Bug bounty and AI: How machine learning is changing the game for cybersecurity

BY ANNA HAMMOND · DECEMBER 22, 2022 · LAST UPDATED ON MARCH 6, 2025

AI presents some fresh opportunities to the bug bounty industry, but caveats apply

You would be hard-pressed to find anyone in the cybersecurity industry who had not heard of ChatGPT over the past few weeks.

Launched by research lab OpenAI as a prototype in November, ChatGPT is an artificial intelligence (AI) chatbot that quickly garnered attention for its detailed responses across many knowledge domains.

Since its beta launch, ChatGPT has enjoyed a seemingly unending list of applications. Poetry, film manuscripts, and architectural design are all included in the discoveries of what this tool can create.

RELATED Bug Bytes #185 – ChatGPT, ChatGPT, and more ChatGPT

As we detailed in our recent Bug Bytes newsletter, the bug bounty and ethical hacking community has already been exploring the possibilities on offer from ChatGPT, from its ability to generate code and bypass security filters.

Of course, this isn't the first time that AI has been associated with cybersecurity. Along with 'blockchain' and 'military-grade encryption', the terms 'AI' and 'machine learning' are frequent buzzwords seen at security conferences, supposedly powering everything from incident response to threat intelligence platforms.

However, it's perhaps the simplicity and raw power of ChatGPT that has turned heads and reignited the discussion around cybersecurity and AI. But how will AI impact the bug bounty industry? Is it really the game-changer many are claiming it to be?

How will AI impact the bug bounty industry?

ChatGPT has shown to be adept in several cybersecurity functions, many of which directly impact the bug bounty industry. This includes creating custom code for vulnerability scanning and assisting with writing penetration test reports, white papers, and vulnerability disclosure communications.

By automating many of the tasks involved in finding and identifying security vulnerabilities, AI may be able to help bug bounty hunters more quickly and effectively locate potential issues. This could save time and effort, and allow bounty hunters to focus on the most critical issues.

AI's impact on the bug bounty industry is likely to be complex and varied

However, it is also worth noting that the use of AI in the security bug bounty industry may also lead to increased competition, as more people and organizations may be able to enter the field. This could make it more difficult for individual bug bounty hunters to stand out and earn rewards.

Moreover, as is always the case with so-called 'dual-purpose' applications, tools used for defensive purposes can also be leveraged by attackers. In the case of ChatGPT, it's not hard to see how such a powerful tool being used to create proof-of-concept exploit code or power highly effective phishing campaigns.

We spoke to Ambuj Agrawal, author of 'Enterprise Automation with Python', who underlined this inherent contradiction between AI and cybersecurity:

> **"Models similar to ChatGPT can structure personalized messages especially for simulated phishing to train employees on how messages can be constructed for phishing attacks. ChatGPT can also answer questions by employees related to security attacks and increase awareness of different information security attacks using its dialog-based user interface.**
>
> **Simulated phishing typically involves sending mock phishing emails to users and asking them to identify the phishing attempts. If a user falls for the simulated phish and clicks on a link or enters their login information, they are provided with feedback and guidance on how to avoid falling for real phishing attempts in the future. Simulated phishing is an effective way to train users to be more vigilant and protect themselves against phishing attacks.**
>
> **The main challenge with AI in infosec is that it is available to malicious hackers as well, and they can create more advanced attacks compared to traditional systems. AI-based adversarial models are a good example of cases where hackers can use AI to come up with new ways of hacking into organizations using AI. "**

The future of AI and cybersecurity

If one thing is certain, AI presents both an opportunity and a potential threat to the security industry. In the future, new exploits – perhaps even entire attack classes – may be created or refined by AI, but the same machine learning engines could also be used to protect networks.

We spoke to Laurent Ach, CTO of Qwant, a privacy-focused search engine, about the current progression in AI and where it will lead with security:

> **"The recent results [with ChatGPT] are so astounding some people suspect their business will never be like before. AI, in general, is certainly changing society but its limitations are often largely overestimated, part of the hype being caused by anthropomorphizing how texts (and images) are generated.**
>
> **In the domain of information security, a few use cases are possible: Large Language Models (LLM) can be used to automatically create phishing content or harmful code, or, at the opposite, to create phishing drills and help secure software development.**
>
> **Overall, LLMs constitute the recent most visible and impressive trend in machine learning and will bring a new layer of automation in many useful tools. However, it's important to understand that they are by design not able to bring anything new beyond a kind of smart interpolation of information. "**

The long-term impact of AI on the security bug bounty industry is likely to be complex and varied. The potential benefits that AI could bring to the security arena are huge, but infosec practitioners must also be aware of the risks and take steps to mitigate them.

The launch of ChatGPT has marked a step change in the public's perception of AI and how it might be the next big disruptor. But for now, at least, it's perhaps best to think of AI and cybersecurity in a similar way to AI in self-driving cars – it's still a good idea to keep hold of the steering wheel and make sure you're doing the driving.

It is likely that AI will play a significant role in the security bug bounty industry in the future. By using AI and machine learning algorithms, companies will be able to automate many of the tasks involved in finding and identifying security vulnerabilities, allowing them to more quickly and effectively identify and fix these issues. This could make the bug bounty process more efficient and allow companies to better protect their systems and data from security threats. Additionally, AI may be able to help companies more accurately assess the severity of a potential security vulnerability and prioritize the fixing of the most critical issues.

Still unsure about the potency of AI? The previous paragraph was written by ChatGPT.

Additional reporting by Sasha Burnside.