



Decálogo LabS IA Responsable e Inclusiva



Índice Documento

Sergi Biosca - CEO de NTT DATA España	5
Francisco Román - Presidente de Fundación SERES	7
Agradecimiento empresas participantes	8
Resumen ejecutivo	10
Decálogo Común para el Uso Responsable de la IA	
• Introducción	14
• Enunciado 1	20
• Enunciado 2	22
• Enunciado 3	24
• Enunciado 4	26
<i>Empresa destacada: Tendam</i>	28
• Enunciado 5	30
• Enunciado 6	32
• Enunciado 7	34
<i>Empresa destacada: El Corte Inglés</i>	36
• Enunciado 8	38
<i>Empresa destacada: GSK</i>	40
• Enunciado 9	42
• Enunciado 10	44
• Enunciado 11	46
• Enunciado 12	48
<i>Empresa destacada: Meliá Hotels International</i>	50
Conclusiones finales	52



NOS RESERVAMOS EL
DERECHO DE ADICIÓN



Sergi Biosca



Sergi Biosca
CEO de NTT DATA España

En el contexto de disrupción actual, **la Inteligencia Artificial es posiblemente la tecnología con mayor potencial para reformular cómo desarrollamos las actividades humanas.** Asistimos a un impacto cuya dimensión supera el concepto de transformación para aproximarse más acertadamente al de redefinición: **en los próximos años, las organizaciones seremos actores destacados en influenciar nuevos modelos sociales fuertemente marcados por el factor tecnológico.**

La Inteligencia Artificial plantea una serie de retos que tienen origen en valores éticos y trascienden el ámbito tecnológico: no hay duda que el impacto creciente en la toma de decisiones automatizadas plantea tantas oportunidades para la generación de valor **como interrogantes acerca del modelo a construir para situar al individuo en el centro del progreso tecnológico.**

En NTT DATA denominamos este recorrido de Platón a Python. Es nuestra forma de expresar el hilo conductor entre principios éticos aparentemente intangibles como la equidad, con su aplicación práctica a iniciativas de Inteligencia Artificial de compañías de todos los sectores.

¿Cómo podemos fomentar que los algoritmos tomen decisiones justas, previniendo sesgos discriminatorios? ¿Cómo garantizar que la IA impulsará aumentar las capacidades de los individuos para generar crecimiento económico y prosperidad? ¿Cómo podemos contribuir a un mejor futuro del mercado laboral a partir de la automatización? ¿Qué criterios aplicaremos para definir el perímetro de la supervisión humana sobre la tecnología?

Como líderes tecnológicos, **formamos una comunidad junto a las compañías a las que damos servicio y sus clientes.** La IA Ética es un nexo común para definir cómo debemos guiar nuestro comportamiento hacia el conjunto de personas en los que la aplicación de decisiones automatizadas va a tener impacto. Y sobre esta idea debemos generar confianza en la sociedad a la que damos servicio, considerando a la IA dentro de un **ecosistema que comprende a las organizaciones y a las comunidades de las que forman parte, a los sectores público y privado, a la investigación y la academia.**

Con la elaboración de este Decálogo, con la Fundación SERES y compañías líderes del mercado español y global, en NTT DATA damos un paso más en nuestro compromiso con el desarrollo de una Inteligencia Artificial Responsable.

En cada uno de los puntos que lo integran hemos buscado reflejar los valores que han de embeberse en este nuevo modelo de progreso impulsado por la tecnología, y que requiere que atendamos tanto las actividades del día a día de una amplia diversidad de perfiles participantes como una **visión socio-económica a largo plazo que nos conduzca a sociedades avanzadas tecnológicamente y prósperas para el conjunto de la población.**





Francisco Román



Francisco Román
Presidente de
Fundación SERES

La irrupción de la Inteligencia Artificial en las compañías genera grandes oportunidades de crecimiento a la vez que plantea importantes desafíos para alcanzar su desarrollo, teniendo en cuenta los beneficios para toda la sociedad, respetando la privacidad de los datos y protegiéndonos ante posibles sesgos.

En plena 4ª Revolución Industrial no podemos permitirnos dejar a nadie atrás. Resulta crucial abordar las cuestiones éticas que plantea su introducción en las diferentes esferas.

“La dimensión ética de la inteligencia artificial no es una característica de lujo o un añadido: tiene que ser parte integral de su desarrollo”, destacan los principios de la Unión Europea para el desarrollo de la IA. Un planteamiento con el que nos alineamos desde Fundación SERES y NTT DATA y bajo el que iniciamos nuestra labor conjunta.

Y es que **el impulso de una Inteligencia Artificial responsable e inclusiva será posible gracias a la creación de un ecosistema de colaboración** en el que las compañías, instituciones públicas, universidades y ciudadanía aborden conjuntamente los nuevos retos y oportunidades que la IA representa para nuestra sociedad.

En Fundación SERES estamos convencidos de que las empresas, en términos *darwinianos*, son los entes mejor preparados para afrontar este tipo de cambios y contribuir enormemente al desarrollo económico y social.

Ante esta transformación digital, es más necesario que nunca poner el acento en uno de los planteamientos estratégicos de la Fundación: la conexión del progreso económico con el social. **La Inteligencia Artificial ha de gestionarse y abordarse bajo una estrategia que permita a la compañía generar valor compartido. Nuestro foco está puesto en la transformación y adaptación del propósito empresarial.** Creemos firmemente que ser responsable puede ser rentable. La generación de valor de la empresa va más allá de los objetivos de negocio y ha de contemplarse también desde la perspectiva de su contribución al progreso de las comunidades.

Este informe es el resultado de la colaboración, junto a NTT DATA, y otras compañías españolas líderes en este ámbito. Un informe cuyo objetivo se centra en el diseño de soluciones tecnológicas que tengan en cuenta la perspectiva social. Un paso más para garantizar el desarrollo de marcos de actuación que puedan permitir a las compañías, investigadores y, a la sociedad en general, comprender mejor y abordar de forma práctica los desafíos sociales asociados a la implementación de la IA, logrando articular el futuro empresarial bajo el prisma de la responsabilidad y la inclusión social.



Agradecimiento Empresas Participantes

Queremos agradecer a todas las empresas que han participado en las sesiones del **LabS IA Responsable e Inclusiva**, en especial a aquellas empresas que han colaborado con sus aportaciones para la elaboración del informe, tanto con entrevistas individuales como en el cuestionario que dio lugar a los statements del decálogo.

Cuatrecasas

El Corte Inglés

Ferrovial

Fujitsu

Fundación Fernando Pombo

GlaxoSmithKline España

IESE Business School

Ilunion (Grupo Social ONCE)

Meliá Hotels International

Pérez-Llorca

Randstad

Santander

Tendam

Vodafone España

Willis Towers Watson



Resumen ejecutivo

Tras un año de trabajo en el LabS de IA Responsable e Inclusiva, hemos logrado profundizar en la perspectiva social de la IA en los diferentes sectores representados en el LabS, así como la difusión de conocimiento y buenas prácticas del uso responsable e inclusivo de la IA colaborando con las empresas vinculadas al laboratorio a través de sesiones de trabajo en grupo, entrevistas y encuestas.

Como hito principal, se ha definido **un decálogo común para el uso responsable e inclusivo de la IA** en las empresas a través de 12 enunciados que se alinean con las directrices éticas establecidas por la Comisión Europea para una Inteligencia Artificial fiable.

Estos enunciados han sido definidos de manera que reúnan una visión compartida de la IA responsable e inclusiva al mismo tiempo que reflejen el compromiso de cada una de las empresas participantes en el LabS por avanzar hacia un uso responsable de la tecnología. De igual manera, este decálogo puede servir como primera aproximación a un planteamiento de principios éticos interno de cada organización, que guíe sus operaciones y esté en línea con la visión estratégica de la entidad.

Tomando en cuenta el impacto de la IA en sus diferentes fases y aplicaciones, hemos considerado elaborar un decálogo que traslade el compromiso que implica un uso responsable e inclusivo de la IA a través de tres grupos de interés: la Sociedad, el Individuo y las Organizaciones.





Sociedad

Desde el punto de vista de la Sociedad, la IA tiene el enorme potencial de fomentar el progreso social y económico. Como punto de partida, reconocemos la importancia de llevar a cabo una labor de concienciación para impulsar el uso responsable e inclusivo de los sistemas de IA. De igual manera, se contempla el diseño de una ruta de transformación digital responsable amparada por la normativa vigente y regulación europea. Finalmente, el empleo de la IA debe favorecer la inclusión de todos los grupos sociales, en especial el de los más vulnerables, convirtiendo la diversidad cultural y disciplinar en una ventaja competitiva. Reflejamos esta visión a través de los siguientes enunciados:



Compartimos una visión de la IA que promueva el progreso social y económico de las personas, las comunidades y la sociedad en su conjunto”.



Estamos comprometidos con el desarrollo de una IA responsable e inclusiva que respete el marco legal de la Unión Europea”.



Atendemos a la oportunidad de fomentar la colaboración con otras organizaciones para impulsar el uso responsable e inclusivo de la IA”.



Defendemos que la IA debe respetar la diversidad y fomentar la inclusión de los individuos en las comunidades, evitando un impacto negativo en los derechos fundamentales de las personas a causa de sesgos”.



Individuo

Desde el punto de vista del Individuo, el empleo de una IA responsable e inclusiva asegura la libertad y autonomía de las personas, características innatas en el ser humano. Además, el correcto desarrollo de las tecnologías debe asegurar una estricta protección de aquellos datos sensibles cuyo uso inapropiado pueda impactar negativamente en las personas. Por ello, consideramos que finalmente la tecnología tiene que responder a un beneficio tangible y transparente de cara a las personas, protegiendo especialmente los derechos fundamentales en todo momento. En consecuencia, hemos definido los siguientes enunciados:



Consideramos que la privacidad de las personas es un derecho que ha de ser protegido especialmente en el desarrollo de iniciativas de IA”.



Promovemos un diseño de la IA que proteja la autonomía de las personas”.



Reconocemos la importancia de diseñar sistemas de IA robustos que prevengan un uso malintencionado, inapropiado o malicioso de la tecnología”.



Sostenemos una aproximación a la Inteligencia Artificial que, desde el diseño, sitúe a las personas en el centro de las oportunidades y beneficios en el uso tecnológico”.



Organización

Y, por último, **desde el punto de vista de la Organización**, entendemos los beneficios de desarrollar una IA sólida y segura atendiendo a los diferentes riesgos implicados en cada fase del ciclo de vida tales como posibles malversaciones, ataques y sesgos. Trasladamos esta visión a través de una transformación cultural que implique a las personas bajo un nuevo modelo colaborativo “humano-máquina”, aumentando las capacidades de supervisión y redundando en nuevas responsabilidades para un mayor desarrollo profesional. Así es como hemos elaborado los enunciados:



Creemos en un modelo de colaboración entre la IA y las personas que redunde en aumentar las capacidades de los empleados y su aportación de valor a las organizaciones”.



Concebimos una cultura organizativa que fomente la responsabilidad a través de la formación tecnológica y la comprensión de las implicaciones éticas de la IA”.



Nos identificamos con un modelo de desarrollo responsable e inclusivo de la IA que integre diversidad de conocimientos y dominios de negocio, así como sensibilidades y enfoques”.



Entendemos la necesidad de orquestar internamente el gobierno de la IA a través de procesos y herramientas, que lo tangibilicen en todo el ciclo de vida de las iniciativas de IA e identifique claramente la responsabilidad de sus participantes”.

Decálogo común para un uso responsable de la IA

Introducción

El auge de la inteligencia artificial (IA) en los últimos años, principalmente impulsada por una creciente disponibilidad de datos y capacidad computacional, se ha traducido en una avalancha de aplicaciones tecnológicas que forman parte de nuestras interacciones diarias tanto en el mundo digital como físico.

A pesar de la aparente omnipresencia de la IA, su característica casi camaleónica, así como la interdependencia con otras tecnologías, dificulta su explicación.

Responder a una definición universal de la IA que incorpore todas las capacidades de esta tecnología, cuyos avances siguen acelerándose y redefiniendo sus límites a un ritmo exponencial, representa un reto para muchas empresas y organizaciones. La propia definición del término “inteligencia” conlleva a una complejidad discutida por expertos y académicos, algunos como Larry Teller concluyendo en que “la inteligencia es lo que las máquinas aún no han hecho”.¹

En 2019, la Comisión Europea a través de un informe elaborado por el Grupo de Expertos de Alto Nivel de IA, conformado en 2018, extiende y enriquece la definición de la IA concluyendo en la siguiente:



Los sistemas de inteligencia artificial (IA) son sistemas de software (y posiblemente también hardware) diseñados por humanos que, dado un objetivo complejo, actúan en la dimensión física o digital percibiendo su entorno a través de la adquisición de información, interpretando los datos estructurados o no estructurados recopilados, razonando sobre el conocimiento, o procesando la información derivada de estos datos y decidiendo la mejor acción a tomar para lograr el objetivo dado².

¹ Cita de Larry Tesler situada en su propio CV alrededor del año 1970. Ver CV de Larry Tesler, apartado “Adages”: http://www.nomodes.com/Larry_Tesler_Consulting/Adages_and_Coinages.html

² A DEFINITION OF AI: MAIN CAPABILITIES AND DISCIPLINES, informe publicado en Abril 2019, elaborado por el Grupo de Expertos de Alto Nivel sobre Inteligencia Artificial conformado por la Comisión Europea.





Dentro de un marco operativo podemos considerar que las técnicas de la IA a través de sus diferentes aplicaciones, se materializan en tres capacidades clave:

1. Generar valor a través del procesamiento de datos.
2. Generar nuevos datos que soporten la toma de decisiones.
3. Generar nuevas interacciones, teniendo la capacidad de adaptarse a nuevos entornos.

Este hecho ofrece nuevas oportunidades tanto para el sector privado como el sector público: favorece el uso eficiente de los recursos, incrementa las capacidades cognitivas y laborales de los trabajadores, aumenta el conocimiento del cliente y facilita el desarrollo de los nuevos modelos de negocio, entre otros.

Desde un punto de vista del sector público, se espera que la IA fortalezca los cimientos sociales, donde se desarrolle un ecosistema que acerque las ventajas de la tecnología a la sociedad. Un estudio publicado por la revista Nature³ ha demostrado que la IA puede impulsar 128 de las 169 metas recogidas en los 17 Objetivos de Desarrollo Sostenible (ODS)⁴, marcados por la organización de Naciones Unidas en su Agenda 2030.

³ El papel de la inteligencia artificial en la consecución de los Objetivos de Desarrollo Sostenible, se trata de un estudio académico elaborado por Ricardo Vinuesa, técnico mecánico de KTH (Centro de Investigación Científico y Tecnológico situado en Estocolmo, Suecia).

⁴ 17 Objetivos de Desarrollo Sostenible (ODS) que aprobó la Organización de las Naciones Unidas (ONU) en 2015.

Entre ellas se encuentran el contribuir al desarrollo sostenible mediante el lanzamiento de nuevos servicios que favorezcan la inclusión financiera y la educación, el fomento de los sistemas bajos en carbono, el impulso en el empleo de las energías renovables y el apoyo al crecimiento de ciudades inteligentes.

Estas ventajas también se extienden al ámbito empresarial, donde las técnicas de IA se fusionan con las necesidades del negocio y la cultura propia de cada entidad, siempre atendiendo al impacto ético y social de su utilización. De este modo, la IA promueve seis oportunidades de mejora basadas en la transformación interna (competitividad operacional, ventaja organizativa e innovación), abre las puertas a nuevos modelos de negocio, da forma a nuevos productos y servicios, y potencia la experiencia con el cliente.



Si bien la IA conlleva beneficios tangibles para los ciudadanos y consumidores, las empresas y el sector público se encuentran bajo una creciente presión por asumir mayor responsabilidad de cara a promover el uso ético de la tecnología.

De acuerdo con los resultados obtenidos en el informe “Eurobarómetro Estándar 92”⁵, España destaca (con un 51%) como segundo país a nivel europeo en el que los usuarios demandan ser informados en todos aquellos servicios o aplicaciones que utilicen inteligencia artificial. La alta sensibilización de los usuarios españoles de cara al uso de la IA se traduce de igual manera en un alto porcentaje de usuarios (50%) que considera que la intervención de políticas públicas es necesaria para el correcto desarrollo y ejecución de las aplicaciones de inteligencia artificial.

⁵ “Euro barómetro Standard 92 - Europeos y la Inteligencia Artificial” es un informe elaborado por la Comisión Europea en el mes de Noviembre de 2019, en el cual se analizan los principales problemas que tanto la ciudadanía española como el conjunto de la población europea identifican como más relevantes, en lo relativo al mercado de trabajo, la economía y el nivel de satisfacción sobre la calidad de vida.

Atendiendo al enorme impacto de la IA y los retos que conlleva, la Comisión Europea promueve la necesidad de ganar confianza entre los agentes económicos y fomenta el respeto a los 4 Principios Éticos (respeto de la autonomía humana, prevención del daño, equidad y explicabilidad) y Derechos Fundamentales (como la dignidad, la igualdad, la protección de la privacidad, la justicia, etc.). Esta visión tiene como objetivo enfocar el uso de la IA a nivel europeo de manera que satisfaga las aspiraciones de los ciudadanos, responda a las necesidades de la sociedad y aumente la competitividad.

A lo largo de un año de trabajo en el LabS de IA Responsable e inclusiva, hemos explorado los potenciales riesgos que conlleva la IA creando un espacio común para fomentar el debate en torno a retos y mejores prácticas. De esta manera consideramos que hemos logrado un objetivo clave al impulsar una mayor sensibilización en diferentes perfiles dentro de las empresas con respecto al uso responsable e inclusivo de la IA.

Como hito de este primer año, hemos diseñado, en conjunto con las empresas participantes del LabS, una declaración común con 12 enunciados que reflejan una visión compartida definiendo la aproximación responsable e inclusiva a la puesta en práctica de la IA en los diferentes sectores.

Estos enunciados han sido elaborados respondiendo a tres grupos de interés clave dentro del desarrollo y uso de la IA: Sociedad, Individuo y Organización. De igual manera, tomamos como base los siete requisitos éticos para una IA fiable establecidos por la Comisión Europea, alineándonos con los objetivos del marco europeo para lograr un ecosistema de excelencia y confianza.



Decálogo común para
el uso responsable
e inclusivo de la IA

ENUNCIADO 1

1

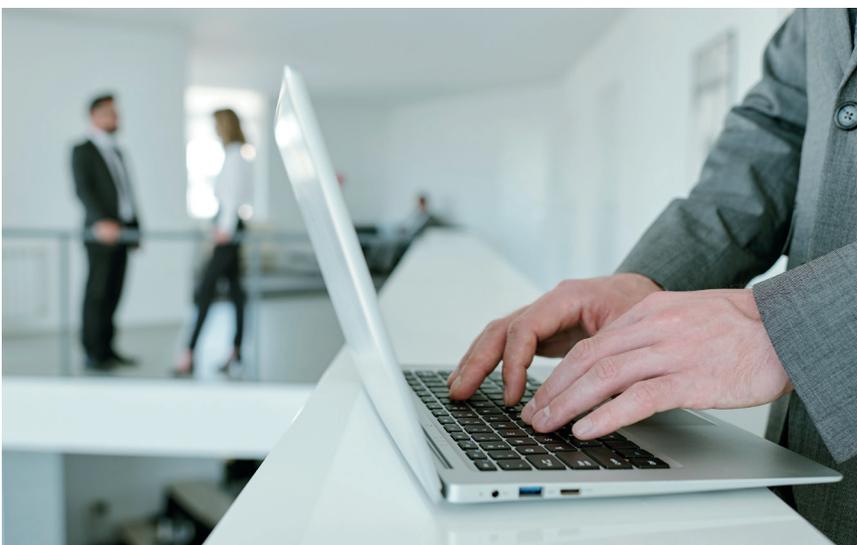
Compartimos una visión de la IA que promueva el progreso social y económico de las personas, las comunidades y la sociedad en su conjunto.

Más que nunca, se reconoce el enorme potencial en las aplicaciones y usos de la IA y se vela por un desarrollo sostenible, que promueva un beneficio **medioambiental y social responsable e inclusivo**. En este sentido, las empresas, siendo agentes de cambio en la sociedad, tienen la oportunidad de liderar una innovación responsable, desarrollando sistemas de IA fiables e inclusivos.

En su máxima expresión, se busca que **aumenten las oportunidades y beneficios, promuevan el desarrollo social e impulsen el progreso económico**.

Si bien es cierto que los sistemas de IA se pueden utilizar para mejorar las habilidades sociales, también pueden **contribuir a su deterioro**. Pudiendo influir negativamente en el bienestar físico y mental de las personas, afectar los derechos fundamentales, restringir el acceso a oportunidades y ralentizar el desarrollo social.

En situaciones en las que existan tales riesgos, deberán llevarse a cabo prácticas de monitorización y control exhaustivo de los sistemas y realizarse una evaluación del impacto y de riesgos, para modificar o justificar su uso según sea necesario, respetando los derechos y libertades de los individuos.





Insights empresas

Todas las empresas participantes coinciden en que la IA entraña beneficios que pueden contribuir al progreso de las personas y de las empresas, impulsar el diseño de ciudades inteligentes y crear oportunidades de crecimiento profesional y sostenibilidad medioambiental. Algunas empresas han comentado ejemplos que redundan en fomentar el progreso social, como mejorar la calidad de vida de las personas, contribuir a un crecimiento sostenible, hacer un uso eficiente de los recursos internos de la empresa para reducir su huella de carbono, etc.

ENUNCIADO 2

2

Estamos comprometidos con el desarrollo de una IA responsable e inclusiva que respete el marco legal de la Unión Europea.

A medida que han ido avanzando las sociedades, se ha configurado un marco normativo común para fomentar el progreso y bienestar de todos los ciudadanos e instituciones. De esta forma, reconocemos los avances por parte de la Unión Europea para crear un “ecosistema de excelencia y confianza”⁶ basado en una estructura legal sólida que integre un conjunto común de normas, directrices y estándares responsables e inclusivos de IA, que regule sus usos y aplicaciones, y defina las responsabilidades de los usuarios.

Un claro ejemplo de ello es RGPD⁷ elaborado por la UE en 2016 en el cual ya se configuran ciertos requisitos legales entorno al uso responsable de la IA estableciendo el derecho de los usuarios a demandar explicabilidad de los sistemas de decisión automatizada según el Considerando 71. Con la continua evolución de la tecnología, estas leyes deben reevaluarse periódicamente y mantenerse actualizadas con los últimos avances tecnológicos. Sobre este punto, la Comisión Europea ha realizado en los últimos años un trabajo ininterrumpido de investigación a través del Grupo de Expertos de Alto Nivel de IA, concluyendo en una primera fase con la evaluación inicial de impacto de cara a una futura propuesta de reglamento prevista para el primer trimestre del 2021.⁸

Diferentes voces dentro de la industria tecnológica han hecho un llamado a los gobiernos en favor de un marco regulatorio que regule el uso de la IA.⁹ Algunos actores incluso consideran favorable crear un **órgano de gobierno global** que haga frente a preguntas clave, como quién será el responsable cuando la IA provoque daños, o bien, si las IA deben tener derechos fundamentales, al igual que sus contrapartes humanas.

⁶ Libro Blanco sobre la inteligencia artificial - un enfoque europeo orientado a la excelencia y la confianza, Comisión Europea, ver: https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_es.pdf

⁷ El Reglamento General de Protección de Datos (GDPR) (Reglamento 2016/679) elaborado por el Parlamento Europeo, el Consejo de la Unión Europea y la Comisión Europea. Este reglamento tiene el objetivo de reforzar y unificar la protección de datos para todos los individuos dentro de la Unión Europea (UE). Así como dar control y unificar el entorno regulador que impacta directamente sobre el almacenamiento, procesamiento, acceso, transferencia y divulgación de los registros de datos de un individuo y afecta a cualquier organización a nivel mundial que procese datos personales de ciudadanos de la Unión Europea.

⁸ Para acceder al portal de la iniciativa de regulación de la Comisión Europea en relación con la IA véase: <https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12527-Requirements-for-Artificial-Intelligence>

⁹ Artículo de The Wall Street Journal, Enero 2020: “Tech Giants’ New Appeal to Governments: Please Regulate Us” (Nuevo llamamiento de los gigantes tecnológicos a los gobiernos: por favor regúlenlos)

Desde la perspectiva empresarial, se pueden establecer mecanismos internos de regulación y control, tales como **un comité de IA responsable e inclusiva** o **un gobierno de la IA**, que garanticen amparar los valores de la empresa y los derechos de los individuos mientras se fomente la innovación responsable de la IA.

Para tener éxito, la **regulación de la IA se vislumbra de carácter internacional**. De lo contrario, se tendrá un mosaico desordenado de diferentes regulaciones extremadamente complejo de cumplimentar y con un alto coste económico y social.



Insights empresas

Se considera que el cumplimiento normativo es esencial para el correcto desarrollo de la actividad empresarial y llevar a cabo prácticas responsables e inclusivas de IA. Sin embargo, algunas empresas creen que actualmente no existe un cuerpo regulatorio riguroso que establezca las pautas y guías a seguir por todos los agentes. Por ello, diversas entidades han decidido fomentar la educación y formación tecnológica, para que sean los propios trabajadores quienes sean responsables de establecer estándares éticos alineados con los principios de la UE. De igual manera, algunas entidades se apoyan en otros cuerpos normativos y éticos dentro de sectores específicos (en Banca tales como “Los Principios de una Banca responsable” o “Los Principios de Inversión Responsable”) cuyas recomendaciones guían las prácticas de IA hacia un uso responsable y sostenible de la misma.

ENUNCIADO 3

3

Atendemos a la oportunidad de fomentar la colaboración con otras organizaciones para impulsar el uso responsable e inclusivo de la IA.

Los avances en el desarrollo responsable e inclusivo de la IA se conciben desde una perspectiva colaborativa entre los diferentes actores de la sociedad. Cada vez son más las empresas que deciden dar un paso al frente y establecer alianzas que permitan la integración de nuevos sistemas y soluciones para generar productos y servicios innovadores bajo un alineamiento de mejores prácticas y consideraciones sociales.

Para ello se están creando **Centros de Excelencia en materia de IA, laboratorios de colaboración interempresarial y grupos de expertos independientes** cuya misión principal sea desarrollar metodologías y mejores prácticas en el uso responsable e inclusivo de la IA. Siempre operando dentro de un marco normativo, estas alianzas impulsan diferentes líneas de acción dentro de la IA responsable e inclusiva: fomentando las capacidades y conocimientos de las personas, incrementando el valor de los resultados desarrollados por los procesos de análisis masivo de datos, o incluso generando propuestas de desarrollo sostenible medioambiental y social.

En este contexto se destacan:

- **Alianza colaborativa entre Google, Facebook, Amazon, IBM y Microsoft** anunciada en septiembre del 2016 en materia de inteligencia artificial (IA). Esta alianza recibe el nombre de **“Partnership on Artificial Intelligence to Benefit People and Society”** o, **“Partnership on AI”**. Entre sus objetivos están dar a conocer las ventajas e investigar en materia de IA y desarrollar unos estándares que garanticen la seguridad, ética y transparencia.
- El IEEE (**Institute of Electrical and Electronics Engineers**) trabaja en una **Iniciativa Global sobre Ética de Sistemas Autónomos e Inteligentes** y lidera la propuesta de estandarización de principios éticos en IA a través de su informe denominado **“Estándares IEEE P7000™”**. Especializándose en áreas como la privacidad de datos, mitigación de sesgos algorítmicos, elaboración de estándares de gobernanza de datos para menores, etc.
- **En el escenario español, destaca la iniciativa independiente OdiseIA.** **OdiseIA** es un observatorio multidisciplinar dedicado a observar, prevenir y mitigar los desafíos del uso de la inteligencia artificial como oportunidad disruptiva. Es un espacio de colaboración (sin ánimo de lucro) formado por empresas, universidades, instituciones y personas físicas que conjuntamente velan por el buen uso de la IA y el empleo más responsable de las tecnologías.



Insights empresas

Cada vez son más las empresas que buscan generar oportunidades mediante la colaboración con otras entidades del sector, potenciando sus ventajas competitivas y beneficios sociales. Estas colaboraciones se conciben como foros abiertos para compartir conocimientos y poner en común buenas prácticas, respetando siempre los secretos comerciales e información confidencial de las entidades.

ENUNCIADO 4

4

Consideramos que la privacidad de las personas es un derecho que ha de ser protegido especialmente en el desarrollo de iniciativas de IA.

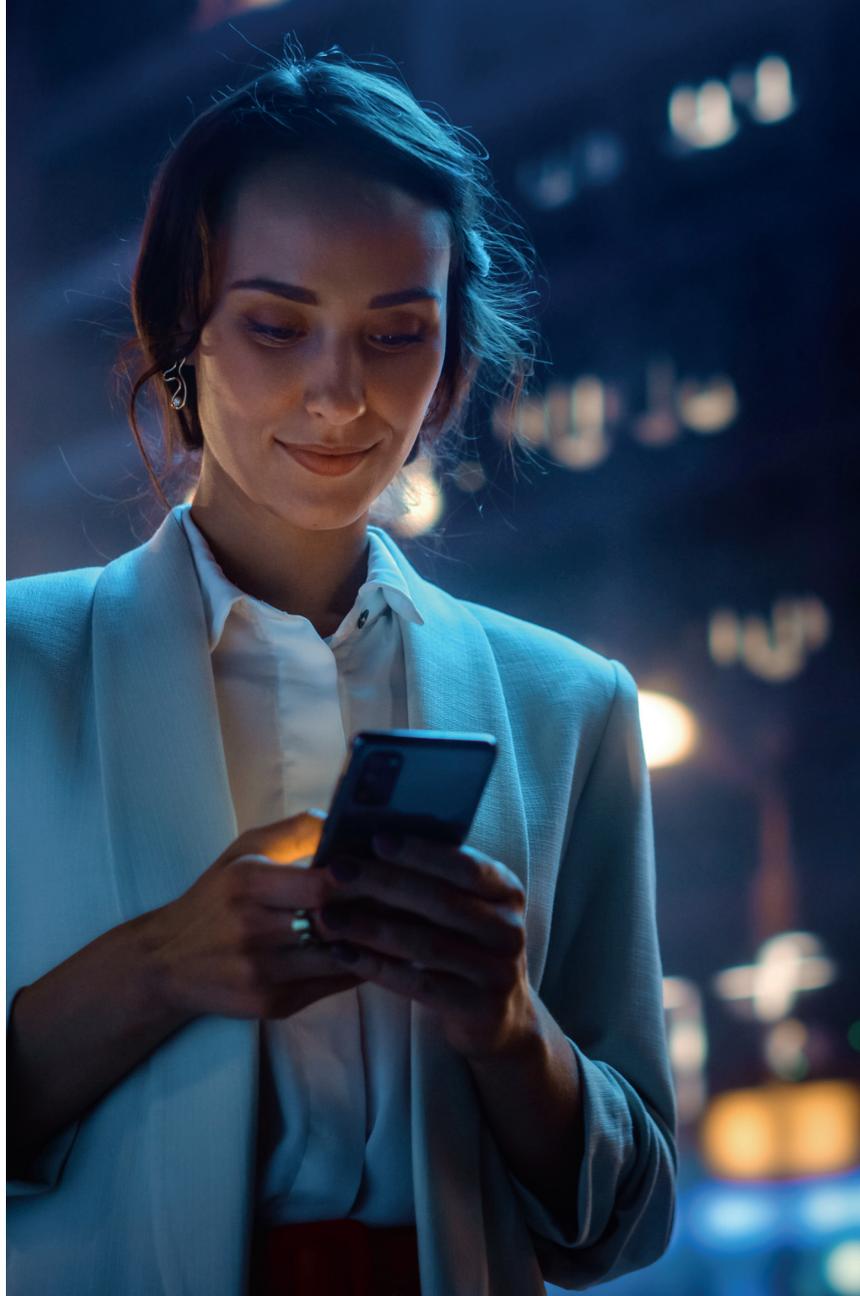
Garantizar la confidencialidad y privacidad de la información del cliente y los trabajadores, así como proteger la propiedad intelectual comercial son de suma importancia para el desarrollo sostenible de la actividad económica de las empresas a medida que se integran e implementan las tecnologías de inteligencia artificial.

El éxito de un algoritmo depende de la cantidad y calidad de los datos con los que entrena. Esto implica la necesidad de acceder a una gran cantidad de información para crear algoritmos efectivos. En los últimos años, las empresas han priorizado sus esfuerzos de cara a crear controles de acceso más seguros y robustos para evitar malversaciones de datos, especialmente cuando se trata de información sensible.

En la misma línea, pueden surgir otros riesgos relacionados a la privacidad a partir del análisis y procesamiento de datos. Por ejemplo, **la IA podría inferir o deducir otra información y hábitos personales que un cliente no desea compartir**. Los registros digitales de transacciones y comportamientos de un usuario en un ámbito concreto pueden permitir a los sistemas de IA inferir no solo las preferencias de las personas, sino también su orientación sexual, edad, género, opiniones religiosas o políticas.

Para evitar que los resultados de la IA deriven en usos no intencionados, las empresas pueden realizar controles a lo largo del ciclo de vida del algoritmo, que permita la detección temprana de posibles riesgos.

A medida que las empresas recopilan, almacenan y procesan una mayor cantidad y diversidad de datos para procesarlos con algoritmos de IA, **garantizar la seguridad de esos datos con una sólida supervisión humana no solo es esencial para avanzar en los objetivos comerciales, sino también es una responsabilidad ética**.



Insights empresas

El manejo de grandes cantidades de información supone un riesgo para la privacidad y seguridad si no se llevan a cabo prácticas y controles exhaustivos que velen por salvaguardar la información. Por ello, la práctica más extendida entre las empresas es dar cumplimiento estricto al Reglamento de Protección de Datos de la UE (GDPR) en toda la materia relativa a la protección de datos. Asimismo, cada vez son más las entidades que apuestan por la formación de equipos especializados en Data Science, que se encarguen de proteger los datos de todos los agentes frente amenazas y riesgos en ciberseguridad.

EMPRESA DESTACADA

Tendam

Tras 140 años operando en el sector textil con un portfolio de marcas diversas, Tendam posee una ventaja competitiva que la posiciona frente a otros competidores del mercado: sus clubs de fidelización con 26 millones de socios, que le permiten hacer un análisis profundo sobre las tendencias y preferencias de sus clientes, y un uso más eficiente de los recursos internos.

Si esta información no se aborda con la debida diligencia, podría suponer un riesgo para la privacidad y seguridad de los datos. Por ello Tendam cuenta con rigurosos protocolos bajo el marco del GDPR para velar por la protección de los datos de sus clientes y no clientes frente a potenciales riesgos. El equipo especializado en Data Analytics trabaja sin acceder a datos personales, sólo se utilizan códigos numéricos que no identifican a los clientes.



El tratamiento de datos se expone tomando en cuenta los principios de explicabilidad de los modelos y el alcance de los mismos para su uso en la toma de decisiones. Este proceso se encuentra bajo vigilancia para supervisar que el tratamiento de los datos dentro los límites de consentimiento.

Carlos Abellán Ossenbach, director del área de Data Analytics puntualiza la importancia de salvaguardar la información, y nos informa de los riesgos potenciales del empleo desafortunado de la IA. Por citar algunos ejemplos textuales:

“

“En Tendam complementamos el desarrollo de modelos predictivos con una analítica descriptiva exhaustiva para entender bien todas las casuísticas y la causalidad del comportamiento de los clientes, de manera que podamos mitigar el potencial sesgo de un modelo “caja negra”.

Finalmente, Ana Fombella, directora de Comunicación y Responsabilidad Corporativa, pone especial hincapié en la labor que está desarrollando Tendam en hacer un uso más responsable de la IA a través de prácticas que aseguren la privacidad y seguridad de los datos y que su tratamiento traslade un beneficio tangible a sus clientes. Esta visión se comparte con el propósito de poder escalar la IA Responsable a todas las áreas de la empresa.

ENUNCIADO 5

5

Defendemos que la IA debe respetar la diversidad y fomentar la inclusión de los individuos en las comunidades, evitando un impacto negativo en los derechos fundamentales de las personas a causa de sesgos.

Estrechamente **relacionado con el principio ético de equidad**, se recomienda llevar a cabo prácticas que favorezcan la inclusión y la diversidad en todo el ciclo de vida del algoritmo, para lograr una IA confiable y robusta. Desde la fase conceptual de los sistemas de IA, esto implicaría garantizar la igualdad de acceso de todas aquellas partes que tengan un interés especial a través de procesos de diseño inclusivos e igualdad de trato.

Los sistemas de IA son en gran parte una reproducción de comportamientos humanos reflejando sesgos, que de manera consciente o no, pueden ser introducidos a través de los datos de entrenamiento o durante las fases de desarrollo del algoritmo.

Por ejemplo, los algoritmos de aprendizaje automático entrenados con datos sesgados o incompletos pueden derivar en la potencial exclusión de ciertos grupos vulnerables al acceso de servicios y productos.

Un estudio llevado a cabo por el **MIT CSAIL**¹⁰ mostró que un CV idéntico tiene un 50% más de probabilidades de resultar en una invitación a una entrevista si el nombre del candidato es europeo-americano, que si es afroamericano. Los últimos resultados sugieren que los algoritmos estarán plagados de los mismos prejuicios sociales, a menos que estén programados explícitamente para abordar estas cuestiones.

Este tipo de ejemplos demuestra que lejos de ser “neutral”, la IA tiene la capacidad de reproducir prejuicios sociales que afecten el progreso de los individuos. Abordando estas preocupaciones y contrarrestándolo con herramientas y metodologías inclusivas y transparentes, es posible convertir a la IA en un aliado clave para ayudar a las personas a tomar decisiones más justas. El diseño responsable de la IA tiene como objetivo evitar que las decisiones de los algoritmos perpetúen y amplifiquen los sesgos y desigualdades sociales ya presentes en nuestra cultura, evitando una discriminación aún mayor.

Para ello existen mecanismos que a lo largo del ciclo de vida ayudan a los científicos de datos a identificar potenciales sesgos tanto en la recopilación como en el desarrollo de los sistemas de IA.

¹⁰ MIT CSAIL. (2004, junio). “Are Emily and Greg More Employable than Lakisha and Jamal?” Un experimento de campo sobre la discriminación en el mercado laboral https://scholar.harvard.edu/files/sendhil/files/are_emily_and_greg_more_employable_than_lakisha_and_jamal.pdf

Los algoritmos deben ser fiables y sólidos en todas las fases del ciclo del sistema. Por tanto, para minimizar el riesgo se pueden establecer **procesos de supervisión** que analicen las limitaciones de los datos, los procesos de desarrollo y validación así como los resultados de los sistemas de IA de manera clara y transparente. Otra línea de acción es implementar una monitorización continua de los resultados asegurando el control de sesgos y el buen funcionamiento de los algoritmos en producción.

Asimismo, **los equipos** involucrados en la construcción, implementación y supervisión de iniciativas **de IA se encuentran en una posición clave para liderar la IA responsable e inclusiva y ser embajadores de prácticas que lleven a minimizar los potenciales daños.**

Por ello, **la multidisciplinaridad y diversidad de culturas en estos equipos generan un valor diferencial al construir algoritmos y soluciones mejor adaptados a los usuarios desde perspectivas más inclusivas.**



Insights empresas

Algunas entidades concuerdan en que controlar los patrones discriminatorios en el algoritmo es una tarea compleja dado a que los sesgos pueden introducirse a lo largo de diferentes fases del ciclo de vida, por lo cual es clave contar con metodologías y mecanismos de control adecuados. Una de las prácticas más empleadas por las empresas para mitigar los sesgos discriminatorios es fomentar la diversidad y la colaboración entre grupos multidisciplinarios en el entorno laboral, obteniendo una IA más objetiva, robusta e inclusiva, cuyos resultados podrán ser escalables a las diferentes áreas de la empresa.

ENUNCIADO 6

6

Promovemos un diseño de la IA que proteja la autonomía de las personas.

La incorporación de robots y de sistemas de IA está cambiando desde ya la forma en que los humanos interactúan entre sí y toman decisiones. En este escenario, la transparencia de los sistemas de IA, la supervisión humana y la reeducación (en inglés - “Reskilling”) pueden ayudar a esclarecer quién toma las decisiones y por qué.

El Principio Ético de Autonomía del Usuario debe tener un papel fundamental que asegure la correcta funcionalidad del sistema y aumente las capacidades del individuo. En este sentido el Grupo de expertos de alto nivel sobre IA señala lo siguiente en la publicación de abril de 2019:

“

Los sistemas de IA no deberían subordinar, coaccionar, engañar, manipular, condicionar o dirigir a los seres humanos de manera injustificada. En lugar de ello, los sistemas de IA deberían diseñarse de forma que aumenten, complementen y potencien las aptitudes cognitivas, sociales y culturales de las personas”.

Las decisiones autónomas e informadas de los usuarios son el resultado de proporcionar el conocimiento y las herramientas para comprender e interactuar con la IA, e incluso, autoevaluar o desafiar razonablemente al programa. **Mediante el empleo de IA, las personas tomarán mejores decisiones que les aproximen más a sus objetivos.**

Otra línea de acción supone implementar mecanismos de “feedback” dentro de los sistemas de IA, permitiendo a los usuarios comunicar errores y solicitar mayor información sobre el resultado del sistema.

De esta manera se busca evitar el uso de aquellos algoritmos diseñados para influir en el comportamiento humano a través de mecanismos difíciles de detectar, aprovechando los procesos subconscientes que amenazan la autonomía individual.



Insights empresas

El principio de autonomía en la voluntad de las personas es cada vez más complejo de diferenciar de la toma de decisiones automatizada de las máquinas, debido a que los procesos decisorios se ven influenciados por los algoritmos de ML. Las empresas destacan que la tecnología, no debe socavar, sino fomentar el desarrollo de las personas y facilitar la toma de decisiones, donde prime la autonomía del individuo.

Un modelo de negocio responsable se logra cuando se cumple con el Principio de Autonomía, que sitúa al ser humano en una posición de libertad frente a la IA. Por ello, las organizaciones impulsan el empoderamiento de su personal, dotándoles de información relevante y necesaria para respaldar una toma de decisiones autónoma en la cual se consideren los límites de aquellas recomendaciones dadas por los sistemas de IA.

ENUNCIADO 7

7 Creemos en un modelo de colaboración entre la IA y las personas que redunde en aumentar las capacidades de los empleados y su aportación de valor a las organizaciones.

Para maximizar el valor de las iniciativas de IA, las empresas necesitan incorporar una estrategia organizativa que permita combinar las capacidades de sus trabajadores con **las capacidades de la tecnología de una manera efectiva y eficiente, creando nuevos tipos de oportunidades laborales.**

El estudio llevado a cabo por el Observatorio de Recursos Humanos¹¹ en materia de transformación tecnológica e inclusión laboral, trata de conciliar una visión empresarial que **rompa con el miedo** de las personas a los cambios que conlleva la transformación tecnológica en el entorno empresarial.

Se demuestra que el rol más importante de las empresas es reconocer y **enfatar el valor que aporta el trabajador** en el ciclo de vida de IA, situando a **la persona en el centro de la estrategia tecnológica**. De manera que se ponga en valor e impulsen sus habilidades, **evitando que su contribución se vuelva invisible u opaca.**

La tecnología debe de servir para mejorar el proceso de toma de decisiones de las personas. Un ejemplo sería la adopción de tecnologías RPA (en inglés - *“Robot Process Automation”*) la cual permite automatizar procesos, mientras da lugar a una mayor dedicación del personal a tareas de supervisión y generación de valor. De forma que los sistemas inteligentes desarrollan tareas más repetitivas y rutinarias; mientras que los trabajadores adoptan nuevos roles que exijan un mayor grado de responsabilidad ética, creatividad y sensibilidad.

Así, la incorporación de la tecnología en el entorno laboral no busca reemplazar al trabajador, sino **“Aumentar la Inteligencia”** en el entorno empresarial (en inglés- *“Augmented Intelligence”*) mediante la combinación de la perspectiva humana y las recomendaciones automatizadas.

¹¹ El estudio “Cómo poner en valor a las personas sobre las máquinas” elaborado por Alicia Roy y publicado por el Observatorio de RH (situado en Madrid, España) manifiesta la necesidad de poner el valor las habilidades de los trabajadores a lo largo del proceso de transformación tecnológica en las empresas.



Insights empresas

Las empresas han puesto en marcha un modelo colaborativo Humano – Máquina que ponga en valor al capital humano, genere ventajas competitivas y fomente la creación de oportunidades de crecimiento profesional. Algunos de los ejemplos colaborativos puestos en práctica son: la presencia de asistentes virtuales para empleados que ayuden a los trabajadores a encontrar la información que necesitan de forma rápida, eficiente y en tiempo real; o bien la creación de un modelo donde sean los sistemas de IA quienes realicen tareas de menor valor para la empresa, mientras que el personal se encargue de desempeñar roles que entrañen una mayor responsabilidad, creatividad y sensibilidad con el cliente. Esto dará lugar a la puesta en marcha de una hoja de ruta que permita emplear eficientemente los recursos internos, crear procesos más eficaces y ágiles, fomentar oportunidades de crecimiento profesional y generar más valor para todos los agentes involucrados (empresa, sociedad e individuo).

EMPRESA DESTACADA

El Corte Inglés

El Corte Inglés apuesta por un uso de la IA centrado en impulsar dos ejes estratégicos. Por un lado, se priorizan aquellas capacidades de la IA que generen nuevas oportunidades para mejorar y fortalecer la relación con los clientes a través de experiencias diferenciales. Y desde el ámbito operacional, se desarrollan aplicaciones que den apoyo en la toma de decisiones traduciéndose en un mayor empoderamiento del personal interno.

El equipo experto en IA, dotado de perfiles matemáticos, físicos e informáticos, está enmarcado por el “Departamento de Big Data Corporativo” el cuál se enfoca en trasladar el estado del arte en IA desarrollando pruebas de concepto que generen valor de negocio. Estas iniciativas son previamente evaluadas y priorizadas por un equipo multidisciplinar, “Equipo Alpha”, formado por responsables de negocio y expertos en IA, fomentando así un diseño colaborativo e inclusivo en el uso de la IA dentro de la organización.

Para garantizar el éxito de esta estrategia, El Corte Inglés se posiciona como una de las compañías pioneras en entender la importancia de implementar iniciativas de formación tecnológica y desarrollo de carrera que permitan aumentar las capacidades analíticas de sus trabajadores para fomentar el talento interno y generar un pool de expertos propios.

Entrevistamos a Gema Hernández, miembro del Departamento BDC, quien destaca el rol clave que ha tenido el Comité de Dirección en impulsar una transformación tecnológica poniendo foco en las capacidades de la IA. Por ello, se ha diseñado una ruta estratégica impulsando un cambio en la cultura organizativa que persiga la implementación de una IA responsable, al mismo tiempo que eleve las capacidades de la organización.



Un ejemplo de ello es la elaboración interna de un Máster Ejecutivo en IA, impartido por la Fundación Ramón Areces, adscrita a la UNED, cuyos contenidos han sido adaptados a las necesidades de El Corte Inglés. Este programa persigue actualizar los conocimientos de personal, quienes habiendo pasado un examen preliminar de conocimientos base, puedan formarse en técnicas de IA avanzadas tales como NLP (Procesamiento de Lenguaje Natural), Redes Neuronales de Deep Learning y MLOps (Machine Learning Operations). Tal es el éxito de la iniciativa, que a fecha septiembre de 2020, ya se han celebrado 3 convocatorias anuales con más de 100 empleados de diferentes áreas (BDC, negocios, viajes, seguros...).

Asimismo, con el objetivo de aumentar las capacidades analíticas del personal interno se apuesta por diseñar soluciones de IA bajo un modelo colaborativo “Humano – Maquina”, también llamado “Inteligencia Aumentada”, que pone en valor a su capital humano generando nuevas oportunidades de crecimiento para los profesionales.

Por ejemplo, se implementan soluciones de Automatización Inteligente de Procesos (IPA, por sus siglas en inglés) que conlleven a procesos más eficientes y una mayor profesionalización de los empleados, poniendo foco en usos donde la IA automatice aquellas tareas que generen un menor valor para la compañía y el personal pueda dedicar su tiempo a desarrollar tareas de supervisión y mayor generación de valor en su día a día.

De esta forma, El Corte Inglés hace frente a uno de los mayores retos que plantea la incorporación de nuevas tecnologías en el entorno empresarial: situar a las personas en el centro de las oportunidades y el uso tecnológico. Para ello, se desarrollan iniciativas de capacitación de los empleados con nuevas técnicas y formación que les permitan adoptar nuevos roles y responsabilidades, y adaptarse para combinar la fuerza humana con los sistemas inteligentes, fomentando modelos colaborativos que generen resultados con un mayor valor.

ENUNCIADO 8

8

Concebimos una cultura organizativa que fomente la responsabilidad a través de la formación tecnológica y la comprensión de las implicaciones éticas de la IA.

Unos de los objetivos de desarrollo sostenible a alcanzar en la Agenda 2030¹² de las Naciones Unidas es promover las oportunidades de aprendizaje para todos los agentes. Ello supone crear oportunidades de crecimiento profesional y fomento de las habilidades, para evitar acudir a un mercado laboral saturado de trabajadores con habilidades limitadas.

Alcanzando un nuevo punto de inflexión laboral, los trabajadores serán los protagonistas de un cambio liderado por la adaptación de sus habilidades y el aprendizaje de nuevos conocimientos (en inglés - *“Reskilling”*) a través de programas educativos que les capaciten para aprovechar las ventajas que proyectan los nuevos puestos de trabajo con una mayor calificación y un mayor reclamo de responsabilidades.



El *“Reskilling”* implica un cambio cultural a nivel organizativo, donde las empresas diseñen una estrategia corporativa que incorpore la responsabilidad de formación tecnológica y desarrollo de las capacidades de los trabajadores favoreciendo su empleabilidad e inclusión en el proceso de transformación tecnológica.

¹² **La Agenda 2030 para el Desarrollo Sostenible** fue firmada en 2015 por los jefes de Estado y de Gobierno de los países miembros de Naciones Unidas. Representa el compromiso internacional para hacer frente a los retos sociales, económicos y medioambientales de la globalización, potenciando el crecimiento económico y solicitando una mayor cohesión y justicia social, con un horizonte medioambiental sostenible.



Insights empresas

Una mayoría de entidades han puesto en marcha la creación de programas de formación tecnológica (“Reskilling”), donde se capacitan a sus empleados para el correcto empleo de nuevas tecnologías, entre ellas la IA. El “Reskilling” supone un cambio cultural en la compañía, donde se busca aumentar la madurez analítica y digital entre los trabajadores, enfocado a conseguir mejores resultados y favorecer su empleabilidad en el mercado laboral. Por ello, tanto los líderes de la organización como los trabajadores deben apostar por situarlo como eje de la cultura de la entidad e incorporarlo en la estratégica corporativa, ya que, de lo contrario, se corre el riesgo de perder un capital humano con un gran potencial.

EMPRESA DESTACADA

GSK

La transformación tecnológica en la industria farmacéutica no es ajena a las oportunidades que genera la IA como ventaja competitiva a corto y largo plazo. En este sentido, GSK traslada su programa de transformación digital en iniciativas que aumenten las capacidades analíticas de su capital humano, enfocándose en desarrollar un proceso de formación cuyo objetivo principal sea empoderar al empleado para que sea protagonista de su propio crecimiento profesional.

De este modo, GSK traslada el desarrollo responsable de la IA situando a su personal en el epicentro del diseño digital y dotándole de herramientas y formación (en inglés- “Reskilling”) que le permitan tomar decisiones más informadas y autónomas respaldadas por datos.

José Luis Parriego, HR head de la fábrica GSK en Aranda de Duero, comenta acerca del programa formativo DataCamp enfocado en promover el “Reskilling”. Este programa está dirigido a empleados de planta y persigue lograr una capacitación general en los ámbitos de programación, estadística, algoritmos, visualización, análisis, metodologías agile, etc. Su objetivo fundamental es asegurar la empleabilidad de los trabajadores y adaptar sus capacidades a las exigencias del proceso de transformación tecnológica.

La compañía cuenta con programas similares de capacitación para otro tipo de tecnologías como Impresión 3D, Robot Process Automation (RPAs), Power Apps o capacitación de mecánicos en automatización, entre otros.



Estos programas de capacitación interna cuentan con un modelo porcentual (10/20/70) en función del tipo de recurso que se emplee para aumentar las capacidades de los trabajadores:

- 10%: plataformas online destinadas al aprendizaje de lenguajes de programación ofertadas por la Universidad de Salamanca y startups.
- 20%: expertos que aconsejen y guíen a los participantes del “DataCamp” en sus proyectos.
- 70%: desarrollo de proyectos reales donde se apliquen los conocimientos aprendidos.

Ana Salgado, BI & Commercial Operations Manager, puntualiza que el “Reskilling” no es solo una práctica más en la carrera global a contrarreloj de transformación tecnológica; sino que además supone un cambio cultural en la compañía, donde se busca aumentar la madurez analítica y digital entre los trabajadores, enfocado a conseguir mejores resultados y favorecer la competitividad de la organización.

Asimismo, esta transformación cultural y profesional persigue la incorporación de nuevos perfiles que se alineen con los valores y principios de GSK, y que busquen impulsar el compromiso de la farmacéutica con la Agenda 2030.



ENUNCIADO 9

9

Nos identificamos con un modelo de desarrollo responsable e inclusivo de la IA que integre diversidad de conocimientos y dominios de negocio, así como sensibilidades y enfoques.

Como hemos mencionado en puntos anteriores, la **multidisciplinariedad y diversidad cultural** en los equipos involucrados en el ciclo de vida de la IA tiene el potencial de trasladar nuevas perspectivas dentro del desarrollo de soluciones, ayudando a prevenir la perpetuación de sesgos y resultando en aplicaciones más justas.

Asimismo, la diversidad en el entorno laboral también puede entenderse como la inclusión de **grupos multidisciplinarios** que favorezcan la integración de una variedad de disciplinas académicas, aportando mayor profundidad, perspectiva y dominios en las áreas de tecnología con el objetivo de crear soluciones más inclusivas, efectivas y robustas.

A lo largo de las fases del ciclo de vida de la IA confluyen una multitud de perfiles que de no ser orquestados a través de modelos de colaboración ágiles pueden actuar en silos limitando el potencial de sus contribuciones: expertos de negocio, ingenieros de datos, expertos en visualización, científicos de datos, arquitectos de sistemas, expertos de gobierno de datos y de IA, expertos legales y ciberseguridad, entre otros.

La armonización de los grupos de trabajo a través de modelos colaborativos que pongan en valor la multidisciplinariedad proporciona un alcance más amplio de información y un **enriquecimiento en las habilidades colectivas**.

De esta manera se persigue evitar sesgos de confirmación causados por el *"pensamiento grupal"* que se conoce como un grupo homogéneo de personas que da solución a un problema de forma rápida y errónea, debido a que comparten una misma mentalidad y perspectiva. Si por otro lado, se formase un grupo heterogéneo con una reserva más rica en experiencias, conocimientos y perspectivas culturales, se obtendrían soluciones más innovadoras y completas que lleguen a integrar todas las esferas de una empresa (esfera legal, administrativa, técnica, ejecutiva...).



Insights empresas

Para muchas empresas, la diversidad de enfoques y perspectivas se traduce en la formación de grupos multidisciplinares en las fases de diseño de iniciativas de IA. Esto posibilita la integración de una diversidad de conocimientos y dominios que se implican desde la ideación de una solución teniendo en cuenta sus implicaciones en diferentes áreas del negocio y su impacto en el usuario final.

ENUNCIADO 10

10

Entendemos la necesidad de orquestar internamente el gobierno de la IA a través de procesos y herramientas que lo tangibilicen en todo el ciclo de vida de las iniciativas de IA e identifiquen claramente la responsabilidad de sus participantes.

Los principios éticos de la IA abordan una serie de requisitos técnicos que como objetivo principal persiguen generar algoritmos de calidad ya sea a través de la implementación de herramientas o de procesos de supervisión que tangibilicen estos principios dentro de las operaciones.

Por ejemplo, los requisitos de reproducibilidad y fiabilidad de los resultados derivados del análisis de datos permiten asegurar el correcto funcionamiento de los algoritmos de IA. Esto implica trasladar la trazabilidad de las decisiones a lo largo de todo el ciclo de vida del algoritmo utilizando diferentes herramientas que documenten los datos del modelo, las técnicas de modelado utilizadas, así como los artefactos creados para su puesta en producción.

Por otra parte, las soluciones de IA implican nuevas fuentes de riesgos que pueden surgir en diferentes etapas del ciclo de vida (diseño, datos, entrenamiento, consumo). Por ello se impera la necesidad de establecer **mecanismos de gobierno que asignen responsabilidades y definan un marco de control** para mitigar los riesgos de la IA de manera transversal en la organización.

De la misma manera, establecer una estructura de gobierno de la IA permite a las organizaciones trasladar un modelo de supervisión humana que facilite la puesta en práctica del principio ético de auditabilidad de los sistemas de IA.





Insights empresas

Las empresas han compartido diferentes perspectivas sobre cómo llevar a cabo las funciones de gobierno de la IA y la operacionalización de la ética dentro de la organización. En función de su visión estratégica, algunas empresas consideran que estas funciones deben de ser descentralizadas y gestionadas por áreas existentes, mientras otras consideran necesaria la creación de grupos específicos para abordar los retos que conlleva el uso de la IA desde un punto de vista responsable e inclusivo. Entre los mecanismos puestos en marcha por algunas empresas, podemos destacar:

- La creación de un departamento interno independiente, que gobierne sobre el uso de la IA desde un punto de vista responsable e inclusivo.
- La configuración de un código ético general que permee transversalmente todas las áreas de la empresa, sin que exista un órgano de control propio que lleve a cabo prácticas específicas de governance.
- La confección de un comité de ética que trabaje conjuntamente con el equipo de Analytics, con el objetivo de sensibilizar a los equipos de IT para que consideren la Responsabilidad Social Corporativa como parte de la cultura en el ámbito tecnológico, fomentar la creación de modelos transparentes y velar por un empleo responsable de la IA.

ENUNCIADO 11

11

Reconocemos la importancia de diseñar sistemas de IA robustos que prevengan un uso malintencionado, inapropiado o malicioso de la tecnología.

Estrechamente vinculada al principio de prevención de daños, la solidez técnica es un requisito fundamental para configurar sistemas de IA robustos, que **se desarrollen con un enfoque preventivo de riesgos**, al comportarse de la manera esperada.

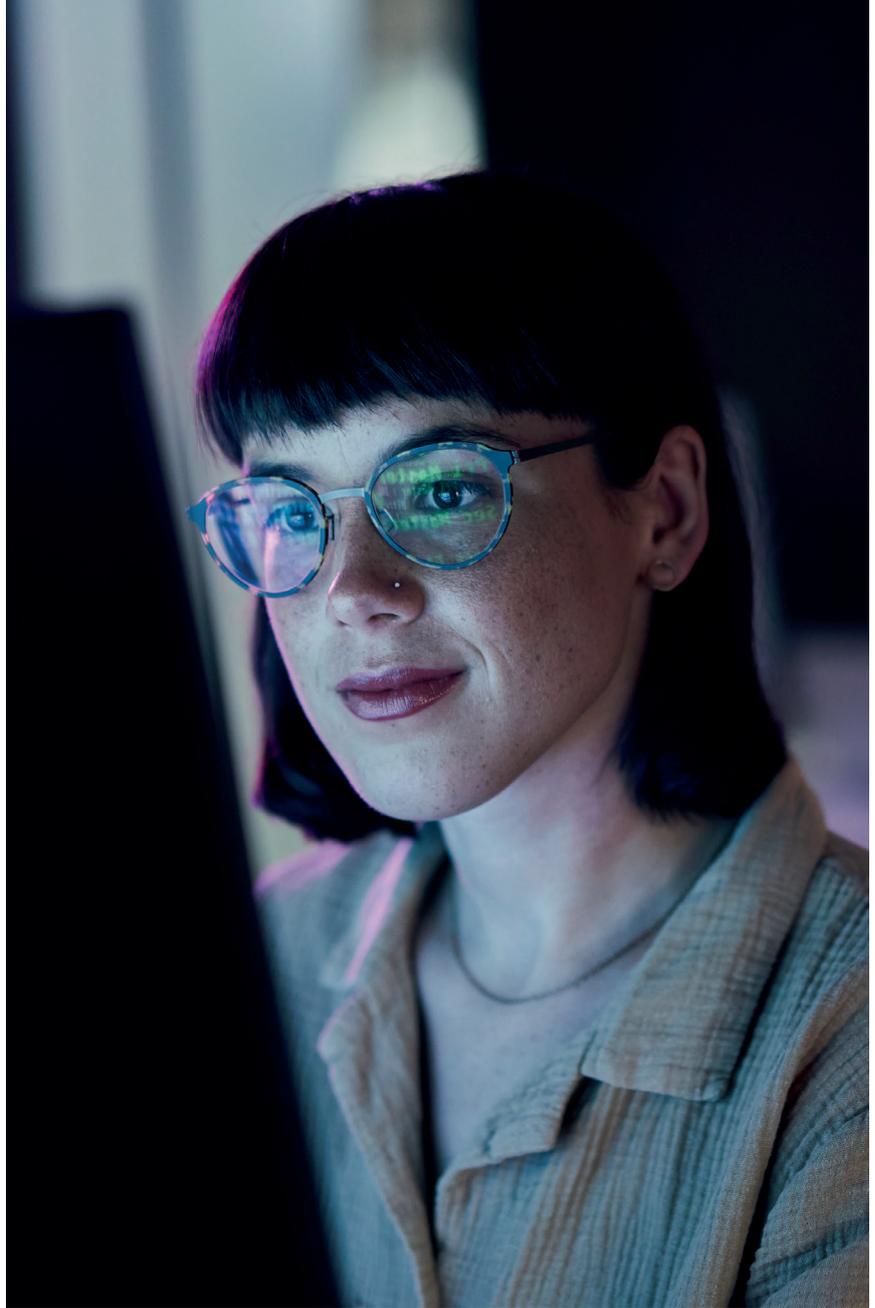
Al igual que los sistemas de software informático, la IA debe **protegerse contra vulnerabilidades** y acciones malintencionadas de terceros. Por ejemplo, ataques cuyo objetivo sean los datos (envenenamiento de datos), el modelo o la infraestructura subyacente.

Para mitigar estos riesgos, se pueden tomar medidas preventivas tales como diseñar planes de respaldo y reversión así como una estrategia de registro robusta que salvaguarde los datos y algoritmos de ML.

El nivel de medidas de seguridad requeridas depende de la criticidad del algoritmo dentro de las operaciones así como el impacto que tenga de cara a su usuario final. Cuando se pueda prever que el proceso de desarrollo o el sistema en sí plantee riesgos particularmente altos, es fundamental que las medidas de seguridad se desarrollen y prueben de forma proactiva y efectiva.

A medida que se incorporan tecnologías de IA en el entorno empresarial, se pueden realizar controles rutinarios que velen por mantener la confidencialidad de los datos del cliente y del empleado, y prevengan potenciales vulnerabilidades.

Por último, con una mayor comprensión de los riesgos que entraña la IA, cada vez más entidades apuestan por el **diseño estratégico de una hoja de ruta tecnológica que incorpore el empleo de herramientas de operacionalización de Machine Learning (MLOps) así como sistemas de ciberseguridad**, que aplaquen las amenazas cibernéticas y garanticen que las soluciones de IA permanezcan seguras.



Insights empresas

Todas las empresas priorizan altos estándares en el manejo de información y tienden a diseñar herramientas internas sólidas e inteligentes de protección de datos que garanticen la privacidad, ayuden a prevenir potenciales vulnerabilidades o ataques, y salvaguarden la información de todos los agentes. Por lo que muchas de las empresas ya consideran que la implementación de una IA robusta es una ventaja competitiva, y un método que permita ganar la confianza de los trabajadores y fidelizar a los clientes.

ENUNCIADO 12

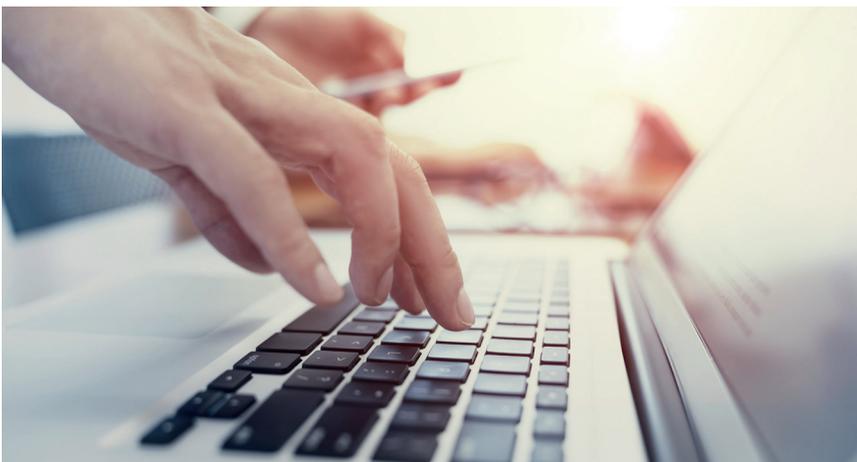
12

Sostenemos una aproximación a la Inteligencia Artificial que, desde el diseño, sitúe a las personas en el centro de las oportunidades y beneficios en el uso tecnológico.

Como se ha mencionado anteriormente, la apuesta por una integración conjunta entre humanos y máquinas (IA) puede generar grandes ventajas competitivas para las empresas, creando nuevas oportunidades profesionales y fomentando el desarrollo sostenible.

Este proceso de transformación tecnológica en las empresas debe situar al hombre en el **centro de las oportunidades y el uso tecnológico**. Un ejemplo de ello podría ser la labor de supervisión llevada a cabo por equipos internos para garantizar que los sistemas inteligentes no socaven la autonomía de la voluntad humana, ni causen efectos adversos en el entorno donde se vienen desarrollando.

La capacidad de supervisión también debe ser proporcional según el área de aplicación del sistema de IA y considerando su impacto en los individuos y en la operación, así como el riesgo potencial que pueda conllevar un resultado erróneo o fallo en el sistema.



De igual manera, el diseño de la IA centrado en las personas traslada de manera clara y transparente las ventajas y límites de la propia **aplicación considerando la autonomía de las personas y la no manipulación** de las mismas como un principio fundamental.

Finalmente, de cara a los consumidores, **los sistemas de IA deben estar centrados en el usuario** y diseñados de manera que permitan soluciones y productos accesibles a todas las personas, independientemente de su edad, género, habilidades o características.



Insights empresas

La visión de la IA centrada en las personas ha sido ampliamente reconocida y compartida por las empresas participantes del LabS. Se puntualiza cómo dentro del propio diseño de las aplicaciones de IA es esencial abordarlas desde un marco alineado con la protección de los derechos fundamentales de las personas.

Para lograr situar a las personas en el epicentro del cambio digital, las empresas saben que es necesario dotarles de la información que necesitan (de forma clara, comprensible y precisa) para que puedan tomar decisiones autónomas e informadas; proteger sus derechos mediante políticas y técnicas no intrusivas; y, por último, permitirles participar proactivamente en el proceso.

EMPRESA DESTACADA

Meliá Hotels International

“Impulsamos una aplicación de la Inteligencia Artificial que, desde el diseño, sitúe a nuestros empleados y clientes en el centro de las oportunidades y beneficios de la transformación digital”

Las capacidades que aportan las nuevas tecnologías como la Inteligencia Artificial, en combinación con una mayor disposición y diversidad de datos, redefinen actualmente las fórmulas para mejorar la experiencia de los clientes.

En este contexto, Meliá Hotels Internacional ha elaborado su plan de transformación digital, pilar clave en su estrategia, basado en tres ejes: potenciar sus canales de venta, personalizar la experiencia de sus clientes y mejorar su modelo operativo y la cultura digital de su organización.

Impulsando un modelo de negocio centrado en las personas, Meliá comparte una visión colaborativa del uso de las tecnologías en las operaciones, con la confianza de asegurar la aportación de valor a sus equipos y promover la mejora del servicio y la experiencia que ofrecen a cada huésped.



La Compañía ha trasladado los beneficios del uso de las capacidades analíticas y sistemas de IA a través de dos enfoques complementarios:

Por un lado, el uso de la IA permite a los trabajadores desempeñar roles que entrañen más responsabilidades y capacidades cognitivas gracias a la automatización de tareas repetitivas de poco valor añadido.

De igual manera, el uso de la IA traslada mayor eficiencia en los recursos internos, dirigiéndolos a optimizar la distribución del personal en función de las necesidades requeridas y a mejorar las interacciones con cada huésped mediante la creación de productos y experiencias personalizadas.

Esta aproximación apunta a generar un valor diferenciado gracias a la IA, potenciando cualidades humanas que impacten positivamente en la experiencia de los clientes en cada interacción.

De esta manera, Meliá pone en marcha una transformación en el ámbito digital y humano, donde los equipos adquieren nuevas capacidades para mejorar en su desempeño y asegurar una entrega de la promesa de marca y de la experiencia del cliente bajo estándares de calidad y excelencia.



Conclusión

En España, los avances en el uso de las técnicas de inteligencia artificial continúan desarrollándose a una gran velocidad, dando como resultado, una mayor competitividad de las empresas y soluciones sofisticadas, productos innovadores y mejoras sustanciales en el bienestar individual y colectivo.

Al mismo tiempo, estos avances han atraído una creciente preocupación pública sobre las consecuencias en el uso de la Inteligencia Artificial poniendo al sector privado en el punto de mira y reclamando la implementación de sistemas de IA **robustos, transparentes y seguros** que, a su vez, salvaguarden la correcta **defensa de los derechos fundamentales y la supremacía de la autonomía del individuo.**

En este sentido, las compañías están desarrollando, con diferentes niveles de madurez, prácticas que atiendan a estas preocupaciones poniendo el foco en implementar una IA centrada en el ser humano, ya sea de cara a mejorar la solución para el usuario final, asegurando la intervención humana gracias a procesos de supervisión o elevando los conocimientos técnicos del personal fomentando la transparencia en los sistemas.

También las empresas comparten preocupaciones similares radicando en la falta de un gobierno global que marque las pautas de auditabilidad y asignación de responsabilidades bajo un marco regulatorio claro y transparente.

Sobre esta línea, la Comisión Europea plantea una primera aproximación legal en el Libro Blanco de IA, publicado en febrero de 2020, incluyendo una serie de requisitos para aquellas aplicaciones que sean consideradas de “alto riesgo” (salud, transporte, energía, sector público, entre otros):

- **Datos de entrenamiento:** asegurar la representatividad de los datos evitando potenciales discriminaciones a grupos vulnerables y usar una cantidad amplia y segura de datos.
- **Gestión documental y conservación de datos:** a través de un registro de modelos y de datos de entrenamiento, donde se incluya documentación detallando las reglas de entrenamiento y validación del modelo.
- **Informar a los usuarios:** de manera clara de que se está interaccionando con una IA y de las posibles recogidas de datos.
- **Robustez y precisión:** asegurar la exactitud y reproducibilidad de los resultados así como la resiliencia ante potenciales ataques y manipulaciones.
- **Supervisión humana:** de tal forma que los resultados dados por la IA tengan que ser aprobados o, como mínimo, supervisados por un ser humano.

A lo largo de este año de trabajo en el LabS las empresas han demostrado un compromiso activo con respecto a hacer un uso responsable de la tecnología, compartiendo una visión en donde la IA responsable e inclusiva se percibe como una palanca de valor.

Por todo ello, las empresas están apoyando el diseño una hoja de ruta que integre un empleo de la IA responsable e inclusiva en la cultura empresarial y que sea escalable transversalmente en la organización; así como **crear una IA confiable, transparente, explicable y segura que pueda ser aceptada por la sociedad**, permitiendo la participación conjunta de todos los grupos de interés para potenciar los beneficios y fomentar las oportunidades de futuro que la IA entraña.

Autores



David Pereira Paz
Head of Data & Intelligence for Europe



Mayte Sanchez Hidalgo
Head of AI Strategy



Andrea Cornavaca Del Pozo
AI Strategy Lead Analyst



Rose Barragan Barranco
AI Strategy Jr. Analyst

Con la colaboración del **Center for Business in Society de IESE**



Acerca de NTT DATA

NTT DATA, perteneciente a NTT DATA Group, es una empresa de consultoría y outsourcing que abarca todos los sectores del ámbito económico, llegando a facturar en el último ejercicio fiscal 1.454 millones de euros, en la que aportan su talento más de 27.000 profesionales repartidos por Europa, USA y Latinoamérica.



Contacto SERES



Lucila García
Subdirectora General lucilagarcia@fundacionseres.org



Myriam Franco
Gestora de Empresas mfranco@fundacionseres.org



Juan Sánchez
Especialista CRM jsanchezvico@fundacionseres.org

Sobre Fundación SERES

Fundación SERES, entidad sin ánimo de lucro, promueve el compromiso de las empresas en la mejora de la sociedad con actuaciones responsables alineadas con la estrategia de la compañía y generando valor para la empresa y la sociedad.

Hoy ya son 150 empresas las que forman parte de este proyecto. El objetivo que persigue la Fundación es impulsar actuaciones sociales de las empresas y transformar la realidad empresarial y conseguir una sociedad mejor y más sana y una empresa más fuerte. SERES se esfuerza en favorecer el aprendizaje individual y colectivo resolviendo problemas sociales reales y creando valor para las empresas.

