Policy Options for AI Tools in Youth Mental Health

Alizea Alvarez-Sookram

Arizona State University

PAF 505 Public Policy

Kenichi Maruyama, PhD, MPA, GPC

07/05/2025

Policy Options for AI Tools in Youth Mental Health

As artificial intelligence (AI) tools become more deeply embedded in everyday life, young people are increasingly encountering them in contexts that touch their mental health. From early-warning systems that monitor mood and behavior to conversational agents offering therapeutic support, AI technologies are marketed as solutions to gaps in youth mental health services. At the same time, concerns about safety, privacy, and efficacy are mounting. Policymakers have begun to grapple with how best to govern this space. This paper outlines three policy alternatives that have been proposed to resolve the challenges of AI in youth mental health: safety standards for AI companions, bans or restrictions on AI therapy, and federal public health guidance. Each represents a distinct governmental tool for addressing the risks and opportunities of this emerging field.

Policy Alternative Narrative

Policy Alternative 1: Safety Standards for AI Companions

California lawmakers advanced an approach that regulates the design of companion chatbots used by children. Assembly Bill 1064, the "Leading Ethical AI Development for Kids Act", requires companies to build guardrails such as robust crisis responses (e.g., triaging to hotlines when self-harm cues are detected), disclosures that the system isn't human, and ongoing reporting on mental-health impacts. The aim is to make companion systems safer without banning them outright. The proponents are state legislators, with support from child-safety advocates; the bill cleared the legislature and awaited action by the Governor as of late September 2025. As a government tool, this is classic regulation/standards-setting (plus information requirements) targeted at vendors whose products reach minors. For my topic, this

approach tries to preserve access to potentially helpful tools (skills practice, psychoeducation) while addressing foreseeable harms in early-warning and therapeutic use cases

Policy Alternative 2: State Bans and Restrictions on AI "Therapy"

Another option is more restrictive: several states have moved to prohibit or sharply limit AI-delivered "therapy," citing safety, privacy, and the risk of substituting machines for clinicians. Recent reporting notes that Illinois and Nevada have enacted bans on using AI to treat mental health, while Utah allows use only with constraints such as strong disclosures and protections for health data. The sponsors vary (state legislators and attorneys general), but the common rationale is precaution, don't let high-risk therapeutic claims into school or clinical workflows until evidence and accountability catch up. As a policy instrument, this is regulation via prohibition/command-and-control, with some informational tools (mandatory disclosure) where use is permitted. For early-warning and chat-based support, this model prioritizes risk containment over innovation, likely slowing integration into schools and youth programs unless paired with research exemptions.

Policy Alternative 3: Federal Public Health Guidance

At the federal level, the U.S. Surgeon General issued an advisory on youth mental health and social technologies. While not a statute, it functions as a public information/education tool and a soft-regulatory signal to platforms, parents, clinics, and schools. The advisory urges stronger safety by design, transparency, independent evaluation, and age-appropriate protections. For my topic, this creates a federal benchmark that local education agencies and vendors can reference when procuring AI-enabled mental-health tools or enabling parental controls. It doesn't directly mandate features, but it can legitimize contract requirements, school board policies, and

grant conditions that dovetail with safer early-warning and therapeutic use (e.g., documented crisis escalation pathways, external audits, and privacy-by-design for minors' data).

Political Context and Competing Problem Definition

These three alternatives illustrate how different actors define the core problem of AI in youth mental health in competing ways. California lawmakers and child advocates frame it as a design problem: AI can be useful, but only if embedded with safety features. States such as Illinois and Nevada define it as a risk problem, concluding that the safest course is prohibition until stronger evidence emerges. Federal health officials, in contrast, frame it as a public-health communication problem, emphasizing the role of information, education, and cultural norms. Each framing leads to a different policy tool, from regulation to prohibition to guidance. This diversity reflects both the novelty of AI and the fragmented nature of U.S. policymaking, where states act as laboratories while the federal government relies heavily on advisories and moral suasion.

Conclusion

AI technologies hold both promise and peril for youth mental health. The three policy alternatives outlined here, safety standards, bans and restrictions, and federal guidance, illustrate the range of governmental tools available to address this issue. While standards-based regulation seeks to balance innovation with safety, prohibitions reflect a precautionary stance, and federal advisories highlight the role of education and norm-setting. Ultimately, policymakers must weigh these alternatives not only for their feasibility and effectiveness but also for their implications for equity and access. The stakes are high: how we govern AI in youth mental health will shape

whether these tools exacerbate risks or become meaningful supports for a generation facing unprecedented mental health challenges.

Policy Memo Outline

Executive Summary

(no elaboration needed)

Problem Statement

Artificial intelligence (AI) tools are increasingly being used by young people in the context of mental health, through early-warning systems that monitor mood and behavior and therapeutic chatbots that provide cognitive-behavioral support. While these tools offer the potential to increase access to support in a youth mental health crisis, they also raise serious concerns about safety, privacy, and effectiveness. Without adequate guardrails, AI companions may mishandle crisis situations, provide inaccurate advice, or expose sensitive data. Policymakers face the challenge of balancing innovation with protection for vulnerable youth.

List of Three Policy Options

Policy Option 1: Safety Standards for AI Companions (California Model)

- *What it is*: Assembly Bill 1064, the *Leading Ethical AI Development for Kids Act*, sets design standards for AI companions used by minors, requiring crisis-response features, disclosure that the chatbot is not human, and regular impact reporting.
- *Proponent(s)*: California state legislators; child-safety advocacy groups.

- *Government tool(s)*: Regulation and standards-setting; information disclosure requirements.

Policy Option 2: State Bans and Restrictions on AI Therapy

- *What it is*: States such as Illinois and Nevada have enacted bans on AI-based therapy, while Utah permits limited use with strong disclosure and privacy protections.

- *Proponent(s)*: State legislators; state attorneys general.

- *Government tool(s)*: Command-and-control regulation (prohibition); mandatory disclosure in permitted contexts.

Policy Option 3:Federal Public Health Guidance

- *What it is*: The U.S. Surgeon General has issued advisories highlighting risks of social technologies and AI, calling for safety-by-design, transparency, and independent evaluation. While not legally binding, advisories influence practice and policy through "soft law."

- *Proponent(s)*: U.S. Surgeon General; Department of Health and Human Services.

- *Government tool(s)*: Public information and education campaigns; soft-regulatory signaling; potential leverage in procurement and grant programs.

Policy Recommendation and Rationale

(no elaboration needed)

References

California Legislative Information. (2025). *AB-1064 Leading Ethical AI Development for Kids Act*. Bill Text - AB-1064 Leading Ethical AI Development (LEAD) for Kids Act.

Kraft, M. E., & Furlong, S. R. (2021). *Public policy: Politics, analysis, and alternatives* (7th ed.). CQ Press.

Office of the U.S. Surgeon General. (2023). *Advisory on social media and youth mental health*. U.S. Department of Health & Human Services. Social Media and Youth Mental Health | HHS.gov

Schoenherr, N. (2025, September 13). *California bill aims to regulate companion AI for kids, amid national debate on AI therapy bans*. The Hill. California AI Child Protection Bill Clears Statehouse