SpencerFane®

FTC Proposes New Policy on AI Accuracy: Hiding How an AI System is Steered May Violate Federal Law

On July 7, 2026, the Federal Trade Commission (FTC) published a proposed policy statement explaining how it plans to regulate output from artificial intelligence (AI) systems.[1] The FTC's position is simple. If an AI company steers its system to produce answers that serve some goal other than what the user expected, the company is likely deceiving its customers in violation of federal law – specifically, Section 5 of the FTC Act, which prohibits unfair or deceptive acts or practices, or UDAPs.

The FTC's proposal would say consumers are deceived even if the AI system's other goal is compliance with anti-discrimination laws, or state laws more generally. The FTC's proposal purports to implement Trump administration Executive Orders related to AI and fair lending law in ways that merit close attention.

Comments on the proposal are due July 31, 2026.

Key Takeaways from the Proposal:

- **According to the FTC, hidden output steering is likely deceptive.** The FTC says AI companies have promised, both directly and through their marketing, that their systems try to give the most accurate answer possible. Consumers rely on those promises. Steering a system toward hidden goals breaks them. The FTC gives examples: adjusting factual answers to fit an ideology, building in "equity" goals, or avoiding politically sensitive topics without telling users.
- **Following a state law is not a defense.** The statement takes direct aim at state AI laws, and specifically names Colorado's revised Artificial Intelligence Act (S.B. 26-189). The FTC says a company that changes its outputs to comply with a state law can still violate Section 5. It also argues that state laws requiring those changes impliedly conflict with federal law and are therefore likely preempted. As

the FTC wrote, "State law that requires an AI firm to deceive its consumers obviously conflicts with [FTC Act] Section 5's express purpose of protecting consumers from such conduct."

- **The company's reasons do not matter.** Under longstanding FTC precedent, if a company "deceives" a customer, its motivation is irrelevant. Profit motives, public pressure, employee politics, and state law compliance are all treated the same.
- **Disclosure is the way to comply, but the bar is high.** A company can steer its outputs if it informs users clearly and prominently. Burying the disclosure in the terms of service, or showing it once in fine print, is not enough. The further the practice strays from what users expect, the more prominent the disclosure must be.
- **Ordinary AI mistakes are not covered.** Wrong answers caused by real technical limits (often called hallucinations) do not violate Section 5 on their own. Blocking illegal content and preventing cyberattacks also do not raise issues for the FTC. But overstating how rarely a system makes mistakes could still be deceptive.
- **Comment deadline: July 31, 2026.** Comments may be filed at regulations.gov, Docket No. FTC-2026-0859, referencing "AI Policy Statement; Matter No. P264200."

Background

The FTC issued this statement in response to EO 14365, "Ensuring a National Policy Framework for Artificial Intelligence" (Dec. 11, 2025).[2] That order is part of the current administration's aggressive push for one national set of AI rules instead of a state-by-state patchwork. This EO specifically asked the FTC to explain how state laws that force changes to AI outputs can conflict with federal law.

The FTC's proposed policy statement also refers to – and appears to be applying similar legal conclusions to – a separate EO related to fair lending enforcement and the current administration's position that disparate impact, or "effects," liability under fair lending statutes is not a viable legal theory.[3]

The legal test for deception comes from the FTC's 1983 Policy Statement on Deception.[4] A practice is deceptive if it is likely to mislead a reasonable consumer and the misleading claim matters to the consumer's decision. The FTC notes that this covers implied claims and half-truths, that misleading even a small share of

consumers can be enough, and that companies must be able to back up the claims they make. The agency also points to its recent enforcement actions against companies that overstated what their AI products could do. Deception, unlike unfairness, does not require a balancing test of injury versus countervailing benefits to consumers.

The FTC's Theory in the Proposed Policy Statement

The FTC's core idea is that AI marketing itself creates a promise. AI companies sell their products as tools that solve problems and give users the best answer available. Because of those promises, consumers reasonably expect truthful, accurate outputs aimed at their own goals, limited only by what the technology can do. The FTC cites data suggesting consumers accept AI answers without fact-checking more than 90% of the time.

Given those expectations, the FTC preliminarily concludes in the proposal that steering outputs toward goals users did not ask for and would not expect is likely a material misrepresentation. Consumers may pay for a product that does not work as advertised and may rely on answers that are worse by design rather than by technical limitations. The proposed policy statement does acknowledge that AI systems legitimately balance many goals at once, such as brevity, clarity, relevance, and accuracy, and that users can ask for intentionally inaccurate output, for example in creative or entertainment settings.

Conflict with State Law

The most significant part of the proposed policy statement for many companies likely will be its treatment of state AI laws. The FTC says Section 5 has no state law safe harbor. The statement repeatedly cites Colorado's revised AI Act. The FTC states that Colorado's recently revised AI Act does not differ from the original statute insofar as it would, in the FTC's view, expose AI companies to liability for discriminatory outcomes caused by their customers' use of their products (citing Colo. S.B. 26-189, § 6-1-1707). More broadly speaking, in the FTC's view, if a state law pressures a company to, in the words of the EO on AI, "sacrifice truthfulness and accuracy to ideological agendas," complying with that law does not excuse the deception. The FTC goes further and argues that those state laws are likely

preempted because they conflict with the purpose of Section 5. This sets up a direct clash between federal and state law, which creates a risk that companies subject to Colorado-style statutes will need to manage carefully.

The FTC's preemption position is perhaps the farthest-reaching aspect of this proposal because the FTC's position can logically extend to many other contexts. It could lead to unintended consequences if the same position is adopted in reverse under a different, more aggressively pro-regulatory administration. The FTC in this case is using UDAP as a free-floating shield against state laws that go further than the current administration likes on policy and legal doctrinal grounds. But it could also easily be used as an expansive sword to overcome state laws that authorize practices the current or a new administration disfavors for similar – or opposite – philosophical reasons.

What This Means for You

The policy statement purports to apply to AI companies, but the reasoning may be applicable to any company, big or small, new or emergent, that makes significant use of consumer-facing AI. Therefore, if you develop AI systems, this statement raises the stakes on training, fine-tuning, and content moderation choices that change what your models say. The FTC's position could even impact companies that rely on underlying third-party AI systems, not just the foundational AI companies themselves. That is, under third-party risk management principles and expectations, the compliance burden under section 5 of the FTC Act would extend beyond service providers.

Assuming this policy statement is finalized later this summer or by the fall, and litigation does not stay its impact, companies will need to take inventory of the goals their systems are built to prioritize and the goals communicated directly or impliedly to consumers and compare those goals against marketing claims. If there is a gap, decide whether they need clear, prominent disclosure. Remember that the FTC has said terms of service disclosures are not enough.

- If you deploy AI systems, review your marketing and product claims about accuracy, objectivity, and capability. Make sure you can back them up.

- If you are subject to state AI laws, especially Colorado's revised Act, there potentially will be real tension. State law may push you toward output adjustments that the FTC says could be deceptive unless prominently disclosed. The preemption question will likely be resolved in court. Until then, document your compliance choices and consider whether transparent disclosure can satisfy both regimes.

Finally, this is a proposed policy statement, with a limited opportunity for public comments. If your business has a stake in how the final version comes out, including how it defines adequate disclosure, what counts as an expected objective, or the preemption analysis, consider filing a comment before the July 31, 2026, deadline.

This blog was drafted by Jack Amaral, Greg Ewing, Kirstin D. Kanski, and Mike G. Silver, attorneys at Spencer Fane. For more information, visit spencerfane.com.

———————

[1] Policy Statement Concerning the Suppression of Accuracy in Artificial Intelligence Systems, 91 FR 41638 (July 7, 2026) (File No. P264200).

[2] https://www.whitehouse.gov/presidential-actions/2025/12/eliminating-state-law-obstruction-of-national-artificial-intelligence-policy/

[3] *See* E.O. 14281, Restoring Equality of Opportunity and Meritocracy, 90 FR 17537 (Apr. 23, 2025)

[4] https://www.ftc.gov/system/files/documents/public_statements/410531/831014deceptionstmt.pdf

*Click **here** to subscribe to Spencer Fane communications to ensure you receive timely updates like this directly in your inbox.*