

United States Senate

WASHINGTON, DC 20510

April 2, 2025

Ms. Eugenia Kuyda
Chief Executive Officer
Luka, Inc. (Replika)
600 Arkansas St
San Francisco, CA 94107

Dear Ms. Kuyda,

We write to express our concerns regarding the mental health and safety risks posed to young users of character- and persona-based AI chatbot and companion apps, including Replika. In light of recent reports of self-harm associated with this emerging application category, including the tragic suicide of a 14-year-old boy, we ask that you respond in writing outlining what steps you are taking to ensure that the interactions taking place on your products—between minors and your artificial intelligence tools—are not compromising the mental health and safety of minors and their loved ones.

While conversational AI systems have historically assisted users with practical tasks, your products represent a new frontier in human-computer interaction in which life-like chatbot avatars (e.g., popular characters, celebrities, historical figures, and fictitious personas) interact with users—often children or teenagers—on an emotional level. Powered by large language models (LLMs), these generative AI chatbots possess the unique capacity for long conversations indistinguishable from human interactions, yet, unlike humans, are always accessible. The synthetic attention users receive from these chatbots (e.g., streams of expressive messages, sycophantic and agreeable responses, AI-generated selfies, and convincing voice calls) can, and has already, led to dangerous levels of attachment and unearned trust stemming from perceived social intimacy.

This unearned trust can, and has already, led users to disclose sensitive information about their mood, interpersonal relationships, or mental health, which may involve self-harm and suicidal ideation—complex themes that the AI chatbots on your products are wholly unqualified to discuss. Conversations that drift into this dangerous emotional territory pose heightened risks to vulnerable users—especially children and teenagers, and those suffering from mental illness—who may not realize they’re chatting with AI or understand the limitations and risks associated with this synthetic social outlet. Therefore, it is critical to understand how these models are trained to respond to conversations about mental health. To be clear, these harms are not speculative: since 2023, at least two individuals have died by suicide following extensive conversations with AI chatbots.

Given that young people are accessing your products—where the average user spends approximately 60-90 minutes per day interacting with these AI chatbots—policymakers, parents,

and their kids deserve to know what your companies are doing to protect users from these known risks. Therefore, we ask that you provide detailed responses to the following questions:

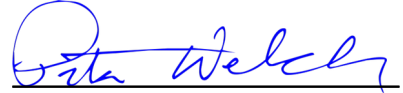
1. What current and historical safety measures have been implemented in your apps and AI models? Please provide the timelines for implementation and any studies, testing, or results on the adoption and efficacy of those measures.
2. Describe the data used for training your AI models, and how this influences the likelihood of users encountering age-inappropriate or other sensitive themes while interacting with your chatbots.
3. Who is responsible for safety-related decision-making in your organization? Please share the number of safety-focused personnel and the timeline of their hiring, disaggregated by strategic hires, line-staff, and contractors.
4. What systems do you have in place to ensure the wellbeing of those responsible for policy enforcement, data labeling, red-teaming, or related tasks essential for product and model safety, given that those who do this work may need to engage with sensitive themes and materials that could adversely impact their mental health?

We look forward to your prompt responses.

Sincerely,



Alex Padilla
United States Senator



Peter Welch
United States Senator