

Principios de Inteligencia Artificial de Telefónica: código de conducta de la IA

26 de junio de 2024

La Inteligencia Artificial (IA) es una tecnología clave para la innovación, que está permitiendo acelerar la transformación digital y que, en definitiva, puede traer mucho progreso a la sociedad y a las industrias. Actualmente, estamos viviendo una etapa de democratización de esta tecnología, gracias a la IA generativa y los modelos de IA de uso general, y este tipo de soluciones ya forman parte de nuestro día a día.

De hecho, la IA ya se usa para mejorar la vida de las personas y de la sociedad, a través de la optimización en los diagnósticos médicos, la gestión y prevención de desastres humanitarios y para avanzar en los objetivos de desarrollo sostenible, entre otros. En el entorno empresarial, la IA permite promover la innovación en servicios y nuevos modelos de negocio, mejorar la interacción y personalización de productos, servicios y experiencias con los clientes, y generar eficiencias.

En Telefónica vemos la Inteligencia Artificial (IA) como una oportunidad para mejorar el bienestar, impulsar el crecimiento económico, tener un impacto social positivo y aumentar nuestra ventaja competitiva, si se desarrolla y use responsablemente, centrado en las personas.

Nuestro fuerte compromiso con los Derechos Humanos hace que trabajemos para identificar, prevenir y mitigar los posibles impactos negativos que la tecnología pueda tener en las personas, en la sociedad o en el medio ambiente. Consideramos que la IA puede contribuir a crear una sociedad más inclusiva y a ofrecer mejores oportunidades para todos.

Reconociendo la naturaleza global de la IA, nuestros principios éticos están alineados con las propuestas de organismos internacionales como la OECD, el Consejo de Europa y la UNESCO, con los que cooperamos para abordar los desafíos que la IA pueda presentarnos.

El presente código de conducta de IA tiene como objetivo promover la adopción de una inteligencia artificial confiable y centrada en el ser humano, así como garantizar un alto nivel de protección de la salud, la seguridad, los derechos fundamentales de las personas y el medio ambiente. Estos Principios nos han guiado desde su definición en 2018, y lo seguirán haciendo con esta actualización que presentamos en 2024.

Para llevarlos a la práctica disponemos de un **modelo de gobierno** aprobado en 2023 y de aplicación en toda la compañía, con el que podemos aprovechar todo el potencial de la IA, al tiempo que protegemos a las personas y minimizamos los posibles riesgos. También materializamos estos objetivos éticos mediante la realización de evaluaciones de riesgos y la creación de herramientas que nos permitan minimizar los sesgos, mejorar la explicabilidad, la accesibilidad y la eficiencia medioambiental.

IA centrada en las personas

En Telefónica sabemos que las conexiones más importantes son las que se producen entre personas y, por eso, queremos que la IA contribuya a hacer de nuestro mundo, un mundo más humano. De esta manera, velamos porque la IA respete y promueva los Derechos Humanos, y que se base en valores como la no discriminación, la diversidad y la justicia social. Nos preocupa el bienestar de las personas y la protección de colectivos vulnerables, por lo que también nos comprometemos a trabajar para preservar la integridad personal y evitar los posibles impactos negativos de la IA en este ámbito.

Creemos que la supervisión humana adecuada a cada caso de uso de la IA es crucial para preservar la dignidad humana y la libertad de decisión. Buscamos desarrollar e implementar sistemas de IA atendiendo a principios de diseño centrado en las personas, que amplíen y complementen las aptitudes cognitivas, sociales y culturales y que generen un espacio de múltiples oportunidades para favorecer y salvaguardar la autonomía humana.

IA transparente y explicable

Nuestro objetivo es lograr una IA confiable, para lo que es fundamental alcanzar una comprensión adecuada de los sistemas de IA. Tratamos de entender la lógica que hay detrás de las conclusiones, recomendaciones y decisiones tomadas por los modelos, para poder evitar posibles errores e incrementar la confianza de los usuarios.

Queremos ser transparentes, facilitando la comprensión del funcionamiento del sistema, así como de asegurar que las personas sean conscientes de la interacción con la IA. Nos esforzamos en alcanzar una IA explicable, que permita rastrear las razones por las que el sistema de IA produce determinados resultados, manteniendo un equilibrio justo entre rendimiento y explicabilidad.

La transparencia y la explicabilidad también nos permiten asegurar que se asumen las responsabilidades correspondientes.

IA justa e inclusiva

Promovemos la precisión de los resultados que producen nuestros sistemas de IA para tomar decisiones justas y confiables. Trabajamos, además, para garantizar que nuestra IA sea representativa y accesible, y para fomentar que las aplicaciones de tecnología de IA produzcan resultados justos, promuevan la inclusión y garanticen la equidad, logrando que los beneficios de la IA sean accesibles para cualquier persona.

Esto implica que los resultados no deberían conducir a impactos discriminatorios por razón de género, raza, origen étnico, religión, orientación sexual, discapacidad u otro tipo de condición. Por ello, trabajamos en aplicar la

tecnología para reducir al mínimo la probabilidad de que los conjuntos de datos de entrenamiento que usemos creen o refuercen un sesgo o una discriminación injusta, así como los que se pueden introducir en el diseño e implementación de los algoritmos. Creemos que la IA puede proporcionar una oportunidad de que las personas puedan tomar decisiones más justas y equilibradas, siempre que adoptemos un enfoque de diseño responsable, justo e inclusivo.

IA respetuosa de la privacidad y la seguridad

Los sistemas de IA se basan en datos, que en muchos casos pueden ser de carácter personal, tanto en las etapas de modelado y entrenamiento de los sistemas, como cuando estos se encuentran en producción. En Telefónica nos comprometemos a respetar el derecho a la protección de datos y a la privacidad de las personas. Al procesar datos personales, y de acuerdo con nuestra política de privacidad, cumpliremos en todo momento con los principios de licitud, lealtad y transparencia, minimización de datos, limitación en su conservación y en su finalidad, exactitud, integridad y confidencialidad, siempre cumpliendo con el principio de responsabilidad activa. Para garantizar el cumplimiento de nuestra política de privacidad, utilizamos una metodología de privacidad desde el diseño.

Al construir sistemas de IA, al igual que con otros sistemas, seguimos el enfoque de seguridad desde el diseño de Telefónica. De acuerdo con nuestra política de seguridad, aplicamos en todas las fases del ciclo de procesamiento y análisis de datos las medidas técnicas y organizativas necesarias para garantizar un nivel de seguridad adecuado al riesgo y, en cualquier caso, de conformidad con las medidas de seguridad establecidas en la legislación vigente en cada uno de los países o regiones donde trabajamos. Así también nos aseguraremos de que nuestros sistemas de IA sean sólidos y robustos frente a ataques o intentos de manipulación, minimizando los posibles riesgos.

Creemos que la trazabilidad es esencial para garantizar la ciberseguridad de los sistemas de inteligencia artificial, ya que permite rastrear y documentar el proceso de toma de decisiones e identificar y mitigar posibles riesgos de seguridad desde el momento del diseño y desarrollo de los sistemas.

IA comprometida con el medio ambiente

Nuestro compromiso con el planeta nos lleva a promover la IA como herramienta diferencial para preservar el medio ambiente. En un mundo en el que la IA tiene el poder de influir en casi todos los aspectos de nuestras vidas, es crucial que aprovechemos esta tecnología de forma que impulse la sostenibilidad ambiental y contribuya, por ejemplo, a fomentar la economía circular y paliar la crisis climática. Mediante la colaboración y el cumplimiento

de estándares aplicables, pretendemos garantizar que los avances de la IA contribuyan de forma positiva y responsable con el planeta.

En coherencia con nuestro objetivo de ser una compañía con cero emisiones netas en 2040, desde Telefónica, trabajamos para minimizar el impacto medioambiental de la IA, fundamentalmente a través de la optimización de la eficiencia energética de los sistemas de IA, los datos en los que se basan, y de los entornos e infraestructuras que los contienen. Para garantizar la eficacia de nuestras acciones, nos esforzamos por evaluar el impacto medioambiental de nuestros sistemas de IA y tomar medidas para reducir su huella de carbono.

IA con responsabilidad y rendición de cuentas en toda la cadena de valor

Creemos firmemente que la supervisión humana, la responsabilidad y la rendición de cuentas son insustituibles por la IA. Nos esforzamos en garantizar una supervisión adecuada al caso de uso y un correcto funcionamiento de los sistemas de IA con los que trabajamos, a lo largo de todo el ciclo de vida, en nuestros desarrollos propios, cuando incluimos componentes de terceros, y cuando los comercializamos. Este enfoque responsable contribuye a la anticipación de posibles impactos negativos para poder mitigarlos y evitarlos.

Somos conscientes de lo compleja que puede llegar a ser la cadena de valor de la IA y la cantidad de actores que pueden estar implicados en un caso de uso. Para evitar la omisión de responsabilidades, trabajamos para asegurar la trazabilidad de las decisiones tomadas en todo el proceso. Asimismo, nuestros canales de consulta nos permiten atender dudas y comentarios de nuestros clientes, responsabilizándonos así por mejorar nuestra oferta de productos, servicios y experiencias.

Nuestro modelo de gobierno de IA define roles y funciones de manera que se materialice la asunción de responsabilidades por todos los actores en la cadena de valor. Además, provee de los mecanismos necesarios para rendir cuentas en caso de que se produzcan impactos negativos, facilitando la auditabilidad mediante la disposición de pruebas y evidencias que documenten el proceso, y que permitan constatar garantías de cumplimiento con la legislación y estándares aplicables.

Nota:

Según el AI Act (regulación europea en materia de IA aprobado en Diciembre 2023), un sistema de IA es un sistema basado en una máquina diseñado para funcionar con distintos

niveles de autonomía, que puede mostrar capacidad de adaptación tras el despliegue y que, para objetivos explícitos o implícitos, infiere de la información de entrada que recibe la manera de generar información de salida, como predicciones, contenidos, recomendaciones o decisiones, que puede influir en entornos físicos o virtuales.



www.telefonica.com