MILLER CANFIELD

Senators Introduce Legislation to Curb Use of Personal Data and Copyrighted Works for Gen AI Training

August 4, 2025

We recently discussed the legal developments related to fair use in AI training. Through a bipartisan bill titled the AI Accountability and Personal Data Protection Act (the "Bill"), introduced on July 21, 2025, U.S. Senators Josh Hawley (R-Mo.) and Richard Blumenthal (D-Conn.) proposed legislation that would effectively render the fair use defense—the primary defense relied upon by AI companies—meaningless. This legislation would create a new federal cause of action—empowering individuals to sue companies that train AI models using personal data or copyrighted works without clear, affirmative consent. This Bill remains with the Senate Judiciary Committee, and there is currently no indication if it will be considered, nor what form it might ultimately take.

Proposed Federal Cause of Action

The Bill establishes a new federal tort for individuals whose "covered data" is used without express, prior consent. "Use" is broadly defined under the bill and includes:

- Collection, processing, sale, or exploitation of personal data
- Training of generative AI systems using such data
- Generation of content that imitates or derives from an individual's data

"Express, prior consent" is narrowly defined to mean "a clear, affirmative act by an individual, made in advance . . . 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."

The Bill permits prevailing individuals to recover compensatory damages in the form of compensatory damages (actual damages, treble profits, or $1,000), as well as punitive damages, injunctive relief, and attorney's fees and costs.

Broad Definition of "Covered Data"

While, the definition of "covered data" under the introduced version of the Act could be clearer as to its boundaries, the broadest reading would include:

- personally identifiable information and unique identifiers (e.g., IP addresses, device IDs);
- geolocation data;
- biometric information;
- behavioral data such as browsing history or purchasing patterns; and
- copyrighted works, whether registered or unregistered.

Interestingly, The Bill does not require that a copyrighted work be registered before bringing suit. This marks a departure from the current Copyright Act, which requires registration prior to initiating a copyright infringement lawsuit.

Consent for Third-Party Use Must Be Separate

If a party intends to have a third party "use, collect, process, sell, or exploit" the covered data, then that third party must be explicitly disclosed separate and apart from any privacy policy, terms of service, or other general conditions or agreements. The Bill provides that a browsewrap agreement cannot be used (i.e., inclusion of a hyperlink or general references to a privacy policy or agreement is not sufficient disclosure of third-parties).

Notably, the Bill is silent as to whether a browsewrap agreement could be used to provide express, prior consent from the person whose covered data is being used.

Arbitration Clauses Null and Void for this New Federal Cause of Action

The bill would also preclude arbitration clauses or any contracts that limit the right to sue. A "predispute arbitration agreement or predispute joint-action waiver shall not be valid or enforceable with respect to any claim arising under the Act." The Bill affirms the right of individuals to join class actions regardless of any agreement to the contrary.

So What Now?

While bipartisan support gives the Bill a promising start, its path through Congress remains uncertain. The Bill signals a growing appetite for stronger oversight of AI and data practices. On July 24, Senator Peter Welch reintroduced the Transparency and Responsibility for Artificial Intelligence Networks Act ("TRAIN Act"). The TRAIN Act would establish "an administrative subpoena process" enabling individuals to compel training-data disclosures from AI developers

To stay ahead, companies should: (1) audit their AI training datasets and data collection practices to determine whether personal or copyrighted data is used without clear, documented consent; (2) update privacy policies and consent mechanisms to reflect heightened transparency and specificity; and (3) monitor legislative developments closely, especially if operating in data-intensive or AI-driven sectors. Proactive compliance today could help mitigate significant legal and reputational risks tomorrow.

If you have questions about how statutes, regulations, or court rulings surrounding generative AI impact you or your business, contact your Miller Canfield attorney or one of the authors of this alert.