

Deloitte.



Ética en la Inteligencia Artificial

Un nuevo imperativo para
empresas, juntas y alta gerencia

Introducción

Para las personas, los líderes, las organizaciones y la sociedad, una pregunta cada vez más urgente es si la inteligencia artificial (IA) puede ayudarnos a avanzar de manera positiva, ¿o nos obstaculizará?

GDPR Reglamento General de Protección de Datos, y la Ley de Privacidad del Consumidor de California se han decretado, al menos en parte, para abordar estas inquietudes.

La aplicación de la IA como herramienta de resolución de problemas ofrece una gran promesa de avance. La IA en medicina, por ejemplo, podría

Los riesgos éticos son una preocupación principal para aproximadamente un tercio de los ejecutivos que esperan una transformación del negocio a corto plazo con la IA.

Cerca del final de su vida, el físico Stephen Hawking, quien habló sobre los beneficios y peligros de la IA, fue coautor de una carta abierta instando a investigadores y desarrolladores de IA

revertir los elevados costos médicos. La IA en la agricultura podría conducir a **duplicar o triplicar los rendimientos de los cultivos** a partir de los insumos de hoy, un logro fundamental en la lucha continua contra la inseguridad de la oferta alimentaria.

Pero la IA también puede ser utilizada contra nuestros intereses humanos a través de medios tales como ataques cibernéticos, **manipulación social o discriminación**, e **incentivos financieros**. Al darse cuenta de esto, los consumidores y los ciudadanos están preocupados por la acumulación de datos que se utilizan para tomar decisiones sin el consentimiento del consumidor o el respeto de los derechos del consumidor. Nuevas leyes como la europea

para enfocar su trabajo en una dirección ética. No mucho después, el 76 por ciento de los ejecutivos que respondieron a una **encuesta de Deloitte** dijeron que esperaban que la IA "transformara sustancialmente" sus empresas dentro de los próximos tres años. Alrededor de un tercio de este grupo agregó que los riesgos éticos eran una preocupación principal sobre la tecnología de IA.

Pero, ¿cómo es una dirección ética en IA y cómo pueden las organizaciones ponerla en práctica? Este es el tema de la discusión a seguir. Comencemos con algunas definiciones básicas.

Ética en la IA, definición

IA es un término amplio que abarca tecnologías que pueden imitar el comportamiento humano inteligente. Actualmente, se utilizan en general cuatro categorías principales de IA:

- **Aprendizaje automático**, capacidad de los modelos estadísticos para desarrollar aptitudes y mejorar su rendimiento a lo largo del tiempo sin la necesidad de seguir instrucciones explícitamente programadas.
- **Aprendizaje profundo**, forma compleja de aprendizaje automático utilizada para el reconocimiento de imagen y de voz que involucra redes neuronales con muchas capas de variables abstractas.
- **Procesamiento del lenguaje natural (PNL)**, tecnología que impulsa interfaces basadas en voz para asistentes virtuales y chatbots, así como para consultar conjuntos de datos, extrayendo o generando significado e intención del texto en una forma legible, estilísticamente neutral y gramaticalmente correcta.
- **Visión por computador**, tecnología que extrae el significado y la intención de los elementos visuales, ya sean personajes (en el caso de la digitalización de documentos) o la categorización de contenido en imágenes como caras, objetos, escenas y actividades.

La ética es "la disciplina que trata qué es bueno y malo, con un deber y obligación moral", así como "los principios de conducta que rige a un individuo o un grupo"⁴. En un negocio, una mentalidad ética apoya la toma de decisiones basada en valores. El objetivo es hacer no solo lo que es bueno para el negocio, sino también lo que es bueno para los empleados, clientes y comunidades de la organización en la que opera.

Consolidando estas dos definiciones,

La "ética en la inteligencia artificial" se refiere a las bases organizacionales que delinean lo correcto y lo incorrecto: piensa en los valores, políticas y códigos corporativos de ética y principios rectores aplicados a las tecnologías de IA. Estas bases establecen objetivos y pautas para la IA en todo el ciclo de vida del producto o servicio; investigación y diseño, construcción y capacitación y, cambios y operación.

Una mentalidad ética apoya la toma de decisiones basada en valores, centrándose en hacer no solo lo que es bueno para el negocio, sino lo que es bueno para los empleados, clientes y clientes de una organización y las comunidades en las que viven y trabajan.

Consideraciones para incluir la ética en la IA

Conceptualmente, la ética en la IA se aplica tanto al objetivo de la solución de IA como a cada parte de la solución de IA. La IA se puede utilizar para lograr resultados del negocio poco éticos, a pesar de que sus partes (aprendizaje automático, aprendizaje profundo, PNL y / o visión por computadora) se diseñaron para funcionar de manera ética. Por ejemplo, un sistema automatizado de solicitud de préstamo hipotecario podría incluir visión por computadora y otras herramientas diseñadas para leer solicitudes de préstamo escritas a mano, analizar la información proporcionada por el solicitante y tomar una decisión basada en los parámetros programados en la solución. Estas tecnologías no procesan dichos datos a través de una lente ética, solo procesan datos. Sin embargo, si la compañía hipotecaria programa inadvertidamente el sistema con objetivos o parámetros que discriminan injustamente por motivos de raza, género o cierta información geográfica, el sistema podría usarse para hacer aprobaciones o denegaciones discriminatorias de préstamos.

En contraste, una solución de IA con un propósito ético puede incluir procesos que carecen de integridad o precisión hacia este fin ético. Por ejemplo, una empresa puede implementar un sistema de inteligencia artificial con capacidades de aprendizaje automático para respaldar el objetivo ético de los procesos no discriminatorios de reclutamiento de personal. La empresa comienza utilizando la capacidad de inteligencia artificial para identificar criterios de rendimiento basados en los mejores resultados del pasado de la organización. Tal muestra de intérpretes anteriores puede incluir sesgos basados en características de contratación anteriores (incluidos criterios discriminatorios como género, raza o etnia) en lugar de simplemente el rendimiento. En otras palabras, la máquina aprende en función de los datos que procesa, y si la muestra de datos no es representativa o precisa, las lecciones que aprende de los datos no serán precisas y pueden conducir a resultados poco éticos.

Para comprender dónde podrían surgir problemas éticos y cómo en el futuro del trabajo podrían evitarse esos problemas, es útil organizar la IA a lo largo de cuatro dimensiones principales de preocupación (**figura 1**):



Tecnología, datos y seguridad.

Observe el enfoque de la organización para el ciclo de vida de la IA desde una perspectiva ética, incluidas las formas en que construye y prueba datos y modelos en soluciones habilitadas para la IA. El liderazgo en esta dimensión proviene de la organización, jefes de información, tecnología, datos, seguridad y privacidad.



Gestión de riesgos y cumplimiento.

Descubra cómo la organización desarrolla y aplica políticas, procedimientos y estándares para soluciones de IA. Vea cómo se relacionan con la misión, los objetivos y los requisitos legales o regulatorios de la organización. Los jefes de riesgo, cumplimiento, legal y ética juegan un papel en esta dimensión.



Personas, habilidades, modelos organizacionales y capacitación.

Comprenda y monitoree cómo el uso de la IA afecta las experiencias de empleados y clientes. Evaluar continuamente cómo los modelos operativos, los roles y los modelos organizacionales están evolucionando debido al uso de IA. Eduque a todos los niveles de la fuerza laboral e implemente iniciativas de capacitación para actualizar o mejorar las capacidades. Establecer protocolos para incentivar el comportamiento ético y alentar decisiones éticas a lo largo del ciclo de vida de la IA. En esta dimensión, la función de recursos humanos comparte la responsabilidad con los equipos de aprendizaje y desarrollo, los oficiales de ética y el liderazgo ejecutivo más amplio.



Políticas públicas, marcos legales y regulatorios, e impacto en la sociedad.

Finalmente, desarrolle un sentido del lugar de la IA en el entorno empresarial. Esto incluye el nivel de aceptación que la IA tiene en el gobierno y la cultura. También incluye la dirección que las leyes y las regulaciones están tomando con respecto a la IA. Aplique esta información al efecto que la IA pueda tener en el tiempo en términos de educación, empleo, ingresos, cultura y otros aspectos de la sociedad.

El CEO, CRO, CCO y CFO tienen roles de liderazgo en las tres primeras dimensiones, mientras que la cuarta dimensión se basa en el liderazgo de la política pública, los organismos de formulación de políticas.

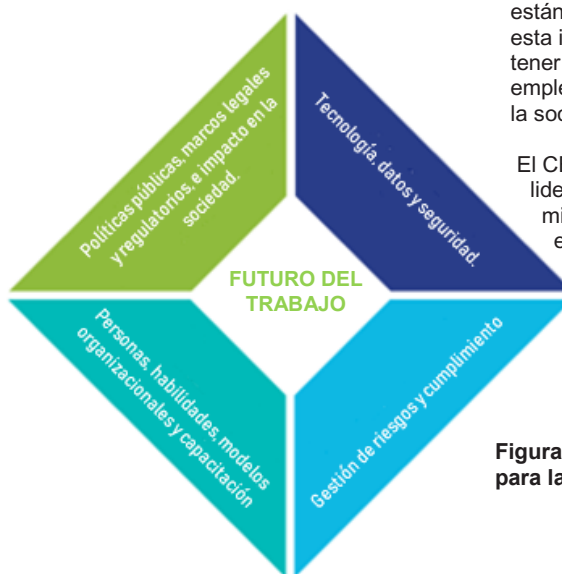


Figura 1: Cuatro dimensiones son vitales para la implementación de la ética IA

Ejemplos de riesgos éticos de IA

Existen consideraciones éticas a lo largo de cada etapa del ciclo de vida de la IA.

Etapa del ciclo de la vida	Ejemplos de riesgos	Herramientas para gestionar el riesgo
Investigación y diseño	<p>Durante la fase de diseño, se pasan por alto los riesgos inherentes de la solución (como una aplicación de visión por computadora que captura y utiliza indebidamente las imágenes de los clientes o empleados u otra información de identificación personal.)</p> <p>El objetivo de la solución no está alineado con la misión y los valores de la organización.</p>	<p>Un marco para lo que define el uso ético de la IA y los datos en la organización</p> <p>Un panel multifuncional con representación de las unidades de negocio, así como de los grupos de tecnología, privacidad, legal, ética y riesgo.</p> <p>Un código de conducta de datos o ética de la IA que los profesionales deben cumplir</p>
Construcción y operación	<p>La organización carece de formas apropiadas para obtener el consentimiento de las personas cuyos datos se utilizan para entrenar el modelo de IA.</p> <p>Un programador construye sesgos (ya sea consciente o inconscientemente) en un modelo destinado a una solución habilitada para IA, como en el ejemplo de reclutamiento de personal descrito anteriormente</p> <p>Los datos utilizados para entrenar un modelo de IA tienen problemas de calidad.</p>	<p>Un proceso para determinar dónde y cómo obtener los datos que entrenan a los modelos.</p> <p>Pautas sobre dónde y cómo el consentimiento del usuario se convierte en una consideración en la fase de capacitación</p> <p>Políticas sobre dónde y cómo construir modelos y si usar tecnología de código abierto</p> <p>Una evaluación de las formas en que una solución de inteligencia artificial puede enseñar comportamientos no sincronizados con la misión o los valores de la organización</p>
Cambios y operación	<p>Un chatbot (una aplicación de IA que puede incluir capacidades de lenguaje cognitivo) aprende nuevos comportamientos que son inapropiados u ofensivos para los clientes</p> <p>La organización no puede evaluar rápidamente a qué nuevas fuentes de datos ha accedido recientemente una solución habilitada para inteligencia artificial</p> <p>La organización carece de la capacidad de probar y monitorear soluciones de IA</p>	<p>Un proceso para que la organización participe en un monitoreo continuo</p> <p>Una evaluación de las formas en que una solución habilitada para IA puede obtener acceso a nuevas formas de datos</p> <p>Un proceso para que la empresa actualice a la junta sobre riesgos o problemas relacionados con la IA</p> <p>Umbrales organizacionales o niveles de tolerancia para ayudar a determinar si se debe retirar una solución habilitada para IA</p>

Primeros pasos en el camino de la ética en la IA

La ética en la IA es un esfuerzo radical con muchas partes en movimiento. Al mismo tiempo, dejando de lado la tecnología, el enfoque inicial debe seguir un camino similar al de otros programas de ética y cumplimiento, que incluyen:

Iniciar en la parte superior. La junta directiva y la alta gerencia establecen el tono para cualquier programa de ética y cumplimiento. Informe al liderazgo superior sobre el uso de herramientas de IA en toda la organización y cómo la IA presenta oportunidades y riesgos para la organización. Luego, revise y, cuando sea necesario, amplíe las políticas, procedimientos y estándares existentes de la empresa (o considere la adopción de otros nuevos) para abordar la ética en la inteligencia artificial como una prioridad clave de la organización.

Comunicar claramente las intenciones. Las bases y mecanismos éticos protegen a la organización al tiempo que abordan los riesgos asociados con el uso de datos y tecnología. Para comenzar, considere desarrollar un "código de conducta" de inteligencia artificial para científicos de datos y otros profesionales de datos mientras establece los canales para escalar los problemas que se presenten. También considere establecer principios para el uso de IA entre la gerencia, los empleados, los inversionistas y los clientes, según corresponda a la organización.

Evaluar los riesgos. La ética de la IA se trata tanto de comprender los riesgos como de establecer un proceso para evitarlos. Sea claro sobre qué tipo de solución de inteligencia artificial está creando y para quién la está creando. Identifique los procesos en el ciclo de vida de la IA que podrían afectar negativamente a las partes interesadas, ordénelos por prioridad y asigne recursos para mitigar los riesgos.

Dar ejemplos. Deje que los equipos de productos sepan qué buscar en las soluciones de monitoreo para la ética en la inteligencia artificial. Una forma es incorporar el uso ético de los datos y las consideraciones tecnológicas en el diseño de la arquitectura de TI (similar al marco de Deloitte [Privacidad en el diseño](#)). Otra es diseñar estructuras de control e integrarlas en soluciones habilitadas para IA.

Poner barreras. Las organizaciones deberían considerar establecer proactivamente barreras de protección para guiar, monitorear y evaluar cómo la organización, los empleados, los proveedores y los clientes utilizan la IA. Dichas barreras pueden ser técnicas u organizativas, similares a los controles internos. Un ejemplo de una barrera técnica es un marco de control integrado en el diseño de una solución de IA que evita que se completen acciones específicas. Otro ejemplo de una barrera técnica es el uso de la IA explicable e interpretable, donde la toma de decisiones detrás de las soluciones habilitadas para IA es transparente y explicable. Un ejemplo de una barrera organizativa es un panel inter-funcional que evalúa todas las soluciones habilitadas para IA antes de que se creen (similar a una [Junta de Revisión Institucional](#)) y considera el impacto de los problemas emergentes relacionados con la ética de la IA en las soluciones de IA existentes y que están por ser implementadas. Dicho panel debe ser responsable y debe tener una carta clara de navegación para impulsar el cambio e influir en las decisiones en toda la organización.

Es fácil quedar atrapado en la complejidad de la IA. Pero comenzar con lo básico puede crear un impacto a corto plazo al tiempo que ofrece el máximo espacio para aprender sobre la marcha. Con el tiempo, la organización puede integrar los matices más finos de la ética en la inteligencia artificial a medida que se conocen sus implicaciones, tanto para la organización como para las partes interesadas.

Aprenda más:

Visite el centro de Notre Dame Deloitte para el liderazgo ético

El [Centro Notre Dame Deloitte para el Liderazgo Ético \(NDDCEL\)](#) es una colaboración entre la Universidad de Notre Dame y Deloitte.

Tras identificar un valor compartido de integridad personal en el mundo empresarial actual, NDDCEL se unió para avanzar en la comprensión e implementación de prácticas de liderazgo ético en el ámbito corporativo.

Hoy, con el uso cada vez mayor de soluciones habilitadas para IA, la ética y el liderazgo están entrando en un territorio desconocido. Actualmente, el Centro se enfoca en desarrollar una comprensión mutua de los problemas éticos, desafíos, oportunidades y soluciones relacionadas con IA y análisis. Como parte de esta iniciativa, el NDDCEL brinda a los ejecutivos y líderes empresariales oportunidades para aprender más sobre este tema en rápida evolución, conectarse con colegas de ideas afines y ayudar a desarrollar la próxima generación de líderes empresariales mediante la traducción de conocimientos de investigación en prácticas líderes.

Para obtener más información, visite el sitio web de NDDCEL



Abordar el imperativo ético en la IA: la responsabilidad de todos

Con sus capacidades de aprendizaje automático, aprendizaje profundo, procesamiento del lenguaje natural (PNL) y visión por computadora, la IA ofrece la emocionante perspectiva de mejorar la condición humana. Pero hay un lado oscuro potencial de la IA que es difícil de ignorar.

El resultado es una nueva frontera en la ética empresarial. Aquellos involucrados en el avance de la IA, incluidos los directorios corporativos, los equipos de gestión, los investigadores y los ingenieros, enfrentan un imperativo creciente de aportar una lente ética a lo que diseñan y construyen. Este enfoque debe articularse a través de bases éticas organizacionales que se apliquen en todo el ciclo de vida de la IA.

Para ser eficaz, las bases deben reflejar una comprensión de las vulnerabilidades relacionadas con la IA a lo largo de las cuatro dimensiones de riesgo descritas anteriormente.

Es demasiado pronto para saber a dónde llevará este viaje a las empresas y sus clientes. Pero el camino a seguir está a la vista. Comienza con un compromiso de alto nivel con el liderazgo ético y un enfoque en la alfabetización técnica y ética. Con barreras éticas para limitar los pasos en falso en el camino, todos en una organización pueden trabajar juntos para producir soluciones que representen el potencial de la IA.





This publication contains general information only and Deloitte is not, by means of this publication, rendering accounting, business, financial, investment, legal, tax, or other professional advice or services. This publication is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action that may affect your business. Before making any decision or taking any action that may affect your business, you should consult a qualified professional advisor.

Deloitte shall not be responsible for any loss sustained by any person who relies on this publication.

About Deloitte

Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee ("DTTL"), its network of member firms, and their related entities. DTTL and each of its member firms are legally separate and independent entities. DTTL (also referred to as "Deloitte Global") does not provide services to clients. In the United States, Deloitte refers to one or more of the US member firms of DTTL, their related entities that operate using the "Deloitte" name in the United States, and their respective affiliates. Certain services may not be available to attest clients under the rules and regulations of public accounting. Please see www.deloitte.com/about to learn more about our global network of member firms.

Copyright © 2019 Deloitte Development LLC. All rights reserved.