Google and Microsoft are bringing AI to office apps. How it could boost productivity for us—and cybercriminals

March 21 2023, by Mohiuddin Ahmed and Paul Haskell-Dowland

Credit: AI-generated image (disclaimer)

Google and Microsoft are on a mission to remove the drudgery from computing, by bringing next-generation AI tools as add-ons to existing services.

On March 16, [Microsoft announced](#) an AI-powered system called Copilot will soon be introduced to its 365 suite apps including Word, Excel, PowerPoint, Outlook and Teams.

The news came about two days after [Google published](#) a blog explaining its plans to embed AI into its Workspace apps such as Docs, Sheets, Slides, Meet and Chat.

Collectively, millions of people use these apps each day. Bolstering them with AI could provide a major productivity boost—as long as security isn't an afterthought.

The advent of generative AI

Until recently AI was mainly used for categorization and identification tasks, such as recognizing a number plate using a traffic camera.

Generative AI allows users to [create](#) new content, by applying deep-learning algorithms to [big data](#). [ChatGPT](#) and [DALL-E](#), among others, have already taken the world by storm.

Now, Microsoft and Google have found a more concrete way to bring generative AI into our offices and classrooms.

Like other generative AI tools, Copilot and Workspace AI are built on large language models (LLM) trained on massive amounts of data. Through this training, the systems have "learned" many rules and patterns that can be applied to new content and contexts.

Microsoft's Copilot is being trialed with just 20 customers, with details about availability and pricing to [be released](#) "in the coming months".

Copilot will be integrated across apps to help expedite tedious or

repetitive tasks. For example, it will:

- help users write, edit and summarize Word documents
- turn ideas or summaries into full PowerPoint presentations
- identify data trends in Excel and quickly create visualizations
- "synthesize and manage" your Outlook inbox
- provide real-time summaries of Teams meetings
- bring together data from across documents, presentations, email, calendar, notes and contacts to help write emails and summarize chats.

Assuming it executes these tasks effectively, Copilot will be a massive upgrade from Microsoft's original Office Assistant, Clippy.

Google's Workspace AI will offer similar capabilities for [paying subscribers](#).

What's under the hood?

Microsoft described Copilot as a "sophisticated processing and orchestration engine working behind the scenes to combine the power of LLMs, including GPT-4 […]."

We don't know specifically which data GPT-4 itself was trained on, just that it was a lot of data taken from the internet and licensed, according to [OpenAI](#).

Google's Workspace AI is built on [PaLM](#) (Pathways Language Model), which [was trained](#) on a combination of books, Wikipedia articles, news articles, source codes, filtered webpages, and social media conversations.

Both systems are integrated into existing cloud infrastructure. This means all the data they are applied to will already be online and stored in

company servers.

The tools will [need full access](#) to the relevant content in order to provide contextualized responses. For instance, Copilot can't distill a 16-page Word document into one page of bullet points without first analyzing the text.

This raises the question: will users' information be used to train the underlying models?

In relation to this point, [Microsoft](#) has said, "Copilot's large language models are not trained on customer content or on individual prompts."

Google [has said](#), "[…] private data is kept private, and not used in the broader foundation model training corpus."

These statements suggest the 16-page document itself won't be used to train the algorithms. Rather, Copilot and Workspace AI will process the data in real-time.

Given the rush to develop such AI tools, there may be temptation to train such tools on "real" customer-specific data in the future. For now, however, it seems this is being explicitly excluded.

Usability concerns

As many people noted following ChatGPT's release, text-based generative AI tools [are prone to](#) algorithmic bias. These concerns will extend to the new tools from Google and Microsoft.

The outputs of generative AI tools can be riddled with inaccuracies and prejudice. Microsoft's own Bing chatbot, which also runs on GPT-4, came [under fire](#) earlier this year for making outrageous claims.

Bias occurs when large volumes of data are processed without appropriate selection or understanding of the training data, and without proper oversight of training processes.

For example, much of the content online is written in English—which is likely [the main language](#) spoken by the (mostly [white and male](#)) people developing AI tools. This underlying bias can influence the [writing style](#) and language constructs understood by, and subsequently replicated by, AI-driven systems.

For now, it's hard to say exactly how issues of bias might present in Copilot or Workspace AI. As one example, the systems may simply not work as effectively for people in non-English-speaking countries, or with diverse styles of English.

Security concerns

One major vulnerability in Microsoft's and Google's AI tools is they could make it much easier for cybercriminals to bleed victims dry.

Whereas before a criminal may have needed to trawl through hundreds of files or emails to find specific data, they can now use AI-assisted features to quickly collate and extract what they need.

Also, since there's so far no indication of offline versions being made available, anyone wanting to use these systems will have to upload the relevant content online. Data uploaded online are at greater risk of being breached than data stored only on your computer or phone.

Finally, from a privacy perspective, it's not particularly inspiring to see yet more avenues through which the biggest corporations in the world can collect and synthesize our data.

This article is republished from The Conversation under a Creative Commons license. Read the original article.

Provided by The Conversation