POLÍTICA DE SEGURIDAD, PRIVACIDAD Y USO RESPONSABLE DE LA INTELIGENCIA ARTIFICIAL

Autor: Arleth Torres Molina Responsable de Seguridad **Organización:** INGENIERÍA INFORMÁTICA EMPRESARIAL, S.L.	**Validado por:** Emilio Bravo García Dirección Técnica **Organización:** INGENIERÍA INFORMÁTICA EMPRESARIAL, S.L.	**Aprobado por:** José Luis Ramírez Dirección General **Organización:** INGENIERÍA INFORMÁTICA EMPRESARIAL, S.L. Firma: Firmado por ***0147** JOSE LUIS RAMIREZ (R: ****7837*) el día 04/02/2025 con un
Fecha: 20/01/2025	**Fecha:** 23/01/2025	**Fecha**: 23/01/2025
Descripción: Establece la política de seguridad, privacidad y uso responsable de la inteligencia artificial de Ingeniería Informática Empresarial, SL.		

Control de Versiones:

Versión	Fecha	Descripción del cambio
1.0	22/01/2025	Primera versión.

1. PRINCIPIOS ÉTICOS, DE PRIVACIDAD Y DE SEGURIDAD DE LA INFORMACIÓN PARA EL USO DE LA INTELIGENCIA ARTIFICIAL

La Dirección de INGENIERÍA INFORMÁTICA EMPRESARIAL, SL (en adelante i2e) concibe y promueve la inteligencia artificial como una tecnología transformadora y una herramienta para la mejora de los procesos internos y del negocio.

Al mismo tiempo, reconoce la necesidad del uso responsable de las soluciones comerciales de la inteligencia artificial (en adelante IA) utilizadas por los usuarios de i2e, así como la importancia de que las soluciones que incorporen algoritmos y tecnologías de IA que i2e desarrolla como proveedor sean éticas, fiables, robustas y cumplan la normativa legal de aplicación.

En consecuencia, la Dirección de i2e ha establecido los principios que regirán de aplicación a todos los empleados para el uso de los soluciones comerciales de IA, así como para la incorporación de las tecnologías de IA en las soluciones que se desarrollan para los clientes que, seguidamente, se resumen.

- **USO ÉTICO DE LA IA**

Se respetarán los derechos y libertades fundamentales de las personas previniendo cualquier efecto discriminatorio en relación con su origen étnico, ideología política, orientación sexual, creencia religiosa, etc.

Asimismo, se comunicará a los usuarios de los algoritmos o tecnologías de IA adoptadas o desarrolladas por i2e aquellas prácticas no autorizadas o que puedan vulnerar el ordenamiento legal, implantando los mecanismos proporcionados de

supervisión de la actividad de los usuarios con objeto de detectar o prevenir un uso no autorizado.

- **USO TRANSPARENTE DE LA IA**

Se dará a conocer y explicará a los usuarios de los algoritmos o soluciones de IA desarrolladas por i2e todos aquellos aspectos necesarios para entender su funcionamiento, las expectativas de sus resultados, así como de las medidas de seguridad de la información implantadas para la protección de la información confidencial y, en particular, de los datos personales tratados.

- **USO INCLUSIVO DE LA IA**

Se planificará e impartirá formación al personal de i2e, capacitándolos en las competencias necesarias para el uso responsable y seguro de los servicios o herramientas de IA, así como para el desarrollo de las soluciones para los clientes que incorporen algoritmos o tecnologías de IA.

Adicionalmente, cuando se desarrollen soluciones para los clientes que incorporen algoritmos o tecnologías de IA se seguirán buenas prácticas de accesibilidad para los usuarios.

- **USO SUPERVISADO DE LA IA**

Se procederá a la verificación por personas expertas del funcionamiento de las herramientas comerciales y sus resultados, actuando preventiva o correctivamente cuando se observen desviaciones.

Asimismo, durante el diseño y el desarrollo de soluciones que incorporen algoritmos o tecnologías de IA se seguirán buenas prácticas de trazabilidad que permitan su auditoría.

- **USO SEGURO DE LA IA**

Las herramientas que i2e ponga a disposición de los usuarios y las soluciones que se desarrollen para clientes formarán parte del alcance del sistema de seguridad y privacidad de la información (SGSPI) de i2e, garantizando la confidencialidad, integridad, trazabilidad y autenticidad de la información manejada y, en especial, de los datos personales tratados, así como la disponibilidad de los servicios para los clientes que incorporen tecnologías de IA.

En particular, se seleccionarán a proveedores que cumplan la presente política y cuyas soluciones o servicios de IA sean conformes con los requisitos de i2e y el respeto del ordenamiento legal.

Asimismo, se gestionarán los riesgos del uso de los servicios y herramientas de IA, así como de las soluciones que se desarrollan para los clientes que incorporen algoritmos o tecnologías de IA.

La Dirección de i2e está comprometida con la asignación de los recursos proporcionales, humanos y materiales, para el logro efectivo de los referidos principios, designando y delegando en la persona responsable de seguridad de la Información la autoridad necesaria para la difusión, supervisión y mejora de la presente política.

INGENIERÍA INFORMÁTICA EMPRESARIAL, S.L.
José Luis Ramirez Terry
Director General
23/01/2025

2. USOS PROHIBIDOS DE LA INTELIGENCIA ARTIFICIAL

No está permitido el uso de soluciones de IA en los casos siguientes:

1. El uso por los empleados de i2e de las herramientas, algoritmos, librerías o tecnologías de IA no autorizadas por la Dirección.

2. El uso por los empleados para fines particulares o beneficio propio de las herramientas de IA que i2e pone a su disposición.

3. El uso de las herramientas de IA que i2e pone a disposición de los empleados para menoscabar o afectar al honor o la dignidad de las personas, así como atentar contra su integridad física o psíquica.

4. El uso de las herramientas de IA que i2e pone a disposición de los empleados para comprometer la seguridad de los servicios propios o de terceros, generando o distribuyendo contenido que facilite lo que sigue:

    a. Amenazas de ingeniería social como la suplantación de identidad, la generación de contenidos engañosos o falsos, etc.

    b. Abuso de la infraestructura o los servicios propios o de terceros, así como daños, interferencias o interrupciones en ellos.

    c. Elusión de los controles o protecciones de seguridad.

5. El desarrollo por i2e de soluciones que incorporen tecnologías IA cuando para su entrenamiento o interacción se haga uso de información confidencial no autorizada.

6. El desarrollo por i2e de soluciones que incorporen tecnologías IA cuando puedan perjudicar a terceros. Los casos de uso de riesgo inaceptable[1] son:

    a. La categorización biométrica e individual para inferir características sensibles (raza, orientación religiosa/ideológica/sexual, etc.).

    b. El reconocimiento de emociones en entornos laborales y educativos, salvo con fines médicos o de seguridad.

    c. La extracción no dirigida de imágenes faciales de internet o circuitos de televisión cerrados (CCTV) para la creación de bases de datos de reconocimiento facial.

    d. La identificación biométrica remota «en tiempo real» o «en diferido» de personas físicas en espacios de acceso público con la finalidad de aplicar la ley, salvo en supuestos expresamente tasados y condicionado a determinados requisitos para fines policiales.

    e. La evaluación social y clasificación de las personas atendiendo a sus circunstancias sociales, características personales o su personalidad, conocidas o predichas, de forma que la clasificación resultante provoque un trato perjudicial, desigual o desfavorable.

    f. La explotación de las vulnerabilidades emocionales de las personas, o de un grupo específico de personas para alterar de manera sustancial el comportamiento de una persona que pertenezca a dicho grupo y provoque o pueda provocar perjuicios físicos o psicológicos a esa o a otra persona.

[1] Artificial Intelligence Act

7. El desarrollo por i2e de soluciones que incorporen tecnologías IA sin el debido análisis y tratamiento de riesgos y la evaluación de impacto en materia de datos personales.

A continuación, se describen algunos casos de uso tales que requieren de un análisis y tratamiento de riesgos, así como la elaboración de una evaluación de impacto en materia de datos personales.

a. la contratación o selección de personas físicas, especialmente para anunciar puestos vacantes, clasificar y filtrar solicitudes o evaluar a candidatos en el transcurso de entrevistas o pruebas.

b. La toma de decisiones relativas a la promoción y resolución de relaciones contractuales de índole laboral, a la asignación de tareas y al seguimiento y evaluación del rendimiento y la conducta de las personas en el marco de dichas relaciones.

c. El perfilado para evaluar determinados aspectos de la vida de una persona (situación económica, salud, preferencias personales, etc.).

d. El inferir las emociones de una persona física en los lugares de trabajo y en los centros educativos, excepto cuando el sistema de IA esté destinado a ser instalado o introducido en el mercado por motivos médicos o de seguridad.

e. La evaluación de la admisibilidad de las personas físicas para acceder a prestaciones y servicios de asistencia pública, así como para conceder, reducir, retirar o recuperar dichas prestaciones y servicios.

f. La evaluación de la solvencia de personas físicas o establecer su calificación crediticia (“social scoring”).

g. El envío o el establecimiento de prioridades en el envío de servicios de primera intervención en situaciones de emergencia, por ejemplo bomberos y servicios de asistencia médica.

8. El desarrollo por i2e de soluciones que incorporen tecnologías IA para la generación, distribución o explotación de ciberamenazas, como las que siguen:

    a. Amenazas de ingeniería social como la suplantación de identidad, la generación de contenidos engañosos o falsos, etc.

    b. Abuso de la infraestructura o los servicios propios o de terceros, así como daños, interferencias o interrupciones en ellos.

    c. Elusión de los controles o protecciones de seguridad.

9. El desarrollo y la comercialización por i2e que vulneren derechos de propiedad intelectual o industrial de terceros.

10. El uso por los empleados de i2e de las herramientas o tecnologías de IA para subir, compartir, divulgar, exponer, procesar información confidencial de la empresa incluyendo, pero no limitándose a:

    a. Código fuente, repositorios, o partes del software desarrollado para terceros sin consentimiento previo del cliente.

    b. Datos confidenciales y datos personales de clientes, proveedores o empleados.

    c. Documentación interna, estrategias, políticas, procedimientos, planes comerciales, etc.

El incumplimiento de lo dispuesto en la presente política por aquellas personas que hagan mal uso, deliberada o involuntariamente, de las soluciones de IA que i2e pone a disposición de los usuarios o bien, aquellas que teniendo conocimiento de prácticas prohibidas no lo notifiquen a la persona responsable de seguridad de i2e, podrán estar sujetos a la imposición de las medidas disciplinarias que resulten de aplicación conforme al régimen de faltas y sanciones estipuladas por la normativa laboral o el convenio colectivo vigente.

La presente política deberá difundirse y ser suscrita por todo el personal y colaboradores de i2e.

Los cambios a la presente Política serán aprobados por la Dirección General de i2e y distribuidos puntualmente por la Responsable de Seguridad de la Información.

3. ACEPTACIÓN

Declaro que he leído y comprendido la presente política de seguridad, privacidad y uso responsable de la inteligencia artificial y mi conformidad expresa con su contenido íntegro.

NOMBRE Y APELLIDOS:

DNI/NIE:

FIRMA Y FECHA:

4. GLOSARIO (INFORMATIVO)

ALGORITMO: Conjunto finito de reglas formales (operaciones lógicas, instrucciones) que permiten obtener un resultado a partir de elementos de entrada. Este conjunto puede ser objeto de un proceso de ejecución automatizado y apoyarse en modelos diseñados mediante aprendizaje automático.

INTELIGENCIA ARTIFICIAL (IA): Conjunto de ciencias, teorías y técnicas cuyo objetivo es reproducir mediante una máquina las capacidades cognitivas de un ser humano. Los avances actuales apuntan a poder confiar a una máquina tareas complejas que antes se delegaba en un ser humano.

Sin embargo, el término inteligencia artificial es criticado por los expertos, que distinguen entre IA “fuerte” (capaz de contextualizar problemas especializados muy diversos de forma completamente independiente) e IA “débil” o “moderada” (que se desempeña extremadamente bien en su campo de entrenamiento). Según algunos expertos, la IA “fuerte” requeriría avances en la investigación básica para poder modelar el mundo en su conjunto y no solo mejoras en el rendimiento de los sistemas existentes.

SISTEMA DE INTELIGENCIA ARTIFICIAL: Un sistema de información basado en una máquina que está diseñado para funcionar con distintos niveles de autonomía y 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 resultados de salida, como predicciones, contenidos, recomendaciones o decisiones, que pueden influir en entornos físicos o virtuales

Los sistemas de inteligencia artificial (IA) son sistemas de software (y en algunos casos también de hardware) diseñados por seres humanos que, dado un objetivo complejo, actúan en la dimensión física o digital mediante la percepción de su entorno a través de la obtención de datos, la interpretación de los datos estructurados o no estructurados que recopilan, el razonamiento sobre el conocimiento o el

procesamiento de la información derivados de esos datos, y decidiendo la acción o acciones óptimas que deben llevar a cabo para lograr el objetivo establecido.

Los sistemas de IA pueden utilizar normas simbólicas o aprender un modelo numérico; también pueden adaptar su conducta mediante el análisis del modo en que el entorno se ve afectado por sus acciones anteriores.

La IA es una disciplina científica que incluye varios enfoques y técnicas, como el aprendizaje automático (del que el aprendizaje profundo y el aprendizaje por refuerzo constituyen algunos ejemplos), el razonamiento automático (que incluye la planificación, programación, representación y razonamiento de conocimientos, búsqueda y optimización) y la robótica (que incluye el control, la percepción, sensores y accionadores así como la integración de todas las demás técnicas en sistemas ciberfísicos).

IA FIABLE: La IA fiable tiene tres componentes: 1) debe ser lícita, es decir, cumplir todas las leyes y reglamentos aplicables; 2) ha de ser ética, demostrando el respeto y garantizando el cumplimiento de los principios y valores éticos, y 3) debe ser robusta, tanto desde el punto de vista técnico como social, puesto que los sistemas de IA, incluso si las intenciones son buenas, pueden provocar daños accidentales.

La fiabilidad de la IA no concierne únicamente a la fiabilidad del propio sistema de inteligencia artificial, sino también a la de todos los procesos y agentes implicados en el ciclo de vida del sistema.

IA ÉTICA: El desarrollo, despliegue y utilización de la IA de un modo que garantice el cumplimiento de las normas éticas, incluidos los derechos fundamentales como derechos morales especiales, los principios éticos y los valores esenciales asociados. Es el segundo de los tres elementos clave necesarios para hacer realidad una IA fiable.

IA ROBUSTA: La solidez de un sistema de IA abarca tanto su solidez técnica (que resulta adecuada en un contexto determinado, como el ámbito de aplicación o la fase del ciclo de vida) así como desde el punto de vista social (garantizando que el sistema de IA tenga debidamente en cuenta el contexto y el entorno en el que opera). Esto es crucial para asegurar que, incluso si las intenciones son buenas, el sistema no provoque daños involuntarios.

La solidez es el último de los tres componentes necesarios para hacer realidad una IA fiable.

TRAZABILIDAD: La trazabilidad de un sistema de IA se refiere a su capacidad para llevar a cabo un seguimiento de los datos, el desarrollo y el proceso de despliegue del sistema, generalmente a través de un proceso de identificación y registro documentados.

CONFIANZA: Pese a que la confianza quizá no sea una propiedad atribuible a las máquinas, la importancia de poder confiar no solo en el hecho de que los sistemas de IA cumplan las leyes y los principios éticos y sean sólidos, sino también en poder confiar en todas las personas y procesos involucrados en el ciclo de vida de los sistemas de IA.

IA CENTRADA EN LA PERSONA: Una IA con un enfoque centrado en la persona se esfuerza por asegurar que los valores humanos ocupen un lugar central en el desarrollo, despliegue, utilización y supervisión de los sistemas de IA, garantizando el respeto de los derechos fundamentales, incluidos los recogidos en los Tratados de la Unión Europea y en la Carta de los Derechos Fundamentales de la Unión Europea; todos ellos constituyen una referencia unitaria a un fundamento común arraigado en el respeto de la dignidad humana, en el que el ser humano disfruta de una condición moral única e inalienable. Esto requiere asimismo tener en cuenta el entorno natural y el resto de seres vivos que forman parte del ecosistema humano, así como un enfoque sostenible que permita la prosperidad de las generaciones futuras.

AUDITABILIDAD: La auditabilidad se refiere a la capacidad de un sistema de IA de someterse a la evaluación de sus algoritmos, datos y procesos de diseño. Constituye uno de los siete requisitos que debería cumplir cualquier sistema de IA fiable.
Esto no implica necesariamente que siempre deba disponerse de forma inmediata de la información sobre los modelos de negocio y la propiedad intelectual del sistema de IA. El hecho de garantizar la existencia de mecanismos de trazabilidad y registro desde las primeras fases de diseño del sistema de IA puede favorecer la auditabilidad del sistema.

CICLO DE VIDA DE LOS SISTEMAS DE IA: El ciclo de vida de un sistema de IA abarca las fases de desarrollo (incluidas las tareas de investigación, diseño, provisión de datos y realización de ensayos limitados), despliegue (incluida la aplicación) y utilización de dicho sistema.

SESGO: Un sesgo es una inclinación que favorece o perjudica a una persona, objeto o posición.
En los sistemas de IA pueden surgir numerosos tipos de sesgos. Por ejemplo, en los sistemas de IA impulsados por datos, como los creados a través del aprendizaje automático, los sesgos en la recogida de datos y la formación pueden dar lugar a sesgos en el sistema de IA. En los sistemas de IA lógicos, como los basados en normas, pueden surgir sesgos como consecuencia de la visión que puede tener un ingeniero del conocimiento acerca de las reglas aplicables en un entorno específico. También pueden aparecer sesgos debido a la formación y adaptación en línea a través de la interacción, o como consecuencia de la personalización en aquellos casos en que se presentan a los usuarios recomendaciones o información adaptadas a sus gustos. Los sesgos no tienen porqué estar relacionados necesariamente con inclinaciones humanas o con la recogida de datos por parte de personas. Pueden surgir, por ejemplo, en los limitados contextos en los que se utiliza un sistema, en cuyo caso no existe la posibilidad de generalizarlo a otros contextos. Los sesgos pueden ser positivos o negativos, intencionados o no. En algunos casos, pueden dar lugar a resultados discriminatorios o injustos, lo que en este documento se denomina «sesgo injusto».

PROVEEDOR: Una persona física o jurídica, autoridad pública, órgano u organismo que desarrolle un sistema de IA o un modelo de IA de uso general o para el que se desarrolle un sistema de IA o un modelo de IA de uso general y lo introduzca en el mercado o ponga en servicio el sistema de IA con su propio nombre o marca, previo pago o gratuitamente.

PARTES INTERESADAS: Todas aquellas partes dedicadas a la investigación, desarrollo, diseño, despliegue o utilización de la IA, así como aquellas que se ven afectadas de forma directa o indirecta por esta, incluidas, con carácter no limitativo, empresas, organizaciones, investigadores, servicios públicos, instituciones, organizaciones de la sociedad civil, gobiernos, autoridades reguladoras, interlocutores sociales, personas físicas, ciudadanos, trabajadores y consumidores.

PERSONAS Y GRUPOS VULNERABLES: Un grupo vulnerable es un grupo de personas que comparten una o varias características de vulnerabilidad.

Debido a su heterogeneidad, no existe una definición generalmente aceptada ni que cuente con un consenso amplio del concepto de «personas vulnerables». Lo que se considera una persona o grupo vulnerable suele depender del contexto. Los sucesos vitales de carácter temporal (como la infancia o la enfermedad), los factores de mercado (como la asimetría de información o el poder de mercado), los factores económicos (como la pobreza), los vinculados a la identidad de las personas (como el género, la religión o la cultura) y otros pueden desempeñar un papel en ese sentido.

La Carta de los Derechos Fundamentales de la Unión Europea recoge en su artículo 21, relativo a la no discriminación, los motivos de discriminación siguientes, que pueden servir como punto de referencia, entre otros: el sexo, la raza, el color, los orígenes étnicos o sociales, las características genéticas, la lengua, la religión o las convicciones, las opiniones políticas o de cualquier otro tipo, la pertenencia a una minoría nacional, el patrimonio, el nacimiento, la discapacidad, la edad o la orientación sexual. En las disposiciones de otras leyes se abordan los derechos de determinados grupos, además

de los enumerados anteriormente. Este tipo de listas nunca pueden ser exhaustivas, y pueden cambiar a lo largo del tiempo.

APRENDIZAJE AUTOMÁTICO: El aprendizaje automático permite construir un modelo matemático a partir de datos que incluyen una gran cantidad de variables que no se conocen de antemano. Los parámetros se configuran a medida que se pasa por una fase de aprendizaje, que utiliza conjuntos de datos de entrenamiento para encontrar vínculos y clasificarlos. Los diferentes métodos de aprendizaje automático son elegidos por los diseñadores en función de la naturaleza de las tareas a realizar (agrupamiento, árbol de decisión).

Estos métodos suelen clasificarse en 3 categorías: aprendizaje supervisado por humanos, aprendizaje no supervisado y aprendizaje no supervisado por refuerzo.

Estas 3 categorías agrupan diferentes métodos, entre los que se incluyen las redes neuronales, el aprendizaje profundo, etc.

METADATOS: Datos utilizados para definir, contextualizar o caracterizar otros datos.

RED NEURONAL (artificial) / NEURONA FORMAL: Sistema algorítmico, cuyo diseño se inspiró originalmente esquemáticamente en el funcionamiento de las neuronas biológicas y que, posteriormente, se aproximó a los métodos estadísticos.

La llamada neurona formal está diseñada como un autómata con una función de transferencia que transforma sus entradas en salidas según reglas lógicas, aritméticas y simbólicas precisas. Reunidas en red, estas neuronas formales son capaces de realizar clasificaciones rápidamente y aprender gradualmente a mejorarlas.

Se utiliza en robótica, traducción automática, etc.

DATOS ABIERTOS: El término se refiere a la puesta a disposición del público, mediante descarga, de bases de datos estructuradas. Estos datos pueden reutilizarse

de manera no lucrativa en las condiciones de una licencia específica, que puede especificar o prohibir, en particular, determinados fines de reutilización.

Los datos abiertos no deben confundirse con la información pública unitaria disponible en sitios de Internet, cuya base de datos no puede descargarse en su totalidad (por ejemplo, las bases de datos de jurisprudencia). No sustituyen la publicación obligatoria de determinadas medidas o decisiones administrativas o judiciales ya adoptadas por determinadas leyes o reglamentos.

Por último, a veces se crea una confusión entre los datos (datos abiertos en sentido estricto) y sus medios de procesamiento (aprendizaje automático, ciencia de datos) para diferentes fines (motores de búsqueda, ayuda a la redacción de leyes, análisis de tendencias jurisprudenciales, anticipación de decisiones judiciales).

DATOS PERSONALES: Información relativa a una persona física identificada o identificable, directa o indirectamente, mediante referencia a uno o varios elementos específicos de dicha persona.

Entre estos, los datos sensibles en el sentido del Reglamento General de Protección de Datos se encuentran los datos personales relativos al origen racial o étnico, las opiniones políticas, las convicciones religiosas o filosóficas, la afiliación sindical, así como datos genéticos, datos biométricos, datos relativos a la salud o relativos a la vida sexual o la orientación sexual.

TRATAMIENTO DE DATOS PERSONALES: Cualquier operación o conjunto de operaciones realizadas o no mediante procedimientos automatizados y aplicadas sobre datos personales o conjuntos de datos, como la recogida, registro, organización, estructuración, conservación, adaptación o modificación, extracción, consulta, utilización, comunicación por transmisión, difusión o cualquier otra forma de puesta a disposición, conexión o interconexión, limitación, supresión o destrucción.

PERFILADO: De acuerdo con el artículo 4 del RGPD, los datos personales se procesan con el fin de evaluar determinados aspectos de la vida de una persona física (situación económica, salud, preferencias personales, etc.).

SEUDONIMIZACIÓN: De acuerdo con el artículo 4 del RGPD, los datos personales ya no podrán atribuirse a un interesado específico sin recurrir a información adicional, siempre que esta información adicional se conserve por separado y esté sujeta a medidas técnicas y organizativas destinadas a garantizar que los datos personales no se atribuyen a una persona física identificada o identificable.

FUENTES:

DIRECTRICES ÉTICAS para una IA FIABLE Grupo de expertos de alto nivel sobre inteligencia artificial
https://digital-strategy.ec.europa.eu/es/library/ethics-guidelines-trustworthy-ai

Consejo de Europa
https://www.coe.int/en/web/artificial-intelligence/glossary