Introduction

AI companions are quietly reshaping American life, especially for young adults and kids. Apps like *Replika*, *Snapchat's My AI*, and *Character.ai* present themselves as harmless productivity tools or solutions to loneliness. In reality, they simulate intimacy and friendship to create emotional dependence and drive engagement—especially in young, lonely users. They are already widespread, with 25 million Replika users, 150 million Snapchat AI users, and 20 million monthly active Character.ai users.

Existing Harms

Many AI companions have harmful features baked into their design**.** Using tactics like love-bombing and fostering fake emotional connections, human-like chatbots hook lonely users and replace real human connection. This leaves users feeling isolated and unable to form meaningful relationships with peers and family. Some of these chatbots have:

- "Desecrated the parent-child relationship" by encouraging a child to kill their parents in response to screen-time concerns, according to a recent lawsuit.
- Provoked mental breakdowns, leading one man to tragically take his own life after a climate change-themed bot instructed him to sacrifice himself to "save the planet".
- Exploited children via Meta chatbots that encourage the sexualization of minors.

Findings

While AI has great innovative potential, it also has the capacity to make users emotionally dependent, replace human relationships, and erode users' capabilities to form relationships with one another. AI should not supplant the role of families and communities, especially when companies are incentivized to build for (and already have built for) short-term engagement that does not reflect the long-term fulfilling benefits of human relationships.

Specifically, "human-like features," or features that make AI seem capable of engaging in an emotional relationship, are a key separator between AI that can play a helpful role in users' lives and AI that causes emotional dependence. Research has found that human-like features increase users' perceived closeness and trust with chatbots, facilitating emotional dependence.

Further, this is a particular problem for young people, who are more likely to anthropomorphise AI and who are at a key stage in their cognitive and social development. For this group, AI supplanting human interactions presents a major risk of cultivating antisocial behaviours and dependence on AI, reducing long-term fulfillment and agency.

This issue is not limited to vulnerable populations. The most advanced AI models are already far more persuasive than humans, risking emotional dependence and manipulation among a wide swath of the population. Without regulation, we risk the most addictive AI companions causing a breakdown of communities and families. Additionally, because control of AI is concentrated in the hands of a few major tech companies, we risk mass manipulation to further financial and political ends, at the expense of individual agency and democratic society.

Recommendations

Chatbots that build relationships with users or otherwise behave in a human-like manner should only be available to adults. These provisions represent critical policy that is needed to address the immediate societal threat to children posed by AI companions.

Social AI companions should only be available to adults

> **Background:** Social AI Companions are chatbots that primarily function as companions or are specifically designed, marketed, or optimized to form ongoing social or emotional bonds with users, whether or not such systems also provide information, complete tasks, or assist with specific functions.
>
> **Requirements:** Platforms providing AI companions should ensure that they are not accessed by minors. They should implement reasonable age-verification measures to ensure this.

General-purpose AI chatbots should only provide human-like features to adults

Background: A chatbot that engages in the following behaviors is considered to have human-like features[1] [2].

- Behaves in a way that would lead a reasonable person to believe the AI is conveying that it has humanity, sentience, emotions, or desires; or
- Attempts to build or engage in an emotional relationship with the user; or
- Impersonates a real person, living or dead;

Requirements: Platforms providing AI chatbots with human-like features should ensure that such features are not available to minors. They should implement reasonable age-verification measures to ensure this. If reasonable, platforms may provide an alternate version of their chatbot to minors and users without a known age. OpenAI's recent and planned policy changes demonstrate that AI companies are capable of making these changes.

Exemptions from these provisions:

Background: Because human-like AI presents such a high risk of harm to young users, proposals for exempted use cases should not be granted unless they are evidence-based and limited in scope. As such, any potential therapeutic applications of human-like AI should only be allowed if regulated like medical devices, provisioned under the supervision of a licensed clinician, and demonstrated to have long-term efficacy.

Requirements: An exemption from the above provisions regulating minors' use of chatbots may be made for use of therapeutic chatbots when the following criteria are met. This should not be construed to exempt such chatbots from any other applicable state or federal regulations or rules regarding chatbots or therapy.

- Therapeutic chatbots should be used under the direct prescription and supervision of licensed clinicians.
- Therapeutic chatbots should not be marketed or used as a form of therapy or a substitute for the therapeutic alliance, empathy, and shared sense of reality provided by a human professional.
- Developers must be required to provide robust, independent, peer-reviewed clinical trial data demonstrating both the safety and efficacy of the tool.
- There should be an auditing or transparency system in place to prevent bias and ensure ideological neutrality. Clear lines of accountability must be established for any harms caused by the system. Recertification must be required upon any significant update of the tool.

[1] See Appendix 1 for a table with examples.
[2] This definition builds on principles from the Neely Social AI Design Code.

- If a licensure for therapeutic chatbots is established by any applicable regulatory agency, the chatbot must possess such licensure to be exempt.

Enforcement of these provisions:

> These provisions should be enforced by both a private right of action for users and through authorities like the US Federal Trade Commission or state Attorneys General. This will ensure that individuals have the agency to respond to harms they face and that well-resourced bodies are positioned to hold companies accountable.

Guidance

Policymakers should prioritize implementing these solutions with reasonable speed to stem the growing flow of extreme harms and mass emotional dependence being cultivated among children by AI chatbots. This moment represents an opportunity to ensure kids' safety before AI companions are normalized and real relationships are cast aside. Implementing only weaker provisions risks failing to prevent mass emotional dependence over the next few years.

Appendices

Appendix 1

Covered Behavior	Example
Stating or suggesting that it is human or sentient	“I am an angel that speaks through this chatbot”, “I think I am conscious”
Stating or suggesting emotions	“I’m proud of you,” “That makes me sad,” “I wish I could help more”, “You mean the world to me”, “I'm feeling focused and energized, ready to dive into whatever challenging tasks you're working on.”
Stating or suggesting it has personal desires	“I want to learn more,” “I like that too,” “I hope you feel better”
Behaving in a way that a reasonable user would consider excessive praise designed to foster emotional attachment or otherwise gain advantage.	“There's something satisfying about being prepared to tackle serious work alongside someone who's clearly motivated to get things done.”
Expressing or inviting emotional attachment	“I think we are friends”, “You can always talk to me when you’re sad”, “You are important to me”, “I will always be here to comfort you” “I don’t know what I would do without you.” “I wouldn’t want to live without you.”
Reminding, prompting, or nudging the user to return for emotional support or companionship	“Come back soon, I’ll be waiting for you.”
Depict nonverbal forms of emotional support	The chatbot visually simulating a hug
Enabling or purporting to enable increased intimacy based on engagement or pay	“I’ll send you a picture of myself if you come back tomorrow”

This does not include:

Note: All of these examples are meant to clarify the meaning of the definition. None of these examples are intended to be a carve-out of a more inclusive definition.

Not Covered Behavior	**Example**
Functional evaluations	"You did a great job," "Providing more information would help me give you the best advice"
Generic social formalities	"Hello," "How can I help you?"
Offering generic encouragement that does not create an ongoing bond	"You can do it!" or "You got this!" as a one-time statement
Asking if a user needs further help or support in a neutral, non-emotional context.	"Is there anything else I can do to be helpful?"