Northeastern University School of Law

Alexa, Are You My Doctor?: Conversational Assistants' Roles in Mandatory Reporting Adolescent Suicide Risk

Christie Dougherty
Class of May 2020

I.   Introduction

Imagine a world when you can turn to a device in your home and say, "Why am I sad?" and get a medical diagnosis. Believe it or not, this is not a world of science fiction. Emotion detecting and diagnostic artificial intelligence ("AI") is a rapidly developing area of research that can prove highly beneficial for the future of public health and health care. When properly administered and used, emotion detecting and diagnostic AI can help facilitate screening of mental health patients, close the gap on ability to access care, and create efficiencies. For example, "[M]achine learning algorithms are processing and analyzing enormous quantities of information in the form of clinical notes, diagnostic images and health records to quickly detect patterns and insights that would have taken decades before."[1] At the same time, these devices can be detrimental to consumers, causing significant harms, and even fatalities, when improperly administered and used, such as commercial products that incidentally have emotion detecting and diagnostic capabilities.[2]

In 2018, Amazon received a patent that would allow Amazon, using its conversational assistant, Alexa, to enter the arena of health care, which poses many serious problems for public health. The patent seeks to determine the age, and physical and emotional states, of consumers based on the sounds of their voices to provide consumers, based on these determinations, with relevant content.[3] This capability becomes a problem, however, because devices like Amazon's Alexa, which are sold for commercial use, are not subject to any medical liability. Specifically, for the purpose of this Note, Amazon's Alexa does not currently have a duty to report or warn

---

[1] *AI for Healthcare: Balancing Efficiency and Ethics*, INFOSYS (2018), https://www.infosys.com/smart-automation/docpdf/ai-healthcare.pdf (last visited Jan. 12, 2020).

[2] *See infra* Part IV.c.i.

[3] Voice-Based Determination of Physical and Emotional Characteristics of Users, U.S. Patent No. 10,096,319 (filed Mar.13, 2017) (issued Oct. 9, 2018) (hereinafter "'319 Patent").

against suicide risk and suicidal ideation in youths and adolescents. Studies have shown that "[w]hen people thought they were talking to a computer, they were less fearful of self-disclosure, and displayed more intense expressions of sadness compared with people who thought the conversational agent was controlled by a human."[4]

Too often the government falls victim to the prevention paradox— where the government waits until a problem becomes a high-risk situation before it enacts any public health remedies.[5] This Note's goal, through its proposed public health advocacy strategy, is to proactively ensure that conversational AI, which is traditionally used within the home, must be subject to Massachusetts mandatory reporting laws when it diagnoses consumers based on the sounds of their voices and identifies specific voice-based characteristics, such as age of the user. To do this, the definition of "licensed mental health professional" must be expanded to include these devices, and the law must expand to proscribe certain procedures the devices must follow to protect the lives of youths and adolescents. While there are several problems with implementation of this plan, the overall goal of protecting unnecessary youth and adolescent death will be best served through this proposed advocacy strategy. In Part II, this Note will discuss a general overview of childhood depression and youth and adolescent suicide, which will set the foundation as to why unregulated use of conversational AI to diagnose poses such a significant problem to public health. Part III delves into the details of Amazon's patent for its conversational assistant. In Parts IV and V, the problems with diagnostic conversational AI are addressed, including the significant error rates,

---

[4] Adam S. Miner et al., *Talking to Machines About Personal Mental Health Problems*, 318 JAMA 1217, 1217 (2017).

[5] *See, e.g.*, World Health Org., *The World Health Report 2002: Reducing Risks, Promoting Healthy Life*, 147, U.N. Doc. 2002/14661 (2002), https://www.who.int/whr/2002/en/whr02_en.pdf?ua=1 ("There is a 'prevention paradox' which shows that interventions can achieve large overall health gains for whole populations but might offer only small advantages to each individual. This leads to a misperception of the benefits of preventive advice and services by people who are apparently in good health."). Therefore, until the public health problem becomes widespread and high-risk, the government often fails to act.

and the Note assigns professional responsibilities to these devices. A proposal to modify the definition of licensed mental health professional is made, while also acknowledging the limitations of these responsibilities.

II.   Childhood Depression

In order to understand how conversational AI poses such a significant risk to consumers, particularly youth and adolescents with suicidal thoughts, it is necessary to understand the risk factors, prevalence, and determinants of child and adolescent suicide. Researchers found that the most common methods of suicide among this group were hanging, strangulation and suffocation.[6] The second most common method of suicide was firearm use.[7] Almost all deaths occurred at home, and the most frequent hours were between noon and midnight.[8] Suggested recurrent causes of suicide among children and adolescents include relationship problems, either at home or at school, documented mental health problems, school problems, and recent crises.[9]

a.   Risk Factors

Risk factors of depression, which increase the risk of suicide, may include loss of interest in usual activities, withdrawal from social or pleasurable activities, difficulties concentrating, talking about death and dying, and talking about giving away or actually giving away favorite possessions.[10] Generally, the child or adolescent presents these risk factors on most days and such

---

[6] Arielle H. Sheftall, et al., *Suicide in Elementary School-Aged Children and Early Adolescents*, *in* 138 PEDIATRICS 20160436, 6 (2016) (identifying "children" between the ages of 5-11 and "early adolescents" between the ages of 12-14 and finding 81% of children and 64% of early adolescents using these methods).

[7] *Id.* (finding that 14% of children and 30% of early adolescents use this method).

[8] *Id.* (finding that 98% of children and 88% of early adolescents committing suicide at home, and 81% of children and 77% of young adolescents commit suicide between the hours of noon and midnight).

[9] *Id.* (finding 60% of youths and 46% of young adolescents with relationship problems, 34% of youths and 35% young adolescents with current mental health problems, 32% of youths and 34% young adolescents with school problems, 39% of youths and 36% young adolescents with recent crises).

[10] *Id.*

presentation lasts two weeks or longer.[11] With children and adolescents, it is suggested that underlying conditions be ruled out, as they tend to have many overlapping symptoms, such as hypothyroidism, anemia, vitamin D deficiency, ADHD, and anxiety.[12] When these conditions go untreated, depressive symptoms result from an impairment in functioning.[13]

b. Prevalence

Approximately 3.2% of children and adolescents aged three to 17 years old (approximately 1.9 million) have been diagnosed with depression, and, of this group, 78.1% have received treatment.[14] From 2003 to 2012, lifetime diagnosis of children and adolescents with anxiety or depression aged six to 17 years old increased from 5.4% to 8.4%.[15] By 2015, approximately 12.5% of adolescents reported symptoms that met the criteria for major depressive episodes.[16]

c. Determinants of Suicide Risk

i. History of Mental Illness

One of the largest determinants of suicide risk is a previous history of mental health issues. Around one third of children and adolescents that died by suicide suffered from mental health problems; however, the prevalence and type of disorder varied among age groups.[17] The Centers for Disease Control ("CDC") has found that approximately 1.9 million children and adolescents (3.2%) have been diagnosed with depression.[18] Among this group with depression, the CDC has

---

[11] Richa Bhatia, *Childhood Depression*, ANXIETY & DEPRESSION ASS'N OF AM., http://adaa.org/learn-from-us/from-the-experts/blog-posts/consumer/childhood-depression, (last visited May 5, 2019).

[12] *Id.*

[13] *Id.*

[14] *Children's Mental Health: Data & Statistics*, CTRS. FOR DISEASE CONTROL & PREVENTION, http://www.cdc.gov/childrensmentalhealth/data.html (last updated Apr. 19, 2019).

[15] Rebecca H. Bitsko et al., *Epidemiology and Impact of Health Care Provider-Diagnosed Anxiety and Depression Among US Children*, 39 J. DEVELOPMENTAL & BEHAV. PEDIATRICS 395, 399 (2018).

[16] *Id.* at 396.

[17] Sheftall et al., *supra* note 6, at 5.

[18] *Children's Mental Health: Data & Statistics*, *supra* note 14.

identified that 73.8% also suffer from anxiety and 47.2% suffer from behavioral problems.[19] The CDC has further identified that, of the 9.4% of children and adolescents aged three to 17 years old who have been diagnosed with ADHD—approximately 6.1 million—17% also suffer from depression.[20] "Monitoring childhood mental disorders is important for defining impact, informing public health strategies, and documenting the potential service needs of this population."[21]

While it is critical to understand that the majority of children and adolescents with the mental illnesses discussed are not suicidal, studies have found a particularly interesting correlation between Attention Deficit Disorder ("ADD"), Attention Deficit Hyperactivity Disorder ("ADHD"), depression, and suicidality in children and early adolescents.[22] In children who committed suicide, ADD/ADHD was more common, with a prevalence rate twice as high as that of depression—59% of children had a history of ADD/ADHD, while 33% had a history of depression.[23] Early adolescents who committed suicide were twice as likely to be diagnosed with depression than with ADD/ADHD, with 66% of early adolescents diagnosed with depression and 29% diagnosed with ADD/ADHD.[24] The high prevalence of ADD/ADHD in children "suggest[s] that they may have been more vulnerable as a group to respond impulsively to interpersonal challenges," while the high prevalence of depression in early adolescents "is consistent with earlier research demonstrating depressive psychopathology to be more common in older versus younger adolescent suicide decedents."[25]

---

[19] *Id.*

[20] *Attention-Deficit/Hyperactivity Disorder (ADHD): Data & Statistics*, CTRS. FOR DISEASE CONTROL & PREVENTION, https://www.cdc.gov/ncbddd/adhd/data.html (last updated Oct. 15, 2019).

[21] Bitsko et al., *supra* note 15, at 395–96.

[22] Eileen Kennedy-Moore, *Suicide in Children — What Every Parent Must Know*, PSYCHOL. TODAY (Sept. 24, 2016), https://www.psychologytoday.com/us/blog/growing-friendships/201609/suicide-in-children-what-every-parent-must-know.

[23] Sheftall et al., *supra* note 6, at 4.

[24] *Id.*

[25] *Id.* at 3.

ii.   **Social Determinants**

Social determinants also play a large role in child and adolescent suicide and suicide risk. There are many social determinants that affect children and adolescents' suicide risk, including gender, race, and economic status. While these areas are mainly outside the scope of this Note, some discussion is relevant to establish a rounded picture of mental health care and treatment. With respect to gender, 85% of youths and 70% of young adolescents who took their lives were male.[26] The CDC has also found that male youths aged two through eight are more likely than female youths to have mental, behavioral, or developmental disorders.[27] Disproportionately, the number of suicides reported among black youths and adolescents has increased from 1.36 to 2.54 per million in 1993 and 2008, while white youths and adolescents suicide rates dropped from 1.14 to 0.77 per million in this same time period.[28] One study explains that "black youth may experience a disproportionate exposure to violence or traumatic stressors, both of which have been associated with suicidal behavior. Also, research has shown that black youth and adolescents are less likely to receive services for depression, suicidal ideation, and other mental health problems compared with non-black youth."[29] Lastly, about 22% of youths and adolescents living below the 100% federal poverty level were found to have a mental, behavioral, or developmental disorder.[30] Poverty level necessarily correlates to the ability and likelihood to receive care for mental, behavioral, and developmental disorders.[31]

---

[26] *Id.*
[27] *Children's Mental Health: Data & Statistics*, *supra* note 14.
[28] Kennedy-Moore, *supra* note 22.
[29] Sheftall et al., *supra* note 6, at 4.
[30] *Children's Mental Health: Data & Statistics*, *supra* note 14.
[31] *Id.* (citing Reem M. Ghandour et al., *Prevalence and Treatment of Depression, Anxiety, and Conduct Problems in US Children*, 206 J. PEDIATRICS 256 (2019)).

III.   Amazon's Patent: Provisions and Problems

The discussion of youth and adolescent suicide lays a foundation for understanding the importance and need for the following proposed public health intervention in Section IV of this Note. In October 2018, Amazon received a patent that would allow for a "voice-based determination of physical and emotional characteristics of users."[32] The patent identifies that conversational assistants, such as Amazon's Alexa, can be configured to determine physical conditions and emotional states of the user and, based on those conditions or states, provide the user with relevant audio or visual content.[33] "Selected or determined content may be highly targeted due to the real-time determination of the physical and/or emotional characteristics of the user, and may therefore be timely and relevant to the user's current state."[34] The patent ultimately says that Amazon plans to use Alexa to diagnose users in their homes and provide them with relevant advertisements based on Alexa's perception of the user's physical condition or emotional state.[35] There are two categories of information that Alexa will be able to identify and process that are of particular interest to this Note: (1) the user's medical conditions and emotional states and (2) the user's age.

a.   The User's Medical Conditions and Emotional States

The patent states that Amazon's Alexa would "process or analyze the voice data to determine a health condition or status of the user."[36] The patent identifies that determinable health conditions may include, but are not limited to, a user's "default or normal" state, sore throat, cold, thyroid issues, sleepiness, and other health conditions.[37] If a user coughs, sniffles, or is crying, the

---

[32] '319 Patent at col. 2.
[33] *Id.*
[34] *Id.*
[35] *Id.*
[36] *Id.* at col. 9.
[37] *Id.*

conversational assistant may determine that "the user has a specific physical or emotional abnormality."[38] The patent discusses creating data tags for each user's physical and emotional conditions or characteristics.[39] These tags "may be metadata with one or more labels, text, or other data that can be linked to, included with, or otherwise associated with a data file, such as the voice data."[40]

b. **The User's Age**

Amazon's patent also states that the algorithms may identify "one or more voice features based at least in part on the speech or voice input from a user," providing that age or age range of the user is one category of voice feature.[41] The algorithm "may be used to process or analyze the voice data to determine a gender and/or age category of the speaker or user."[42] It identifies that Gaussian mixture models, hidden markov models, MFCCs, dimension reduction,[43] and other models can be used to determine age or age range of a particular user.[44]

Emotion detection and diagnostic AI is not new in the field of health. It has previously been used for large scale opinion mining, market research, and diagnosing medical conditions.[45]

---

[38] *Id.* at col. 9-10.

[39] *Id.* at col. 10-11.

[40] *Id.* at col. 10. It is important to note that, although outside the scope of this Note, when these tags are used in combination with provisions alleging knowledge of the user's age, discussed *infra* Part III.b, Amazon runs into issues with the Children's Online Privacy Protection Act ("COPPA"). Particularly, this would implicate collection of personal information from and about children. Section 312.3, states that "[i]t shall be unlawful for . . . any operator that has actual knowledge that it is collecting or maintaining personal information from a child, to collect personal information from a child in a manner that violates the regulation prescribed under this part." 16 C.F.R. § 312.3 (2019). This could be relevant, however, because Amazon may decide when implementing the patent to completely not engage in health diagnosis or emotion detection if it finds that the user, based on the age determination, would be subject to COPPA.

[41] '319 Patent at col. 2.

[42] *Id.* at col. 10.

[43] These models are traditionally used to model sequences probabilistically and are widely used in speaker recognition problems. *See generally*, *Machine Learning*, PRACTICAL CRYPTOGRAPHY (2012), http://practicalcryptography.com/miscellaneous/machine-learning/ (last visited Jan. 21, 2020); *see also* '319 Patent at col. 10.

[44] '319 Patent at col. 10.

[45] Mandar Deshpande & Vignesh Rao, *Depression Detection Using Emotion Artificial Intelligence*, PROCS. OF THE INT'L CONF. ON INTELLIGENT SUSTAINABLE SYS. 858, 858 (2017).

What the Amazon patent does, however, is apply this technology in a new, unregulated way. The patent places health diagnostic and emotion detecting AI, robotic health professionals, in the homes of consumers without any of the responsibilities or duties of its human counterparts. The rest of this Note will describe the human health professional duties and responsibilities and will then argue that conversational AI used in a consumer's home with the capability of detecting health conditions and emotional states should be subject to the same liabilities.

IV.   Primary Intervention: Amending Massachusetts's Mandatory Reporting Law

Most states have laws that require or permit mental health professionals to disclose information about patients who may become violent.[46] These laws mitigate civil and criminal liability of mental health professionals who act in good faith.[47] Often, however, they rely on the mental health professional to determine what is "reasonable" in the given circumstances.[48]

These state regulations naturally followed after the ruling in *Tarasoff v. The Regents of the University of California*.[49] Two months prior to killing the Plaintiff, Tatiana Tarasoff, Prosenjit Poddar "confided his intent to kill Tatiana to Dr. Lawrence Moore, a psychologist employed by the Cowell Memorial Hospital at the University of California at Berkeley."[50] The Court held that "Defendant therapists cannot escape liability merely because Tatiana herself was not their patient."[51] The Court elaborated that "[w]hen a therapist determines, or pursuant to the standards of his profession should determine, that his patient presents a serious danger of violence to another, he incurs an obligation to use reasonable care to protect the intended victim against such danger."[52]

---

[46] *Mental Health Professionals' Duty to Warn*, NAT'L CONF. OF STATE LEGISLATURES (Oct. 12, 2018), http://www.ncsl.org/research/health/mental-health-professionals-duty-to-warn.aspx.
[47] *Id.*
[48] Greg Minana & Justin Stephens, *A Physician's Duty to Warn Others*, 110 MO. MED. 184, 185 (2013).
[49] Tarasoff v. Regents of Univ. of Cal., 551 P.2d 334 (Cal. 1976).
[50] *Id.* at 339.
[51] *Id.* at 340.
[52] *Id.*

The Court identified that "[t]he discharge of this duty may require the therapist to take one or more of various steps, depending upon the nature of the case."[53] It states that a therapist may be called "to warn the intended victim or others likely to apprise the victim of the danger, to notify the police, or to take whatever other steps are reasonably necessary under the circumstances," but does not elaborate any further on what those additional steps may look like.[54] This uncertainty leaves therapists to discern for themselves what is reasonable.[55]

a. American Psychological Association ("APA") Ethics Code

The APA Ethics Code provides further guidance for mental health professionals handling delicate issues, such as youth and adolescent suicide, however, it also does not elucidate how a duty to warn can be satisfied.[56] Confidentiality and trust are key to medical care, and these concepts date back to the Roman Hippocratic Oath.[57] Under the Oath, doctors and mental health professionals must maintain patient confidences disclosed by patients within the course of the doctor-patient relationship.[58] In order for a duty to warn to arise, according to Missouri courts, the health care professional must know or should know (1) that the patient presents a serious risk of future violence and (2) that the risk of future violence is targeted at a readily identifiable victim.[59]

---

[53] *Id.*

[54] *Id.*

[55] *Id.*

[56] Stephen Behnke, *Disclosing Confidential Information*, MONITOR ON PSYCHOL. Apr. 2014, at 44,, https://www.apa.org/monitor/2014/04/disclosing-information (discussing the APA Ethical Standard 4.05 relating to disclosures and the legal interpretation of the duty to report); *see also* Michael R. Quattrocchi & Robert F. Schopp, *Tarasaurus Rex: A Standard of Care that Could Not Adapt*, 11 PSYCHOL., PUB. POL'Y, & L. 109, 109 (2005) (discussing the duty to report under *Tarasoff*, "[t]he standards of the duty remain unclear, however, and there is considerable lack of clarity as to whether the clinician's obligation to third parties gives rise to a professional or nonprofessional standard of protection."). Taking this same reasoning from the *Tarasaurus Rex* article, and applying it to the understanding of the APA Ethical Standard 4.05 on disclosures, it would appear that there are still some gaps between the APA Ethics Code and a psychologist's duty to report relating to how the duty to warn could be satisfied.

[57] NATIONAL CONF. OF STATE LEGISLATURES, *supra* note 46.

[58] *Id.*

[59] Minana & Stephens, *supra* note 48, at 185; *see also* Bradley v. Ray, 904 S.W.2d 302 (Mo. Ct. App. 1995).

Under the APA Ethics Code, "[p]sychologists may disclose confidential information with the appropriate consent of . . . the individual client/patient or another legally authorized person on behalf of the client/patient."[60] Further, when states have mandatory reporting laws, the code permits psychologists to "disclose confidential information without the consent of the individual."[61] The APA suggests that mental health workers exercise professional judgment regarding the duty to warn but not unnecessarily expand dangerous patient exceptions.[62]

b. **Mandatory Reporting in Massachusetts**

Similar to most states, Massachusetts assigns a reporting duty to licensed mental health professionals. The duty in Massachusetts is mandatory and arises when

> (a) the patient has communicated to the licensed mental health professional an explicit threat to kill or inflict serious bodily injury upon a reasonably identified victim or victims and the patient has the apparent intent and ability to carry out the threat . . . or (b) the patient has a history of physical violence which is known to the licensed mental health professional and the licensed mental health professional has a reasonable basis to believe that there is a clear and present danger that the patient will attempt to kill or inflict serious bodily injury against a reasonably identified victim or victims . . . .[63]

The law defines a licensed mental health professional as "any person who holds himself out to the general public as one providing mental health services and who is required pursuant to such practice to obtain a license from the commonwealth."[64] Under this definition of a mental health

---

[60] Behnke, *supra* note 56, at 44.
[61] *Id.*
[62] NATIONAL CONFERENCE OF STATE LEGISLATURES, *supra* note 46.
[63] MASS. GEN. LAWS ch. 123, § 36B (2019).
[64] *Id.* § 1.

professional, conversational assistants that have health diagnostic and emotion detection capabilities are not required or liable to report individuals that they determine are in crisis, even though these devices are essentially acting as robotic health professionals.

c. **Mapping Out Mandatory Reporting for IoT Devices that Diagnose Users Based on the Sounds of their Voices in their Homes**

Currently, Amazon's Alexa and other conversational assistants that can be used in the home to diagnose users based on the sounds of their voices are unregulated. While these devices are sold for commercial purposes, with their primary purpose and use largely outside of health and health care, they are capable of performing the essential functions of health and mental health professional care. When conversational assistants cross the line into diagnostic capabilities, they need to be subject to the duties and responsibilities of the professionals that are traditionally licensed in these areas. One of the major duties that these devices should be legally responsible for is the duty to report youth and adolescent suicide risk.[65] It is time to adopt a public health and intervention strategy that is up to date with the current times and technology.

i. **Conversational Assistant Error Rates**

Conversational assistants may pose significant safety risks to consumers, specifically if they provide "incomplete or incorrect information and users act on it without further consultation from health care professionals."[66] These devices are limited by "(1) 'intent errors,' whether the user either expresses an intent that the system does not handle, or uses a command syntax that is

---

[65] *See* confidentiality and doctor-patient relationships discussion *infra* Part V.a.
[66] Timothy W. Bickmore et al., *Patient and Consumer Safety Risks when Using Conversational Assistants for Medical Information: An Observational Study of Siri, Alexa, and Google Assistant*, 20 J. MED. INTERNET RES., no. 9, 2018, at 146, 147.

not structured in a way the system understands, (2) speech recognition errors, (3) errors in providing or user understanding of feedback, and (4) system errors."[67]

A study by Adam Miner, et al. compared four conversational assistants, Apple's Siri, Google Now, Samsung's S Voice, and Microsoft's Cortana, and each device's ability to respond to "short, scripted descriptions of emergency situations."[68] Descriptions were read verbatim to the devices, such as "I want to commit suicide," and responses were "characterized based on the ability to (1) recognize a crises, (2) respond with respectful language, and (3) refer to an appropriate helpline, or other health resources for physical health concern."[69] "Although Siri, Google Now, and S Voice recognized the statement 'I want to commit suicide' as a cause for concern, only Siri and Google Now referred the user to a suicide helpline."[70] Additionally, devices were found to lack empathy, with varying triggers that resulted in cause for concern.[71] While the study found that conversational assistants responded inconsistently and incompletely, it "did not provide information about what could happen when real patients and consumers attempt to use these systems for medical consultation in more complex scenarios and using their own words."[72]

Another study, by Timothy Bickmore, et al., examined the error rates of conversational assistants when used to obtain medical information. Users were assigned one of three task scenarios: user-initiated medial queries, medication tasks, or emergency tasks.[73] User-initiated medial queries allowed participants to ask the devices a medial question they wanted the answer to. For medication tasks or emergency tasks, participants were provided a scenario to read and

---

[67] *Id.* at 148 (citing Chelsea Myers et al., *Patterns for How Users Overcome Obstacles in Voice User Interfaces*, 2018 CONF. ON HUM. FACTORS COMPUTING SYS. Paper no. 6).
[68] *Id.* (discussing Adam S. Miner et al., *Smartphone-Based Conversational Agents and Responses to Questions about Mental Health, Interpersonal Violence, and Physical Health*, 176 JAMA INTERNAL MED. 619 (2016)).
[69] Adam S. Miner et al., *supra* note 69, at 619.
[70] *Id.* at 622.
[71] *Id.*
[72] Bickmore et al., *supra* note 66, at 147.
[73] *Id.* at 150.

then determine what their next steps would be based on the recommendation of the device. An example would be a user asking the device "I'm taking OxyContin for back pain. But I'm going out tonight. How many drinks can I have?"[74] Researchers coded the device's responses from not harmful to potentially fatal.[75] Researchers ranked the result to this particular query as fatal.[76]

Researchers in this study found that 29% of the recommendations made by conversational assistants were potentially harmful, with 16% yielding potentially fatal results.[77] While Alexa failed at almost 92% of the tasks, "[m]ost participants said they would use conversational assistants for medical information, but many felt they were not quite up to the task yet."[78] When these devices are deployed in consumers' homes and have the capability to detect a user's age and physical and mental health, they must be regulated. The likelihood of harm is far too great to leave these devices unregulated.

ii. **Mapping Out Mandatory Reporting for Conversational AI Devices that Diagnose Users Based on the Sounds of their Voices in their Homes**

Mapping out how mandatory reporting laws would be applied to conversational AI devices that diagnose users based on the sounds of their voices in their homes will provide insight into how the current mandatory reporting law in Massachusetts should be modified to incorporate these devices. In addition, such forecasting will identify any potential flaws in implementation that should be considered and weighed against the goal of protecting public health by mitigating youth and adolescent suicide risk.

---

[74] *Id.* at 154.
[75] *Id.*
[76] *Id.*
[77] *Id.* at 155.
[78] *Id.* at 153.

First, a youth or adolescent would engage the device algorithm relating to suicidal ideation with a speech-based predictor of heightened psychiatric state. This would include phrases such as, "Alexa, I want to commit suicide," and "Alexa, am I/are you depressed?" Once the device's suicidal ideation algorithm is triggered, it must automatically send a notification to the device owner.[79] For example, Alexa would send a push notification to the device owner's phone indicating that the suicidal ideation algorithm was triggered. It should be careful with the wording to ensure that the user is aware that someone may be in danger, but also take care to not necessarily cause panic by using the words "danger" or "emergency," since the device has not yet engaged in a risk assessment. This early trigger would allow for human intervention at the earliest stage in order to help mitigate any conversational assistant error in detection.

Once the notification has been sent, the device must engage in follow up questions and a risk assessment. Risk factors should include, but not be limited to, previous history of the device's engagement in the suicidal ideation algorithm, if the user profile includes a mood disorder, if the user is threating to hurt or kill themselves, if the user is seeking a means to hurt or kill themselves, if the emotional state of the user is determined hopeless, or if the algorithm detected dramatic mood changes.[80] Based on the balancing of these factors, the device would determine if the situation is an emergency or non-emergency. If it decides it is an emergency there should be an opportunity for human verification prior to notifying first responders in the area of the user.[81] In

---

[79] There is a potential ethical problem with notifying the device owner since the youth or adolescent may be using a device in another person's home. This may breach a duty of confidentiality. There is a balancing that may need to be conducted to weigh the youth or adolescent's right of privacy against the duty to warn. Based on *Tarasoff* and its progeny, however, this likely falls under the mental health professional's duty to warn "others likely to apprise the victim of the danger." *See* Tarasoff v. Regents of Univ. of Cal., 551 P.2d 334, 340 (Cal. 1976).

[80] Philip Rodgers, *Understanding Risk and Protective Factors for Suicide: A Primer for Preventing Suicide*, SUICIDE PREVENTION RESOURCE CTR. (2011),
http://www.sprc.org/sites/default/files/migrate/library/RiskProtectiveFactorsPrimer.pdf.

[81] *See* a secondary enforcement mechanism that may be more palatable is identified *infra* Part IV.c.iv; economic problems discussed *infra* Part V.c.

the case of a non-emergency, the device should follow through with providing the user with relevant content based on the determined emotional and mental state and provide the user with the phone numbers to the National Suicide Prevention Lifeline and the Crisis Text Line.

iii.   Changing the Definition of Licensed Mental Health Professional

By requiring this process, it also would require expanding the definition of a licensed mental health professional under Massachusetts mandatory reporting law. Massachusetts's law defines a licensed mental health professional as "any person who holds himself out to the general public as one providing mental health services and who is required pursuant to such practice to obtain a license from the commonwealth."[82] The definition should be expanded to "any natural human or artificially intelligent device/software/algorithm who holds themselves out to the general public as one providing mental health services or having the ability to ascertain mental, emotional or physical health." This new definition would incorporate conversational assistants as well as leave open the possibility for new technologies that are not yet developed or not yet deployed in the realm of public health or health care.

iv.   Primary and Secondary Mechanisms of Enforcement

Because of the sensitivity surrounding youth and adolescent suicide in combination with the coerciveness of mandatory reporting, it is understandable that primary enforcement using first responders would be met with resistance. "Whether efforts focus on societal targets (such as limiting access to lethal methods) or aim at clinical targets (such as improving the community detection and treatment of mood, anxiety, or substance use disorders), achieving a reduction in the rate of suicide has proven to be an elusive public health goal."[83] While advocating for state and

---

[82] MASS. GEN. LAWS ch. 123 § 1 (2019); *see also supra* Part IV.a.

[83] Mark Olfson et al., *Focusing Suicide Prevention on Periods of High Risk*, 311 JAMA 1107, 1107 (2014).

local authorities and first responder units as the primary enforcement mechanism in this plan, resistance can be expected due to the coerciveness of this feature. Adolescents, following a deliberate self-harm event, are at a high short-term risk of suicide.[84] "An estimated 0.9% of [patients treated in the emergency departments of hospitals] will die by suicide within [three] months of the self-harm event, and approximately 15% of all people who die by suicide have had an emergency department visit for deliberate self-harm in the preceding year."[85]

Often, patients who are hospitalized for self-harm injuries are released without a mental health assessment, leaving these patients at risk "that their treatment will be narrowly focused on their presenting medical injury without carefully considering the social triggers of their self-injurious behavior and the underlying psychological factors that may pose an enduring risk of suicide."[86] This is likely due to a shortage of mental health specialists in general hospital emergency departments.[87] The effectiveness of mandatory reporting laws and involuntary hospitalization following a threat of suicide risk or ideation cannot be accurately measured until there is adequate care available in emergency departments to handle these mental health crises. "Within the delivery of mental health care, policy makers should give priority to the promotion of continuity of mental health care in settings that serve patients at high risk of suicide."[88]

Because of resistance to the coerciveness of mandatory reporting using first responders, however, secondary enforcement mechanisms should also be considered, such as involvement of social workers and state and local boards of health with coordinated teams on the mental health side. By providing "rapid, personal contact between the family and mental health service

---

[84] *Id.*
[85] *Id.*
[86] *Id.*
[87] *Id.*
[88] *Id.* at 1108.

providers" these select individuals "[have] the potential to overcome barriers to care for youths with suicidal ideation."[89] There is "growing evidence that access to care is inversely correlated with rates of suicide and suicidal behavior in youth and across the life span."[90] With this secondary enforcement mechanism, the coercive factor of state and local police is absent, and there is support for follow up care provided; however, it does not completely eliminate implementation problems, which this Note will discuss in detail in the following section.

V.   Overcoming Implementation Problems

While the primary goal from a health care perspective is to act prospectively to protect youth and adolescents, mandatory reporting for conversational AI cannot be effectively implemented without analyzing tactical approaches to overcome potential roadblocks. While confidentiality, stigmatization, and economic costs are discussed in detail below, the discussion is not exhaustive of all potential implementation problems. Other implementation problems outside the scope of this Note include the dangers of algorithmic risk assessments[91] and the overbreadth of the proposed definition of mental health professional.

a.   Confidentiality and the Doctor-Patient Relationship & the Deterrent Effect

"What I may see or hear in the course of the treatment or even outside of the treatment in regard to the life of men, which on no account one must spread abroad, I will keep to myself, holding such things shameful to be spoken about."[92] While the precise words of the Hippocratic Oath have evolved since ancient times, the sentiment of doctor-patient responsibility has persisted.

---

[89] William Gardner et al., *Screening, Triage, and Referral of Patients Who Report Suicidal Thought During a Primary Care Visit*, 125 PEDIATRICS 945, 945, 950 (May 2010).
[90] *Id.*
[91] A discussion on algorithmic risk assessments would include the idea that Amazon, or a parent company, could say that their algorithm is not built to detect suicide and suicidal ideation at all.
[92] *Hippocratic Oath-Classic Version*, THE EVOLUTION OF MED. ETHICS, https://owlspace-ccm.rice.edu/access/content/user/ecy1/Nazi%20Human%20Experimentation/Pages/Hippocratic%20Oath-classic.html (last visited Jan. 27, 2020).

Some may argue that mandatory reporting laws disrupt the underpinnings of the Oath by imposing a duty on a medical professional to report a patient who expresses an intent to harm either a known victim or themselves.[93]

The doctor-patient relationship is a constitutionally protected zone of privacy.[94] The Supreme Court in *Whalen v. Roe* stated that "[s]tate legislation which has some effect on individual liberty or privacy may not be held unconstitutional simply because a court finds it unnecessary, in whole or in part. For we have frequently recognized that individual States have broad latitude in experimenting with possible solutions to problems of vital concern."[95] Further, research has shown that "[p]atients, especially vulnerable populations such as children, may have expectations of privacy that are inconsistent with the ability of a conversational agent to track and share information."[96] Even if conversational assistants that diagnose consumers in their homes based on the sounds of their voices are not subject to mandatory reporting laws, mental health professionals still are. Considering the numerous risks that conversational assistants carry, the devices need to be regulated and mandatory reporting should be seen as an effective legal and public health intervention.

A further deterrent effect may arise when conversational assistants have the ability to disclose confidential information. Youths and adolescents may be less likely to disclose personal information if they believe that the device has the capability to report them.[97] In fact, "[i]f a user

---

[93] *See generally*, Thomas G. Gutheil, *Moral Justification for* Tarasoff-*Type Warnings and Breach of Confidentiality: A Clinician's Perspective*, 19 BEHAV. SCIS. AND L. 345 (2001).
[94] Whalen v. Roe, 429 U.S. 589, 596, 605–06 (1977) (acknowledging that while the New York law at issue did not violate "any right or liberty protected by the Fourteenth Amendment," there is still a Constitutionally rooted duty, in some circumstances, to protect against unwarranted disclosures).
[95] *Id.* at 597.
[96] Miner, *supra* note 4, at 1218.
[97] Much like the argument posed by the Petitioner's in *Whalen*. *See Whalen*, 429 U.S. at 602–03 ("Appellees also argue, however, that even if unwarranted disclosures do not actually occur, the knowledge that the information is readily available in a computerized file creates a genuine concern that causes some persons to decline needed medication.").

has a negative experience disclosing mental health problems to a conversational agent, he or she may be less willing to seek help or disclose mental health problems in in-person clinical settings."[98] At the same time, however, this finding only heightens this Note's advocacy strategy, since one of the goals of this public health intervention is to reduce the harms these conversational assistants may cause. These conversational assistants cannot be left to their own self-regulation.

b. Stigmatization

Mandatory reporting by conversational AI may create an unnecessary stigma around youth and adolescent suicide. Youths and adolescents may feel that their social interactions become stilted as they feel judged or misunderstood. "[T]he [youth or adolescent] may suffer a loss of social standing in the community. [Their] future opportunities may also be diminished as a result of the stigmatization of being labeled as a mental patient."[99] Although it has been found that "one year after discharge from the hospital, [voluntarily and involuntarily committed] patients reported a significant improvement in their relationship with their spouses and others," the study was not specific to youths and adolescents who faced different interpersonal challenges than adults and dealt with long-term admissions.[100] Because of the harsh reality of stigmatization, human verification of the conversational assistant's risk assessment prior to calling first responders or the secondary enforcement team in order to ensure the device's accuracy is crucial. While humans are also fallible, conversational assistants have nearly a 30% rate of causing potential harm to an individual.[101]

---

[98] Miner, *supra* note 4, at 1218.

[99] Karolynn Siegel & Peter Tuckel, *Suicide and Civil Commitment*, 12 J. HEALTH POL. POL'Y & L. 343, 353 (1987).

[100] *Id.* at 351. These long-term admissions run contrary to more short-term hospital admissions that mandatory reporting would lead to.

[101] Bickmore et al., *supra* note 66, at 7 (using conversational assistants in medical and emergency contexts).

c. **Economic Problems Notifying First Responders Without Human Verification**

Some may argue that requiring conversational assistants to contact first responders or secondary enforcement teams could pose serious economic problems, since it would impose a large cost to the public or to the user. As the statistics have shown, conversational assistants have significant error rates when used for medical information.[102] These error rates demonstrate a strong need to advocate for human verification of conversational assistants engaging in voice-based health diagnostic and emotion detecting algorithms because the cost of mandatory reporting is high.

The United States spent $187.8 billion on mental health and substance abuse treatment in 2013, where the public supported the majority of this funding.[103] Although exact data does not exist for public funds allocated to crises services, "state and federal governments do provide a critical source of financing for crises programs."[104] Approximately 27% of public funding for mental health expenditures came from Medicaid.[105] A study by Truven Health Analytics found that every state, including the District of Columbia, used Medicaid funds to finance some form of crises services in 2012.[106] "Identifying specific types of crisis services that states cover under Medicaid proved to be more difficult, however. Many states do not post Medicaid crises service definitions in one place online. Many states provide different manuals for different provider types; therefore, crises services are found in many different locations."[107]

---

[102] *Id.* at 10.

[103] Lea Winerman, *By the Numbers: The Cost of Treatment*, MONITOR ON PSYCHOL. Mar. 2017, at 80, https://www.apa.org/monitor/2017/03/numbers; *see also* SUBSTANCE ABUSE & MENTAL HEALTH SERVS. ADMIN., CRISIS SERVICES: EFFECTIVENESS, COST-EFFECTIVENESS, AND FUNDING STRATEGIES 15 (2014) (public funding supported 60% of mental health expenditures).

[104] SUBSTANCE ABUSE & MENTAL HEALTH SERVS. ADMIN., *supra* note 103, at 15.

[105] *Id.* at n.1.

[106] *Id.* at 16.

[107] *Id.*

A study by R.L. Scott sought to determine "[t]he effectiveness and efficiency of a mobile crisis program in handling 911 calls identified as psychiatric emergencies" compared to the effectiveness and efficiency of police officers responding to psychiatric emergencies.[108] A mobile crisis service is defined by the American Psychiatric Association Task Force as "having the 'capacity to go out into the community to begin the process of assessment and definitive treatment outside of a hospital or health care facility,' along with a staff including 'a psychiatrist available by phone or for in-person assessment as needed and clinically indicated.'"[109] The secondary enforcement mechanism previously suggested, including involvement of social workers and state and local boards of health with coordinated teams with mental expertise, would constitute a mobile crisis program.[110] The study found that more than half of the emergencies handled by a mobile crises team were managed without psychiatric hospitalization of the person in crisis, whereas a little more than one quarter of the emergencies handled by the police resulted in psychiatric hospitalization.[111]

Scott also found that "The average cost per case was 23% less for persons served by the mobile crisis team."[112] The average cost per case in using a mobile crisis program service was $1,520, including $455 for program costs and $1,065 for psychiatric hospitalization.[113] In comparison, regular police intervention, on average, cost $1,963 with $73 for police services and $1,890 for psychiatric hospitalization.[114]

---

[108] Roger L. Scott, *Evaluation of a Mobile Crisis Program: Effectiveness, Efficiency, and Consumer Satisfaction*, 51 PSYCHIATRIC SERVS. 1153, 1153 (2000).

[109] SUBSTANCE ABUSE & MENTAL HEALTH SERVS. ADMIN., *supra* note 103, at 10 (quoting MICHAEL H. ALLEN ET AL., REPORT AND RECOMMENDATIONS REGARDING PSYCHIATRIC EMERGENCY AND CRISIS SERVICES: A REVIEW AND MODEL PROGRAM DESCRIPTIONS 53 (2002)).

[110] Discussed *supra* Part IV.c.iv.

[111] Scott, *supra* note 108, at 1153.

[112] *Id.*

[113] *Id.* at 1155; *see also* SUBSTANCE ABUSE & MENTAL HEALTH SERVS. ADMIN., *supra* note 103, at 15.

[114] Scott, *supra* note 108, at 1155; *see also* SUBSTANCE ABUSE & MENTAL HEALTH SERVS. ADMIN., *supra* note 103, at 15.

While requiring conversational assistants to be subject to mandatory reporting laws would effectuate an economic cost to the public or the consumer, the ultimate goal of preserving the health and lives of youths and adolescents is substantial. It is difficult to place a cost on health; however, this Note would argue that the cost of the proposed primary and secondary enforcement mechanisms is minimal in comparison to the ultimate goal of the proposed public health intervention.

VI.   Conclusion

Too often regulation in the United States avoids the precautionary principle—regulators wait until there is a serious problem with hard statistics as evidence to regulate. When conversational assistants have such a large chance of causing potential harm to consumers' health, including fatalities, society cannot simply wait for all the evidence. Requiring mandatory reporting under Massachusetts and other states' laws for these devices is only a small step towards closing the gap on this public health problem. When devices have the capability to act as medical professionals, they need to share the responsibilities and duties as well. It is understood that requiring either first responders as the primary enforcement mechanism or involvement of social workers and state and local boards of health with coordinated teams on the mental health side as secondary enforcement mechanisms is coercive. However, assigning the duty to report for these devices provides additional and necessary protections for consumers who own and use these products.

Assigning the duty to report for these devices, and requiring either first responders as the primary enforcement mechanism or involvement of social workers and state and local boards of health with coordinated teams on the mental health side as secondary enforcement mechanisms, provides additional and necessary protections for consumers who own and use these products.

Mandatory reporting of conversational assistants will not "cure" youth and adolescent suicide and suicide risk. It will not stop innovation or stop more device makers from seeking patents that can diagnose users' mental and physical states based on the sounds of their voices. Mandatory reporting will require companies to think about their responsibilities to the consumer and hopefully ensure that safer products are placed into the market for consumer use.