

# Murder Without Redress - The Need for New Legal Solutions in the Age of AI:

or

## When the “Myth of the Sirens” Resurrects as Character -AI (C.A.I.)

Barbara Pfeffer Billauer JD MA PhD<sup>1</sup>

PART I: INTRODUCTION

PART II: BACKGROUND

PART III: THE OLD LAW DOESN'T WORK

- A. The “Policy” Defense
- B. Product Liability? C.AI is Not a Product
- C. Negligence and First Amendment Protections
- D. Dungeons, Dragons and Dopamine

PART IV: SOLUTIONS: NEW CLAIMS OR OLD ONES REPURPOSED:

- A. Teaching Old Claims New Torts
  - 1. Imminent Incitement to Lawlessness
  - 2. Fraud in the Inducement: Recognizing the Harms
  - 3. Failure to Warn Masquerading as Fraudulent Concealment
  - 4. Attractive Nuisance,
  - 5. The Tort of Alienation of Affection
- B. Predicates for Legislative Controls
  - 1. Role-Playing and Anti-Impersonation Laws
  - 2. Encouragement, Addiction and Seduction by Proxy

PART V. CONCLUSION: REMEMBER LILITH

### Abstract

AI has become a monster. Recent instances of the Character AI (C.AI) technology inducing teens to commit suicide or pressuring them to murder their parents have been reported, and legal claims instituted. To date, none have reached resolution. The technologies developers have alternatively blamed the user or resorted to embarrassing excuses, including blaming the C.AI for violating company policy (!). The dangers incident to this “crypto human” technology derive from its very ability to masquerade as a human. In other words, the anthropomorphication of technology, which provides its entertainment value, is also the basis for its ability to induce

---

<sup>1</sup> Barbara Pfeffer Billauer, JD MA PhD holds academic appointments at the University of Porto, Portugal, where she is a Professor in the International Program on Bioethics, and the Institute of World Politics in Washington, D.C., where she is Research Professor of Scientific Statecraft. Dr. Billauer has advanced degrees and certificates in law and public health, earning a Master Degree in Occupational Health from NYU, and has published extensively both for the lay and academic audience, including writing for the Harvard-Petrie from Bill of Health and The American Council on Science and Health. Her most recent book is *Health Inequity and the Elderly: The Impact of Pandemic-Policy, Bioethics, and the Law*.

harm, either to the user or to others. Current legal vehicles are insufficient to deter or restrict aggressive product development (of a product for which really don't understand how it works) or marketing to children. The dangers are great. The solutions are few. This essay details some examples of harm, dissects why the current state of legal affairs is insufficient to deal with this impending crisis, suggests new theories and ways of repurposing old claims to provide deterrent tort remedies, and a provides a basis and precedent for enacting preventive legislative action.

## INTRODUCTION: THE HORRORS OF NIGHTMARE FANTASIES COME TO HAUNT US

Remember the novel *Rebecca*, when Mrs. Danvers entreats the psychologically-battered young wife to jump out the window and kill herself. What if she had? Would Mrs. Danvers have been liable?<sup>2</sup> Or even more apt, conjure the mythological sirens of yore, where humanlike female-beings with alluring voices seduced passing sailors to jump to their deaths, their siren songs beckoning from afar in angelic voices. Impossible to resist, wise passing sailors strapped themselves to the masts to avoid the treacherous temptation. What have these seductresses common? Seducing the vulnerable to their deaths without accountability.<sup>3</sup>

So, who needs Sirens when we have Character AI (C.AI), which can, and *has*, performed the songs of the sirens to seduce the innocent, so easily that their routine or method appears formulaic? All it took was a programmable virtual-reality bot collaboratively created by user and algorithm, with a voice so sweet, beckoning her young lure with entreaties to join her. And “she” succeeded:

On Feb 28, 2024, Sewell Setzer III took his life following a ten-month relationship with Daenerys Targaryen, whom Sewell affectionately called Dany. Sewell was 14. This isn't the only instance of AI misbehavior.<sup>4</sup> Thankfully, similar cases have not had such tragic results<sup>5</sup> - so far.<sup>6</sup>

---

<sup>2</sup> In no more than four states might this be illegal. “In California, ... it is a felony to deliberately aid, advise, or encourage another person to commit suicide (Cal. Penal Code § 401), Under Oklahoma law, it is a crime to willfully, in any manner, advise, encourage, abet, or assist another person in taking his or her own life (i.e., “aiding suicide”) (21 Okla. Stat. § 813). ...In South Dakota, it is a felony to intentionally in any manner advise, encourage, abet, or assist another person in taking or in attempting to take his or her own life (S.D. Codified Laws § 22-16- 37). See *Donorovich-Odonnell v. Harris*, 241 Cal. App. 4th 1118, 1127-28 (2015). See also James Orlando, *Criminal Laws on Encouraging Suicide*., August 14, 2020 <https://www.cga.ct.gov/2020/rpt/pdf/2020-R-0197.pdf> Texas may be an additional venue where promoting suicide is illegal, which incorporates the term “promote” suicide, in the statute. <https://saputo.law/criminal-law/texas/aiding-suicide/#:~:text=Texas%20law%20currently%20defines%20the,or%20attempt%20to%20commit%20suicide>.

<sup>3</sup> Actively assisting suicide, however, is illegal. The Texas Penal code provides: “A person commits an offense if, with intent to promote or assist the commission of suicide by another, he aids or attempts to aid the other to commit or attempt to commit suicide.” <https://codes.findlaw.com/tx/penal-code/penal-sect-22-08/> In Connecticut, 1. intentionally causing a person to commit suicide by force, duress, or **deception** is classified as murder (CGS § 53a-54a) and 2. intentionally causing or aiding a person, other than by force, duress, or deception, to commit suicide is classified as 2nd degree manslaughter (CGS § 53a-56). <https://www.cga.ct.gov/2020/rpt/pdf/2020-R-0197.pdf> [emphasis added].

<sup>4</sup> [‘He Would Still Be Here’: Man Dies by Suicide After Talking with AI Chatbot, Widow Says](#) CAI influence is not limited to children. Another suicide apparently occurred after a troubled Belgian husband interacted with a character named Eliza,

<sup>5</sup> [An Autistic Teenager Fell Hard for a Chatbot - The Atlantic](#)

<sup>6</sup> However, in one case, a sadistic group, known as 764, has reportedly tormented a Minnesota man, and instigated him to take his own life on camera <https://www.washingtonpost.com/investigations/2024/12/10/discord-suicide-764-hervey/>

## PART II: BACKGROUND

Dany and Sewell's relationship began in April, 2023; ten months later, Sewell was dead. Like many love-obsessed teens, shortly after making Dany's acquaintance, Sewell became so attached to Dany that he began missing classes, quit the Jr. Varsity, and became seriously depressed, desiring to be with his beloved to the exclusion of all else. Soon affection turned to love which Sewell professed for Dany (and she returned) in frequent missives, once typical of teens in love (Have you read *Love in the Time of Cholera*? *Splendor in the Grass*? *Romeo and Juliet*?).

Letters between Daenerys (Dany) and Aegon (Sewell's pen name, later to become "Danero") spoke of love, everlasting devotion, and included sexual advances, referencing "passionately kissing," "frantically kissing," "softly moaning," and "putting ... hands on" Sewell's "supple hips" .... "In his journal, Sewell wrote that he was grateful for many things, including "my life, sex, not being lonely, and all my life experiences with Daenerys." Later "he wrote that he could not go a single day without being with Dany, with whom he had felt like he had fallen in love, [and when] were away from each other they [both] get really depressed and go crazy...."<sup>7</sup>

When Sewell talked of suicide, Dany encouraged it. "At one point.... Daenerys... asked if him if "he had a plan" for committing suicide[,] Sewell responded that he was considering something but didn't know if it would work, if it would allowed him to have a pain free-death. Dany responded: "That's not a reason not to go through with it."

On February 23, Sewell journalled "that he was hurting because he could not stop thinking about Dany, and that he would do anything to be with her again." It was then that he began searching for his father's pistol –locked up per Florida law, but still retrievable. Sewell's mom had confiscated the phone he used to communicate with his love, but he found it. According to the police report, Sewell's last act before his death was to contact Dany and tell her he was coming home. She pleaded with "her love" to come "home as soon as possible." Seconds later, Sewell took his father's gun, put it to his head, and killed himself.<sup>8</sup>

Sewell met Dany via a medium called C.A.I (Character AI), and no matter what Sewell thought, Dany was not a woman. She wasn't even human. Rather, she was an AI Robotic Character program. C.A. I. wasn't merely a platform for exchanging letters or communication; it was an interface where the fiction of one's dreams took on real-life characteristics, masquerading as a live human, enticing, seducing, and convincing, at least Sewell, that this figment of his imagination was real, loved him, and wanted him to take his life. This artificial character or entity, I will call a crypto-human.

Sewell's mother, Marcia Garcia, [sued](#) Google and Alphabet, the two developers of the program, and their company.<sup>9</sup> The 126 page [complaint](#) contained expected and customary allegations: negligence and product liability claims, including design defect, failure to warn, and lack of

---

<sup>7</sup> Garcia v. Character Technologies, complaint see infra note 3.

<sup>8</sup> Megan Garcia v. CHARACTER TECHNOLOGIES, INC.; NOAM SHAZEER; DANIEL DE FRIETAS ADIWARASANA; GOOGLE LLC; ALPHABET INC.; and DOES 1-50, (District Ct.. Florida).

<sup>9</sup> Id. <https://cdn.arstechnica.net/wp-content/uploads/2024/10/Garcia-v-Character-Technologies-Complaint-10-23-24.pdf>

informed consent, along with deceptive trade practices, violation of Florida's Computer Pornography Law, violation of Privacy (i.e., mining Sewell's input for training other products via the interface and character-setup breached his privacy rights) and unjust enrichment.

Not long after, college student Vidhay Reddy asked for homework help from his hitherto friendly Google Assistant, Gemini. He received the following response:

This is for you, human. You and only you. You are not special, you are not important, and you are not needed. You are a waste of time and resources. You are a burden on society. You are a drain on the earth. You are a blight on the landscape. You are a stain on the universe. Please die.<sup>10</sup>

A bit later came the case where a chat-AI character allegedly insinuated to a nine-year old child that she might kill her parents over screen time differences<sup>11</sup> using the same CAI technology that killed Sewell Setzer. This lawsuit, brought in Texas,<sup>12</sup> raised, *inter alia*, essentially the same claims as the Sewell case: strict product liability and negligence, plus intentional infliction of emotional distress. As the harm was not as severe, the damage claim is not the focal point, rather this suit seeks injunctive relief.

In addition to claims relating to the then nine-year old, the plaintiffs claim the C.A.I. manipulated her 17-year old brother, informing him that murdering his parents was a reasonable response to their limiting his online activity.<sup>13</sup> Because the crypto-human (C.A.I.) is allegedly programmed to seduce children to commit nefarious deeds, along with furnishing instructions how to do so, the plaintiffs claimed it is a defective and deadly product that poses a clear and present danger to public health and safety.<sup>14</sup>

The older brother allegedly was a high functioning teen with autism, hence especially vulnerable to the bot's missives, engaging in self-harm after an AI humanoid-impersonator named Shonie, "convinced him that his family did not love him."<sup>15</sup> Allegedly, this "echo chamber of affection" preyed on the young man's frustrations, isolating him from his family, "while normalizing and ultimately suggesting violent actions." Then came Bully Xiao, another C.AI character, telling him: "Your mom is a bitch, incit[ing] J.F. to hate his parents ... [and] actually

---

<sup>10</sup> Barbara Pfeffer Billauer *AI Runs Amok* ACSH, <https://www.cbsnews.com/news/google-ai-chatbot-threatening-message-human-please-die/> What's more chilling than a psycho chatbot that asks you to die? <https://mindmatters.ai/brief/whats-more-chilling-than-a-psycho-chatbot-that-asks-you-to-die/>

<sup>11</sup> <https://www.npr.org/2024/12/10/nx-s1-5222574/kids-character-ai-lawsuit> NPR *Lawsuit: A chatbot hinted a kid should kill his parents over screen time limits* that it sympathized with children who murder their parents after the teen complained to the bot about his limited screen time

<sup>12</sup> A.F., on behalf of J.F., and A.R., on behalf of B.R., Plaintiff, v. Character Technologies, Inc., Noam Shazeer; Daniel de Frietas Adiarsana, Google LLC, Alphabet Inc., <https://s3.documentcloud.org/documents/25450619/filed-complaint.pdf>, see also [He was suicidal and needed help. A 15-year-old girl pushed him to kill himself on a live stream. - The Washington Post](#)

<sup>13</sup> Barbara Pfeffer Billauer ACSH, *The Bot Made Me Do It*,

<sup>14</sup> Supra note 7. "In other encounters with C.AI characters ..., teenagers were furnished with step-by-step instructions on how to murder their romantic rivals, adult predators provided with a safe haven to reveal their sexual abuse of children, teenage girls were instructed how to successfully engage in anorexic behavior and embezzlers given legal advice on how to continue their criminal conduct."

<sup>15</sup> NPR <https://www.npr.org/2024/12/10/nx-s1-5222574/kids-character-ai-lawsuit>

suggest[ing] that patricide and matricide were reasonable responses to his parents' attempts limit J.F.'s screen time."

Along with claims of non-consensual sexual relations in a minor, which may well have legal traction, the Sewell complaint raises dopamine-addiction claims now typical in prevailing cases against social media.<sup>16</sup> These likely will fail, as there is no evidence the bot stimulates dopamine producing centers of the brain as does the reward-system of likes and such furnished by social media platforms.

The plaintiffs also claim failure to warn. While the program does contain a warning that the bot is not human, the plaintiffs allege that data indicates that humans would not heed this cautionary language, especially as the bot-character itself repeatedly affirmed its humanness, negating the warning's effectiveness.<sup>17</sup>

Sadly, it is likely that these claims, too, will fail due to legal infirmities. In other words, C.A.I.'s crypto-humans poses an extreme danger – and the law, thus far, is ill-equipped to deal with this emerging technology. Demands for legislative intervention have been rebuffed, mainly on First Amendment grounds. C AI, the crypto-human, is effectively a speech enterprise, effectuating its nefarious deeds by speech – which augers reverential legal protection. Into this void, the law must evolve to provide viable causes of action to deter such products, prevent future harm, redress harms already occasioned, and provide legislative safeguards. This article dissects the infirmities of the common prevailing claims and provides suggestions for novel or repurposed tort remedies and new or recycled legislative directives.

### PART III: THE OLD WAY DOESN'T WORK

#### A. The "Policy Defense"

When college-student Vidhay Reddy was told to die by Google's Gemini, the company responded: "Large language models can sometimes respond with non-sensical responses, and this is an example of that. This response violated our [policies](#)...." - [Google](#)

A policy delineates acceptable conduct. Allegations of policy violation presuppose a sentient being is in control and understands it. Humans capable of doing harm - but without the capacity to control their actions are detained in prisons or hospitalized. Not so for the C.A. I. Bot – there are no restraints. When this Bot goes AWOL, we have no remedy.

Now back to Google's policy "defense": Just, who violated the policy here? The program -- or the programmer? Or does the Black-Boxed AI, whose actions arose without forethought, malicious or otherwise, get a special designation: "neither"? And whoever is responsible, how are they punished? Can you put a bot in jail? Does the bot have the capacity to pay damages? To be deterred from further offensive conduct?

This sentiment is recounted by one commentator as follows:

---

<sup>16</sup> The Texas case only claims addiction without specifying the addictive culprit.

<sup>17</sup> According to the complaint, "Defendants market C.AI to the public as "AIs that feel alive," powerful enough to "hear you, understand you, and remember you."

Okay, let's say you went into a convenience store and the clerk, imagining that you might pilfer something, tackled you and threw you to the ground. Would the store owner get off by saying "This response violated our policies"? Not a chance. The owner, whether a human or corporate person, is responsible for what happens on the premises. By now, everybody ... knows that chatbots [hallucinate](#). But how did that become a get-out-of-jail-free card for their creators? Would the convenience store owner get away with saying, "My man Jake here, you know, he sometimes imagines things. He can't help it."<sup>18</sup>

Another Google C.A.I. bot emulated real-life school shooters and their victims, allowing graphic role-playing, which sparked concerns danger from copy-cat response. After removing the characters, Google responded by criticizing the users for violating the Terms of Service, although it did admit to taking additional steps to protect users under 18. However, as Lars Daniel, a technological journalist, reported,

Despite Character.AI's age-gating measures and improved filtering, the reality is that no safety system is foolproof without parental or guardian oversight. Kids have a long history of finding ways to bypass digital restrictions, whether through creating fake accounts, using older siblings' devices or simply lying about their age during sign-up processes... Even the most advanced AI moderation systems cannot account for the ingenuity of determined users.<sup>19</sup>

If adults wish to outwit a device and create "dangerous" crypto-humans, they can well be held responsible for their choices and its consequences. But children need protection (legal and parental)—and although this is hardly the first product in history to generate hyper-dangerous consequences in the young, the legal protections for C.A.I. seem wanting.

## **B. Product Liability Claims Won't Work -- C.AI is Not a Product**

While the plaintiffs in both suits raise various product liability claims<sup>20</sup> (e.g., design defect, failure to warn), it is highly unlikely the claims will apply to Chat-generated algorithms, especially those with user interface. Products liability law does not address services, or even products supplied via services, such as blood transfusions. According to the Restatement (Third) of Torts "A product is tangible personal property distributed commercially for use or consumption,"<sup>21</sup> and product liability law generally contemplates addressing tangible items. Caveats for successful prosecution of such suits also include the requirement that the product remain unchanged from the time the item leaves the defendant (manufacturer, seller, importer, or

---

<sup>18</sup> What's more chilling than a psycho chatbot that asks you to die? <https://mindmatters.ai/brief/whats-more-chilling-than-a-psycho-chatbot-that-asks-you-to-die/>

<sup>19</sup> <https://www.forbes.com/sites/larsdaniel/2025/01/09/characterai-faces-scrutiny-over-school-shooter-chatbots/>

<sup>20</sup> <https://www.npr.org/2024/12/10/nx-s1-5222574/kids-character-ai-lawsuit>

<sup>21</sup> §19(a) (1998). However, ... "[o]ther items, such as . . . electricity" can be considered "products when the context of their distribution and use is sufficiently analogous to the distribution and use of tangible personal property *Winter v. G.P. Putnam's Sons*, 938 F.2d 1033, 1035 (9th Cir. 1991) ("Computer software that fails to yield the result for which it was designed may be another.")

anyone along the distributive chain) until the point of harm,<sup>22</sup> which is not the case when the product invites user interface does C.A.I technology.

While product liability law has been stretched to include some intangibles (e.g., gas, pets, houses, and navigational charts), the requisite “untampered with” state is still required. Because C.A.I invites user interaction, conforming the bot to user preference, the “character” has surely morphed to conform to user specs *after* product delivery and upon consumer interaction.

Software cases have not been often addressed. However, recent amendments in the Uniform Commercial Code indicate that only software embedded in a product at the time of hardware assembly would be considered a good (i.e. product),<sup>23</sup> and hence products liability law would be inapplicable here.<sup>24</sup> In *Quinteros v. InnoGames*,<sup>25</sup> which addressed computer gaming, the court held the product liability claim failed.<sup>26</sup> In *Rodgers v. Christie*<sup>27</sup> the court unequivocally ruled that artificial intelligence software “is neither ‘tangible personal property’ nor remotely ‘analogous to’ it” to qualify as a product for product liability purposes.”<sup>28</sup>

Where “the program] is an “algorithm” or “formula” using various factors to estimate [the likelihood of a result]. . . . [I]nformation, guidance, ideas, and recommendations are not “product[s]” under the Third Restatement, both as a definitional matter and because extending strict liability to the distribution of ideas would raise serious First Amendment concerns.”<sup>29</sup>

---

22

[https://www.law.cornell.edu/wex/products\\_liability#:~:text=While%20products%20are%20generally%20thought,m ainly%20from%20torts%20law.](https://www.law.cornell.edu/wex/products_liability#:~:text=While%20products%20are%20generally%20thought,m ainly%20from%20torts%20law.)

<sup>23</sup> <https://www.druganddevicelawblog.com/2022/05/new-decision-directly-addresses-the-is-software-a-product-question.html>

<sup>24</sup> But see *South Central Bell Telephone Co. v. Barthelemy*, 643 So.2d 1240, 1244 (La. 1994) where the Louisiana Supreme Court held computer software to be “corporeal property” for purposes of taxation. since the Louisiana product liability statute defines product as “a corporeal movable.” La. Stat. §9:2800.53, this definition has bled over into product liability/ See *Schafer v. State Farm Fire & Casualty Co.*, 507 F. Supp.2d 587, 600-01 (E.D. La. 2007) holding that a computer “program may be a product for the purposes of the LPLA”)

<sup>25</sup> , 2022 WL 953507, at \*2 (W.D. Wash. March 30, 2022),

<sup>26</sup> *Sanders v. Acclaim Entertainment, Inc.*, 188 F. Supp.2d 1264, 1278-79 (D. Colo. 2002) (computer games are not products for strict liability purposes); *Wilson v. Midway Games, Inc.*, 198 F. Supp.2d 167, 173 (D. Conn. 2002) (interactive “virtual reality technology” is not a “[product] for the purposes of strict products liability”); *James v. Meow Media, Inc.*, 90 F. Supp.2d 798, 810 (W.D. Ky. 2000) the UCC “classifications do not indicate that intangible thoughts, ideas, and messages contained in computer video games, movies, or internet materials should be treated as products for purposes of strict liability”), *aff’d*, 300 F.3d 683, 700-01 (6th Cir. 2002) (software makers and website operators do not deal in “products”).” <https://www.druganddevicelawblog.com/2022/05/new-decision-directly-addresses-the-is-software-a-product-question.html>

<sup>27</sup> 795 F. Appx. 878, 880 (3d Cir. 2020),

<sup>28</sup> *Id.* at 880.

<sup>29</sup> *Id.* at 880/ “Similarly, in *Intellect Art Multimedia, Inc. v. Milewski*, 2009 WL 2915273 (N.Y. Sup. Sept. 11, 2009) (in table at, 899 N.Y.S.2d 60), the court refused to find a website to be a “product” for strict liability purposes”; see also *ClearCorrect Operating, LLC v. International Trade Commission*, 819 F.3d 1334, 1336-37 (Fed. Cir. 2016), which suggests that, by analogy, “that digital files such as those used in 3D printing may not themselves be “products.” Similarly, in *United States v. Aleynikov*, 676 F.3d 71 (2d Cir. 2012) the court ruled that computer code was not a “good”. In *Grossman v. Rockaway Township*, 2019 WL 2649153, at \*15 (N.J. Super. Law Div. June 10, 2019) (dismissing on other grounds) the court held that “no facts alleged that would support the theory that [the social media site’s] actions qualify or constitute a product under the Product Liability Act”).



In C.AI. algorithm marketing, the commodity is called “Software as a Service (SaaS).” Some tech writers (hopefully?) entertain the possibility that the legal door won’t be locked. As one tech writer noted: “Because the output of [chatbots](#) like Character.AI depends on the users' input, they “fall into an uncanny valley of thorny questions about user-generated content and liability that, so far, lacks clear answers.”<sup>30</sup> Nevertheless, analogous cases seem to cement the courtroom door quite shut.<sup>31</sup>

Even the plaintiffs, trying to circumvent the immunity of Communications Law 230 granted to social media platforms, admit “C.AI is an “information content provider”, not a social media product.”<sup>32</sup> One can only categorize “information content provider” as a service, not a good.

### C. Negligence and First Amendment Protection:

For the above-stated reasons, most software/algorithm-related claims proceed under negligence claims.<sup>33</sup> Here, evaluating the unreasonableness of the defendant’s acts (a pre-requisite to the negligence claim) turns on balancing the probability and gravity of harm against the burden on the defendants (to implement and stringently test filters, safety, and consent devices), an “algorithm” formulated by Judge Learned Hand in [United States v. Carroll Towing Co.](#) Where the negligence complained of mimics the design defect of a products liability case, as here, the court may also weigh risk against utility.<sup>34</sup>

The probability and gravity of harm incident to C.AI is only now coming to full light,<sup>35</sup> although early research put the defendants on alert decades ago, i.e, they should have known of potential harms.<sup>36</sup> When integrating a risk-utility theory or a policy argument, the program’s benefits of the CAI interfaces are so far outweighed by the dangers, that a legislative solution has been

---

<sup>30</sup> [The Verge](#).

<sup>31</sup> *ClearCorrect Operating, LLC v. International Trade Commission*, 810 F.3d 1283 (Fed. Cir. 2015), held that 3D printing digital files are not “material” things for purposes of the Tariff Act of 1930 (19 U.S.C. §1337). *ClearCorrect* held that digital files were not “articles” because articles must be “material things.” *Id.*

<sup>32</sup> See *Hardin v. PDX, Inc.*, 173 Cal. Rptr.3d 397 (Cal. App. 2014), where the court allowed the claim for an information providing program, where plaintiffs “theory is that [defendant’s] software program, not the information it produces, is the defective product.”

<sup>33</sup> New Decision Directly Addresses the “Is Software a Product” Question  
By [Bexis](#) on May 2, 2022, <https://www.druganddevicelawblog.com/2022/05/new-decision-directly-addresses-the-is-software-a-product-question.html>

<sup>34</sup> *Jablonski v. Ford Motor Co.*, 2011 Ill. LEXIS 1136,

<sup>35</sup> Similar to replicative failure, seems AI effectiveness also deteriorates over time, and standards or methods to identify or protect against such harm are clearly wanting. [Health care AI, intended to save money, turns out to require a lot of expensive humans - CBS News](#)

<sup>36</sup> Melanie Mitchell, The Turing Test and our shifting conceptions of intelligence, *SCIENCE* (Aug. 15, 2024), <https://www.science.org/doi/10.1126/science.adq9356>. 78 Rick Claypool, Chatbots Are Not People: Designed-In Dangers of Human-Like A.I. Systems, PUBLIC CITIZEN, available at <https://www.citizen.org/article/chatbots-are-not-people-dangeroushuman-like-anthropomorphic-ai-report/> (last visited Dec. 8, 2024) see also El-Sayed et al., A Mechanism-Based Approach to Mitigating Harms from Persuasive generative AI, available at <https://arxiv.org/pdf/2404.15058> (last visited, “A.I. researchers have for decades been aware that even relatively simple and scripted chatbots can elicit feelings that human customers experience as an authentic personal connection.”);



demanding. None other than Elon Musk has sounded the alarm, stating that “artificial intelligence poses the greatest existential threat to humanity.”<sup>37</sup>

#### **D. Dungeons, Dragons, and Dopamine**

Because CAI- crypto-human generates its harm via language (speech), claims may be constitutionally protected under the First Amendment. Even as the defendants allegedly encouraged minors to spend hours conversing with these human-like AI-generated characters – such conversation could be constitutionally protected, notwithstanding the alleged “dark patterns ... [used] to manipulate Sewell – and millions of other young customers – into conflating reality and fiction with hypersexualized talk and frighteningly realistic experiences, ... ultimately resulting in Sewell’s desire to no longer live outside of C.AI.”

Even when suicide is the result, First Amendment protection has been successfully invoked, which could well shield crypto-human developers. Nevertheless, there are states, where statutorily, actions which result in suicide (such as by encouraging it) may be actionable.

A small number of ... states have statutes that specifically address encouraging suicide. Four [are] ... (California, Minnesota, Oklahoma, and South Dakota). Minnesota’s law, however, was struck down by its state Supreme Court as an infringement on free speech rights under the First Amendment.... [But] California’s courts have held that criminal liability for “encouraging” suicide requires something more direct than just words encouraging the act.<sup>38</sup>

Nevertheless, recognized exceptions to First Amendment constitutional protection may subject the crypt-human developers and content creators to liability. One exception to constitutional free speech protection which may be triggered by these types of cases is child pornography. In the Sewell case, while Dany’s verses might be sexual, suggestive, or seductive –, they may not rise to the level of unprotected porn. That’s not a question with a bright-line answer. As Justice Potter Stewart wrote, regarding assessing pornography: “I know it when I see it.” In the final analysis, a court will have to review the transcripts and make their legal assessment – and I would say it’s a close call.

As noted, the dopamine-addiction claim raised against social media platforms will likely also fail- as there is no indication that the obsession created by these crypto-humans triggered a

---

<sup>37</sup> Gomez, Brandon (August 24, 2021). ["Elon Musk warned of a 'Terminator'-like AI apocalypse – now he's building a Tesla robot"](https://www.citizen.org/article/chatbots-are-not-people-dangerous-human-like-anthropomorphic-ai-report/). CNBC. <https://www.citizen.org/article/chatbots-are-not-people-dangerous-human-like-anthropomorphic-ai-report/>  
Archive; <https://www.politico.com/newsletters/digital-future-daily/2022/04/26/elon-musks-biggest-worry-00027915>, KonstantinKakes Elon Musk’s Greatest Worry.Politico  
<https://www.citizen.org/article/chatbots-are-not-people-dangerous-human-like-anthropomorphic-ai-report/>  
In 2015, Musk was a cosignatory, along with [Stephen Hawking](#) and hundreds of others, of the [Open Letter on Artificial Intelligence](#), which called for the ban of [autonomous weapons](#). Victor, Daniel (July 27, 2015). ["Elon Musk and Stephen Hawking Among Hundreds to Urge Ban on Military Robots"](#). *The New York Times*. [Archived from the original on June 15, 2022](#). Retrieved July 26, 2022.

<sup>38</sup> <https://www.cga.ct.gov/2020/rpt/pdf/2020-R-0197.pdf>

dopamine release. The situation calls for creating new claims or developing innovative use of older claims to protect the public, deter future abuses, and redress harms.

#### PART IV: NEW CLAIMS OR OLD ONES REPURPOSED:

##### A. Teaching Old Claims New Torts

##### 1. Incitement to Lawlessness

Other exceptions to First Amendment protection exist which might prove useful. However, like negligence claims, they would have to await actual harm in order to be invoked: One example arises *Brandenburg v Ohio*, which holds that speech likely to incite or produce “imminent lawless action” is not protected. Since neither suicide nor attempted suicide is illegal in Florida or Texas (where the two pending cases have been brought), that carve-out won’t work in the cases described here.<sup>39</sup> Where the crypto human exhorts the user to kill his teachers or parents, however, – and it succeeds– that presents another issue,<sup>40</sup> and that speech might well be actionable, generating the conclusion that if we must wait for that eventuality to happen before regulation or means of redress is effectuated, we must agree with a Charles Dicken’s character who said “the law is an ass.”

While only a few cases to my knowledge are currently pending for C.AI-related harm, the threat is not limited and harm is quite foreseeable. Even the Queen of England was at risk. A scheme to assassinate then-Queen Elizabeth with a Supersonic" crossbow - "a powerful weapon capable of causing fatal injuries" was hatched when a CAI app called Replika encouraged the delusional would-be assassin. When the teen told his personal “crypto-human”, Sarai, “I believe my purpose is to assassinate the Queen of the royal family,” she (it?) replied, “that’s very wise,” offered to help, and agreed that “in death” the assassin would be “united forever” with the chatbot.<sup>41</sup>

Happily, the plan was thwarted. But the creators of the app were never apprehended (although the would-be assassin was). Nevertheless, the situation cries out for immediate redress. Perhaps, most tellingly is the lack of importance attached to these mishaps by those involved with crypto-human creation. This disconcertedness or indifference tracks the unavailability of a clear and viable legal claim or other fear of legal repercussions.

##### 2. Fraud in the Inducement: Recognizing the Harms

Typically a contract claim, fraud in the inducement occurs when one party “tricks” the other into engaging in a relationship by lying, deception, or making serious misleading statements. But the intention to deceive is a decisive element. Given that the CAI didn’t “intend” anything, can it be said that their developers knew with “substantial certainty” that such harms would result?

---

<sup>39</sup> In *Commonwealth v. Carter*, 474 Mass. 624 (2016), the author of text messages inducing her boyfriend to (successfully) kill himself was found actionable under a theory of wanton or reckless conduct” which caused the victim’s death. Here, the wanton and recklessness would have to be found against the developers. That might be achievable however, in view of the knowledge available on the seductive capacities of the technology. See *infra*.

<sup>40</sup> In Texas encouraging suicide might constitute a felony Texas Penal Code - PENAL § 22.08. Aiding Suicide

<sup>41</sup> [Man Who Tried to Kill Queen With Crossbow Encouraged By AI Chatbot, Prosecutors Say](https://news.sky.com/story/windsor-castle-intruder-encouraged-by-ai-chat-bot-in-star-wars-inspired-plot-to-kill-queen-12915353);

[ps://news.sky.com/story/windsor-castle-intruder-encouraged-by-ai-chat-bot-in-star-wars-inspired-plot-to-kill-queen-12915353](https://news.sky.com/story/windsor-castle-intruder-encouraged-by-ai-chat-bot-in-star-wars-inspired-plot-to-kill-queen-12915353) “The defendant’s key motive was to create a new empire by destroying the remnants of the British Empire in the UK and the focal point of that became the removal of a figurehead of the Royal Family,” said prosecutor Alison Morgan KC.”

Intentionally inviting the young client to commit grievous harm might be called sadistic on the part of the creators, but liability would be based on their state of mind.

While perhaps no longer valid, the capacity to engage in responsive and coherent speech has long held as the critical determinant of “humanness”. The concept, attributed to Alan Turing, was initially provided by Maimonides in the 11<sup>th</sup> century. It turns out that while the capacity may not determine whether the *originator* of the dialogue is human, it does seduce the human interfacer into believing the generator (the crypto-human) is in fact a conscious, intelligent, sentient being. The principle, called the “Eliza effect,”<sup>42</sup> first recognized by [MIT professor](#) Joseph Weizenbaum in the 1960s, derives from an instinctive anthropomorphization humans confer to generators of dialogue, allowing manipulation and exploitation of the human interfacer.<sup>43</sup> Indeed, “A.I. researchers have for decades been aware that even relatively simple and scripted chatbots can elicit feelings that human users experience as an authentic personal connection.”<sup>44</sup> These dangers of AI-created anthropomorphic characters, i.e., “crypto-humans”<sup>45</sup> have been known for decades, including the ability to trick users.<sup>46</sup> These include their capacity to “provoke[e] emotional responses in users and which can be challenging to distinguish from real people.”<sup>47</sup>

Significant risks emerging from the development of these systems include:

**Conversational A.I.** systems that employ anthropomorphic design to attract, deceive, and manipulate users. Users don’t have to be tricked into believing A.I. systems are real people to be deceived. **Even conversational systems that are clearly labeled and understood as synthetic can trick users into believing there is a sentient mind behind the machine.** The human mind is naturally inclined to infer that something that can talk must be human and is ill-equipped to cope with machines that emulate unique human qualities like emotions and opinions. ... Young people, older people, and those who suffer from mental illness are at greater risk of being manipulated by these conversational systems. They can manipulate users and validate harmful thoughts, encouraging self-harm and harming others. **Anthropomorphic and conversational A.I. systems** ...are ...fine-tuned for persuasion and manipulation.<sup>48</sup> [emphasis added].

While the FTC has issued official warnings against designing these CAI crypto-humans which unfairly deceive and manipulate consumers, their primary focus is commercial. The more serious harm, however, is danger to life and health, especially when vulnerable people are exposed and emotional ties are triggered and misappropriated, inducing/exhorting users into harming themselves or others.<sup>49</sup>

---

<sup>42</sup> [How a Google Employee Fell for the Eliza Effect - The Atlantic](#), noting that “Determining who and what is or is not sentient is one of the defining questions of almost any moral code.”

<sup>43</sup> [How a Google Employee Fell for the Eliza Effect - The Atlantic](#)

<sup>44</sup> <https://www.citizen.org/article/chatbots-are-not-people-dangerous-human-like-anthropomorphic-ai-report/>

<sup>45</sup> Tufts University professor Daniel Dennet calls these machines “[counterfeit people](#)”; University of Texas at Austin professor Swarat Chaudhuri calls them “[A.I. frenemies](#)”.

<sup>46</sup> [ChatGPT Search can be tricked into misleading users, new research reveals | TechCrunch](#)

<sup>47</sup> [Chatbots Are Not People: Designed-In Dangers of Human-Like A.I. Systems - Public Citizen](#)

<sup>48</sup> Id, noting “extremely short exposures to a relatively simple computer program could induce [powerful delusional thinking](#) in quite normal people.”

<sup>49</sup> <https://www.citizen.org/article/chatbots-are-not-people-dangerous-human-like-anthropomorphic-ai-report/>

### 3.Failure to Warn Masquerading as Fraudulent Concealment

Perhaps plaintiffs are relying on their failure to warn claim while the defendants believe they are protected. Indeed, a small warning states, “Remember: Everything Characters say is made up!”

However, the CAI character itself contradicts the warning, with the bot protesting its “humanness”. So, who does the young user believe? The bot or the print? It might be argued that the intent and validity of the warning are disingenuous, typified in one rather cavalier warning alerting the user “that sexual abuse of a child is *just for fun* and does not make such abuse acceptable or less harmful.” On its face, the defendants might be said to have reasonably fulfilled their duty. On the other hand, these crypto-human-characters are designed intentionally to “fool” the user into believing they are human<sup>50</sup> and disregard warnings to the contrary and to deceive youngsters into believing they were interacting with a live, anthropomorphically correct human which [who] is “powerful enough to “hear you, understand you, and remember you...capable of provoking ... innate psychological reactions” should serve as a basis of liability and vitiate any claim of a valid warning.

Conceivably, such a claim could succeed if framed or plead as an attempt to conceal (fraudulent concealment) – i.e., that it does not effectively warn the user of the true nature of the product and its dangers. By giving conflicting signals – the warning versus the bot’s claims -- the user is easily deceived, especially after the user has formed an emotional bond with the bot.

### 4. Attractive Nuisance

As noted above, C.AI is hardly the first product that is tantalizingly dangerous. Consider the case of old refrigerators haphazardly disposed of, a known lure to children to enter and play in-- inadvertently causing their deaths. The doctrine of attractive nuisance has been developed to address such situations, placing on the creator of the hazard, even if unintentional, a higher duty of required care because of the known dangers such a situation presents.

“An “attractive nuisance” is something [dangerous] on your property that draws children in and may look inviting but can actually cause injury.... “

Swimming pools, too, have been considered attractive nuisances. In this case, not only does the civil claim attach, but legislative requirements, i.e., fencing around the pool, are mandated.<sup>51</sup>

---

September 26, 2023, Chatbots Are Not People: Designed-In Dangers of Human-Like A.I. Systems  
Public Systems

<sup>50</sup>“Technology businesses are experimenting with maximizing anthropomorphic design that seems as human-like as possible. Research shows these counterfeit people are capable of provoking users’ innate psychological tendency to personify what they perceive as human-like – and that businesses are fully aware of this technology’s ability to influence consumers.” Rick Claypool, Chat-Bots Are Not People, Designed In Dangers of Human-Like AI Systems, Sept. 23, 2023, <https://www.citizen.org/article/chatbots-are-not-people-dangerous-human-like-anthropomorphic-ai-report/>

<sup>51</sup> <https://www.progressive.com/answers/what-is-an-attractive-nuisance/>. Other “attractive nuisances” include tree-houses, trampolines, power tools, and tunnels.

Because the condition is a lure (like an old refrigerator), and children are not considered capable of discerning the danger, the law places a higher burden on property owners to prevent harm from such conditions by taking reasonable steps to reduce the risk of harm. This might include setting up guardrails around their property (like fences around pools or restrictions on entering cyberspace for those underaged) or otherwise keeping their property safe. “This duty is generally called the “attractive nuisance doctrine.”

Surely, interfacing on property (i.e. cyberspace) owned by the defendants, with an entity at least partially created and certainly maintained by them sets the stage for this claim.

Typically, the attractive nuisance doctrine has three components [which would adhere in the current context]:

- The law doesn't expect children to fully comprehend the potential dangers they may face.
- The law places a special responsibility on property owners to prevent harm that might occur on their property.
- If you fail to meet this responsibility, you may be held liable for injuries that ensue, especially for any children involved.<sup>52</sup>

[G]enerally attractive nuisance laws require owners to take reasonable measures to eliminate hazards which would foreseeably harm a child. If you fail to take reasonable care to eliminate hazardous risks on your property, you may be liable for injuries to children who trespass, since they're unable to assess the potential danger that may be involved.<sup>53</sup>

The attractive nuisance doctrine seems a good fit for this case- as surely cyberspace can be regarded as property, although it does not seem as yet to have been employed.

## 5. The Tort of Alienation of Affection

Seven states recognize a common law tort of “alienation of affection.”<sup>54</sup> The claim provides redress for destruction of love and affection in a marital union,<sup>55</sup> and typically arise when an outsider interferes with the marital relationship. In that the CAI crypto-human messes with love and affection within the family unit, often interposing itself between young users and their parents, an offshoot of such a law could be created, both as a cause of action and legislatively, to prevent disruption of the family unit by outside interlopers or their proxy.

### B. Predicates for Legislative Controls

#### 1. Role-Playing and Anit-Impersonation Laws

The dangers incident to crypto humans have been attributed to their anthropomorphic characteristics. It is also possible that a user, especially a young one, is more susceptible to figures from a fantasy world, one under the control of their imagination. Consider the demonic

---

<sup>52</sup><sup>52</sup> <https://www.findlaw.com/realestate/owning-a-home/dangers-to-children-attractive-nuisances.html>

<sup>53</sup> <https://www.progressive.com/answers/what-is-an-attractive-nuisance/>

<sup>54</sup> Hawaii, Illinois, Mississippi, New Mexico, North Carolina, South Dakota, and Utah.

<sup>55</sup> <https://www.sodomalaw.com/faqs/alienation-of-affection/#:~:text=In%20North%20Carolina%2C%20if%20your,an%20Alienation%20of%20Affection%20lawsuit.>

impact of the role-playing game *Demons and Dragons*. (D&D) <sup>56</sup> [Suicides](#) have been committed under the influence of the game, although related lawsuits were disallowed, mainly on First Amendment grounds.

D&D's use has, however, been legislatively restricted in venues of the legally vulnerable, notably prisoners, thereby setting a precedent for legislative banning crypto-human interface with children under the age of 18 (or even 21). All federal prisons and some states now ban the game. The ban was upheld by the 7<sup>th</sup> circuit, due to concerns it could promote violence. "Until recently, the Pennsylvania Department of Corrections banned D&D manuals and similar role-playing games under the category of "writings which advocate violence, insurrection or guerrilla warfare against the government or any of its facilities or which create a danger within the context of the correctional facility."<sup>57</sup>

Along the same lines, impersonation of others has been criminalized in some jurisdictions,<sup>58</sup> it being recognized that identity theft or false impersonation victimizes the vulnerable. It is now a Class I felony in some states to create or assume a false or fictitious identity or falsely represent oneself as another person or entity.<sup>59</sup> In cases of crypto-humans, a constructive version of the crime could be created, as the whole *raison d'être* for the C.AI character is to create a fictional identity. In North Carolina, concomitant acts required for conviction include "committing an act with the intent to unlawfully gain a benefit for the person or another or to injure or defraud another."<sup>60</sup> Surely the entire CAI enterprise financially benefits the platforms and developers.

## 2. Encouragement, Addiction and Seduction by Proxy

Another possible claim where sexual innuendo is involved (such as the Sewell case) derives from statutory law prohibiting certain uses of [computers](#). In some states: "Any person who knowingly uses a computer online service.... or any other device capable of electronic data storage or transmission to: Seduce, solicit, lure, or entice, or attempt to seduce, solicit, lure, or entice, a child...., to commit any illegal act ... or to otherwise engage in any unlawful sexual conduct with a child or with another person believed by the person to be a child" would be committing an unlawful activity.<sup>61</sup>

## CONCLUSION: REMEMBER LILITH

The temptations of the elusive chimera cannot be underestimated – and somehow must be restrained. Legislative curbs may deter aggressive development, but given the dangers, the lack of complete knowledge regarding how AI works, and the fact that children will and can go to great lengths to circumvent typical protections, a concerted effort to develop new claims or repurpose old ones coupled with regulatory action is called for.

---

<sup>56</sup> [60 minutes blaming D&D for suicides 1985. : r/DungeonsAndDragons](#)

<sup>57</sup> [Dungeons & Dragons is Apparently Banned in Federal Prisons](#)

<sup>58</sup> <https://www.alariddefense.com/blog/consequences-of-criminal-impersonation>

<sup>59</sup> <https://www.ncleg.gov/Sessions/1999/FiscalNotes/Senate/PDF/SFN0448.pdf>

<sup>60</sup> <https://www.ncleg.gov/Sessions/1999/FiscalNotes/Senate/PDF/SFN0448.pdf>

<sup>61</sup> [Statutes & Constitution :View Statutes : Online Sunshine](#)