+44 345 094 0945
www.waterstons.com

News article

Mar 2023

# Preparing Australia For A Safer AI-Bot Future

In partnership with BlackBerry, Charlie Hales engaged in a panel discussion at the AISA Cybercon this past Tuesday. Discussing AI, ChatGPT and more. Keep reading to find out what they discussed & what this means for you.




Since November 2022, the world has been captivated by AI-bot market entrant – ChatGPT. In just a few short months, enter Google's Bard – just as the world is already grappling with the benefits and risks of AI-bots in different areas of education, work and society – and at all levels of government.

Among the many benefits, AI chat-bots introduce many risks and challenges, far beyond cheating on exams. Recent research revealed Australia could be less than a year away from a cyberattack credited to ChatGPT. It's no surprise people with malicious intent are testing the waters on how to successfully use tools like these for nefarious purposes.

**The panel at the AISA Cyber Conference in Canberra, was hosted by BlackBerry Australia. The featured speaker at the panel were...**

- Jonathan Jackson, Director of Engineering, Asia Pacific & Japan (Moderator)
- Dr Jacob Wallis, Head of Program, Information Options and Disinformation at the Australian Strategic Policy Institute (ASPI)
- Adam Heywood, CIO, Australian Transport & Safety Bureau (ATSB)
- Charlie Hales, Managing Director, Waterstons Australia

**What did the panel discuss?**

- How does it work?  A brief introduction to 'Artificial General Intelligence' and 'Large Language Models' – in comparison to the human brain
- Some of the uses of AI-bots like Chat GPT in key industry sectors (government, education, financial advice and – the software industry)
- The type of challenges and cybersecurity risks it presents – including disinformation, phishing and malware campaigns and coding skills development
- Key concerns among technology / cybersecurity leaders in Australia, according to a recent survey of IT professionals by BlackBerry
- How we can better prepare for the potential cyber threats it poses – including the use of AI to fight fire with fire
- Improving threat detection and prevention – how to include AI in different threat classification processes and cybersecurity strategies

**Here are some excerpts from the talk, what was discussed and how the above questions were addressed.**

"It's a very powerful tool. But obviously it can't. It's built on what it's learned. So it's still learning for reference at the moment still in development. There's lots of people testing it like ourselves to see what it can do. So yes, exciting space where it comes. But obviously, with that comes the risks, which obviously, we're going to talk about a bit more do."

"You know, so it's one of the fastest growing piece of technology we've ever seen. So in terms of, you know, that that evolution it's a big, big problem because, up until now, Chad GPT and other large, large vendors was basically been a texting text, right. So there's kind of this capability, but now it's evolved to multimodal. So some of the emerging technologies that we need to look out for as we go forward in 2023 and beyond, you know, reveal told us that we need to be the most sub secured nation by 2050. Looking at technology like this with eyes wide open is a big challenge for us all in industry and partnerships because we were facing anniversary that that is quite significant. So in terms of evolution, many Charlie from your perspective things like multimodal what does that mean? text images video we talked about? You mentioned deep fakes. Let's have a chat about that, because that's a big problem. What's your perspective on the evolution of this thing?"

"Multimodal is basically as you mentioned before previously, it's text and text. Now use imagery video, to be able to scrim images. A critical spot personas that we're fishing can help with setting up these tacks and creating really realistic fix, as we were just talking about and use that obviously for attacks. So it's no longer just the person developing the phishing email and turning it into good English rather than their native language with the stick that we've seen in the past. It's building our social profiles and imagery and what we need to do so it's really really powerful to know and I think it's something you know, it's great to hear how was recognition, we're behind where we need to be. There's a lot of learning out there we can use as well as they'll be good practices from around the world to make sure that we aren't getting to where we want to be by 2030. But this can keep going on but I don't want to talk about all of them."

"AI can be used to monitor or sample the new environment in simple terms and spot actually, when that happens, they're doing something they shouldn't you know, when Joe Bloggs goes on Facebook every day does the same thing. We do other things as well but actually when they come on and do because we're different one site or do something else in the environmental try know their privileges which is an obvious pick up on things like that. So as long as is the collaboration and things like that are getting developed to be able defenders as well as enabling the hackers it's it's it can be used for good as well."

"What do you see as the best way to combat our driven disinformation campaigns? All governors are industries, taking the necessary steps to limit this kind of complex questions. First of all, they step back and on this, let's look at high adversaries use these kinds of tools in order to attack strategic advantage. So we need to understand what kind of political warfare looks like. High adversaries understand our own vulnerability. So that's kind of the first step because I think there's been a degree of complacency in that we've we have for a long time not had to think about how our adversaries really invest. in shaping the information environment in influencing international international political debate, because we've had such a long period since perhaps the Cold War, where we just had this level of strategic competition. So that's that's the first step understand how our adversaries see these kinds of tactics. The next step, I think, is understanding where the interests of governments particularly the Australian Government, and in our context, and the industry, align where those interests align or where they where they don't diverge. So particularly in the context of disinformation, we're obviously dealing with platforms that are US governed the US based companies that are converting based in Silicon Valley, Puffin would also also Chinese so that their interests simply diverge from the Australian Government's there are points of alignment, but there are points of divergence. So how do we develop the collaborative partnerships given that context? How do we deliver develop the collaboration that is going to provide sufficient resilience for us to defend against this threat? And then it is about kind of building off the significant advantages these kinds of technologies offer to the defender community as much as to the to the attacker community. So that involves understanding the TTPs developing the the predictive capabilities that will allow us to develop algorithms that learn from generative AI that will detect the patterns of data that emerge from the TTPs of the attackers. Simply just develop information sharing mechanisms across government and industry, the build up resilience because it is that gap between kind of the critical infrastructure infrastructure that the industry managers and government that kind of defines policy and articulates kind of national security priorities. So it's important kind of filling that gap, finding the connective tissue in that gap that allows us to defend against these threats."

There were many more interesting points made by the members of the panel. Some important discussions for all organisations in Australia to be aware of and begin getting ahead of the curve before these AI generated attacks happen.

Thanks to Charlie Hales for being part of this panel & thank you to BlackBerry for partnering with Waterstons Australia on this one.

**If you would like to know more, you can get in touch with Charlie. charlie.hales@waterstons.com**