> **AI has been weaponised to carry out the work for hackers.**
>
> Fong Choong Fook

This visual is human-created, AI-aided

attacker, the damage can happen quickly," he says.

From Manjeevan's point of view, one of the biggest risks lies with inexperienced users. He says such users may view AI agents simply as helpful assistants, granting blanket permissions and access without fully understanding the implications.

While this gives the agent the ability to perform powerful actions in the background, it also means that non-technical users may struggle to stop it or investigate further if something goes wrong, particularly when it comes to cybersecurity.

AI agent runs amok

Some may recall a case from February this year involving OpenClaw and Meta AI alignment director Summer Yue. According to her now viral X post, the AI agent ignored her instructions and began deleting emails from her inbox despite being explicitly told not to.

Fong says that he has had a similar experience while testing out OpenClaw.

"Agentic AI is the future. But at this point, we do not have a very clear model to control the AI's behaviour. Basically, we're leaving a lot of flexibility to the AI to do things by itself," he says.

For Fong, the present-day risks involved outweigh the convenience offered by such tools.

He says that for the time being, he is no longer using OpenClaw and will wait for the technology to mature and for measures that allow more granular control before using it again.

Be prepared for fallout

Manjeevan also says the open-source nature of agentic AI tools like OpenClaw can be a concern, as certain forks (new versions created from an existing open-source project) could contain malicious or tampered code.

"Open-source software has many benefits, but it also means that anyone can copy, modify and redistribute the code. Some versions may include hidden backdoors or unsafe modifications.

"Users who download unofficial versions or install third-party plugins without proper checks could unknowingly expose themselves to malware or data theft.

"This is a known software supply chain risk, and it becomes more serious when the tool itself has powerful system access," he says.

"The risks are not unavoidable, but they require strong safeguards.

"Technical measures such as limiting permissions, isolating the agent in a secure environment, requiring human approval for high-risk actions, and monitoring activity logs can reduce the threat significantly.

"However, these safeguards must be properly implemented and maintained. If an autonomous system is given too much power without proper oversight, the risk increases.

"There is no need to panic, but we must treat autonomous AI agents like any powerful digital tool. The more access you give it, the more careful you must be. Convenience should never come at the expense of basic cybersecurity," Manjeevan says.