# O avanço e os perigos da Inteligência Artificial: o ChatGPT e a cibersegurança

dinheirovivo.pt/opiniao/o-avanco-e-os-perigos-da-inteligencia-artificial-o-chatgpt-e-a-ciberseguranca-16114526.html

3 de abril de 2023

A Inteligência Artificial (IA) é um dos campos de desenvolvimento tecnológico mais importantes da atualidade. Como muitos referem, a Inteligência Artificial pode ser aplicada para melhorar a qualidade de vida de todos os seres humanos, em vários aspetos.

Por meio de computadores, robôs, dispositivos móveis e outros meios tecnológicos, a Inteligência Artificial pode ser usada para ajudar a resolver problemas humanos e aumentar a produtividade, tornando o mundo mais conectado.

Além disso, e mesmo sendo este um dos campos mais controversos, a Inteligência Artificial pode ser também utilizada para ajudar na redução da criminalidade, aumentando a segurança pública. Será também uma aliada na melhoria da educação, na saúde e na assistência social, bem como útil na melhoria dos sistemas de transporte, tornando-os mais seguros e eficientes.

No entanto, face ao desenvolvimento atual que temos assistido da Inteligência Artificial, esta levanta-nos sérios dilemas. Não só éticos e morais, mas também securitários.

Veja-se que, recentemente, mais de mil especialistas ligados à IA e empresários de renome das tecnologias - incluindo Elon Musk -, mas não só, assinaram uma carta que pede "uma pausa de seis meses no desenvolvimento de sistemas gigantes de Inteligência Artificial". Os signatários argumentam que é necessária esta pausa para que "os potenciais riscos à segurança sejam estudados e controlados".



Isto levanta-nos várias questões tais como: que problemas poderemos encontrar, que ameaçam a segurança das nossas comunidades, à medida que desenvolvemos a Inteligência Artificial para o benefício da sociedade?

O tema é muito amplo e complexo. Por isso, vamos tentar desconstruí-lo, focando-o. Concentremo-nos apenas num caso - que, diga-se, é o mais paradigmático: o "novo" *chatbot* da Open-AI, o ChatGPT. E foquemo-nos apenas na dimensão que a VisionWare opera: a cibersegurança. Assim, quais são os riscos do ChatGPT para a segurança cibernética?

A emergência da tecnologia da IA foi sempre recebida com um certo ceticismo e incerteza, e não é difícil perceber porquê. Quando apresentada com uma forma de tecnologia tão avançada que pode fazer o seu próprio pensamento, temos de dar um passo atrás

necessário mergulhando diretamente para dentro. Embora tornando as nossas vidas muito mais fáceis e ágeis em muitos aspetos, a tecnologia da IA que possuímos hoje e continuamos a melhorar pode ter consequências terríveis para o futuro da cibersegurança - daí a existência do *malware* ChatGPT.

Falámos sobre os riscos da utilização do ChatGPT e que um programa melhorado como este pode ser perigoso nas mãos erradas. O programa de IA pode escrever código instantaneamente e de acordo com dados recentes, o ChatGPT também pode elaborar um programa malicioso bastante convincente. O *malware* é basicamente código malicioso. Muitas redes subterrâneas na *dark web* já levaram à utilização do *chatbot* para eliminar *malware* e facilitar ataques de *ransomware*.

Estas preocupações são ainda mais preocupantes e prementes, quando os gigantes da indústria estão dispostos a investir fortemente em tecnologia de IA. A Microsoft alargou recentemente a sua parceria com a OpenAI num investimento multibilionário para "acelerar a descoberta de IA".

O *malware* gerado por IA é significativamente mais perigoso do que o *malware* tradicional já que: é mais fácil de criar; é acessível a todos; é capaz de produzir resultados automaticamente; pode ser manuseado com maior facilidade.

De acordo com a revista Forbes, vários utilizadores do ChatGPT alertaram anteriormente para o facto do programa poder codificar um *malware* capaz de espionar as teclas digitadas pelo utilizador ou que poderia ser utilizado para criar *ransomware*.

Os termos de serviço da OpenAI proíbem especificamente o uso do programa ChatGPT para criar qualquer tipo de *malware* - definido pela empresa como "conteúdo que tenta criar *ransomware*, *keyloggers*, vírus ou outro *software* com a intenção de causar algum dano". Foi também proibido produtos de construção que visam o "uso indevido de dados pessoais" e "indústrias ilegais ou prejudiciais".

Embora o programa forneça todos esses avisos, quando solicitado a criar *malware*, muitas pessoas ainda encontraram uma maneira de contornar esses mesmos avisos. Vários fóruns da *dark web* encontram-se com inúmeros utilizadores a partilhar tutoriais e a vangloriarem-se por utilizar o *software* ChatGPT para criar malware e ameaças cibernéticas. A velocidade com que os hackers adotaram o programa foi alarmante por si só, com o programa lançado em novembro de 2022 e evidências de scripts de malware aparecendo apenas um mês após este lançamento.

Por todos estes motivos, parece que vivemos tempos muito desafiantes no campo da (inovação da) segurança cibernética, quando a nossa aliada inteligência artificial acaba por se revelar a principal inimiga de quem nos protege.

*Fundador e CEO da VisionWare*