI Governance: Insights from State-Level Regulations Enacted in 2024, Varnum, (Jan. 6, 2025), https://natlawreview.com/article/emerging-trends-ai-governance-insights-state-level-regulations-enacted-2024.

[130]Hesam Sheikh Hassani, AI Governance: Frameworks, Tools, Best Practices, Datacamp, (Sept. 4, 2024), https://www.datacamp.com/blog/ai-governance.

[131]Emerging Trends in AI Governance: Insights from State-Level Regulations Enacted in 2024, Varnum, (Jan. 6, 2025), https://www.varnumlaw.com/insights/state-level-ai-regulations-enacted-in-2024/.

[132]Emerging Trends in AI Governance: Insights from State-Level Regulations Enacted in 2024, Varnum, (Jan. 6, 2025), https://www.varnumlaw.com/insights/state-level-ai-regulations-enacted-in-2024/.

[133]Jeremy Werner, Building Proactive Governance for AI: Best Practices and Key Frameworks to Mitigate Regulatory Risk, Babl, (Jan. 23, 2025), https://babl.ai/building-proactive-governance-for-ai-best-practices-and-key-frameworks-to-mitigate-regulatory-risk/.

[134]SB24-205 Consumer Protections for Artificial Intelligence, https://leg.colorado.gov/bills/sb24-205.

[135]Emerging Trends in AI Governance: Insights from State-Level Regulations Enacted in 2024, Varnum, (Jan. 6, 2025), https://www.varnumlaw.com/insights/state-level-ai-regulations-enacted-in-2024/.

[136]Risk-Based Approach, EU AI Act, https://www.euaiact.com/key-issue/3#:~:text=The%20EU%20AI%20Act%20introduces,health%2C%20safety%20and%20fundamental%20rights.

[137]Artificial Intelligence Risk Management Framework (AI RMF 1.0), National Institute of Standards & Technology, https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf.

[138]AI Risk Management Framework, https://www.nist.gov/itl/ai-risk-management-framework.

[139]Jeremy Werner, Building Proactive Governance for AI: Best Practices and Key Frameworks to Mitigate Regulatory Risk, Babl, (Jan. 23, 2025), https://babl.ai/building-proactive-governance-for-ai-best-practices-and-key-frameworks-to-mitigate-regulatory-risk/.

[140]Ten Recommendations for Global AI Regulation, Centre for Information Policy Leadership, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_ten_recommendations_global_ai_regulation_oct2023.pdf.

[141]Brian Judge, Mark Nitzberg, Stuart Russell, When code isn't law: rethinking regulation for artificial intelligence, Oxford Academy Policy and Society, (May 29, 2024), https://academic.oup.com/policyandsociety/advance-article/doi/10.1093/polsoc/puae020/7684910.

[142]Ten Recommendations for Global AI Regulation, Centre for Information Policy Leadership, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_ten_recommendations_global_ai_regulation_oct2023.pdf.

[143]Ten Recommendations for Global AI Regulation, Centre for Information Policy Leadership, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_ten_recommendations_global_ai_regulation_oct2023.pdf.

[144]Tatiana Rice, Jordan Francis, Keir Lamont, U.S. State AI Legislation: How U.S. State Policymakers Are Approaching Artificial Intelligence Regulation, Future of Privacy Forum, (Sept. 2024), https://fpf.org/wp-content/uploads/2024/09/FINAL-State-AI-Legislation-Report-webpage.pdf.

[145]Tatiana Rice, Jordan Francis, Keir Lamont, U.S. State AI Legislation: How U.S. State Policymakers Are Approaching Artificial Intelligence Regulation, Future of Privacy Forum, (Sept. 2024), https://fpf.org/wp-content/uploads/2024/09/FINAL-State-AI-Legislation-Report-webpage.pdf.

[146]Tatiana Rice, Jordan Francis, Keir Lamont, U.S. State AI Legislation: How U.S. State Policymakers Are Approaching Artificial Intelligence Regulation, Future of Privacy Forum, (Sept. 2024), https://fpf.org/wp-content/uploads/2024/09/FINAL-State-AI-Legislation-Report-webpage.pdf.

[147]Hope Anderson, Nick Reem, Juliann Susas, Automated Decision Making Emerges as an Early Target of State AI Regulation, White & Case, (Mar. 7, 2025), https://www.whitecase.com/insight-alert/automated-decision-making-emerges-early-target-state-ai-regulation.

[148]Hope Anderson, Nick Reem, Juliann Susas, Automated Decision Making Emerges as an Early Target of State AI Regulation, White & Case, (Mar. 7, 2025), https://www.whitecase.com/insight-alert/automated-decision-making-emerges-early-target-state-ai-regulation.

[149]Hope Anderson, Nick Reem, Juliann Susas, Automated Decision Making Emerges as an Early Target of State AI Regulation, White & Case, (Mar. 7, 2025), https://www.whitecase.com/insight-alert/automated-decision-making-emerges-early-target-state-ai-regulation.

[150]Jeffrey A. Sonnenfeld, Stephen Henriques, The Right Approach to State Regulation of AI, Yale Insights, (Feb. 4, 2025), https://insights.som.yale.edu/insights/the-right-approach-to-state-regulation-of-ai.

[151]Maureen Fulton, Mikaela Witherspoon, State Attorneys General Provide Guidance on Artificial Intelligence Under Existing Data Privacy Laws, Lexology, https://www.lexology.com/library/detail.aspx?g=dd10b21f-c084-47ab-8470-d86bf2505a8c#:~:text=The%20advisory%20stresses%20that%20AI,the%20Massachusetts%20Consumer%20Protection%20Act.

[152]Ten Recommendations for Global AI Regulation, Centre for Information Policy Leadership, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_ten_recommendations_global_ai_regulation_oct2023.pdf.

[153]Brian Judge, Mark Nitzberg, Stuart Russell, When code isn't law: rethinking regulation for artificial intelligence, Oxford Academy Policy and Society, (May 29, 2024), https://academic.oup.com/policyandsociety/advance-article/doi/10.1093/polsoc/puae020/7684910.

[154]Brian Judge, Mark Nitzberg, Stuart Russell, When code isn't law: rethinking regulation for artificial intelligence, Oxford Academy Policy and Society, (May 29, 2024), https://academic.oup.com/policyandsociety/advance-article/doi/10.1093/polsoc/puae020/7684910.

[155]Jonathan Taplin, Move fast and break things? Not again, and not with AI., The Hill, (Sept. 22, 2024), https://thehill.com/opinion/technology/4891654-move-fast-and-break-things-not-again-and-not-with-ai/.

[156]Florence G'sell, Regulating Under Uncertainty: Governance Options for Generative AI, Stanford Cyber Policy Center Freeman Spogli Institute, https://cyber.fsi.stanford.edu/content/regulating-under-uncertainty-governance-options-generative-ai.

[157]Esmat Zaidan, Imad Antoine Ibrahim, AI Governance in a Complex and Rapidly Changing Regulatory Landscape: A Global Perspective, Nature Humanities & Social Sciences Communications, (Sept. 1, 2024), https://www.nature.com/articles/s41599-024-03560-x.

[158]Everything You Need To Know About Regulatory Sandboxes, State Policy Network, (Oct. 12, 2021), https://spn.org/articles/what-is-a-regulatory-sandbox/.