What 'Free' AI Chatbots Are Really Costing You

Devi Ganesan, Data Scientist & Artificial Intelligence Practitioner

AI chatbots have surged in popularity recently. Who wouldn't love a cost-effective, always-available buddy that can provide instant responses on almost any topic of interest?

As a regular user of AI chatbots, I believe these tools have the potential to transform our working style and become indispensable for personal assistance, customer service, and business operations. Many companies have recognized this potential and are releasing incredibly useful AI chatbots for free or paid public use, each with its own strengths and specializations such as ChatGPT, Meta AI, Google Bard, Jasper Chat, Perplexity and so on.

Despite their numerous benefits, it is crucial to understand the significant risks related to data privacy and security [1]. Being aware of these risks is essential for protecting your data while leveraging the advantages of AI chatbots.

What are LLMs

When an AI chatbot receives your prompt/question, it not only generates a response but also stores your question in its memory so that it can use that information as context for answering your follow-up questions with more sensible and human-like responses. But why would tech companies let us use such sophisticated chatbots for free? Here comes the age-old wisdom that is hidden in the idiom "there is no free lunch".

Data is the new fuel and Large Language Models (LLMs) that power up these chatbots are voraciously data-savvy. LLMs are a type of deep neural networks that are initially trained using a large corpus of text data, such as web pages, books, online forums, reviews and Wikipedia. This initial training phase can take weeks, months, or even years depending on various factors. Once trained and deployed as conversational chatbots, LLMs are designed to continuously learn from conversations and improve their performance over time. Just as much as you learn from these chatbots, they learn from you too. So, are these chatbots really 'free'? Perhaps it's a win-win for both us and the tech companies.

Why You Need to Be Cautious

Conversations with AI chatbots are anonymised and stored for a particular duration. These conversations may be randomly sampled and used for further training of the LLMs that power up the chatbots. While the tech companies take measures to protect anonymised data, there is always a risk of data breaches or unauthorised access. If a chatbot service is hacked, any stored data—including private conversations—could be compromised. Though it is true that stored data are anonymised, researchers have shown that it is feasible to re-identify individuals using patterns and sophisticated statistical analyses [2].

The leakage of sensitive data can lead to severe consequences, such as identity theft, financial loss, and damage to personal and professional reputations. Malicious actors can exploit your personal details for targeted phishing attacks, scams, or even blackmail. For businesses, the risk is even higher, as proprietary information, trade secrets, and confidential strategies could be exposed, leading to

significant financial and competitive disadvantages. So, beware of the hidden costs behind the ‘free’ chatbots!

Protecting Your Privacy (on the example of ChatGPT)

Be it individual information or business information, it is always better to be vigilant and preventive rather than becoming a victim. Since ChatGPT is widely used, here are some guidelines for protecting your privacy when interacting with it [3][4]:

- Never provide ChatGPT with details like your name, address, phone number, financial information, or login credentials. Stick to general, non-sensitive topics and queries.
- ChatGPT has an incognito mode that prevents your conversations from being used to train its language model. Activate it by going to Settings > Data Controls and turning off the "Improve the model for everyone" toggle.
- Even in incognito mode, ChatGPT stores your conversations for 30 days for abuse monitoring. Periodically clear your chat history by clicking your username and selecting "Clear conversations."
- When creating a ChatGPT account, use a dedicated email address instead of your primary one to prevent linking your conversations to your identity.
- Exercise caution when using third-party apps or plugins that integrate with ChatGPT, as they may have different privacy practices. Review their permissions, policies, and backgrounds before use.
- Do not input confidential information, such as medical records or proprietary data, into ChatGPT, as it may be stored and used for training purposes.
- As a general practice, we should be cautious by making our conversations less indicative of personal details. For example, while using a chatbot for drafting an email, instead of saying, “Draft an email to my boss, Jane Doe, at jane.doe@company.com, about my upcoming leave for a family emergency”, you could say “Draft an email to my supervisor about my upcoming leave for a personal matter”.
- Stay informed about any changes to ChatGPT's privacy policy and data practices, as these may evolve over time.

As AI chatbots become increasingly integrated into our daily lives and business operations, safeguarding our data privacy and security is paramount. By following best practices and being aware of the potential risks, we can enjoy the benefits of these advanced technologies without compromising our personal or business information.

References

1. Generative AI Models: Opportunities and Risks for Industry and Authorities.
2. https://www.edps.europa.eu/system/files/2021-04/21-04-27_aepd-edps_anonymisation_en_5.pdf
3. https://www.mcafee.com/blogs/internet-security/chatgpts-impact-on-privacy-and-how-to-protect-yourself/
4. https://www.perplexity.ai/ and https://chatgpt.com/