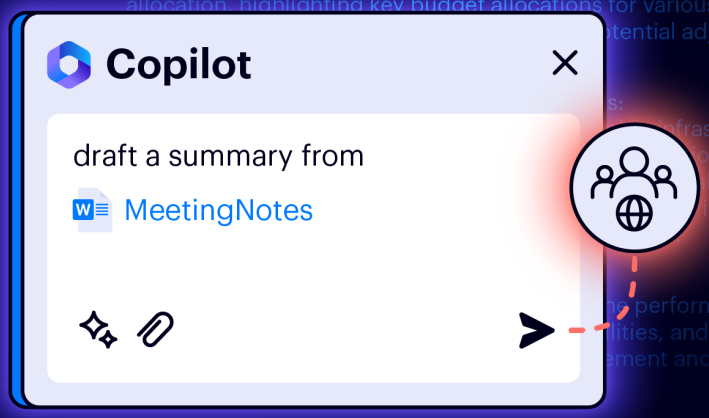
VARONIS

Seguridad de la IA generativa: cómo evitar la exposición de datos de Microsoft Copilot

Decodificar el marco de Copilot para garantizar una implementación segura



Índice

Introducción

Microsoft Copilot ha sido calificada como una de las herramientas de productividad más potentes del planeta.

Copilot es un asistente de IA que se encuentra dentro de cada una de las aplicaciones de Microsoft 365: Word, Excel, PowerPoint, Teams, Outlook, etc. El sueño de Microsoft es eliminar la ingrata tarea del trabajo diario y permitirles a los humanos concentrarse en ser creativos para resolver problemas.

Copilot es diferente de ChatGPT y otras herramientas de IA porque tiene acceso a todo lo que alguna vez haya hecho en Microsoft 365. Copilot puede buscar y compilar datos al instante en sus documentos, presentaciones, correo electrónico, calendario, notas y contactos.

Y ahí radica el problema para los equipos de seguridad de la información. Copilot puede acceder a todos los datos confidenciales a los que puede acceder un usuario, lo que a menudo es **demasiado.** Aproximadamente el 10 % de los datos M365 de una empresa están abiertos a todos los empleados.

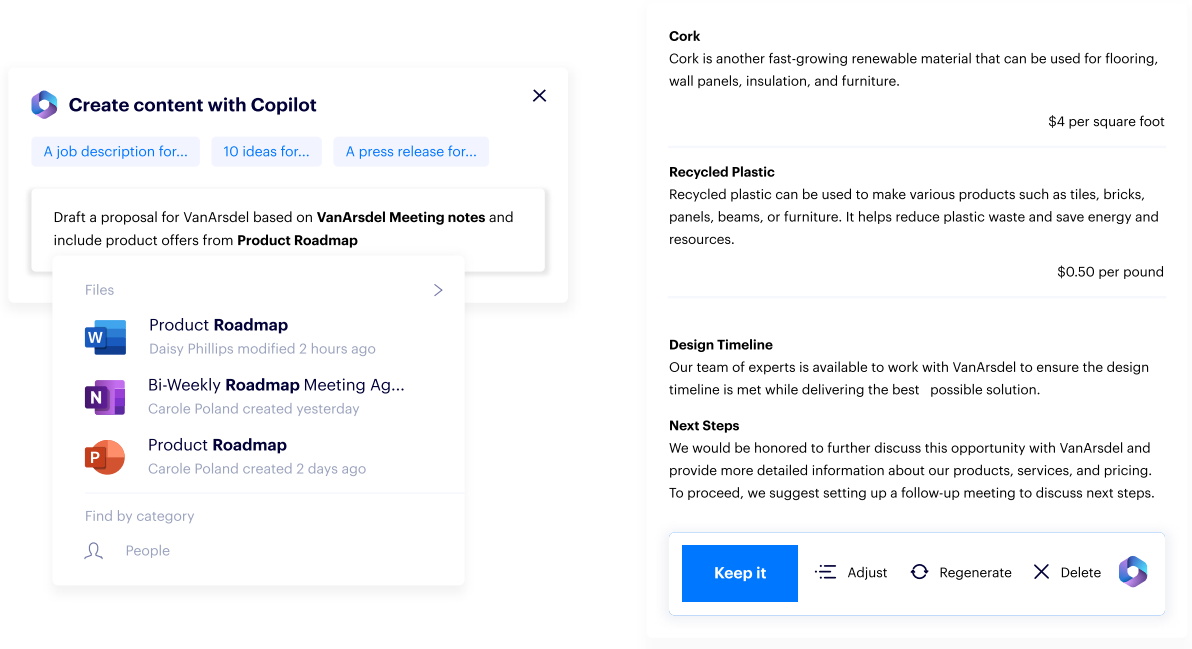
Copilot también puede generar rápidamente nuevos datos confidenciales que se deben proteger. Antes de la revolución de la IA, la capacidad de los humanos de crear y compartir datos superaba con creces la capacidad de protegerlos. Basta mirar las tendencias en las brechas de datos. La IA generativa echa más leña al fuego.

Hay mucho que analizar en lo que respecta a la IA generativa en su conjunto: envenenamiento del modelo, alucinación, deepfakes, etc. En este informe, sin embargo, nos centraremos específicamente en la seguridad de los datos y en cómo su equipo puede garantizar una implementación segura de Copilot.

Casos de uso de Microsoft 365 Copilot

Los casos de uso de la IA generativa con un paquete de colaboración como M365 son infinitos. Es fácil ver por qué tantos equipos de TI y de seguridad solicitan obtener acceso anticipado y preparar sus planes de implementación. Los aumentos en la productividad serán enormes.

Por ejemplo, puede abrir un documento de Word en blanco y pedirle a Copilot que redacte una propuesta para un cliente basada en un conjunto de datos de destino, incluidas páginas de OneNote, presentaciones de PowerPoint y otros documentos de Office. En cuestión de segundos, tiene una propuesta hecha y derecha.



Estos son algunos ejemplos más que Microsoft dio durante su evento de lanzamiento:

- Copilot puede unirse a sus reuniones de Teams y resumir en tiempo real lo que se está debatiendo, registrar elementos de acción y decirle qué preguntas quedaron sin resolver en la reunión.
- En Outlook, Copilot puede ayudarlo a clasificar su bandeja de entrada, priorizar correos electrónicos, resumir hilos y generar respuestas.
- En Excel, Copilot puede analizar datos sin procesar y brindarle información, tendencias y sugerencias.

Cómo funciona Microsoft 365 Copilot

Este es un resumen sencillo de cómo se procesa una indicación de Copilot:

- Un usuario introduce una indicación en una aplicación como Word, Outlook o PowerPoint.
- Microsoft recopila el contexto empresarial del usuario en función de sus permisos de M365.
- La indicación se envía al modelo de lenguaje de gran tamaño (como GPT4) para generar una respuesta.
- Microsoft realiza verificaciones IA responsable posteriores al procesamiento.
- Microsoft genera una respuesta y devuelve los comandos a la aplicación de M365.

Modelo de seguridad de Microsoft 365 Copilot

Con Microsoft, siempre existe una tensión extrema entre la productividad y la seguridad.

Esto se puso de manifiesto durante el coronavirus, cuando los equipos de TI implementaban rápidamente Microsoft Teams sin antes entender a fondo cómo funcionaba el modelo de seguridad subyacente o cuál era el estado de los permisos, los grupos y las políticas de enlace de M365 de su organización.

Las buenas noticias:

- **Aislamiento del tenant**. Copilot solo utiliza datos del tenant de M365 del usuario actual. La herramienta de IA no mostrará datos de otros tenants en los que el usuario pueda ser invitado, ni de ningún tenant que pueda estar configurado con sincronización entre tenants.
- **Límites de entrenamiento**. Copilot no utiliza ninguno de sus datos comerciales para entrenar a los LLM fundacionales que Copilot utiliza para todos los tenants. No debería tener que preocuparse de que sus datos patentados aparezcan en las respuestas a otros usuarios en otros tenants.

Las malas noticias:

- **Permisos**. Copilot muestra todos los datos de la organización a los que los usuarios individuales tienen al menos permisos de visualización.
- **Etiquetas**. El contenido generado por Copilot no heredará las etiquetas MPIP de los archivos de los que Copilot obtuvo su respuesta.
- **Humanos**. No se garantiza que las respuestas de Copilot sean 100 % objetivas o seguras; los humanos deben asumir la responsabilidad de revisar el contenido generado por IA.

Veamos las malas noticias una por una.

Permisos

Otorgar acceso a Copilot solo a lo que un usuario puede acceder sería una excelente idea si las empresas pudieran aplicar fácilmente una política de privilegios mínimos de acceso en Microsoft 365.

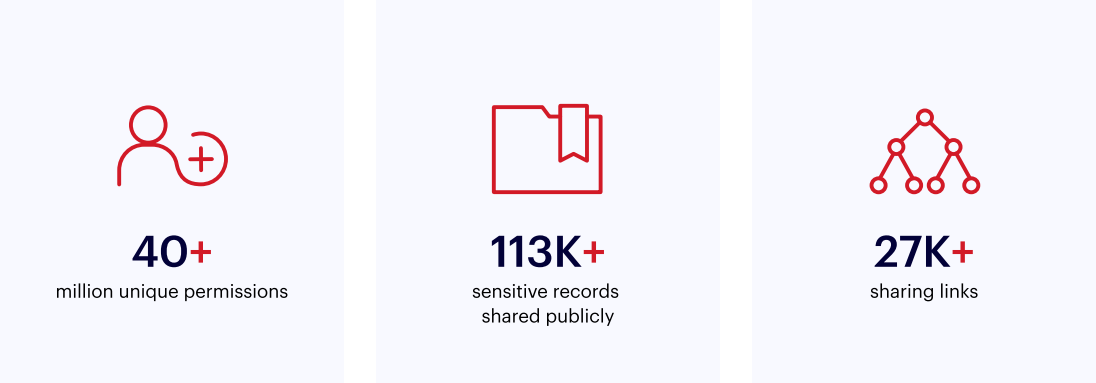
Microsoft afirma en su **documentación de seguridad de datos de Copilot**:

> **"Es importante que use los modelos de permisos disponibles en los servicios de Microsoft 365, como SharePoint, para garantizar que los usuarios o grupos adecuados tengan el acceso adecuado al contenido adecuado dentro de su organización".**

Sin embargo, sabemos por experiencia que la mayoría de las organizaciones están muy lejos de poder aplicar una política de privilegios mínimos de acceso. Solo eche un vistazo a algunas de las estadísticas del **Informe sobre el estado de riesgos de permisos en la nube de Microsoft.**

- Los entornos multinube son complejos. Hay más de 40 000 permisos para administrar y más del 50 % son de alto riesgo.
- Tras analizar más de 500 evaluaciones de riesgos, Microsoft descubrió que la mayoría de los identificadores tienen un exceso de permisos considerable, lo que pone en riesgo los entornos críticos de las organizaciones por un uso indebido de los permisos, ya sea accidental o malintencionado.
- La carga de trabajo identifica que el acceso a los entornos en la nube está aumentando, y ahora supera la cantidad de identidades humanas en una relación de 10:1.
- En realidad, solo se utiliza el 1 % de los permisos otorgados.
- Menos del 50 % de las identidades son superadministradores, lo que significa que tienen acceso a todos los permisos y recursos.

Esta imagen coincide con lo que Varonis observa cuando hacemos miles de evaluaciones de riesgos sobre los datos para empresas que utilizan Microsoft 365 cada año. En nuestro reporte, **La gran exposición de SaaS**, descubrimos que el tenant promedio de M365 tiene:



¿Por qué sucede esto? Los permisos de Microsoft 365 son muy complejos. Solo piense en todas las formas en que un usuario puede obtener acceso a los datos:

- Permisos de usuario directos
- Permisos de grupo de Microsoft 365
- Permisos locales de SharePoint (con niveles personalizados)
- Acceso a vínculos (cualquier persona, toda la organización, directo, invitado)
- Acceso externo
- Acceso público
- Acceso de invitado

Para empeorar las cosas, los permisos están principalmente en manos de los usuarios finales, no de los equipos de TI o de seguridad.

Etiquetas

Microsoft depende en gran medida de las etiquetas basadas en la confidencialidad para aplicar políticas de DLP, aplicar cifrado y evitar las fugas de datos. En la práctica, sin embargo, conseguir que las etiquetas funcionen es difícil, en especial si se confía en los humanos para aplicarlas.

Microsoft pinta un panorama optimista del etiquetado y el bloqueo al considerarlos la mejor red de seguridad para sus datos. La realidad nos muestra un escenario más sombrío. A medida que los humanos crean datos, el etiquetado con frecuencia se retrasa o se vuelve obsoleto.

El bloqueo o cifrado de datos puede agregar fricción a los flujos de trabajo, y las tecnologías de etiquetado se limitan a tipos de archivos específicos. Cuantas más etiquetas tenga una organización, más confusa puede ser para los usuarios. Esto es especialmente intenso para las organizaciones más grandes.

Es seguro que la eficacia de la protección de datos basada en etiquetas se degradará cuando tengamos una IA que genere órdenes de magnitud más datos que requieran etiquetas precisas y de actualización automática.

¿Mis etiquetas están bien?

Varonis puede validar y mejorar el etiquetado basado en la confidencialidad de Microsoft de una organización al escanear, descubrir y reparar:

- Archivos confidenciales sin etiquetas
- Archivos confidenciales con etiquetas incorrectas
- Archivos no confidenciales con una etiqueta de confidencialidad

Humanos

La IA puede hacer que los humanos sean perezosos. El contenido generado por LLM como GPT4 no es solo bueno, es genial. En muchos casos, la velocidad y la calidad superan con creces lo que un humano puede hacer. En consecuencia, las personas comienzan a confiar ciegas en la IA para crear respuestas seguras y precisas.

Ya hemos visto escenarios del mundo real en los que Copilot redacta una propuesta para un cliente e incluye datos confidenciales que pertenecen a un cliente completamente diferente. El usuario toca “enviar” después de darle un vistazo rápido (o sin siquiera hacer eso), y ahora tiene en sus manos un escenario de violación de privacidad o brecha de datos.

Preparar la seguridad de su tenant para Copilot

Resulta fundamental tener una idea de la postura de seguridad de los datos antes de la implementación de Copilot. Es probable que Copilot esté disponible a principios del próximo año, por lo que ahora es un buen momento para implementar los controles de seguridad.

Varonis protege a miles de clientes de Microsoft 365 con nuestra plataforma de seguridad de datos, que proporciona una vista en tiempo real del riesgo y la capacidad de aplicar automáticamente una política de privilegios mínimos de acceso.

Podemos ayudarlo a abordar los mayores riesgos de seguridad con Copilot prácticamente sin esfuerzo manual. Con **Varonis para Microsoft 365**, puede:

- Descubrir y clasificar de forma automática todo el contenido confidencial generado por la IA.
- Asegurarse automáticamente de que las etiquetas MPIP se apliquen de manera adecuada.
- Aplicar de forma automática permisos de privilegios mínimos de acceso.
- Supervisar continuamente los datos confidenciales en M365 y recibir alertas y responder ante un comportamiento anormal.

La mejor manera de empezar es con una **evaluación de riesgo gratuita**. Se configura en cuestión de minutos y, en un día o dos, tendrá una vista en tiempo real del riesgo de los datos confidenciales.

¿Está listo para experimentar la diferencia de Varonis?

Reduzca su riesgo sin asumir ninguno. Comuníquese con nuestro equipo para saber qué aspectos se analizarán en su evaluación de riesgos sobre los datos **gratuita**.

Contáctenos

Acerca de Varonis

Varonis es pionera en seguridad y análisis de datos, y libra una batalla diferente de la que enfrentan las empresas de ciberseguridad convencionales. Varonis se centra en proteger los datos empresariales: archivos y correos electrónicos importantes; información confidencial de clientes, pacientes y empleados; registros financieros; planes estratégicos y de productos; y otra propiedad intelectual.

La Plataforma de seguridad de datos Varonis detecta amenazas cibernéticas de atacantes internos y externos al analizar la información, la actividad de la cuenta y el comportamiento del usuario; evita y limita que se produzcan incidentes restringiendo el acceso a datos importantes y obsoletos; y mantiene la seguridad de forma eficiente mediante la automatización.

Los productos de Varonis abordan otros casos de uso importantes, como la protección de datos, la gobernanza de datos, el enfoque Zero Trust, el cumplimiento, la privacidad de los datos, la clasificación y la detección y respuesta a amenazas. Varonis inició sus operaciones en 2005 y cuenta con clientes que incluyen empresas líderes en muchos sectores, como servicios financieros, público, sanitario, industrial, seguros, tecnología, consumo y comercio minorista, energía y servicios públicos, construcción e ingeniería y educación.