

An Interview with Chat GPT

On the risks of AI to Cyber Security



Audience: General



Reading Time: 5 Mins



At the time of writing this article there are a lot of conversations about the risks and benefits of AI and its potential impacts on society. We thought it would be interesting to interview OpenAI's ChatGPT for its own thoughts on the potential risks of AI to Cyber Security.

GMCF: Thank you for being here, ChatGPT. To begin, can you explain some of the risks that AI could pose to cybersecurity?

ChatGPT: Of course. The first major risk is the potential use of AI by malicious actors. Advanced AI can be used to automate hacking attempts, making them faster and more effective. These attacks can use advanced techniques like machine learning to adapt and evolve, making them more difficult to detect and prevent. They can analyse vast amounts of data faster than a human could, identify patterns and weak points, and exploit them.

AI can also be used for social engineering attacks, such as generating convincing phishing emails or impersonating trusted individuals. Traditionally, phishing attacks might have relied on broad, scattergun approaches or manually crafted spear-phishing emails. However, with AI, it

becomes possible to automate the process of personalizing phishing emails or messages, making them much more convincing. An AI could analyse large amounts of data about an individual, their social media posts, their company's public details, and then craft a phishing email that is tailored to them and thus more likely to be successful.

Furthermore, there is the risk of adversarial attacks [manipulating AI system's data inputs to trick it into behaving unexpectedly or undesirably] on AI systems themselves. Adversarial attacks aim to exploit vulnerabilities in the way an AI processes information. For example, tiny, almost imperceptible alterations to an image can cause an AI image recognition system to completely misinterpret it. The same principles could be applied to attacks on AI systems used for cybersecurity, causing them to miss actual threats.



European Union
European Regional
Development Fund

The Cyber Foundry project is part funded by the European Regional Development Fund



**Manchester
Metropolitan
University**



**University of
Salford
MANCHESTER**

Lastly, there is the issue of AI's dependency on data. AI needs vast amounts of data to learn and function, and this data often includes sensitive information. The storage and processing of this data can create opportunities for data breaches if not properly secured.

GMCF: Thank you for that detailed answer. Considering these risks, how can organisations protect themselves?

ChatGPT: There are several ways to mitigate these risks. For the risk of AI-powered attacks, organizations should continually update their security systems to detect and prevent them. They can also use AI in their defence, with systems that can learn and adapt to new threats.

Adversarial attacks can be mitigated by implementing defences into AI models such as adversarial training, where the model is trained with adversarial examples to improve its robustness.

“

For the risk of AI-powered attacks, organizations should continually update their security systems to detect and prevent them.

”

Regarding data security, organizations need strong data governance policies. This includes anonymizing data when possible, encrypting sensitive data, and controlling access to data.

Moreover, the principle of least privilege should be applied, which means each user should have the minimum levels of access necessary to perform their tasks.

GMCF: Okay – so, what about the other side of the coin? Can AI be used to enhance cybersecurity defences?

ChatGPT: Absolutely. Just as AI can be used to carry out sophisticated attacks, it can also be used to develop advanced defences. For instance, AI can help automate the process of threat detection and response, freeing up valuable time for human analysts to focus on strategic decision-making and complex incident response tasks. It can also help in predicting attacks, identifying anomalies, and enhancing user behaviour analytics.



GMCF: Surely all this would be very expensive for SMEs to adopt?

ChatGPT: Adopting robust AI-powered cybersecurity measures can indeed require significant resources, which might be challenging for small to medium enterprises (SMEs). However, there are some strategies and tools that can make it more accessible and affordable:

- **Managed Security Services:** SMEs can outsource their cyber security needs to managed security service providers (MSSPs) that specialize in providing enterprise-grade security at a lower cost.
- **Shared Threat Intelligence:** By participating in cybersecurity communities or subscribing to threat intelligence feeds, SMEs can benefit from the collective knowledge

and resources of larger entities. This can include AI-generated threat intelligence.

- **Scalable Solutions:** Many cyber security solutions, including AI-based ones, are scalable. This means that SMEs can start small and increase their investment as their business grows.
- **Cloud-Based AI Solutions:** With the advent of cloud technology, many AI-powered cybersecurity solutions are delivered via the cloud. This can significantly reduce costs associated with infrastructure and maintenance.
- **Cyber Insurance:** While not a preventative measure, cyber insurance can protect SMEs from the financial impacts of cyber-attacks.
- **Cybersecurity Training:** Human error is a significant factor in many cyber-attacks. Training staff on basic cybersecurity practices can be a cost-effective way to improve an SME's security posture.



Finally, while there is a cost associated with implementing cybersecurity measures, the cost of not doing so can be far greater.

A cyber-attack can result in lost revenue, recovery costs, regulatory fines, and damage to an SME's reputation. As such, investment in cybersecurity, including AI-powered solutions, should be seen as a necessity rather than a luxury.

GMCF: That's insightful. And lastly, what's your vision on the future of AI in cybersecurity?

ChatGPT: AI will continue to play an increasingly significant role in both cybersecurity threats and defences. We will see more sophisticated AI-powered attacks, but also more advanced AI-driven defences. We should expect the development of autonomous response systems capable of detecting and neutralizing threats in real-time.

Ethical AI use will become a focal point as well, with more emphasis on transparency, accountability, and privacy. Lastly, the intersection of quantum computing and AI could revolutionize cybersecurity, potentially creating new vulnerabilities but also ushering in powerful new defence mechanisms.

Now, it is essential to remember that AI is a tool, and like any tool, its impact depends on how it is used. This is why a combination of technical advances, thoughtful policies, and ethical practices is crucial for navigating the future of AI in cybersecurity.

GMCF: Thank you ChatGPT.

--- End

This writer must confess that as a human used to communicating with other humans using text format, I very soon started to forget that this was indeed an artificial intelligence, or to be technical, a natural language processing tool, at the other end of my probing.

Overall an interesting take on both the potential risks that AI might cause, and the benefits that AI might bring in the realm of cyber security. Both the risks and the potential benefits that were mentioned are quite plausible, and yet there are also many unknowns as to where this technology will take us.

What is clear is that cyber security will continue to grow as a highly important factor as society continues to innovate and adopt new digital technologies.