**Using AI Software to Detect At-Risk Students and Offer Assistance**

Assignment: Extra, Extra!: Irresponsible Data Science in the News

Cesar Hernandez

University of Pittsburgh

CMPINF 2140: Responsible Data Science

Matt Burton & Nora Mattern

1/27/2025

Between 2009 and 2019, the proportion of American hospitals seeing an “increase in mental health hospitalizations due to children with suicidal or self-harming behaviors grew by 25.8%, costing $1.37 billion” (Barry 2023). This troubling rise in youth suicide behavior led many communities to search for ways to protect children better and provide timely interventions. In response, companies like GoGuardian, Gaggle, Lightspeed, Bark, and Securly developed AI-based monitoring software to help prevent suicides by tracking students' computer activity and identifying those at risk (AAP 2025; Barry 2024; CDT 2023; Time 2023).

GoGuardian is one of the key players in providing these AI algorithms to schools across the country. While the Children's Internet Protection Act (CIPA) “requires schools and libraries to filter harmful internet content”, GoGuardian saw an opportunity to also help address rising suicide rates by monitoring students’ online activities for signs of self-harm or suicidal thoughts (Barry 2024; FCC 2019). Neosho High School in Missouri uses GoGuardian’s software, installed on school-issued Chromebooks and other devices (Barry 2024). When a student searches for terms related to self-harm or suicide, the algorithm triggers an alert to school staff for immediate action.

During school hours, staff review these alerts to determine if there's a real threat. If they suspect something, they pull the student out of class for further assessment using a standardized questionnaire or involve a counselor (Barry 2024). After hours, alerts go directly to law enforcement, who may visit the student’s home to ensure their safety (Barry 2024). This system is part of a more significant effort to prevent self-harm and suicide among youth, combining AI monitoring, school staff review, law enforcement intervention, and collaboration with families to ensure a comprehensive response.

However, implementing AI-based monitoring software raises several important challenges, including concerns about privacy, the potential to deepen existing inequalities, the lack of standardized processes for addressing flagged alerts, and the risk of unnecessary law enforcement involvement, which could inadvertently worsen the situation (AAP 2025; Barry 2024; CDT 2023; Time 2023).

A primary concern is privacy. The software runs in the background of students' devices, collecting vast amounts of data about their lives. Right now, there is no national regulation in place that governs how much data can be collected, how it's stored, or whether it's shared (AAP 2025; Time 2023). There’s also the issue of informed consent—families might not fully understand what they’re agreeing to, how they can opt out, or how the software works (AAP 2025; Time 2023). If a family chooses to opt out, they may be required to provide their own devices, which can be an added burden (AAP 2025; Time 2023).

Another significant challenge is how AI-based monitoring may worsen existing inequalities. Reports show that AI software sometimes flags LGBTQIA-related or race-related content, introducing bias into the monitoring process as this can lead to unintended consequences, like outing LGBTQIA students to school officials without their consent (AAP 2025; CDT 2023; Time 2023). There is also a lack of transparency about the algorithm’s logic, which raises concerns about whether the companies are open to feedback or criticism. Although GoGuardian claims to be partnering with a research institution to assess its product's efficacy, no results have been published yet (Barry 2024).

A third issue is the absence of standardized protocols for handling flagged alerts, determining when escalation is necessary, and what type of intervention is needed. With a lack of a standard protocol, there have been reports of AI-based monitoring used for disciplinary purposes, such as suspensions (AAP 2025; Time 2023). Ryan West, Chief of Neosho’s Police Department, mentioned that about three times a year, a school officer has to visit a student’s home between 11 p.m. and 2 a.m. to intervene in a suicide attempt (Barry 2024). When school staff aren’t available to review alerts on weekends and holidays, law enforcement is contacted directly and follows their procedures for assessing alerts on a case-by-case basis. By visiting students of color outside of school hours, and mostly at night, the fear is that violent interactions could increase (AAP 2025; Time 2023). Without clear protocols, this can result in bias, false positives, and over-policing, which could unfairly target certain students and create an unsafe environment for everyone.

Despite the good intentions behind AI-based monitoring software, aiming to keep children safe and intervene when needed, the situation feels like naïve, irresponsible data science. There doesn't appear to have been enough research into how effective these methods are or enough reflection on how they may affect individuals in the long run. The focus seems to be on meeting the immediate goal without fully considering the potential consequences. If I were to rethink this approach, I’d create an advisory board that includes mental health professionals, community members, school districts, case workers, and law enforcement to provide input on the process—how alerts are triggered and how care is provided both immediately and long-term. I’d also conduct internal and external validation to ensure transparency and objectivity in evaluating the software.

From a personal perspective, I’ve learned that with a problem like this, you can’t rush to find a solution. You must think beyond what’s happening now and carefully consider the entire process and its long-term impact. In this case, the core principles of people-centered, responsible data science were overlooked. Collaboration with various professionals and asking more focused questions could've helped the creators improve their algorithm.

But it’s not only the creators' responsibility—the data users must also ask more questions. My biggest realization from this case is that people need to be more defensive in thinking about their work. What are the positive consequences? And now, what terrible consequences could happen? Companies like GoGuardian are working to drive positive change, but sometimes that positive change has unintended negative consequences. I plan to make sure that I question myself thoroughly throughout my work, reflect on my processes, and document my processes. I hope to uphold the principles of people-centered data science.

**References:**

1. Monitoring Software for Suicide Risk in Schools. (2025). Aap.org. <https://www.aap.org/en/patient-care/media-and-children/center-of-excellence-on-social-media-and-youth-mental-health/qa-portal/qa-portal-library/qa-portal-library-questions/monitoring-software-for-suicide-risk-in-schools/?srsltid=AfmBOoo4urFXqVBCwxgQGvOnS5MvV5WRpUK7PR2yZFbxfRBeDKt8JL1E>
2. Barry, E. (2023, March 28). Hospitals Are Increasingly Crowded With Kids Who Tried to Harm Themselves, Study Finds. The New York Times. <https://www.nytimes.com/2023/03/28/health/pediatric-mental-health-hospitalizations.html>
3. Barry, E., & Dickie, G. (2024, December 9). The Brave New World of A.I.-Powered Self-Harm Alerts. The New York Times. <https://www.nytimes.com/2024/12/09/health/suicide-monitoring-software-schools.html?searchResultPosition=2>
4. *Report - Off Task: EdTech Threats to Student Privacy and Equity in the Age of AI. (2023, September 20). Center for Democracy and Technology.* [*https://cdt.org/insights/report-off-task-edtech-threats-to-student-privacy-and-equity-in-the-age-of-ai/*](https://cdt.org/insights/report-off-task-edtech-threats-to-student-privacy-and-equity-in-the-age-of-ai/)
5. Engle, J. (2024, December 18). Should Schools Spy on Student Devices to Prevent Self-Harm? The New York Times. <https://www.nytimes.com/2024/12/18/learning/should-schools-spy-on-student-devices-to-prevent-self-harm.html>
6. Federal Communications Commission. (2019, December 30). Children’s Internet Protection Act (CIPA). Federal Communications Commission. <https://www.fcc.gov/consumers/guides/childrens-internet-protection-act>
7. AI Monitoring in Schools Can Come at a Great Cost. (2024, February 15). TIME. <https://time.com/6694425/ai-monitoring-school-suicide-cost-essay/>