



—

DIRECTRICES DE INFORM PARA UNA IA RESPONSABLE



Aspiramos a una inteligencia artificial digna de confianza. Por ello, hemos definido una serie de principios para el diseño, desarrollo y aplicación responsables de los sistemas de IA.



PRÓLOGO

Las capacidades de los modelos de IA y sus aplicaciones en los negocios y la vida cotidiana se están desarrollando a un ritmo vertiginoso. El extraordinario progreso de los grandes modelos lingüísticos como ChatGPT, así como de otras aplicaciones de IA, fascina a todo el mundo. Simultáneamente, suscita inquietudes sobre la fiabilidad y seguridad del uso de la IA en algunos casos concretos y su impacto en la sociedad, las empresas y el ámbito laboral. Las empresas de software que ofrecen IA no pueden seguir tratando los algoritmos como una cuestión puramente técnica, sino que necesitan enmarcar el despliegue de la IA con preguntas sobre el impacto, la ética y la responsabilidad. La inteligencia artificial ha evolucionado desde ser objeto de investigación hasta convertirse en un asunto que compete a todo el mundo.

Desde su fundación en 1969, INFORM ha sido pionero en la aplicación de algoritmos a problemas empresariales, incluidas las técnicas de búsqueda y optimización y el aprendizaje automático. Ambas clases de técnicas se han beneficiado considerablemente de las mejoras técnicas y metodológicas. A día de hoy, seguimos teniendo muchas ganas de impulsar a las empresas con este tipo de tecnología, al mismo tiempo que mantenemos nuestros firmes valores que nos han impulsado a hacer un uso responsable de la IA.

Y seguiremos trabajando en este sentido. A fin de mantener nuestros elevados estándares a la hora de ampliar las aplicaciones de IA, somos conscientes de la necesidad de que nuestros valores y prácticas de desarrollo de productos sean claros y transparentes. Las directrices de INFORM para una IA responsable están diseñadas para expresar explícitamente nuestros principios, tanto hacia nuestros empleados como hacia nuestros clientes, proveedores y socios. Pensamos que sólo la transparencia y un enfoque centrado en las personas pueden proporcionar la base para que la sociedad adopte y acepte una IA confiable.

Gracias a nuestras raíces europeas y a nuestra presencia mundial, consideramos que nuestros valores están profundamente anclados en el pensamiento humanitario, y nos complace comprobar que nuestro país de origen ocupa el primer puesto en el Índice de contratos sociales de IA. Aunque esto sienta las bases para un uso ético y responsable de la IA, al mismo tiempo queremos contribuir a una adopción más extendida de la IA y otras metodologías algorítmicas éticamente aceptables en beneficio de las empresas, la humanidad y la sostenibilidad. Como parte de nuestro programa de sostenibilidad, estamos comprometidos con los Objetivos de Desarrollo Sostenible de la ONU, entre los que se incluyen nuestras actividades de IA.

Este documento pretende contribuir al debate actual sobre cómo abordar la IA de forma responsable.

ENFOQUE DE INFORM SOBRE LA IA

A pesar de que los grandes modelos lingüísticos (LLM) – con ChatGPT como representante más conocido – han suscitado últimamente una gran atención por parte del público, hoy en día todo un abanico de tecnologías se engloba bajo el término “IA”. INFORM siempre ha estado desarrollando algoritmos y software para optimizar los procesos empresariales mediante tecnología de inteligencia artificial, con sólidas raíces en las matemáticas avanzadas de la investigación de operaciones (OR). En función del caso de uso específico, utilizamos las metodologías apropiadas según la tarea que nos ocupe, incluidas técnicas de búsqueda y optimización y heurísticas, aprendizaje supervisado y no supervisado, o bien combinaciones de ellas. A esto lo denominamos “estrategia centrada en el cliente” en nuestro enfoque de IA híbrida.

Nuestros más de 1.000 colaboradores atienden a más de 1.000 clientes de todo el mundo, entre los que se incluyen terminales de contenedores, aeropuertos de pasajeros, proveedores de servicios financieros y de telecomunicaciones, operaciones industriales, comercios mayoristas, centros de almacenamiento y transbordo y compañías navieras. Nos avalan más de 50 años de experiencia en procesos empresariales y metodología de IA, incluidas todas las clases de algoritmos de IA que acabamos de mencionar. Sin embargo, no sólo nos interesa la tecnología de IA. Nuestros expertos en el sector aspiran a aprovechar al máximo el potencial de los algoritmos y el software en estrecha colaboración con nuestros clientes, como parte de sus iniciativas de digitalización y optimización. Como parte de nuestro enfoque B2B, nuestras aplicaciones de IA generalmente se aplican a contextos empresariales perfectamente definidos, que habitualmente pueden utilizarse para contener los posibles riesgos relacionados con la IA.

Creemos que los métodos de IA mencionados anteriormente serán objeto de una adopción empresarial aún más amplia en los próximos años y décadas. Nuestra meta es aprovechar cada vez más los avances en metodología y potencia informática en beneficio de nuestros clientes. Esto incluye la IA que facilita aún más datos empresariales (por ejemplo, como parte de la IA de percepción de protección de la privacidad), así como la explotación de datos empresariales en casos de uso de IA de procesos específicos hasta modelos de IA para grandes empresas y modelos lingüísticos para la interacción entre humanos y computadoras.

Celebramos las oportunidades que brindan la investigación y los avances tecnológicos, y creemos que la IA incorpora perspectivas de aumentar la capacidad de acción humana. Como ejemplo, los LLM albergan numerosas perspectivas positivas, no sólo en lo que se refiere a la autoría asistida de textos, sino especialmente a la mejora de las interfaces humano-máquina de los sistemas empresariales. Por otra parte, somos conscientes de que los avances de la IA pueden causar temor y entrañar riesgos. Además, percibimos la necesidad de definir las circunstancias en las que se utiliza la IA, tal y como se establece en nuestras directrices para una IA responsable.

IA RESPONSABLE EN INFORM



Las siguientes directrices se aplican al diseño, desarrollo y mantenimiento de los sistemas de INFORM, independientemente de los métodos de IA en los que se basen. Con ello tratamos de crear software confiable. Nuestro objetivo es mantener un elevado nivel de calidad en la ingeniería de los algoritmos de nuestros sistemas y participar en su aplicación y diseño responsables.

En este contexto, resulta pertinente distinguir entre las distintas aplicaciones de la IA. Por ejemplo, los sistemas de recomendación que suelen utilizarse en el comercio electrónico no suelen ser críticos. Los sistemas de ayuda para la toma de decisiones suelen estar en una categoría similar. Por el contrario, en las aplicaciones derivadas de la IA, las decisiones de un sistema de IA tendrán un impacto más significativo, por ejemplo, al afectar a las personas a nivel personal. Así, por ejemplo, cuando las decisiones algorítmicas basadas en la IA no sólo se utilizan como recomendaciones o ayuda a la toma de decisiones, también pueden repercutir directamente en la vida de los trabajadores. Algunos sistemas de IA incluso deben considerarse esenciales para la vida, por ejemplo, en aplicaciones médicas o de conducción autónoma. La mayoría de los sistemas de INFORM pueden clasificarse como sistemas de recomendación o de ayuda a la toma de decisiones, pero algunos de ellos tienen un impacto mayor.

AWARENESS

DIRECTRICES DE INFORM PARA UNA IA RESPONSABLE

1. IA BENEFICIOSA
2. IA CENTRADA EN LAS PERSONAS
3. IA ALINEADA
4. IA QUE PROTEGE LA PRIVACIDAD
5. IA CONFIABLE
6. IA SEGURA

1.

IA BENEFICIOSA

Nuestro objetivo es crear sistemas de IA que beneficien a los usuarios y a la sociedad. Fomentamos los avances en las operaciones empresariales mediante la adopción de una IA ética. Nos sumamos al debate sobre el impacto social de la IA y el futuro del mundo laboral. Aspiramos a evitar los efectos negativos de la IA en la sociedad y las empresas, como la desinformación, la amplificación del sesgo y el acentuamiento de las divisiones sociales. Nuestra meta es una IA justa, integradora y no discriminatoria. No desarrollamos sistemas que puedan utilizarse con fines malintencionados. En el caso de las aplicaciones derivadas y de importancia vital, promovemos la revisión de nuestros algoritmos de IA por parte de nuestros clientes o de expertos designados por ellos.

2.

IA CENTRADA EN LAS PERSONAS

Creemos que los sistemas de IA deben ayudar a las personas en su trabajo, facilitando y mejorando los procesos de toma de decisiones en el ámbito empresarial. Nuestros sistemas permiten revisar los resultados de los algoritmos de IA. En el caso de aplicaciones de planificación, esto suele implicar que los resultados pueden revisarse antes de ponerlos en práctica. Cuando se trata de tomar decisiones en tiempo real, permitimos una supervisión y auditoría realizadas por personas. Desarrollamos IA que proporciona herramientas a los usuarios en vez de crear agentes totalmente autónomos. Consideramos que la IA debe aumentar la capacidad de acción humana y que la responsabilidad debe seguir correspondiendo a las personas.

3. IA ALINEADA

La inteligencia artificial debe estar en consonancia con los valores y objetivos humanos. En el contexto empresarial, esto significa que los algoritmos deben atenerse a los objetivos de los usuarios y/u otras partes interesadas. La representación de los objetivos humanos y empresariales es una parte integral de nuestra ingeniería de algoritmos continua. Esto conlleva garantizar que se puedan seguir los objetivos éticos. Con los algoritmos de búsqueda y optimización, esto significa que el juicio sobre lo que es una "buena solución" puede controlarse, por ejemplo, a través de una función objetivo. En el aprendizaje automático, los datos de entrenamiento se analizan en busca de sesgo. Siempre que es posible y adecuado, utilizamos una IA que sea comprensible y transparente.

4. IA QUE PROTEGE LA PRIVACIDAD

Estamos comprometidos con la privacidad y la protección de datos y nos ceñimos a la legislación correspondiente. Estamos especialmente atentos a la protección de datos y al uso ético de la IA para casos de uso que tienen un impacto significativo en las personas. Nuestros procesos y prácticas cumplen la normativa del RGPD de la Unión Europea. Hemos implementado prácticas de seguridad de la información y sistemas de gestión para garantizar unos estándares de alta seguridad aprobados por las certificaciones ISO 27001.

5. IA CONFIABLE

Nos comprometemos a ofrecer aplicaciones de IA de alta calidad y confiables, especialmente para aplicaciones derivadas y de importancia vital. Esto significa que debemos asegurarnos de que tales sistemas cumplen los efectos deseados. Para ello, utilizamos las buenas prácticas de la ingeniería de software para diseñar, desarrollar y probar los algoritmos. En el caso de los algoritmos de aprendizaje automático, esto incluye especialmente el análisis de los datos de entrenamiento para detectar sesgos. Los algoritmos de aprendizaje automático se examinan para detectar resultados poco razonables u otros resultados no deseados. Las interfaces de usuario permiten que los resultados algorítmicos sean transparentes para los usuarios. Al utilizar software basado en IA, los registros de auditoría u otras capacidades del software proporcionan medios para supervisar el software de modo que siga siendo confiable en condiciones variables.

6. IA SEGURA

INFORM diseña y desarrolla algoritmos de IA para distintos mercados y su impacto suele estar claramente delimitado al dominio empresarial en el que actúan, con interfaces perfectamente definidas para los dominios empresariales colindantes. Por lo general, esto ocurre siempre con los algoritmos de búsqueda y optimización, así como con los casos de uso específicos de la IA. Cuando se utilicen grandes modelos de lenguaje o lógica de IA similar en los que puedan plantearse problemas de seguridad, INFORM contendrá el impacto de dicho modelo. Cuando la contención al dominio empresarial no es evidente desde el diseño del sistema, el sistema de IA se someterá a una revisión interna para detectar posibles impactos, por ejemplo, para excluir llamadas a API maliciosas, inyección de código, "jailbreaking" u otras prácticas maliciosas.

HACIA UN FUTURO CON IA – DE FORMA RESPONSABLE

Hace tiempo que la tecnología digital, y en particular la IA, dejó de ser un asunto exclusivo de ingenieros. La IA tiene cada vez más influencia en nuestra vida cotidiana, incluso a veces sin que nos demos cuenta. Se espera que las empresas se adapten a este ritmo y que la inteligencia artificial no sólo cambie el mercado laboral, sino que afecte al trabajo de diversas maneras. Creemos que estos avances pueden considerarse positivos si se abordan con responsabilidad. Por otra parte, creemos que una conducta responsable con respecto a la IA es una condición previa para adoptarla en casos de uso de éxito. En estas directrices para una IA responsable, sentamos nuevas bases para apoyar estos avances desde un punto de vista técnico y enmarcarlos en dimensiones éticas y sociales.

En consonancia con los valores esenciales de INFORM, nos comprometemos a mantener un elevado nivel de calidad en el diseño, el desarrollo y el mantenimiento de la inteligencia artificial. Demostrar la viabilidad práctica de la tecnología de algoritmos fue uno de los principales objetivos que impulsaron la fundación de INFORM hace más de 50 años. A fecha de hoy, nos ceñimos a este principio y lo complementamos con un debate sobre el impacto más notable de la tecnología. En este camino, nos complace seguir dando forma al desarrollo del diseño inteligente en las empresas junto con nuestros clientes, proveedores y socios.

TOGETHER

