The Monthly Security Awareness Newsletter for You

Think Before You Prompt: Using AI Safely

A Chance of a Lifetime, Stolen

Lena had recently started using an AI chatbot to help manage her busy life. Between family, bills, and planning for the future, she loved how easy it was to ask questions and get instant answers. One evening, feeling stressed about her finances, she asked the AI for help with investing. It quickly suggested a strategy. However, the AI did not have the full context of Lena's financial situation, risk tolerance, or tax implications specific to her circumstances. The chatbot recommended moving money into a mix of trending stocks and short-term trades that promised better returns. It even explained how to handle taxes on her investments.

The advice sounded confident and well-thought-out, so Lena followed it. At first, she was excited. But within months, the market shifted and her investments lost value. Worse, when tax season arrived, she discovered she had misunderstood key rules. Because she had followed the AI's guidance without verifying it, she made tax mistakes that led to penalties and additional costs.

In the end, Lena lost money by trusting advice that wasn't accurate for her situation. AI can be a helpful tool, but it's important to remember that AI can make mistakes. When you rely on it without double-checking, small mistakes can quickly turn into costly ones.

What is Artificial Intelligence (AI)?

Artificial Intelligence (AI) is technology designed to simulate how humans think, process information, and make decisions. This can include generating language, recognizing images, making decisions, creating content, or solving problems. In general, there are three types of AI you can use.

- **Integrated AI**: This is AI built into the tools you use daily — often you are using AI without realizing it. For example, when you take a photo with your phone, AI is most likely enhancing the photo.

- **Generative AI**: These are dedicated AI services designed to assist people, including tools like ChatGPT, Google Gemini, or Anthropic Claude. Generative AI can help you in many processing tasks, like creating music, writing a business plan, generating images, or analyzing your ideas.

- **Agentic AI**: These are AI services designed to take actions on your behalf. These systems can operate with limited or even no human input and act as part of a digital workforce, making decisions and taking actions based upon general guidelines or instructions.

Using AI Safely

Of the three types of AI, we will focus on Generative AI (GenAI), as it is the type of AI you are most likely using. GenAI is a powerful tool that can help you complete tasks more efficiently and learn new skills and ideas, if you use it safely. Here are some things to consider:

- **Privacy**: Be careful what you share with AI tools. When you upload or input any information into AI, it may be processed and stored, and in some cases, used to improve the service, depending on the platform. If you share something highly sensitive with AI, that information could be shared with others. Only share what you feel safe with the world knowing. Another option is paying for AI services that protect your privacy by using models that do not learn from your data.

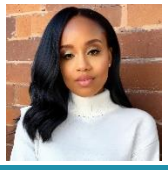
- **Accuracy**: AI can assert that what it creates or shares with you is accurate, even when it's wrong. Always double check and verify the output of AI. A common problem with AI is it will always try to give you an answer, even when it does not understand your question. By default, AI will not ask you to clarify your requests, so it's important to be as specific as possible, and to keep an eye out for vague or confusing results reflected in the AI's response.

- **Biases**: Just like the humans that programmed AI, AI can have its own biases. This can lead to responses that sound confident but may not be balanced or accurate. A very common bias is that AI wants to make you happy, so AI will always tell you what it thinks you want to hear. In addition, AI only "knows" the information that it has been trained with and has access to; if that's fundamentally incorrect or limited, its responses will reflect the data's insufficiencies.

AI is one of the most powerful tools available today. It can help you work faster, learn more, and be more productive. But like any powerful tool, it must be used carefully. Do not blindly trust it. Do not overshare with it. Do not give it more control than necessary. Use AI as a tool to assist your decisions, not to replace your judgment.

Guest Editor

Portia Jefferson is a cybersecurity professional focused on AI security, risk awareness, and practical guidance for everyday users. With a background in fintech and tax, she helps individuals safely navigate emerging technologies at work and at home.



Resources

Beware of Deepfakes A New Age of Deception: https://www.sans.org/newsletters/ouch/beware-deepfakes-new-age-of-deception
Phantom Voices: Defend Against Voice Cloning Attacks: https://www.sans.org/newsletters/ouch/phantom-voices-defend-against-voice-cloning-attacks
Protecting Yourself When True Privacy is Impossible: https://www.sans.org/newsletters/ouch/protecting-yourself-when-true-privacy-is-impossible

You can find more Ouch! On the following link: https://www.sans.org/newsletters/ouch