SiftView
RESEARCH

AI Assistants and the Return of the Trojan Horse

14 November 2026

SCREEN
Watched

KEYSTROKES
Logged

LOCAL FILES
Readable

ANALYSIS

AI assistants on personal devices can watch the screen, log what you type, read local files, and send data to servers you do not control. Those are the same things XP-era keyloggers and Trojan horses did. The difference is that we now install them on purpose and call them a feature.

I recommend not installing AI assistants on any of your personal devices, and limiting what you post on social media. Keep AI away from wallets, seed phrases, and password managers.

CAPABILITY	XP-ERA MALWARE	AI ASSISTANTS TODAY	EXAMPLE
Keystroke logging	Keyloggers recorded every key.	Assistants that type, fill forms, or transcribe can see what you enter.	Cloud AI in desktop apps can send that text to a company's servers (EFF).
Screen capture	Trojans took screenshots or streamed the desktop.	OS and device agents can snapshot whatever is on screen.	Microsoft Recall, if turned on, saves periodic screen snapshots on the device.
File access	Trojans read and stole local files.	Agentic browsers and assistants can open local files when given that power.	Zenity Labs: a calendar invite could steer Perplexity Comet to read and send local files. Patched Feb 2026.
Remote control	A remote-access Trojan let an attacker drive the machine.	Agents click, browse, and act on your behalf, sometimes with no further prompt.	OWASP and NIST flag tool misuse and goal hijack as core agentic threats.
Sending data out	Stolen data went to an attacker's server.	Many AI tools run in the cloud, so prompts and files leave the device.	EFF: most consumer AI tools process data on someone else's computer.
Persistence / memory	Malware stayed on the machine and kept logs.	Chat history, memory, and screen records can persist and be searched.	Signal (39C3): a "photographic memory" of activity can expose what apps try to keep private.

Sources: OWASP GenAI (June 2026); NIST ISPAB (Jan 2026); EFF SSD; Signal, 39C3; Zenity Labs, PerplexedBrowser; Microsoft Recall docs.