JUNE 2024

AI LEGISLATION SCORECARD

A RUBRIC FOR EVALUATING AI BILLS

epic.org

ELECTRONIC
PRIVACY
INFORMATION
CENTER

ABOUT EPIC

The Electronic Privacy Information Center (EPIC) is a 501(c)(3) non-profit public interest research and advocacy center in Washington, D.C. EPIC was established in 1994 to focus public attention on emerging privacy and civil liberties issues. Our mission is to secure the fundamental right to privacy in the digital age for all people through advocacy, research, and litigation. For more information about EPIC, please visit www.epic.org.

AUTHOR

Kara Williams, Law Fellow

CONTRIBUTIONS BY

Caitriona Fitzgerald, Deputy Director
Calli Schroeder, Senior Counsel and Global Privacy Counsel
Enid Zhou, Senior Counsel
Grant Fergusson, Equal Justice Works Fellow
Becca Downes, Executive Assistant

ACKNOWLEDGEMENTS

EPIC would like to thank the Rose Foundation for Communities and the Environment for providing funding that supported the work on this project. Special thanks to Ben Winters for providing his expertise and support for the development of this scorecard and to Quinn Anex-Ries, Grace Gedye, and Alicia Solow-Niederman for providing valuable insight and feedback throughout the drafting process.

We also wish to thank our generous donors, who make it possible for us to stand up to Big Tech as an independently funded organization.

Introduction

The United States is facing a growing wave of AI legislation at both the state and federal levels. Hundreds of bills seeking to regulate AI were introduced in at least 40 states this legislative session. Dozens of federal regulations have also been proposed. These bills varied widely in their approaches to regulating AI—some tried to set out comprehensive frameworks, some created task forces or commissioned further study, and some focused on regulating narrow or sector-specific AI uses.

EPIC set out to create a tool for evaluating AI bills: EPIC's **AI Legislation Scorecard.** The scorecard provides a rubric for lawmakers, journalists, advocates, and academics to use to evaluate the strength of AI bills. The scorecard lays out key provisions that effective AI legislation should contain, including data minimization requirements, impact assessment and testing obligations, prohibitions on particularly harmful AI uses, and robust enforcement mechanisms.

This scorecard lays out *minimum standards* for the responsible use of commercial AI and automated decision-making systems. AI developers and deployers can and should supplement these guardrails as necessary to protect the rights and safety of the public. While this scorecard is primarily aimed at evaluating laws regulating commercial use of AI, government agencies should be held to these same standards where applicable. Responsible government development, procurement, and use of AI systems should be a model for the safe and effective adoption of AI in commercial settings.

EPIC developed the AI Legislation Scorecard through expert consultations, internal bill analysis, and evidence-backed policy research. We relied heavily on leading AI proposals and frameworks, including the National Institute of Standards and Technology's AI Risk Management Framework, the White House's Blueprint for an AI Bill of Rights, and the Online Civil Rights Act from Lawyers' Committee for Civil Rights Under Law.

EPIC's AI Legislation Scorecard sets forth a model for effective comprehensive AI legislation. But any AI legislation should serve as a complement to current laws and regulations. AI developers and deployers must still follow the same laws as everyone else. Many laws and regulations already provide meaningful recourse for AI harms—such as with AI-facilitated civil rights violations or consumer fraud. Enforcement agencies should turn to these laws in addition to any AI-specific legislation to address harms resulting from the use of AI. Similarly, AI legislation should supplement existing and forthcoming privacy legislation. Enacting strong privacy laws is a key step in mitigating many of the worst AI risks today.

Above all, this scorecard provides a guide for what robust comprehensive AI legislation should look like at either the state or federal level. If you have questions or would like EPIC to review a specific piece of state or federal AI legislation, please contact Kara Williams at williams@epic.org.

AI Legislation Scorecard

This scorecard can be used to evaluate the strength of both state and federal AI legislation, but it was developed with comprehensive AI bills in mind. Not every provision will be relevant to narrow bills aimed at regulating specific sectors or use cases.

1. STRONG DEFINITIONS

- **Algorithms/artificial intelligence/automated decision-making systems** definitions should focus on the function of the system (replacing or impacting human decision-making) and should cover both sophisticated AI models and simpler algorithms/automation processes.
    - No broad carve-outs for government use/national security.
- **Algorithmic discrimination** should be defined as use of an AI system in a manner that discriminates, in treatment or effect, or otherwise makes unavailable the equal enjoyment of goods, services, or opportunities on the basis of a protected class (with exceptions for use to identify/prevent discrimination or increase diversity and inclusion).
    - **Protected classes** should include (at minimum): race, color, ethnicity, national origin, religion, sex, status as pregnant, gender identity, sexual orientation, familial status, disability, biometric or genetic information, income source or income level, or any other classification protected by law.

- **Consequential use of an AI system** definitions should include: (1) uses of AI that have the potential to meaningfully impact any person's safety or well-being (including mental, emotional, and reputational well-being); (2) uses of AI in which the outputs facilitate decisions with legal or similarly significant effects on any person's civil rights, civil liberties, privacy, or equal opportunities; or (3) uses of AI in which the outputs can impact any person's access to or significant change in the price of critical benefits, resources, or services.
    - *Bonus if the definition also includes uses of AI that have the potential to meaningfully impact the safety of the climate or environment, critical infrastructure, voting or elections integrity infrastructure, sensitive/classified government information, or intellectual property.*
    - *Note: Some bills may refer to this concept as "high-risk AI" or in terms of AI that is used in a "consequential decision." This concept should be defined similarly to the above definition, regardless of the term the bill uses.*
- **Developers** are the entities designing, creating, maintaining, modifying, or updating the AI system that is then provided to the deployer. Developers and deployers may be the same party, and there may be multiple developers for one AI system.
- **Deployers** are the entities using the AI system to assist in decision-making or offering the AI system to the end user. Developers and deployers may be the same party.

2. BASELINE REQUIREMENTS

- All consequential uses of an AI system should be covered, not only uses that are the “sole” or “controlling” factor in consequential decisions.
- Developers and deployers should be required to disclose when content is AI-generated.
- Unless the fair use doctrine applies to the use, developers and deployers should be required to obtain affirmative express consent from content creators before using their content to train AI systems, and they should be required to compensate creators fairly.

3. PROHIBITS ALGORITHMIC DISCRIMINATION

- Developers and deployers should be prohibited from developing or deploying AI systems that causes algorithmic discrimination.
    - A duty of care alone is not sufficient.

4. PROHIBITS PARTICULARLY HARMFUL USES

- Particularly harmful uses of AI systems should be prohibited, including (at minimum): emotion or attribute recognition, social scoring, one-to-many facial recognition, and nonconsensual deepfakes.
- The bill must lay out a process (or assign rulemaking authority) for designating additional technologies or uses of AI systems as worthy of moratorium or ban if they either (1) do not work as intended/advertised or (2) can work but are more likely than not to cause harm.

5. DATA MINIMIZATION

- Developers and deployers should be prohibited from collecting, processing, retaining, or transferring personal data unless the collection, processing, retention, or transfer is necessary and proportionate to develop, train, or maintain a specific product or service, to the extent the individual gives affirmative express consent for such use and the use is compatible with the context in which the personal data was collected.
    - Affirmative express consent requires an affirmative act by an individual that clearly communicates their freely given and unambiguous consent in response to a standalone request. Affirmative express consent cannot be obtained through the use of dark patterns.
- Developers and deployers should be required to delete source data used to train AI systems once the training purpose is complete or within 2 years, whichever is shorter.

6. TRANSPARENCY AND ACCOUNTABILITY

A. CONTENT OF AUDITS AND IMPACT ASSESSMENTS

- The pre- and post-deployment impact assessments and audits required in this section must include the following information (at minimum):
    - provenance and quality of training data and inputs, including ensuring the data and testing are sufficient to address the real-world inputs for which the AI system will be used and to evaluate impacts on various demographic groups;
    - how errors in data entry or machine processing are measured and limited (including dangers of relying on AI-generated data as training data);

  - inputs and logic on which the AI system operates;
  - how the AI system was developed and tested;
  - intended uses and foreseeable misuses of the AI system;
  - process and results of regular validation studies and comparative output assessments over time that test for AI model drift or degradation;
  - types of outputs generated by the AI system;
  - any downstream uses of AI system outputs beyond intended purposes;
  - results of any bias audits or testing for discriminatory impacts and impacts on protected classes;
  - data management policies and procedures, including data security evaluations and/or AI red-teaming procedures;
  - procedures for human review or redetermination; and
  - results of risk-benefit analyses conducted in deciding to use the AI system.
- Audits are required to be completed by independent auditors, who must meet certain, appropriate qualifications relevant to the sector they are evaluating.
- Developers and deployers must retain documentation and results of audits and impact assessments for a minimum of 5 years.

B. PRE-DEPLOYMENT REQUIREMENTS

- Before the AI system can be deployed, developers must provide to deployers: (1) proof of positive impact assessment results that meet the criteria in this section, including proof that use of the AI system does not cause algorithmic discrimination and (2) testing results proving the AI system functions appropriately given the intended use.

- Before the AI system can be deployed, deployers must create and implement a testing and evaluation regime to ensure the post-deployment requirements in this section will be met.

C. POST-DEPLOYMENT REQUIREMENTS

- Deployers should be required to perform audits and impact assessments at least once a year and whenever there is a material change to the context in which the AI system is used, how the AI system operates, or the data that is used as part of ongoing AI system operations.
- Developers and deployers of AI systems for consequential uses should be required to affirmatively provide documentation and results of audits and impact assessments to the appropriate enforcement authority (rather than requiring documentation to be provided only upon request).
- If a risk assessment uncovers signs of bias or algorithmic discrimination, deployers must pause use of the AI system until that bias or algorithmic discrimination can be mitigated—or decommission the AI system if the bias or algorithmic discrimination cannot be addressed.
    - Deployers must ensure that pausing or decommissioning the AI system does not disrupt the provision of essential goods, services, or opportunities to impacted groups.
- If developers or deployers fail an audit for an AI system, they must report the failure to government regulators and stop use of the AI system until that failure is mitigated.
- Developers must inform any downstream deployers of failed audits or signs of bias or algorithmic discrimination discovered through impact assessments.

D. NOTICES AND TRANSPARENCY

- Developers and deployers should be required to make public (on their own websites and in a central repository) a plain-language summary of the results of required audits and impact assessments, details about the data on which the AI system was trained, how the risks were weighed against potential benefits and how risks were mitigated, procedures for human review of the AI system, and procedures for ongoing testing and evaluation of the AI system.
- Exemptions to disclosure should be limited to trade secrets and should not include overly broad or vague terms like “proprietary,” “confidential,” or “business” information.
- Deployers should be required to provide notice to individuals that they are being subject to an AI system for a consequential use **prior to** the AI system being used (with enough advance notice that individuals can decide to exercise their right to opt out). This notice must include a summary of (or a link to) a disclosure meeting the requirements set forth in this section.
- All disclosures must be clearly displayed, accessible, and in plain language understandable to a reasonable person.
- If a developer or deployer makes a material change to its public disclosures, it must provide individuals with notice of the change and an opportunity to withdraw any previously given consent.
- If a bill includes federal funding for state adoption of AI systems, funding should be conditioned on state compliance with the transparency and accountability standards set forth in this section.

7. DATA SECURITY

- Developers and deployers should have a duty of care to protect personal data against unauthorized access, use, destruction, modification, or disclosure.
- Data should not be stored, held, or transferred in plain text form.
- Developers are encouraged to use privacy enhancing techniques to minimize the use of personal data in developing AI systems.
- Developers and deployers should be required to create and implement a standard operating procedure for detecting and responding to security incidents and breaches, which should include reporting the incident to relevant government regulators and affected individuals.
- All employees of developers and deployers that work on AI development or deployment must be appropriately trained on how to protect personal data.

8. PROHIBITS UNFAIR PRACTICES

- Developers and deployers should be prohibited from retaliating against consumers who exercise their individual rights.
  - Retaliation includes increasing the price or decreasing the quality of a good or service.
- Developers and deployers are prohibited from taking adverse action against a whistleblower for engaging in lawful whistleblower activities.
- Use of manipulative design or dark patterns to subvert individuals' decision-making, including in providing/withdrawing consent and in exercising individual rights, should be prohibited.

9. INDIVIDUAL RIGHTS

- Unexplainable/uninterpretable AI systems should not be deployed for consequential uses.
- Individuals subject to a consequential use of an AI system should have the right to be notified of the outputs, determinations, or decisions produced and to an explanation of how the AI system produced that output, determination, or decision.
    - The notification and explanation must be written in plain language and must be specific enough that the individual can identify whether the decision was based on inaccurate or incomplete information.
- Individuals subject to a consequential use of an AI system should have the right to access information about the logic of the system.
- Individuals subject to a consequential use of an AI system should have the right to correct any incorrect or incomplete personal information used in generating an output, determination, or decision.
- Individuals subject to a consequential use of an AI system should have the right to opt out of the use of AI system that does not put them in a worse position than they were in before opting out, including disqualifying them from any opportunity/decision.
    - The alternative for individuals who exercise the right to opt out must be a human decision.
- Individuals subject to a consequential use of an AI system should have the right to request redetermination and/or human review.

10. ENFORCEMENT AND RULEMAKING

- Enforcement mechanisms should include a private right of action.
  - Statutory damages should be available.
  - *Bonus if the bill allows additional damages for intentional/repeat violations.*
- The attorney general or other relevant agency/government body should have investigative and enforcement authority.
  - Disgorgement authority (including deletion of algorithms and datasets developed with unauthorized content and forfeiture of profits earned from their use) should be explicitly authorized.
  - Injunctive relief and the ability to impose additional requirements on developers and deployers should be available.
- Cure periods should be available only for violations discovered through internal testing/processes (including red teaming) and not those discovered through an investigation or consumer complaint.
- Government regulators and consumers should be able to sue both developers and deployers for violations.
- A bill should not preempt more protective local regulations (or more protective state regulations if the bill is federal).
- The relevant consumer protection authorities should have rulemaking authority.
- Adequate funding and staffing for enforcement and rulemaking should be appropriated.