Have you given away secrets on ChatGPT? Be careful about giving away proprietary details, cautions a professor

April 18 2023, by Rob Nicholls

Credit: Pixabay/CC0 Public Domain

Whether it's mild curiosity or the business imperative of starting to use artificial intelligence (AI), as part of a unique selling proposition, many people now use ChatGPT, either at work or at home.

When it launched, the user-friendly AI large language model went from zero to 100 million users in only 60 days. Since then, Microsoft has invested U.S. $10 billion in startup Open AI, and embedded GPT4 into Bing. Dedicated Bing apps with GPT4 are now available on the App Store and the Play Store.

Part of the user experience of OpenAI's ChatGPT is that AI tools can generate really useful text from a specific prompt, making for possible opportunities to save time in the day-to-day at work with machine learning—for example, when writing emails.

You can even use the job description provided in an advertisement to get ChatGPT to write the "perfect" cover letter. Or a staff member from human resources can generate the "perfect" job advertisement for LinkedIn from the scratchy brief provided by the business unit.

But what are the risks in trying ways ChatGPT can be used at work?

As a starting point, let's look at the example of the job ad.

If the role is one which is commonly advertised, then little is lost by sharing the form of the job description with a couple of hundred million other uses of ChatGPT.

But if the job description includes information that could be used by a competitor to identify your business, then the risks are significantly higher (especially if recruitment is an important part of insider business strategy in your workplace).

With companies like Samsung having recently been stung by staff members inadvertently giving away material via ChatGPT, it is

important to consider the risks carefully before using it at work.

And just like that, the code was gone

This issue is particularly challenging if staff members have started using ChatGPT as part of the code development process.

It's really appealing to do so. One of the great aspects of ChatGPT is its automation use case in reducing coding time in software development projects for programmers. This can be done at a design level ("How do I sort job application letters using Natural Language Processing in Python?") or at the code level ("How do I use Gaussian Naive Bayes in scikit-learn?").

One really useful feature is the ability to paste code in as a prompt to ask ChatGPT to improve it. This is as simple as "What is wrong with the following code?" ChatGPT can even recognize the coding language that you're using.

The problem is that ChatGPT can then include the material that you have used as a prompt to improve its answers in the future, by training the algorithms. This is precisely what happened to Samsung developers who used ChatGPT to both improve code and keep meeting notes.

Material that would have been regarded as some of the most sensitive by Samsung was available to developers outside of Samsung, simply because Samsung engineers used ChatGPT to decrease their development time.

Will GPT4 make using ChatGPT at work riskier?

The latest in the GPT series after GPT3, GPT4 has an incredibly

accurate voice-to-text feature. But the risk of the text becoming part of the training set to improve the generative AI is high.

It's a simple way to both transcribe work meetings, and even better, to generate the minutes before the end of the meeting. However, the transcription will also be part of the GPT4 ecosystem before the end of that meeting. There have been news headlines about Italy's decision to "ban" ChatGPT over privacy concerns.

Essentially, the argument made by Italian authorities is that the data collected by ChatGPT was in breach of the European General Data Protection Regulation. However, consistent with other European countries, it seems likely that Italy will walk back from this approach by the end of April. The only change required will be to have an age verification (over 18) check on users.

Generative AI uses billions of data points in order to be able to create text on a predictive basis. It improves in response to user feedback. The challenge faced by businesses that employ curious people is that this feedback may include company confidential material.

The solution is simple in theory but much harder in practice. If material would not normally be disclosed outside of the business, it should not be used as a prompt for ChatGPT or for Bing.

The practical difficulty with this rule is that search engines, including Google with its generative AI called Bard, are an essential business tool. The issue may be to decide whether search engines are there to provide information or to provide answers.

So, I should avoid using ChatGPT at work?

Not sure what you should (and shouldn't) share with our friend,

ChatGPT? Try this simple test:

Is the output of the ChatGPT session a document that would normally be regarded as confidential by your business? Then it should not be shared on ChatGPT.

If you did write your cover letter or resume using ChatGPT, the AI system used to filter applicants could also run your cover letter through GPTZero. This online tool from Open AI can detect whether text was written by a generative AI by examining that text's "perplexity" (a measurement of the randomness of the text) and "burstiness" (a measurement of the variation in perplexity).

But of course, the improvement in ChatGPT's text output is challenging these tools. So, who is to say how this will change in the future with developing AI technology?

Provided by University of New South Wales

Citation: Have you given away secrets on ChatGPT? Be careful about giving away proprietary details, cautions a professor (2023, April 18) retrieved 23 April 2024 from https://techxplore.com/news/2023-04-secrets-chatgpt-proprietary-cautions-professor.html