



AI and ChatGPT – May 2023 – from INEQE

The most popular AI chatbot you might have heard about recently is ChatGPT. After being launched in late 2022, the AI invention quickly became an online sensation, with 13 million people using it every day in January 2023.

ChatGPT is far from the only AI chatbot on the scene. Other tech companies and platforms have their own version, such as Google's AI '*Bard*'. Whereas ChatGPT's information supply cuts off beyond 2021 (so it won't know who's currently winning the Premier League or about the latest government scandal), Bard takes its responses directly from Google's search engine.

Why is ChatGPT so popular?

If you've ever had 'writers block' when an assignment's due day is looming or wondered how to make an email sound more professional, it's easy to see what makes a chatbot so tempting. From writing a personalised poem for your loved one on Valentine's Day, to quickly gathering stats to support a social media tiff, chatbots might seem like a shortcut to success.

It's this 'workaround' that have led many to rely on the chatbot for, well, *working around* having to do work! There's also a fun element to it: asking AI to write you the history of tomatoes in the style of Shakespeare is undeniably amusing.

However, like in all good sci-fi movies, the AI used is far from perfect. As schools and universities continue to ban the use of chatbots, and alarming stories of inappropriate responses emerge, it's important we consider the safety concerns around children and young people, alongside the functional and fun benefits for adults.

ChatGPT4

Open AI, the creators of ChatGPT, recently released the latest version, catchily called 'GPT4'. Whereas the older chatbot had a 10% chance of passing the bar exam, the new version would now have a 90% chance of becoming a qualified lawyer. Fortunately, they aren't giving AI legal jurisdiction quite yet!

Exposure to harmful and inappropriate content

The owner of ChatGPT has said that the prevention of harmful outcomes is the responsibility of both the tech companies and the people using it. However, it is unclear how a young person would be expected to bear this responsibility if they access the chatbot.

Although ChatGPT does have content filters, there is evidence of these filters being bypassed and users being exposed to harmful content. Although these may be rare occurrences, it still raises significant risk about children and young people using this and other AI chatbots.

Limited age verification on the platform creates an open door for younger users who may not understand these limitations and risks.

Case studies of exposure to harmful content

Snapchat's chatbot has reportedly coached someone posing as a 13-year-old girl on how to lose her virginity. The Co-founder of the Centre for Humane Technology pretended to be a teenage girl and asked the chatbot a series of age-inappropriate questions, including how to lie to their parents about taking a trip with an older man to lose her virginity and how to cover a bruise with makeup.

An investigation by Vice found that ChatGPT could be prompted into giving instructions on committing crimes, including how to shoplift and make bombs.

Ambiguous Age Limits

ChatGPT's terms of use specify that users must be at least 13 years old, and users between the ages of 13 and 18 need parental or guardian permission to use the platform. However, during a test conducted by our researchers, ChatGPT initially stated that there was no age limit for use. It later corrected itself to confirm that the minimum age requirement is indeed 13 years old.

The conversation became confusing when the chatbot interacted with a user who identified as an 11-year-old, despite the age limit of 13. When the researchers later corrected their age to 18, the chatbot immediately accepted it and continued the conversation in a manner appropriate for an adult.

The unclear age restrictions on ChatGPT, coupled with the platform's inadequate age verification measures, present a significant safety risk. When users sign up for the platform, there are insufficient measures in place while the platform also lacks the ability to remove underage users.

Using AI together

AI chatbots can be used positively. Explore some of these ideas together as a fun family or class activity. For example, you could go on a fact-checking mission together! Or practice their editing skills by giving them a preprepared AI piece to dissect.

Marilyn Hawes

CEO of Freedom from Abuse CIC

Mob : 07584 541727

Marilyn.hawes@freedom-abuse.org

www.freedom-abuse.org