ADVERTISEMENT

Microsoft error sees confidential emails exposed to AI tool Copilot

2 days ago

Share Save

Liv McMahon
Technology reporter

Microsoft has acknowledged an error causing its AI work assistant to access and summarise some users' confidential emails by mistake.

The tech giant has pushed Microsoft 365 Copilot Chat as a secure way for workplaces and their staff to use its generative AI chatbot.

But it said a recent issue caused the tool to surface information to some enterprise users from messages stored in their drafts and sent email folders - including those marked as confidential.

Microsoft says it has rolled out an update to fix the issue, and that it "did not provide anyone access to information they weren't already authorised to see".

ADVERTISEMENT

However, some experts warned the speed at which companies compete to add new AI features meant these kinds of mistakes were inevitable.

Copilot Chat can be used within Microsoft programs such as Outlook and Teams, used for emails and chat functions, to get answers to questions or summarise messages.

"We identified and addressed an issue where Microsoft 365 Copilot Chat could return content from emails labelled confidential authored by a user and stored within their Draft and Sent Items in Outlook desktop," a Microsoft spokesperson told BBC News.

"While our access controls and data protection policies remained intact, this behaviour did not meet our intended Copilot experience, which is designed to exclude protected content from Copilot access," they added.

"A configuration update has been deployed worldwide for enterprise customers."

The blunder was first reported by tech news outlet **Bleeping Computer**, which said it had seen a service alert confirming the issue.

It cited a Microsoft notice saying "users' email messages with a confidential label applied are being incorrectly processed by Microsoft 365 Copilot chat".

The notice added that a work tab within Copilot Chat had summarised email messages stored in a user's drafts and sent folders, even when they had a sensitivity label and a data loss prevention policy configured to prevent unauthorised data sharing.

Reports suggest Microsoft first became aware of the error in January.

Its notice about the bug was also shared on a support dashboard for NHS workers in England - where the root cause is attributed to a "code issue".

A section of **the notice on the NHS IT support site** implies it has been affected.

But it told BBC News the contents of any draft or sent emails processed by Copilot Chat would remain with their creators, and patient information has not been exposed.

'Data leakage will happen'

Enterprise AI tools such as Microsoft 365 Copilot Chat - available to organisations with a Microsoft 365 subscription - often have stricter controls and security protections in place to prevent sharing of sensitive corporate information.

But for some experts, the issue still highlights risks of adopting generative AI tools in certain work environments.

Nader Henein, data protection and AI governance analyst at Gartner, said "this sort of fumble is unavoidable", given the frequency of "new and novel AI capabilities" being released.

He told BBC News organisations using these AI products often lack tools needed to protect themselves and manage each new feature.

"Under normal circumstances, organisations would simply switch off the feature and wait till governance caught up," Henein said.

"Unfortunately the amount of pressure caused by the torrent of unsubstantiated AI hype makes that near-impossible," he added.

Cyber-security expert Professor Alan Woodward of the University of Surrey said it showed the importance of making such tools private-by-default and opt-in only.

"There will inevitably be bugs in these tools, not least as they advance at break-neck speed, so even though data leakage may not be intentional it will happen," he told BBC News.

ADVERTISEMENT

Amazon shares fall as it joins Big Tech AI spending spree
Xbox producer tells staff to use AI to ease job loss pain
Microsoft rolls out AI screenshot tool dubbed 'privacy nightmare'

Tech Decoded
The world's biggest tech news in your inbox

Sign up for our Tech Decoded newsletter to follow the world's top tech stories and trends. **Outside the UK? Sign up here**.

Artificial intelligence | Microsoft | Cyber-security