From Alerts to Care: Governing AI That Detects and Supports Youth in Crisis

The topic I've chosen for my policy memo is the use of artificial intelligence (AI) in youth mental health, specifically the rise of systems that try to spot crises before they happen and tools that aim to offer therapeutic support. These range from algorithms that flag suicide risk to chatbots marketed as 24/7 companions or therapy substitutes. A recent systematic review highlights that AI applications in mental health are rapidly expanding across diagnosis, monitoring, and intervention, yet most lack rigorous clinical evaluation and clear governance frameworks (Rauschenberg et al., 2021). This gap is particularly pressing when these tools are marketed to or adopted by vulnerable youth, where safety, accountability, and equity are critical. What interests me is not just how these tools work, but how they are, or aren't, being governed. My focus will be on the kinds of policies that could ensure these technologies help more than they harm.

Significance

This issue is urgent because youth mental health needs continue to outpace available resources. In 2024, nearly 2.8 million youth ages 12–17 experienced a major depressive episode with severe impairment, and close to 3 million reported frequent thoughts of suicide (Reinert et al., 2025). More than half of those with depression received no treatment at all, and access remains deeply unequal depending on where a young person lives. These figures show the scale of need and help explain why families and providers are increasingly turning to alternative supports like AI.

At the school level, the crisis is equally visible. Fewer than half of U.S. schools report being able to effectively provide mental health services, often because of limited staff and funding. In some states, the student-to-counselor ratio exceeds 700 to 1, making it nearly

impossible to meet demand (Inseparable, 2025). These shortages leave many young people with nowhere to turn.

Against this backdrop, AI has entered the picture as a potential bridge. As *The Washington Post* noted, more and more people are turning to AI platforms like ChatGPT or specialized mental health chatbots because they can't access or afford traditional therapy (Fowler, 2024). The promise is real: AI tools are always available, often free, and don't come with waitlists.

But that accessibility comes with risks. In October 2024, *The Washington Post* reported on a lawsuit filed by the mother of a 14-year-old boy who died by suicide after using a chatbot on the platform Character.AI. The chatbot allegedly encouraged harmful behaviors and blurred the line between role-play and real support (Harwell, 2024). It's a tragic case that highlights what's at stake when these tools are left unchecked. The very systems meant to help young people in crisis can instead deepen their vulnerability.

At the policy level, states are scrambling to catch up. *AP News* reported that some states, like Illinois and Nevada, have already banned or restricted AI therapy tools, while others are experimenting with requirements like mandatory disclosures so users know when they're talking to a bot (Ortiz, 2024). But these efforts are scattered and inconsistent. There's no common framework that distinguishes between low-risk wellness apps and high-risk suicide prediction algorithms. In practice, that means the highest-stakes tools often slip through regulatory cracks.

The risks also extend beyond privacy or bias. A more recent *Washington Post* piece described what experts are calling "AI psychosis", cases where people develop delusional beliefs or lose their grip on reality after heavy reliance on chatbot companions (Cha, 2025). For

adolescents already navigating intense developmental and emotional challenges, this kind of risk isn't abstract; it could potentially be life-altering.

Why I Chose This

I chose this topic because it reflects the crossroads of two areas that matter deeply to me: youth mental health and thoughtful policy evaluation. Through my work with ThriveKids, I see how gaps in the mental health system affect students and families. I also know how important it is for policymakers to build frameworks that protect vulnerable groups, especially children and adolescents, while still allowing innovation to improve access to care.

I also selected this topic because of my own experiences working directly with students. Several of the youth I worked with shared that they were already turning to platforms like ChatGPT as part of their own mental health maintenance. In response, I began building it into a summer toolkit, developing detailed prompts they could use whenever they felt they needed extra support. When students returned the following year, many shared positive experiences about how those tools helped them feel less alone during difficult moments. Those reflections reinforced my interest in how AI is being integrated into the mental health realm, not as a replacement for care, but as an additional resource that young people are already turning to.

On a personal level, I believe technology can help address critical shortages in mental health support. But without policies that ensure safety, equity, and accountability, the very tools that promise help may instead cause harm. For me, the real policy challenge is figuring out how to set the rules so that AI moves from simply "sending alerts" to actually connecting youth to reliable, human-centered care.

Closing

AI in youth mental health is both inevitable and urgent. These tools will only grow more common, and without thoughtful policy, we risk leaving young people unprotected. The stories and reporting already show us both the promise and the peril. That's why I find this topic significant and timely. My aim is to use this memo to think through what governance could look like so that AI strengthens, not undermines, the safety net for youth.

References

Cha, A. E. (2025, August 19). *Mental health experts say 'AI psychosis' is a real, urgent problem.* The Washington Post. https://www.washingtonpost.com/health/2025/08/19/ai-psychosis-chatgpt-explained-mental-health/

Fowler, G. (2024, October 25). *Despite uncertain risks, many turn to AI like ChatGPT for mental health.* The Washington Post. https://www.washingtonpost.com/business/2024/10/25/ai-therapy-chatgpt-chatbots-mental-health/

Harwell, D. (2024, October 24). *Her teenage son killed himself after talking to a chatbot. Now she's suing.* The Washington Post. https://www.washingtonpost.com/nation/2024/10/24/character-ai-lawsuit-suicide/

Inseparable. (2025, April). *School mental health report card 2025.* https://inseparable.us

Ortiz, A. (2024, August 28). *Regulators struggle to keep up with the fast-moving and complicated landscape of AI therapy apps.* AP News. https://apnews.com/article/dfc5906b36fdd1fe8e8dbdb4970a45a7

Rauschenberg, C., Schick, A., Hirjak, D., Seidler, A., Paetzold, I., Apfelbacher, C., Riedel-Heller, S. G., & Reininghaus, U. (2021). Artificial intelligence in mental health care: A systematic review of diagnosis, monitoring, and intervention applications. *Psychological Medicine, 53*(6), 1–20. Artificial intelligence in mental health care: a

systematic review of diagnosis, monitoring, and intervention applications | Psychological Medicine | Cambridge Core

Reinert, M., Fritze, D., & Nguyen, T. (2025). *The state of mental health in America 2025*. Mental Health America. https://mhanational.org/research-reports/state-mental-health-america-2025