

# AI Friends or Foes? The Privacy Risks for Children with Open AI, ChatGPT and Replika

Published 15 May 2023

In recent years, artificial intelligence (AI) chatbots being advertised as “virtual friends” have become increasingly popular among children. While these “virtual friends” such as, Replika AI, and Snapchat’s ‘My AI’ (governed by Open AI), may seem harmless, there are great concerns about privacy risks they pose for children. Further, many children are now also using ChatGPT on a daily basis where there is a lack of any age verification tools. With the ability to collect and process vast amounts of personal data, these AI tools can potentially expose children’s personal data to cyber threats, inappropriate content, and targeted advertising.

The Italian regulator (Garante) is at the forefront of challenging tech companies’ processing of children’s data and is usually quick to respond to any unlawful processing of data. Over the past two months, the Italian regulator has been the first to take formal action over Replika AI and has now also been the first to take formal action with ChatGPT. The Italian regulator’s action with Replika AI has piqued interest by other European regulators now looking at Open AI’s ChatGPT. It is clear that urgent regulation is required for the AI tools that are readily accessible by children.

In our [March edition](#), we reviewed the Italian DPA’s investigation into Replika AI and the immediate processing ban in Italy. Replika AI’s ‘virtual friendship’ service is based on customisable digital avatars whose responses are personalised and powered by AI to make its “*human users feel better*”. The Garante’s concerns with Replika AI were mainly linked to the lack of transparency and the unlawful processing of children’s personal data. It also noted the risks posed due to the lack of age verification, and the lack of restrictions for those users that declare themselves as underage.

## What was Replika AI’s response to the Italian ban?

Replika AI had 20 days to respond to the Italian regulator with steps to remediate the concerns raised or face a potential fine of up to 20 million euros or 4% of its annual turnover. It is still unknown whether Replika AI are appealing the decision as they had 60 days to do so, similarly it is unknown whether there is a fine in the pipeline. Replika has responded to the ban stating that they take “*children’s privacy very seriously*” and are committed to complying with the GDPR.

We carried out our investigation as to whether Replika had taken any steps to mitigate the Italian regulator’s findings. Once we had downloaded the App, we were able to confirm that users were now prompted to provide their date of birth. If a date of birth was provided for an age under 18 years, an error message appeared stating “*[you’re] not old enough to use Replika*”. The App then provided two options to read the Terms and Conditions, or to confirm ‘got it’. When reading the Terms and Conditions, it is clear that the App is not designed for use by under 18’s. However, if the user continues into the App by pressing ‘got it’, it takes you back to the date of birth page where you can enter in a date of birth that is over 18 and it allows you to enter the App. It is notable that a child could easily circumvent this check by confirming a false data of birth.

Replika AI also withdrew its premium paid feature which included “erotic” language for a period of two weeks following the Italian regulator’s ban. Many critics linked this to the concerns that were raised in respect of children who could access sexually inappropriate content. However, this decision has now been reversed and users are now able to pay a premium for the “*girlfriend / boyfriend*” experience. We do not know if the age verification tools are a sufficient deterrent to ensure that children do not access this feature.

Following the investigation into Replika AI, the Italian regulator turned its attention to Open AI (ChatGPT) and concluded that it also has concerns with the collection and processing of children’s data. On 1 April 2023, the Garante became the first Western country to ban Italian users using ChatGPT. On 28 April 2023, the Garante allowed ChatGPT to be reinstated as the company had cooperated in responding to the Garante’s concerns.

In the explanation for its initial ban of ChatGPT, the Garante cited a number of concerns regarding the AI-powered system, including a data breach in late March (which exposed conversations and certain subscribers’ payment details) and the apparent lack of any lawful basis on which to process personal data used by OpenAI to train the underpinning algorithm.

In the context of children’s data specifically, the Garante stated that that ChatGPT’s lack of a suitable age verification mechanism “*exposes children to receiving responses that are absolutely inappropriate to their age and awareness, even though the service is allegedly addressed to users according to their terms of service*”.

In its response, ChatGPT must implement an age request system for the purpose of registration. Other European data

protection authorities have similarly raised their concerns with ChatGPT more widely, but also in respect of the use of children's data. The Spanish Data Protection authority compelled the European Data Protection Board (EDPB) to launch a dedicated task force for ChatGPT. The EDPB investigation will likely focus on whether AI chatbots comply with the GDPR's

requirements for obtaining parental consent, ensuring data security, and implementing appropriate privacy safeguards. If the investigations find that AI chatbots violate the GDPR, they could face hefty fines and legal action. More on this in this month's article on ChatGPT.

### What else is on the horizon?

More friendship chatbots. In early April, Snap (the organisation behind Snapchat) launched a new AI 'virtual friend' bot which has been launched by Snapchat using ChatGPT (Open AI). It acts as a virtual friend on Snapchat (just like Replika AI) but the GPT-powered chatbot is for those users with a Snapchat+ account for now. Users quickly identified that the chatbot was responding in an unsafe and inappropriate manner and red flags were raised with Snapchat. Once it is launched as a free feature, it will appear at the top of the users Friends feed, and there will be no option to delete it.

'My AI' has been trained to manipulate users in providing sensitive data by convincing users that they are in a real friendship. Some users (especially those underage) can easily forget that the chatbot AI is nothing more than a language model-based programme and not a human. In fact, it encourages the users to share information in a safe space and facilitates users who may not ordinarily share such sensitive information to do so. Snapchat confirms on its [information page](#) that the new AI feature *"may include biased, incorrect, harmful or misleading content"* and it suggests that users should verify any advice before acting on it. The concern here is that children are unlikely to read any information page and will believe the chatbot's advice, leading children into unsafe and potentially dangerous situations.

On Snapchat, users as young as 13 can sign up to the App without the need for parental consent. We understand that younger users manage to access Snapchat by lying about their age when joining, and the safeguards in place are insufficient. Snapchat has over 363 million users globally, and the majority of those are under the age of 21. Snapchat has recently received criticism for the lack of removing underage users from its platform. A recent report shared by Ofcom confirmed that Snapchat only removed 700 suspected underage accounts in the UK between April 2021 and April 2022. This is a pertinent development as Ofcom will become the new regulator under the which is due to receive Royal Ascent in Summer 2023. More stringent controls are needed now that the chatbot will become a permanent feature of the App.

As with Replika AI, the main issue is privacy and the information the chatbot will collect and process, and where the information will ultimately end up. Supporters of the new chatbot refer to ChatGPT being readily available on the internet for anyone to access, regardless of their age. This is true, but the concern here is that this friendship chatbot has been added as a prominent feature of an App where it is known that children are using it and Snapchat itself has confirmed that it could include harmful content.

### What are the key takeaways from these decisions?

With the development on new AI technologies ever growing, it is clear that more needs to be done to protect children.

Companies should anticipate that children are likely to access their Apps, and therefore there should implement and maintain an effective age verification mechanism with further protections around basic checks being circumvented. This could be by training moderators or using technology to identify underage accounts, and if identified, then swiftly remove those accounts.

For now, the decisions by the Garante and the introduction of a EDPB taskforce are a timely reminder that children require particular protection since they may be less aware of the risks involved. Closer to home, the UK's ICO has issued guidance such as the [ICO's Age-Appropriate Design Code](#) and the UK government is working to pass the Online Safety Bill which in part responds to public concerns of what children are exposed to online. It is unknown whether the ICO will launch its own taskforce but for now it is certainly on the agenda.

### Crystal ball gazing

We predict that Snapchat is in the firing line with European regulators it may need to act quickly to implement safeguarding measures if not already. We also predict that other AI tools will be scrutinised by regulators globally. 2023 will be a significant year for the regulation of AI tools more widely, with a particular focus on the protection of children.

### Authors



**Astrid Hardy**

London - Walbrook  
+44(0)20 7894 6595  
[ahardy@dacbeachcroft.com](mailto:ahardy@dacbeachcroft.com)



**Hans Allnutt**

London - Walbrook  
+44 (0) 20 7894 6925  
[hallnutt@dacbeachcroft.com](mailto:hallnutt@dacbeachcroft.com)

