A FEDERAL TORT FOR THE AGE OF ALGORITHMS: THE AI ACCOUNTABILITY AND PERSONAL DATA PROTECTION ACT

Romain Catala*

Abstract: The AI Accountability and Personal Data Protection Act (AAPDPA) was introduced by U.S. Senators Josh Hawley and Richard Blumenthal on July, 21 2025. It marks a significant legislative initiative aimed at establishing a federal civil liability for the unauthorized use of personal data and copyrighted material in the development or deployment of artificial intelligence systems. This article analyzes the bill's scope, definitions and remedial framework. It also situates it within broader trends in U.S. privacy and AI regulation and assesses its implications for compliance, litigation risk, and federalism. This Senate bill reflects a profound evolution in the United States toward a more rights-centered approach with regards to the mass commercial and technological exploitation of personal data.

Keywords: AI; Personal Data; Digital Law; Federal Law; Comparative Legal Systems

* Sciences Po Paris, France.

Table of Contents

Introduction 196

I. Legislative Context and Purpose 196

II. Scope and Definitions 196

III. Liability 198

IV. Relationship to Existing Law and Federalism 199

V. Critical Evaluation 199

Conclusion 200

INTRODUCTION

On July 21, 2025, senators Josh Hawley and Richard Blumenthal introduced the AI Accountability and Personal Data Protection Act, a bipartisan bill that would create a federal private right of action against entities that exploit personal data or copyrighted works in artificial intelligence (AI) systems without prior, express consent.

While the U.S. has traditionally approached privacy and technology regulation in a fragmented, sector-specific manner, the AAPDPA proposes a cohesive and enforceable liability structure for the use of sensitive data in emerging AI technologies.[1]

I. LEGISLATIVE CONTEXT AND PURPOSE

The AAPDPA emerges amid increasing public concern over the opaque and largely unregulated data practices of generative AI platforms.[2] The bill's stated aim is to hold accountable "any entity that collects, processes, sells, or uses personal data or expressive works without express consent"[3], particularly in the development or deployment of artificial intelligence. Unlike earlier American privacy frameworks, this bill centers civil enforcement by individuals, rejecting the traditional model of agency-led oversight. The legislation reflects a notable convergence of civil liberties, consumer protection, and copyright discourse—particularly in light of ongoing litigation against firms such as OpenAI, Meta, and Stability AI for alleged unauthorized data use.[4] Some might see this as another Brussels effect.

II. SCOPE AND DEFINITIONS

The bill outlines a set of definitions that shape its scope and reflect the lawmakers' intent to address personal data use comprehensively within the digital ecosystem.

It defines "artificial intelligence system" as "any data system, software, hardware, application, tool, or utility that operates, in whole or in part, using artificial intelligence."[5] This broad scope intentionally captures a wide spectrum of generative and predictive models and intends to encompass all emerging technologies that could exploit personal data for the purpose of automated content generation.

In terms of scope, the bill is meant to apply to all "covered data".[6] In a few words, this term includes all information, data or material that directly or indirectly "[…] identifies, relates to, describes, is capable of being associated with, or can reasonably be linked, directly or indirectly, with a specific individual"[7]; it constitutes an exceptionally broad definition, which is both novel and legally consequential for

---

[1] Daniel J. Solove & Woodrow Hartzog, *The FTC and the New Common Law of Privacy*, 114 Colum. L. Rev. 583 (2014).
[22] Int'l Ass'n of Privacy Pros., *Consumer Perspectives on Privacy and Artificial Intelligence* (2024), https://iapp.org/resources/article/consumer-perspectives-on-privacy-and-artificial-intelligence (summarizing Pew Research Ctr. & KPMG findings).
[3] S. 2367, § 2(b), 118th Cong. (2025).
[4] Complaint, *Tremblay v. OpenAI, Inc.*, No. 3:23-cv-03223 (N.D. Cal. July 5, 2023).
[5] S. 2367, § 2(3).
[6] S. 2367, § 2(4).
[7] Id. § 2(4)(a)(i).

the American legal framework. Also, While the bill's primary aim is data privacy, it also applies to material that "is generated by an individual and is protected by copyright, regardless of whether the copyright has been registered with the U.S. Copyright Office or any other registration authority."[8] This broad definition applies not only to registered copyrighted works, but also to any work that is capable of copyright protection. While not explicitly defined *per se* further in the bill, the term may include expressive content (such as text, images, or audio), behavioral data, and other forms of human-derived input scraped or collected from public and private sources.[9] It emphasizes the human origin of the data that fuels AI systems and marks a legislative effort to distinguish such data from purely synthetic or machine-generated alternatives.

By emphasizing "human-generated" content, the bill expands the scope of liability to cover training datasets that include creative or personal information originally produced by individuals. This may encompass publicly available web content, unregistered works, or metadata associated with human behavior. As such, even non-identifiable or anonymized content may fall within the Act's protections if it originated from a human source.

The term introduces several legal and operational complexities. For instance, AI developers may struggle to determine what portion of their datasets qualifies as human-generated, and whether sufficient consent has been obtained. In cases where models are trained on mixed datasets (comprising both synthetic and human-generated inputs), the legal exposure may hinge on the presence and proportion of human-origin data.

Key terms include:

- "Covered data": encompassing personal identifiers, behavioral data, and biometric information. As explained earlier, this term is defined to encompass all information that identifies, or is reasonably capable of identifying, an individual, whether directly or indirectly. This includes personally identifiable information such as, for example: behavorial data, biometric information, browsing history, geolocation data, and even copyrighted content, such as artistic and intellectual works.

- "Express, prior consent ": The term "express, prior consent" means a clear, affirmative act by and individual, made in advance of any appropriation, use, collection, processing, sale, or other exploitation of covered data, indicating a freely given, informed, and unambiguous consent to the specific appropriation, use, collection, processing, sale, or other exploitation of covered data of the individual.[10] The consent required under the bill is characterized as "express and prior". this notion is very close to the notion of consent developed by the General Data Protection Regulation (GDPR) and moves away from popular concepts in American data-protection such as opt-out.[11] Probably inspired by the European conceptions (notably the General

[8] Id. § 2(4)(a)(iii).
[9] Margot E. Kaminski, *The Right to Explanation, Explained*, 34 Berkeley Tech. L.J. 189 (2019).
[10] S. 2367, § 2(5).
[11] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 Apr. 2016, art. 4(11), 2016 O.J. (L 119) 1 [hereinafter GDPR].

> Data Protection Regulation); it means a clear, affirmative, and unambiguous manifestation of the individual's will. This positioning expressly excludes any form of implied consent or consent obtained by default (for example, the consent by "opt-out"), which are popular concepts in the current American data-protection legal frameworks.

Regarding this bill's personal scope of application, it is broad and includes both natural persons and legal entities. "Any person who, in or affecting interstate or foreign commerce, appropriates, uses, collects, processes, sells, or otherwise exploits the covered data of an individual, without the express, prior consent of the individual, shall be liable to the individual."[12] Put differently, this proposed legislation would impose civil liability on anyone who, "in or affecting interstate or foreign commerce," uses covered data — including for training AI systems or reproducing it as an AI-generated output — without the individual's express, prior consent.

Notably, the bill makes no express exemption for academic, nonprofit, or research-based use, suggesting that all uses of personal or expressive data are presumptively unlawful without affirmative authorization. This is particularly noteworthy given ongoing debates within the scholarly community about the potentially unfair use of academic work by generative AI models. The rationale behind it must be that these kinds of exceptions could be a loophole beneficial for actors of the private sectors (some foreign systems who implemented such exceptions are currently struggling with their application, especially private-public research partnerships that allow private actors to process the data for commercial purposes).

III. LIABILITY

At the heart of the AAPDPA is a newly established federal tort aimed at addressing harms arising from the use of covered data by any natural or legal person without obtaining the requisite consent.

This liability also encompasses entities that knowingly assist, facilitate, or derive benefit—direct or indirect—from such unauthorized exploitation. This includes, for example, those who provide tools, platforms, or services that enable or amplify the misuse of personal data.

The bill creates a private right of action in federal court, authorizing individuals to seek:

- Compensatory damages;
- Statutory damages (minimum $1,000 per violation);
- Treble profits derived from unauthorized data use;
- Punitive damages for egregious conduct;
- Injunctive relief;

[12] S. 2367, § 3(a).

- Attorney's fees and litigation costs.[13]

Further, the bill explicitly prohibits predispute arbitration agreements and class action waivers related to claims brought under the Act. This approach positions AAPDPA as a robust consumer-protective statute, one that mirrors the structural remedies found in antitrust or civil rights laws.

IV. RELATIONSHIP TO EXISTING LAW AND FEDERALISM

Unlike prior federal privacy proposals, the AAPDPA adopts a non-preemptive posture. The bill allows state-level protections—such as California's CCPA/CPRA and Illinois's BIPA—to remain in force where they offer equal or greater protection.[14]

The Act also operates in parallel with existing federal statutes, including the Copyright Act and sector-specific laws such as HIPAA and COPPA. It does not require proof of copyright registration, thus enabling claims based solely on unauthorized use of expressive works. This structure reflects a growing federalist compromise: setting national baselines for rights, while allowing states to maintain or expand upon them. For AI practitioners, this implies a multi-jurisdictional compliance challenge.

V. CRITICAL EVALUATION

The AAPDPA represents a shift toward rights-based data protection in the U.S., aligning in key respects with the GDPR. It enshrines express consent as a prerequisite for data processing, mirroring GDPR arts. 6(1)(a) and 9(2)(a) and imposes transparency obligations akin to GDPR arts. 12–14.

The AAPDPA offers several notable strengths. Most importantly, it empowers individuals to challenge unauthorized uses of their data directly, rather than relying exclusively on agency enforcement. By introducing statutory damages and treble profits, the bill creates powerful incentives for accountability and deterrence. Its decision not to preempt state law preserves robust protections in jurisdictions like California and Illinois, while also encouraging compliance mechanisms such as consent management and data provenance systems that may become industry standards.

At the same time, the proposal presents weaknesses that could complicate its effectiveness. The bill's reliance on private litigation raises concerns about uneven access to enforcement and potentially inconsistent outcomes. Its broad definitional reach risks sweeping in protected speech, fair use, or non-commercial activity, raising constitutional and policy questions. Moreover, the prospect of increased exposure to class actions—including "no-injury" suits premised on technical violations—creates significant litigation risk for companies developing or deploying AI systems.

For general counsel and compliance officers, these features of the bill translate into concrete operational challenges. They will have to determine how consent should be meaningfully captured and documented, what constitutes adequate due diligence when acquiring or licensing third-party datasets, and how can AI developers mitigate

---

[13] Id. § 3(b)(2)(A).
[14] Jack M. Balkin, *Information Fiduciaries and the First Amendment*, 49 U.C. Davis L. Rev. 1183 (2016).

legal risks across overlapping (state and federal and foreign) regimes. These questions illustrate both the promise and the complexity of embedding a private right of action at the heart of federal AI governance.

CONCLUSION

The AAPDPA signals a clear legislative intent to regulate AI by reinforcing individual autonomy and control over personal and creative data.

This newly proposed bill could significantly affect which materials can be used as training datasets for AI systems (including LLMs) and, thus, have a profound impact on current methods of LLM training. Beyond prohibiting the use of personally identifiable information in model development, the bill would also impose liability for the use of potentially copyrighted content. Clarifying the liability issues associated with the use of AI systems is a major step toward ensuring the responsible deployment of this technology. Consequently, this Senate bill may represent a pivotal step in shaping the future legal framework governing AI.

Whether or not the bill passes in its current form, it represents a critical milestone in the evolution of American AI law. Legal departments, technology firms, and public interest advocates should prepare for a new era in which data accountability becomes enforceable not merely as policy—but as federal tort law.