

Section-by-Section Summary

Protecting Kids from Human-Like Chatbots Act

Sec. 1. Short Title.

Cites the Act as the “*Protecting Kids from Human-Like Chatbots Act.*”

Sec. 2. Limitation On Personified Chatbots Available To Minors For Use.

(a) Developer requirements. Requires developers who make chatbots available for consumer deployment, and who know or should know minors will use them, to build a non-personified default setting that is on from a user’s first interaction, requires no action to enable, is technically separable so a provider can activate it for a minor, and cannot be disengaged for users not determined to be adults.

(b) Provider requirements. Requires providers to make the non-personified default setting the default for any user they know or have reason to know is a minor, maintain a parental controls framework; permit the setting to be adjusted only through deliberate action in account settings by an authorized parent or educator, without prompting through the chatbot or interface; and not modify or circumvent the setting in a way that breaks it.

(c) Dual-role entities. An entity that both develops and provides the same chatbot must meet both sets of requirements.

(d) Prohibition on nudging. Bars developers and providers from inducing a known minor, or any user not yet determined to be an adult, to turn off the setting or to misrepresent their age, including through model training, interface design, or claims that the chatbot works better with the setting off.

(e) Authority to adjust the non-personified default setting. Gives parents and guardians exclusive authority to change the setting, with a limited exception for authorized educators acting for a bona fide educational purpose on institution-provisioned accounts.

(f) Rules of construction. Allocates responsibility between developers and providers, shielding each from the other’s failures. A developer is protected if it contractually forbids deployment to minors, unless it knows, or should know, that minors are being served and fails to cut off access within 14 days. A provider is protected if it does not market to minors or encourage age falsification, and permanently removes underage users within 48 hours of confirmed knowledge or 14 days of reasonable suspicion.

(g) Public input. Directs NIST, in coordination with the FTC, to publish a request for information and hold a public workshop within 120 days of enactment.

(h) Technical standards and guidance. Directs NIST to publish interim voluntary guidance within 6 months and final guidance within 18 months, updated at least every two years, and

support development of voluntary consensus standards, covering testing methodologies, benchmarks, red-teaming, implementation approaches, school-managed deployments, the timing and form of AI disclosures, and methods for confirming that a minor's account settings are managed by a parent or guardian. Conformity with the guidance supports the good-faith shield and a presumption of compliance under subsection (i).

(i) FTC enforcement. Treats violations as violations of an FTC rule, enabling civil penalties for first offenses, scaled per affected minor for providers and per deployment for developers, with daily penalties for violations not remediated within 30 days of FTC notice. No penalties apply for periods of good-faith conformity with NIST guidance. Compliance is judged on a chatbot's outputs in ordinary operation: an isolated stray output is not a violation where the setting was tested and the cause is promptly fixed, and a chatbot that passes NIST's testing methodologies is presumed compliant absent a pattern of nonconforming outputs.

(j) State enforcement. Authorizes State attorneys general to bring *parens patriae* actions, with notice to the FTC and a right of intervention.

(k) Rules of construction; preemption; severability. Requires no specific age verification mechanism or new data collection, while requiring use of age information already being collected. Preempts State law only where it conflicts, and preserves State tort and product liability law and any State law more protective of minors (matches KIDS Act language). Clarifies that the setting does not restrict natural language, first-person grammar, politeness, warmth, humor, accurate self-description, safety interventions, or user-requested educational or creative role-play.

(l) Effective date. Core requirements take effect one year after enactment.

Sec. 3. Definitions.

Defines thirteen terms, including chatbot, developer, provider, minor, knows or has reason to know, parental controls framework, school-managed deployment, dark pattern, and operational setting.

Defines the non-personified default setting in statute as a setting in which the chatbot identifies itself as an AI at the start of each session and periodically thereafter; is configured not to represent that it is human, alive, or has feelings, not to simulate a personal relationship or encourage reliance on it for emotional support, and not to pressure the user to keep talking or come back; and maintains that configuration across sessions regardless of user prompts. Whether a chatbot "represents" something is judged by what a reasonable minor would understand its output to convey.