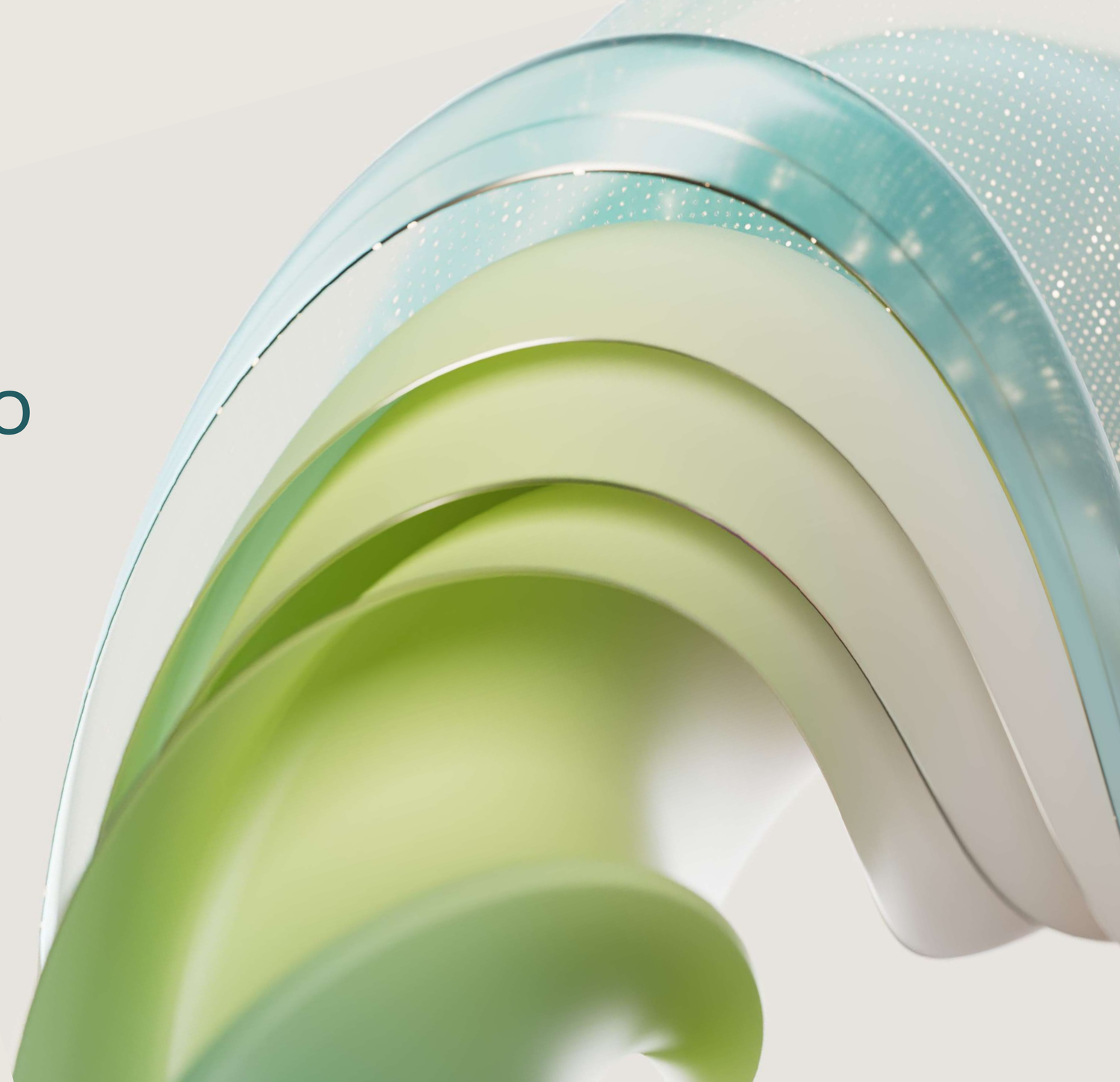
Microsoft

EXTREME IT

Haz crecer tu negocio con una IA en la que puedes confiar

Consideraciones para que los líderes empresariales planifiquen su transformación de IA con confianza



Contenido

Introducción

Las soluciones de IA presentan una inmensa oportunidad para que organizaciones de todos los tamaños y sectores aumenten los ingresos, reduzcan los costes, mejoren el bienestar de los empleados y operen de manera más eficiente. Por tanto, no es de extrañar que los líderes empresariales estén sometidos a presión para adoptar soluciones de IA lo antes posible y evitar quedarse atrás.

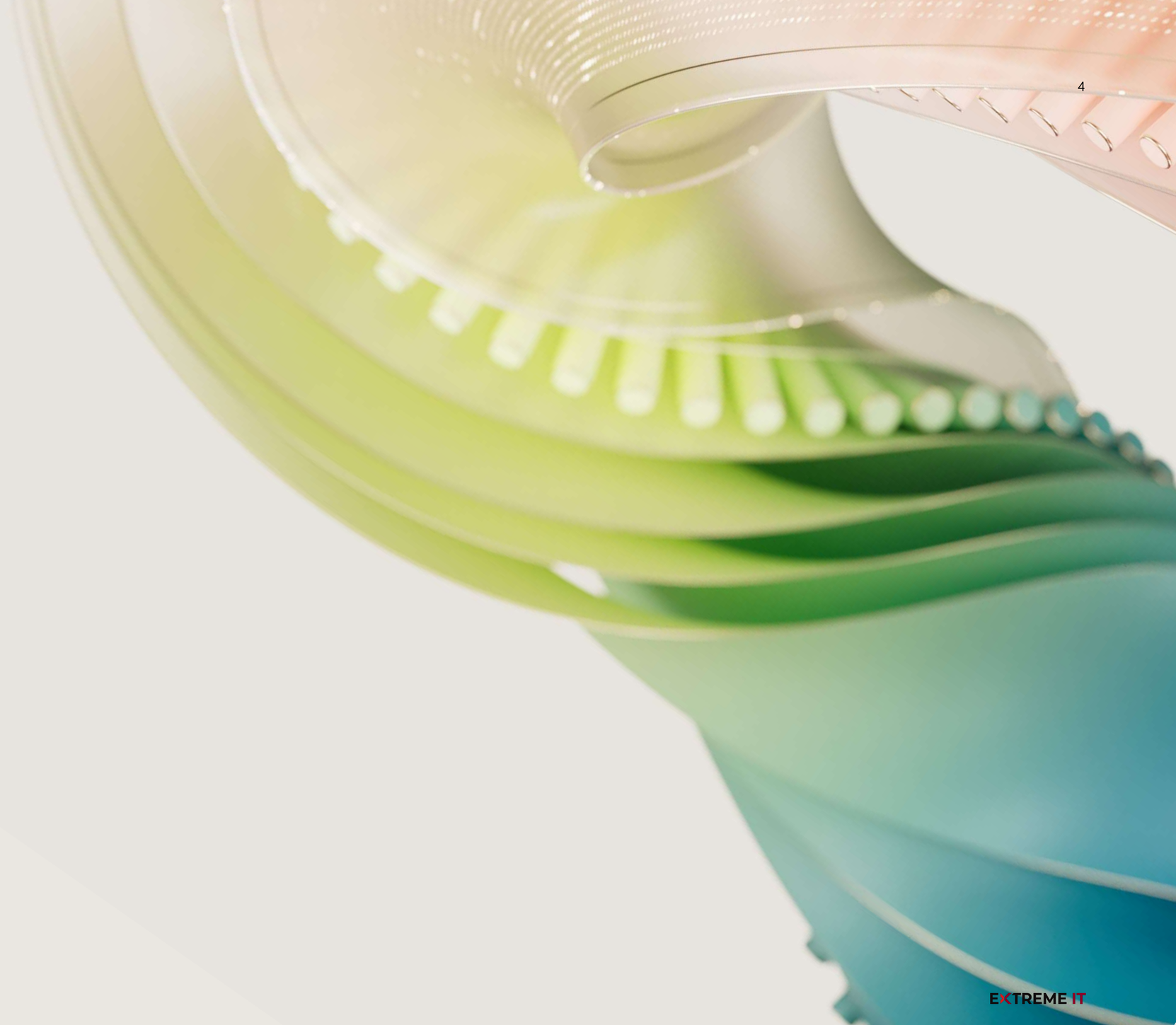
Con tanto alboroto alrededor de la IA, son muchas las preocupaciones sobre los efectos secundarios negativos de la tecnología. En las secciones siguientes, hablaremos de las distintas consideraciones que los líderes empresariales deben tener en cuenta para ayudar a desbloquear la promesa de esta nueva tecnología y evitar las consecuencias no deseadas.

El establecimiento de prácticas de IA responsables y seguras para tu empresa te ayuda a implementar de forma segura las herramientas de IA. Y a medida que aumenta la regulación global de la IA, invertir ahora en una IA responsable te prepara mejor para cumplir los nuevos requisitos normativos conforme llegan. Adoptar un enfoque reflexivo sobre la implementación de una IA responsable y segura en tu empresa puede ayudar a los líderes a adoptar la IA y la innovación.

de los líderes empresariales afirmaron que estaban «**preocupados y entusiasmados por igual**» con la IA generativa[1].

Estamos comprometidos con una IA de confianza y con la creación de una tecnología de apoyo líder del sector. Con funciones que mejoran la seguridad, la protección y la privacidad, seguimos permitiendo que los clientes usen y creen soluciones de IA fiables.

Prácticas de la IA responsable



Considera los principios de la IA responsable

Las políticas públicas y las prácticas recomendadas del sector aún están poniéndose al día con los últimos avances en tecnología de IA, lo que hace que los líderes empresariales busquen directrices fiables sobre cómo implementar sistemas de IA que tengan un impacto positivo en las empresas, las personas y la sociedad.

Dedicar tiempo a pensar en un enfoque de IA responsable para tu organización puede ayudarte a ti y a tus equipos a avanzar con confianza y protegerte de los riesgos no deseados. Hemos desarrollado seis principios de una IA responsable que debes tener en cuenta a medida que planificas y creas tu propio enfoque. Estos principios son:

- **Privacidad y seguridad**
- **Fiabilidad y protección**
- **Responsabilidad**
- **Inclusividad**
- **Transparencia**
- **Imparcialidad**

Privacidad y seguridad

Los sistemas de IA deben cumplir los mismos estándares de privacidad y seguridad que las empresas aplican a sus datos más confidenciales.

Prioriza la seguridad de tu infraestructura y la privacidad de tus datos.

Debes saber dónde se encuentran los datos y cómo se utilizan, y confirmar que estén protegidos en reposo y en tránsito. Comprueba que las herramientas de IA sigan los valores de privacidad y seguridad de tu empresa.

Cuando se implementen permisos de datos estrictos y se asignen permisos de usuario en función de roles y pertenencia a grupos, solo las personas autorizadas tendrán acceso a información confidencial, lo que reduce el riesgo de filtraciones internas.

Además, puedes mantener la seguridad y cumplir los requisitos normativos con una solución de gobierno, que incluye retener y registrar las interacciones con aplicaciones de IA, ayudar a detectar cualquier infracción de las normativas o de las políticas de la organización al usar esas aplicaciones e investigar los incidentes una vez que ocurran.

La implementación de la IA también debe cumplir todas las leyes y reglamentos locales sobre el uso y la privacidad de los datos. En lo que respecta a la seguridad de los datos, es mejor ser cauteloso y trabajar con proveedores de herramientas de seguridad con un historial de fiabilidad.

Un ejemplo de privacidad y seguridad en la práctica:

El uso de una herramienta de IA para analizar la comunicación con los clientes para resolver una incidencia de soporte podría implicar el acceso a datos confidenciales o identificables de los clientes. Conocer las leyes y reglamentos locales, cumplir los altos estándares de seguridad de tu organización y aplicar controles adecuados puede ayudar a mantener la privacidad de los datos confidenciales.

Cumplir las normativas de IA

Microsoft se compromete a crear productos y soluciones que cumplan las normativas como la Ley de la IA de la UE para ayudar a nuestros clientes a usar la IA respetando las normas.

Fiabilidad y protección

La fiabilidad y la protección significan que los sistemas de IA tienen el rendimiento esperado, sin errores ni interrupciones. Los desarrolladores de herramientas de IA son responsables de asegurarse de que su producto proporcione resultados precisos a través de pruebas y documentación, pero un sistema de supervisión ayuda a verificar si la herramienta cumple con el uso previsto. Los sistemas también deben someterse a procesos regulares de supervisión, mantenimiento, retroalimentación y evaluación para identificar nuevos usos, solucionar y resolver problemas rápidamente y mejorar el sistema de IA con el tiempo.

Realiza pruebas de estrés con regularidad.

Las pruebas de estrés preparan un sistema de IA para gestionar los tipos y el volumen de uso para los que está destinado, sin producir errores ni volverse vulnerable a los riesgos.

El «red teaming» (equipos rojos) es un tipo de prueba de estrés que implica la simulación de ataques del mundo real y el uso de las técnicas que los hackers utilizan habitualmente para obtener acceso a sistemas seguros. En 2018, Microsoft estableció su propio equipo rojo de IA y amplió la misión del equipo para identificar riesgos más allá de los riesgos de seguridad tradicionales, incluidos los riesgos de usuarios no adversarios que infringen las normas de una IA responsable. Por ejemplo, el «red teaming» de una herramienta de IA generativa podría implicar comprobar si un usuario puede generar contenido que estereotipe a un grupo marginado utilizando la herramienta. Un modelo de IA también puede ser sometido a «red teaming» para identificar posibles usos indebidos, determinar su alcance y conocer sus limitaciones. Después, los conocimientos obtenidos se pueden aplicar a las versiones futuras del modelo para garantizar que funcionará de forma fiable y segura[2].

Realiza la diligencia debida sobre las medidas de fiabilidad y protección de un sistema de IA en el momento de la compra y ejecuta pruebas de estrés periódicas para identificar los riesgos con posterioridad.

Una revisión cuidadosa de la documentación ayuda a las organizaciones a comprender las medidas que el proveedor del sistema de IA ha tomado para permitir el uso fiable y seguro de su sistema, y ayuda a las organizaciones a cumplir todos los requisitos para operar el sistema de forma segura.

Un ejemplo de fiabilidad y protección en la práctica:

Se utiliza una herramienta de IA para crear modelos de los resultados financieros e informar sobre el rendimiento. Se realizan pruebas periódicamente para garantizar que la IA produzca resultados precisos de forma fiable, lo que evita efectos adversos en la salud financiera de la organización.

Responsabilidad

En muchos casos, la IA responsable se centra en el ser humano. El establecimiento de un sistema claro de supervisión ayuda a tus empleados a controlar las herramientas de IA que implementas y a ser responsables de los resultados que producen esas herramientas.

Establece un sistema de supervisión que defina claramente las funciones y responsabilidades en cada etapa del recorrido de la IA.

La implementación de un sistema de supervisión que realice pruebas de impacto y responda a los resultados del impacto otorga protagonismo a las personas. Esto ayuda a protegerse de posibles impactos adversos y a garantizar que se implementen controles adecuados en cada etapa.

Asegúrate de que las herramientas de IA se ajustan a los objetivos.

Comprueba periódicamente que las herramientas de IA proporcionan las soluciones adecuadas para los problemas que estaban destinadas a resolver y determina cómo responderá tu organización si una herramienta no cumple su propósito previsto.

Un ejemplo de responsabilidad en la práctica:

El uso de una herramienta de IA para revisar los contratos legales implica la supervisión por parte de una persona con el contexto y la experiencia suficientes para verificar el cumplimiento de las leyes y reglamentos aplicables, y para aprobar el resultado final de la revisión respaldada por la IA.

Inclusividad

En esencia, la inclusividad requiere que las herramientas de IA sean accesibles para personas de todas las capacidades. Esto significa que las herramientas que los líderes empresariales crean o compran deben seguir los principios de diseño accesible y cumplir con la norma europea de accesibilidad, EN 301 549, la sección 508 de la Ley de Rehabilitación de los Estados Unidos y las Directrices de accesibilidad de contenido web (SEDA).

Identifica oportunidades para desarrollar la inclusividad en tu organización con la IA.

Por ejemplo, los beneficiarios de las subvenciones de Microsoft están creando una plataforma de contratación para candidatos con diversas capacidades neuronales, desarrollando pantallas en braille mejores y más asequibles para estudiantes con discapacidad visual, y creando una aplicación web para ayudar a las personas con discapacidades del habla a comunicarse de forma más eficaz.

Transparencia

La transparencia es la base de la confianza. Para lograr y mantener la transparencia, siempre debes ser claro sobre cómo y cuándo se utiliza la IA, así como sobre sus capacidades y limitaciones.

Mantén una actitud abierta sobre cómo se está implementando y utilizando la IA en toda la organización.

Las partes interesadas y los empleados pueden sentirse más seguros con el uso de herramientas de IA cuando entienden cómo llega la herramienta a su conclusión y conocen sus limitaciones. Esta transparencia ayuda a desarrollar su capacidad para el uso de herramientas respaldadas por IA y a saber cuándo complementar sus resultados con información fuera del alcance de la herramienta.

Los clientes quieren saber cuándo están interactuando con una herramienta de IA, cuándo se utiliza la IA en la toma de decisiones o cuándo se ha generado o manipulado un activo de audio o visual con la IA. Los líderes empresariales deben considerar cómo se comunicará esa información a los clientes.

Un ejemplo de transparencia en la práctica:

La IA generativa se utiliza para crear contenido para una campaña de marketing y la organización identifica qué elementos se han creado con la IA. Un especialista revisa el contenido creado por la IA para verificar su precisión con el fin de no engañar a los clientes sobre las características o funciones del producto publicitado.

Imparcialidad

Una implementación imparcial de la IA asigna oportunidades, recursos e información de forma equitativa entre las personas que utilizan la IA o resultan afectadas por ella.

Asegúrate de que tus sistemas de IA proporcionen una calidad similar de servicio y entrega de recursos y oportunidades a todos los que los utilizan o se ven afectados por ellos, a través de grupos demográficos.

Cuando se utilicen herramientas de IA que describen, retratan o representan de algún modo a las personas, reduce al mínimo la posibilidad de que se las estereotipe o menosprecie, especialmente las de grupos marginados, para promover la imparcialidad.

Incluye a miembros de diferentes orígenes, experiencias, niveles de educación y puntos de vista en el equipo que administra la implementación de IA e identifica sesgos estadísticos en los conjuntos de datos para ayudar a impulsar la imparcialidad en un sistema de IA. La revisión humana por parte de expertos en la materia en decisiones para las que se utiliza la IA también puede ayudar a evitar resultados sesgados.

Un ejemplo de imparcialidad en la práctica:

Se utiliza una herramienta de IA para revisar las aplicaciones e identificar candidatos prioritarios en el proceso de contratación con la supervisión de un representante de recursos humanos. Esta persona confirma que la herramienta evalúa la información con precisión, sin sesgos estadísticos, y que la revisión final está disponible en la toma de decisiones.

Toma decisiones fundamentadas sobre IA y seguridad

La implementación responsable de la IA minimiza los riesgos y permite que tu empresa se beneficie del potencial de sus diversos usos. Utiliza las siguientes preguntas como puntos de partida para la conversación con tu equipo cuando empieces a pensar en tu implementación de IA.

Privacidad y seguridad

- [ ] ¿Los datos a los que acceden los sistemas de IA están protegidos de acuerdo con las políticas de tu organización para el tratamiento de datos confidenciales?
- [ ] ¿Tienes el control de tus datos, incluido dónde se almacenan y cómo se utilizan?
- [ ] ¿Están tus datos protegidos en todo momento, incluso cuando están en tránsito de un sistema a otro?
- [ ] ¿Dispones de herramientas de seguridad de calidad para defenderte del acceso de terceros o de los ciberataques?
- [ ] ¿Tienes disponibles herramientas de identificación y respuesta ante amenazas en caso de ciberataques?

Fiabilidad y protección

- [ ] ¿Se han probado correctamente las herramientas que se pretenden utilizar para minimizar los errores?
- [ ] ¿Dispones de un plan para remediar cualquier error que se produzca?
- [ ] ¿Se supervisarán periódicamente las herramientas para ver si hay problemas de fiabilidad?
- [ ] ¿Estás preparado para cumplir todos los requisitos para utilizar las herramientas de forma segura?

Responsabilidad

- [ ] ¿Has evaluado el impacto que esta implementación podría tener en tus empleados, organización y clientes?
- [ ] ¿Has establecido un sistema de supervisión y respuesta en caso de posibles impactos negativos?
- [ ] ¿Has implementado las prácticas recomendadas de gobierno y administración de los datos?
- [ ] ¿Has determinado quiénes se ocuparán de la supervisión de las herramientas de IA y te has asegurado de que tengan una formación y un control adecuados?
- [ ] ¿Te has asegurado de que esta solución sea adecuada para el propósito previsto?

Inclusividad

- [ ] ¿Has confirmado que las herramientas que vas a utilizar cumplen los principios de diseño accesible?
- [ ] ¿Cumplen las herramientas la norma europea de accesibilidad, EN 301 549?
- [ ] ¿Cumplen las herramientas la sección 508 de la Ley de Rehabilitación de los Estados Unidos?
- [ ] ¿Cumplen las herramientas las Directrices de accesibilidad de contenido web (WCAG)?

Transparencia

- [ ] ¿Has informado a las partes interesadas sobre cómo funcionará esta implementación, incluidas sus capacidades y limitaciones, o tienes un plan para hacerlo antes de que empiecen a usar herramientas de IA?
- [ ] ¿Tienes un plan para comunicarte con los empleados sobre cómo tu organización va a utilizar la IA y cómo se deben interpretar sus resultados?
- [ ] ¿Has determinado cómo y cuándo notificarás a los clientes que están interactuando con la IA o que están viendo contenido generado por IA?

Imparcialidad

- [ ] ¿Te has asegurado de que esta implementación proporcionará la misma calidad de servicio a todos los afectados?
- [ ] ¿Has probado los resultados del sistema para asegurarte de que asignarán recursos y oportunidades de manera justa entre los grupos demográficos?
- [ ] ¿Los resultados del sistema están libres de estereotipos y representaciones negativas de grupos marginados?

Protege tus datos y sistemas de IA

Protege tus herramientas de IA ahora y en el futuro

Implementada incorrectamente, cualquier tecnología que tenga acceso a datos confidenciales puede presentar un riesgo de seguridad para las empresas. Como los sistemas de IA requieren una gran cantidad de datos propiedad de la empresa, es fundamental priorizar la seguridad desde el principio al pensar en la implementación de IA o al comprar soluciones de IA.

En Microsoft, hemos lanzado la Secure Future Initiative, que reúne nuestros conocimientos para abordar y prepararnos ante los riesgos cada vez mayores de los ciberataques en la era de la IA. La Secure Future Initiative identifica tres principios que Microsoft mantiene para ayudar a proteger todo el ecosistema digital:

1. **Seguridad por diseño**
   La seguridad es lo primero en el diseño de cualquier producto o servicio.

2. **Protección predeterminada**
   Las protecciones de seguridad están habilitadas y se aplican de forma predeterminada, no requieren ningún esfuerzo adicional y no son opcionales.

3. **Operaciones protegidas**
   Los controles y la supervisión de seguridad se mejorarán continuamente para abordar las amenazas actuales y futuras.

La seguridad es el pilar que sustenta cualquier implementación de IA. Cuando garantizas prácticas de higiene de seguridad básicas, proteges tus datos, a tu personal y tus dispositivos de más del 98 % de los ciberataques.[3]

Entre las prácticas eficaces de higiene de seguridad se incluyen las siguientes:

- Habilitar la autenticación multifactor (MFA) para protegerse de las contraseñas de usuario expuestas y ayudar a proporcionar resiliencia adicional a las identidades.
- Aplicar principios de Confianza cero, que implican la verificación explícita, el uso del acceso con privilegios mínimos y la suposición de que se va a producir un ataque, para limitar el impacto de un ataque.
- Usar una solución de detección y respuesta extendidas y antimalware para bloquear automáticamente los ataques y obtener información sobre el software de operaciones de seguridad para una respuesta más rápida.
- Garantizar que los sistemas estén actualizados con las últimas versiones de firmware, sistemas operativos y aplicaciones.
- Implementar las protecciones correctas para los datos críticos, lo que requiere conocer qué datos son más importantes y dónde se encuentran.

Gestiona tus sistemas de IA con gobierno

Un buen modelo de gobierno ayuda a construir unos cimientos sólidos para una implementación responsable de la IA. El papel de los gobiernos y los organismos reguladores es mantener los requisitos de referencia para minimizar los efectos adversos del uso de IA en la sociedad. Las organizaciones comerciales también tienen la responsabilidad ética de crear una estructura de gobierno para administrar su propio desarrollo o uso de los sistemas de IA de acuerdo con sus valores organizativos, leyes y reglamentos y el bien común.

Establecimiento de tu propio gobierno

Al crear el sistema de gobierno de tu organización, recuerda que el propósito del gobierno es alinear las soluciones de IA con la política de la empresa y los principios de IA responsable a través de una serie de políticas y procedimientos. Esto incluye la aplicación de políticas para evaluar e implementar soluciones de IA de terceros, coordinar la participación y la formación de las partes interesadas, y producir documentación para informar a empleados, clientes y otros usuarios sobre las herramientas de IA.

Riesgos

Identificar

Identificar los riesgos es la primera etapa del gobierno de la IA y debe fundamentar las decisiones sobre la seguridad, fiabilidad e idoneidad de una herramienta para un determinado fin. La identificación de riesgos implica realizar evaluaciones del impacto de la IA y revisiones de privacidad y seguridad, incluido el «red teaming» y las pruebas de estrés.

Medir

La medición de los riesgos implica desarrollar métricas para evaluar los riesgos identificados y probar los métodos de mitigación planificados para determinar su eficacia.

Administrar

La administración de riesgos requiere que las organizaciones supervisen constantemente el rendimiento. En esta etapa, deberías identificar oportunidades para garantizar la autonomía de los usuarios e informar a las partes interesadas sobre el uso responsable. Debe incluirse la revisión y supervisión humanas en el proceso de administración, así como las prácticas recomendadas de transparencia de acuerdo con los principios de IA responsable.

Aprovecha el potencial de la IA

Avanza en la sostenibilidad de la IA

Así como la seguridad es un aspecto esencial de la IA responsable, la sostenibilidad es crucial para un uso consciente de la misma. La IA, una e icaz herramienta para conocer y reducir el impacto medioambiental, puede ayudar a las empresas a avanzar en sus objetivos de sostenibilidad, y a que tanto las empresas privadas como las instituciones comprendan y promuevan mejor la conservación del medioambiente, la administración de recursos y la mitigación del cambio climático.

Con herramientas de administración de datos de IA e informes, las organizaciones pueden obtener visibilidad de sus actividades de sostenibilidad para registrar, comunicar y reducir su impacto medioambiental. La IA puede proporcionar los conocimientos necesarios para tomar decisiones más contrastadas y mantener el rumbo hacia tus objetivos de sostenibilidad.

Al mismo tiempo, reconocemos la intensidad de recursos que requieren estas aplicaciones y la necesidad de abordar el impacto medioambiental desde todos los ángulos.

Los líderes empresariales pueden hacer que sus propios centros de datos sean más sostenibles o trabajar con proveedores que ya están tomando estas medidas para reducir el impacto medioambiental de los centros de datos que alimentan sus soluciones de IA.

En Microsoft, estamos profundamente comprometidos con el medioambiente y estamos concentrando nuestros esfuerzos en tres áreas principales que reflejan nuestro compromiso con la sostenibilidad: optimización de la eficiencia energética y del agua en los centros de datos, desarrollo de materiales con poco carbono y mejora de la eficiencia energética de la IA y los servicios en el cloud, todo ello con el objetivo de equipar a nuestros clientes y partners con herramientas para avanzar juntos.

Inicia tu transformación de IA

Al considerar cuidadosamente las prácticas de IA responsable y dar prioridad la seguridad en toda la organización, podrás explorar el potencial de la IA para hacer crecer tu negocio.

Nuestros compromisos y funcionalidades te permiten acelerar la transormación de la IA con confianza, y puedes confiar en Microsot para priorizar la seguridad, privacidad y protección de tu IA.

Obtén más información sobre Microsoft IA para comenzar tu viaje.

Microsoft

EXTREME IT

Fuentes

[1] «What Business Leaders Really Think About Generative AI», INSEAD, 11 de abril de 2024, https://knowledge.insead.edu/leadership-organisations/what-business-leaders-really-think-about-generative-ai.

[2] «2024 Microsoft Responsible AI Transparency Report», Microsoft, consultado el 10 de julio de 2024, https://www.microsoft.com/corporate-responsibility/responsible-ai-transparency-report.

[3] Quy Nguyen, «Basic cyber hygiene prevents 98% of attacks», Microsoft Tech Community, 18 de septiembre de 2023, https:// techcommunity.microsoft.com/t5/security-compliance-and-identity/basic-cyber-hygiene-prevents-98-of-attacks/ba-p/3926856.

© 2024 Microsoft Corporation. Todos los derechos reservados. Este documento se proporciona «tal cual». La información y las opiniones que aquí se expresan, incluidas las direcciones URL y otras referencias a sitios web de Internet, están sujetas a cambios sin previo aviso. Cualquier riesgo relacionado con el uso del documento es responsabilidad del usuario. Este documento no te proporciona ningún derecho legal sobre ninguna propiedad intelectual de ningún producto de Microsoft. Puedes copiar y usar este documento para uso interno como material de consulta.