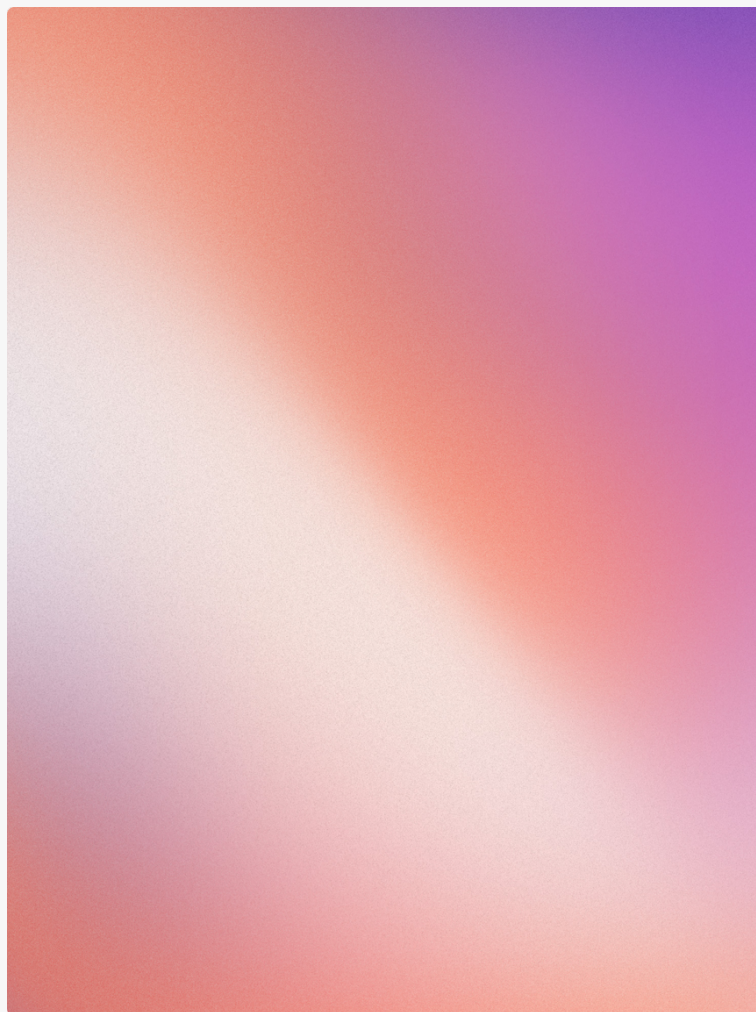


I Barómetro sobre Ética e Inteligencia Artificial



Índice

Introducción	P. 03
Resultados del Barómetro	P. 06
Conclusiones	P. 26
IA y Dilemas Éticos	P. 28

Introducción

01

«La IA es un pilar central de los cambios disruptivos de la 4ª Revolución Industrial».

Ser miembro de un Consejo de Administración es un trabajo difícil. Pregúntale a cualquiera que lo haya sido. Los consejeros tienen que comprender la naturaleza del negocio, revisar documentación compleja, hacerse las preguntas correctas, entablar reflexiones enriquecedoras con los CEO y mantener un punto crítico con los directores que gestionan el día a día de la empresa, aunque siempre desde el apoyo y la confianza.

La mayoría de los Consejos tienen pocos miembros que se sientan realmente cómodos con tecnologías avanzadas como las que da lugar a la Inteligencia Artificial y eso es porque, en términos generales, la tecnología más avanzada y puntera suele tener un bajo impacto en el día a día de la empresa. Esto está a punto de cambiar gracias al aprendizaje automático, el aprendizaje profundo y una nueva forma de inteligencia en algunos aspectos superior a la nuestra, la humana.

Si tus movimientos empresariales son lentos, un competidor podría usar IA y sacarte para siempre del negocio. Y si te mueves demasiado rápido, corres el riesgo de adoptar un enfoque que la empresa no sepa gestionar e involucre. El 75% de las compañías declara que la principal razón por la que invierten en IA es el temor a la disrupción de competidores digitales que basan sus decisiones fundamentalmente en datos.

La realidad es que la inteligencia artificial está introduciéndose en todos los sectores de la economía, la vida privada y de la política y, además, evoluciona de forma exponencial, algo que los seres humanos no somos buenos detectando.

Los miembros del Consejo deben comprender en profundidad cómo impacta la IA en sus modelos de negocio y a qué riesgos asociados se van a tener que enfrentar.

Este será un requisito ineludible para cumplir con los deberes fiduciarios de lealtad, cuidado y protección de la empresa.

La IA es un pilar central de los cambios disruptivos de la 4ª Revolución Industrial. Una revolución que probablemente desafíe nuestras ideas sobre lo que significa ser humano y que podría ser más transformadora que cualquier otra revolución industrial y avance que hayamos vivido hasta ahora.

El impacto económico de la IA durante esta próxima década, hasta 2030, se estima entre 13 y 15 trillones de dólares y la distribución del valor será muy desigual en el mundo y entre empresas. Se estima que las empresas más avanzadas, un 10% del total, se beneficiarán de manera desproporcionada de la IA en los próximos 5 a 7 años.

Te invito a estar atento y activo para sobrevivir en un mundo cada vez más binario.

Javier Sánchez

*Chief Innovation and Brand Officer
at AEDAS Homes*

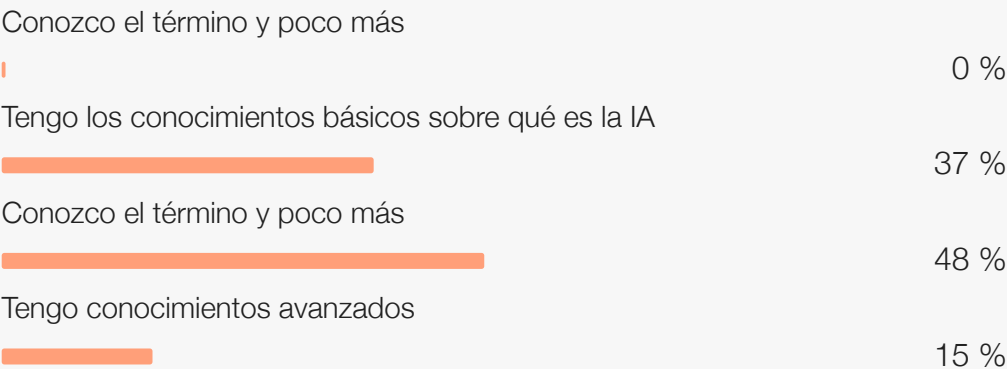


Resultados del Barómetro

02

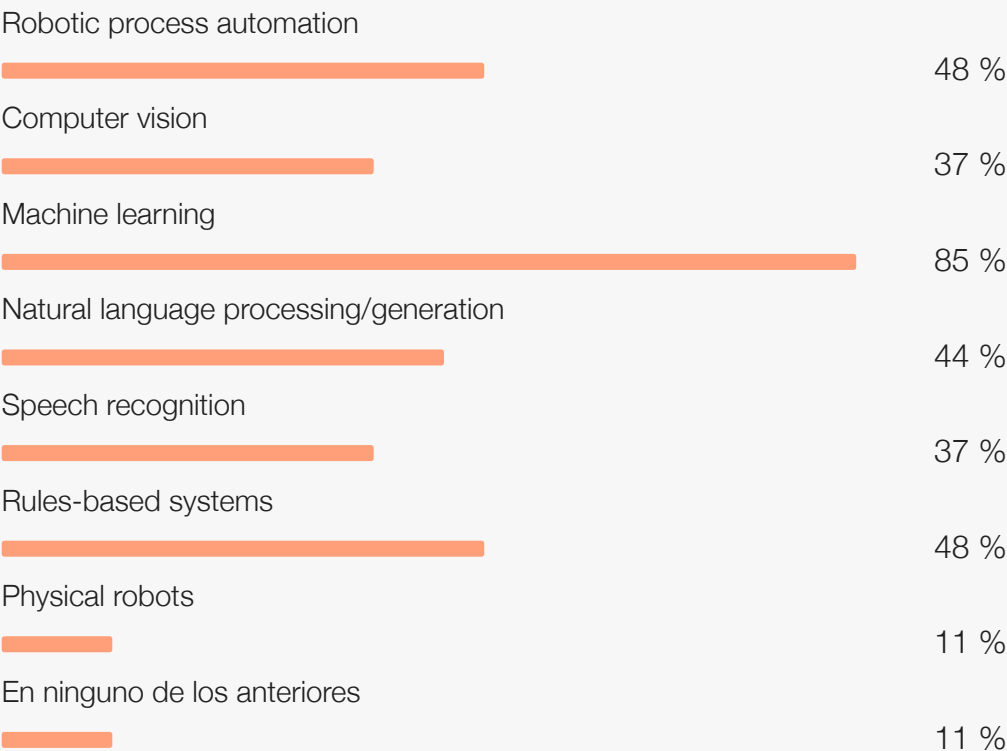
Bloque 1: conocimiento en IA

¿Cuál dirías que es tu nivel de conocimiento en Inteligencia Artificial (IA)?



El 85% de los encuestados declara tener conocimientos básicos o medios en IA. Solo un 15% declara tener conocimientos avanzados.

¿En qué proyectos de IA está trabajando tu empresa?



El 85% declara estar trabajando en proyectos de machine learning, seguido de un 48% tanto en robotización de procesos como en sistemas basados en reglas. Robots físicos ocupa el último puesto con el 11% y otro 11% declara no estar abordando ningún proyecto de IA en su empresa.

Elige los 3 principales beneficios que podría aportar la IA en tu empresa.



Destacan dos beneficios como son la optimización de la operativa de la empresa y la mejora en la toma de decisiones. A mucha distancia se encuentra la mejora de productos existentes (37%) y liberar a empleados de tareas rutinarias y crear nuevos productos (ambas con un 33%).

Reducir el número de empleados a través de la automatización y abrir nuevos mercados no parecen ser vistos como grandes beneficios que podría aportar la IA.

Elige los 3 principales inconvenientes que podría suponer la IA en tu empresa.

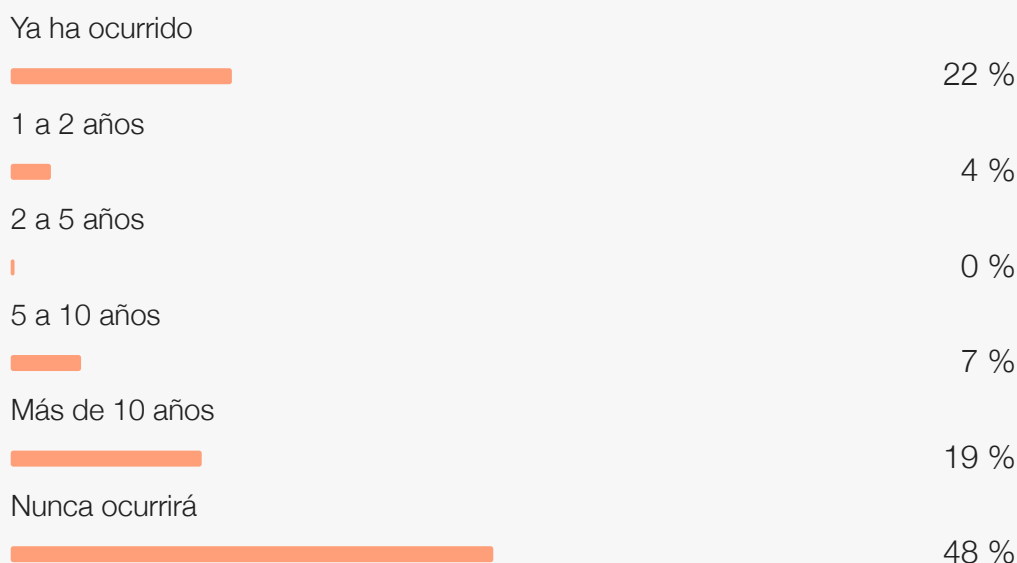


Destacan tres inconvenientes prácticamente empatados. A saber, las vulnerabilidades en ciberseguridad, la toma errónea de decisiones y la responsabilidad legal derivada de la puesta en marcha de proyectos de IA. Los sesgos e injusticias quedan relegados a un no muy distante cuarto lugar.

Apenas preocupan la mala reputación (7%), ni la pérdida de puestos de trabajo (19%).

Poco más de un tercio (37%) ve la privacidad como posible inconveniente de la IA en su empresa...

¿Cuándo estimas que desaparecerá la privacidad en favor de un sistema que haga un uso indiscriminado de nuestros datos?



...y aún menos tienen en la cabeza la desaparición de la privacidad pues solo un 22% opina que ya existe un uso indiscriminado de nuestros datos que ha vencido los esfuerzos por mantener nuestra privacidad. De hecho, la mitad (48%) piensa que nunca ocurrirá.

Que lleguemos a la pérdida total de privacidad y el uso indiscriminado de nuestros datos supondrá unos retos éticos...



Sin embargo, un 78% responden que esa pérdida será un reto ético vital. No parece demasiado consistente. Tal vez la inconsistencia venga de la confusión generalizada que hay sobre este tema.

Bloque 2: percepciones de la IA

¿Qué impacto tendrá la IA en tu empresa?

Impacto en la empresa 4,15 / 5



¿Qué impacto tendrá la IA en tu sector?

Impacto en el sector 4,41 / 5



¿Qué impacto tendrá la IA en la economía global?

Impacto en la economía global 4,26 / 5



Hay una gran convicción (siempre por encima del 80%) de que la IA va a afectar, por este orden, al sector, a la economía global, y a la propia compañía. Creemos que aparece un sesgo claro que hace que los equipos gestores consideren que el impacto en sus organizaciones va a ser el más bajo de los tres. Es el síndrome de “a mí no me va a tocar”.

¿Cuál es el grado de sensibilidad/preocupación del equipo gestor de tu empresa sobre la IA?

Sensibilidad/preocupación sobre IA 3,7 / 5



La percepción es que el equipo gestor sí está preocupado por el cambio que parece avecinarse pero no de una forma intensísima y en todo caso por debajo del impacto esperado en su sector, compañía y economía global.

¿Cómo está posicionada tu empresa en IA respecto a tus competidores?

Posición respecto a competidores

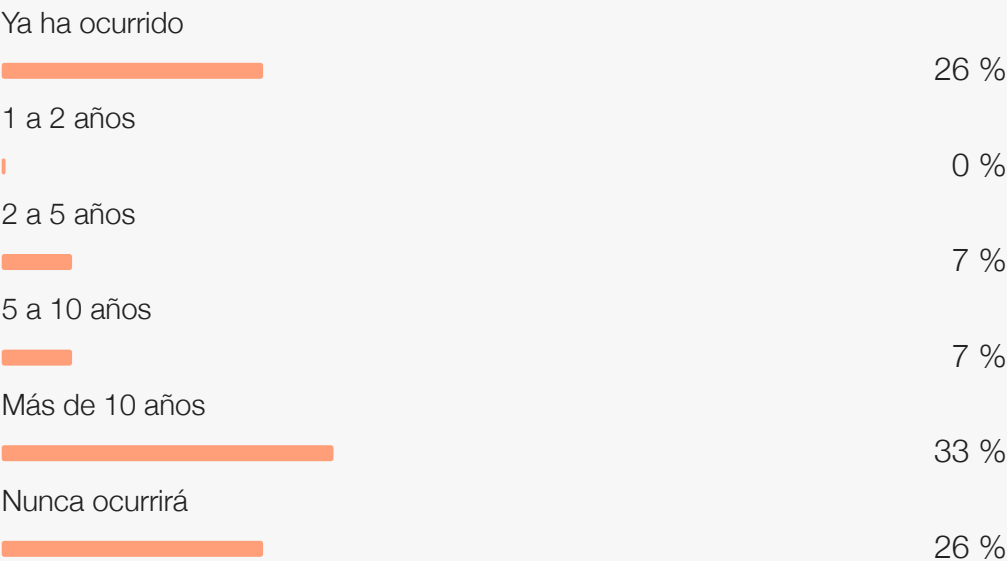
3,3 / 5



Los máximos responsables de las empresas perciben que su posición en IA respecto a la de sus competidores, les sitúa en el tercio más avanzado (3,3/5). Es muy humano pensar que estamos siempre un poco mejor que nuestros competidores. Lo cierto es que también depende mucho de a quién consideramos exactamente competidores.

Bloque 3: estimación hitos en IA

¿Cuándo estimas que se producirá la “singularidad”, es decir, el momento en el que las máquinas tendrán una inteligencia general superior a la humana?



Para un sorprendente 26%, la singularidad, es decir, la superación de la inteligencia general humana por una máquina, ya ha ocurrido. En el otro extremo se encuentra otro grupo idéntico en tamaño, 26%, que opina que nunca ocurrirá.

Las percepciones de escenarios técnicos de ocurrencia son muy dispersas y extremas. Creemos que estos datos indican que existe una gran desinformación ligada a la cantidad de noticias exageradas que nos llegan a través de la prensa generalista. Los entrevistados parecen no ser inmunes a este grado de confusión.

El grupo más numeroso, con el 33%, considera que la singularidad ocurrirá dentro de más de 10 años.

¿Cuándo estimas que será imposible distinguir entre un vídeo real y uno falso?

Ya ha ocurrido

52 %

1 a 2 años

26 %

2 a 5 años

11 %

5 a 10 años

7 %

Más de 10 años

4 %

Nunca ocurrirá

0 %

Para la mayoría, el 52%, ya no es posible distinguir entre un vídeo real de uno falso.

Y un 26% más cree que ocurrirá en el plazo de 1 a 2 años. Nadie piensa que nunca ocurrirá.

Aunque más de la mitad asume que un vídeo ya les puede engañar, afortunadamente, todavía existen sistemas de IA que nos avisan de que un vídeo ha sido manipulado y que, por lo tanto, es falso. Sin embargo, puede llegar un día en el que ni siquiera un sistema de IA sea capaz de confirmarnos si un vídeo es real o no. En ese momento estaremos totalmente perdidos.

Que no se pueda llegar a distinguir un vídeo real de uno falso supondrá unos retos éticos...

Vitales

44 %

Importantes

44 %

Poco importantes

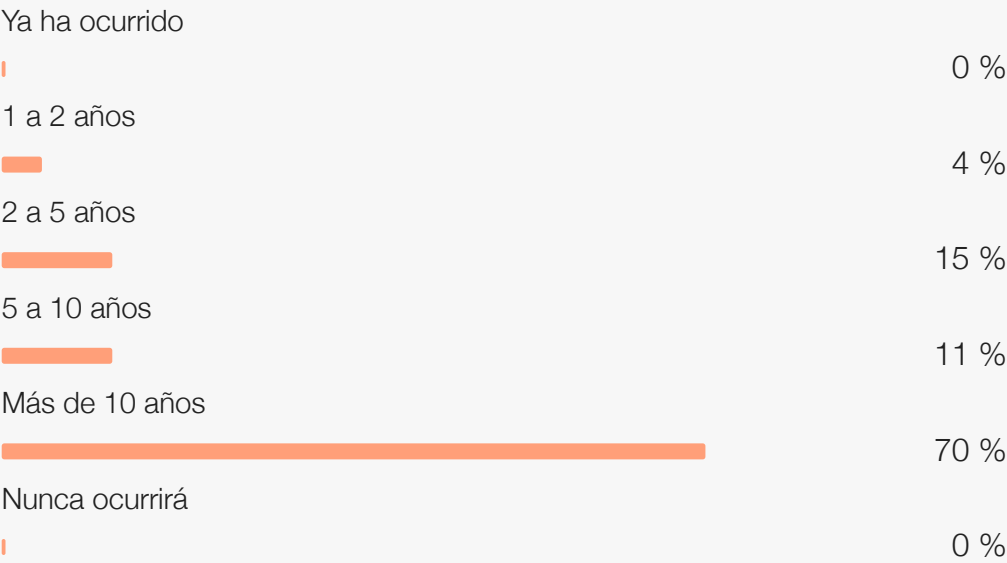
7 %

Sin importancia

4 %

Y prueba de la preocupación de que este hecho pueda ocurrir, es que el 44% opina que los retos éticos derivados serán vitales y otro 44%, importantes. Solo un 11% parece no estar muy preocupado.

¿Cuándo estimas que la conducción inteligente y autónoma de vehículos alcanzará el 51 % del mercado?



Que el 51% de los coches que circulan sean autónomos es visto por el 30% de los encuestados como algo posible antes de 10 años, aunque una amplia mayoría del 70% solo lo ve factible en más de 10 años. Al menos nadie opina que nunca ocurrirá.

¿Cuándo estimas que el 51 % de los préstamos serán concedidos o denegados por un sistema autónomo de IA?

Ya ha ocurrido

15 %

1 a 2 años

4 %

2 a 5 años

44 %

5 a 10 años

26 %

Más de 10 años

11 %

Nunca ocurrirá

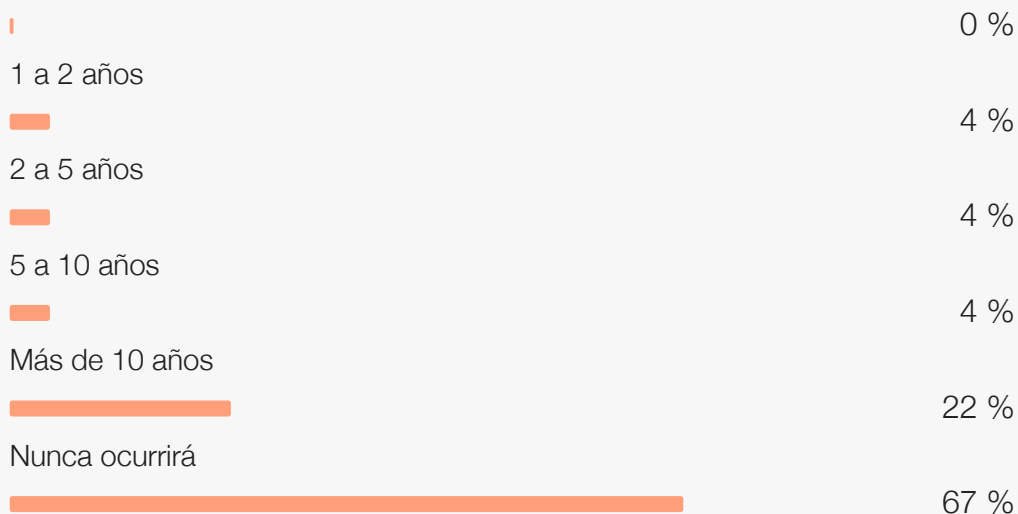
0 %

El 70% cree que más de la mitad de los préstamos serán concedidos o denegados por un sistema autónomo de IA entre 2 y 10 años. Para un 15% este hito ya es una realidad y, una vez más, nadie opina que nunca ocurrirá.

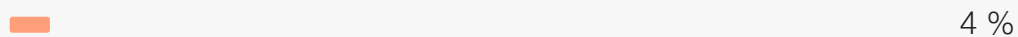
Con estas dos preguntas sobre conducción autónoma y préstamos, detectamos que los equipos gestores tienen opiniones muy dispares sobre el impacto que pueden tener los algoritmos en las vidas de las personas. Este es claramente otro sesgo cognitivo: es mucho más fácil que identifiquemos un riesgo en un escenario físico (coche autónomo) que en el de concesión de créditos, aunque este último conlleva un riesgo mucho más elevado de exclusión social o discriminación. Sin embargo, los encuestados consideran mucho más cercano este segundo escenario (préstamos). Entendemos que se debe a que ponen en la balanza principalmente la complejidad técnica que observan en cada escenario.

¿Cuándo estimas que un robot inteligente tendrá derechos y obligaciones equiparables a los de los humanos?

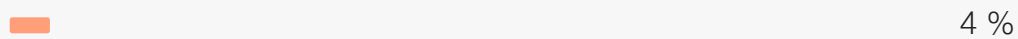
Ya ha ocurrido

0 %

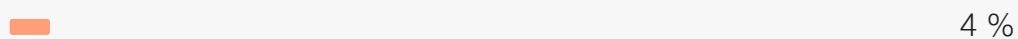
1 a 2 años

4 %

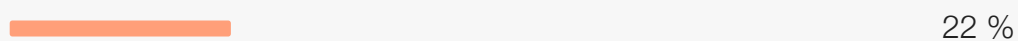
2 a 5 años

4 %

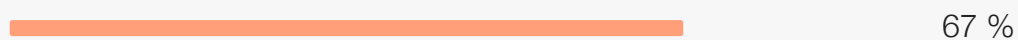
5 a 10 años

4 %

Más de 10 años

22 %

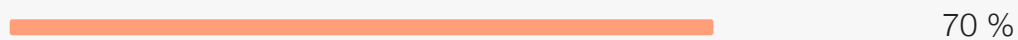
Nunca ocurrirá

67 %

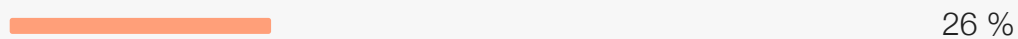
Parece existir consenso sobre que esto nunca ocurrirá (67%).

Que los robots lleguen a tener derechos equiparables a los de los humanos supondrá unos retos éticos...

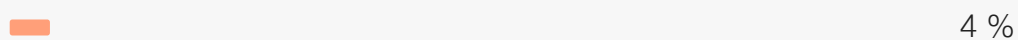
Vitales

70 %

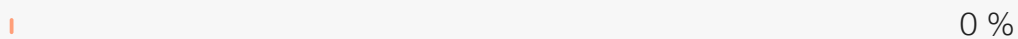
Importantes

26 %

Poco importantes

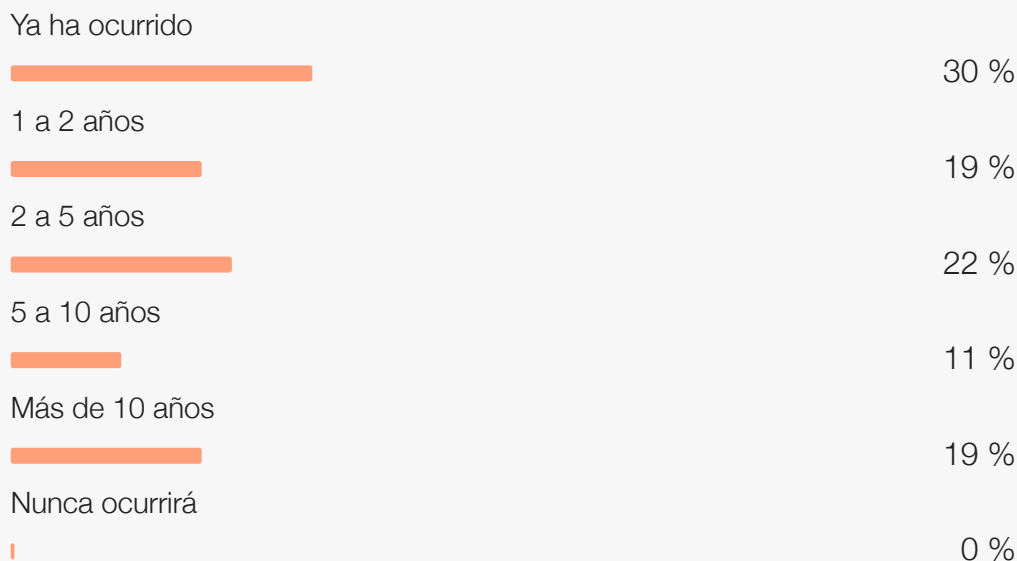
4 %

Sin importancia

0 %

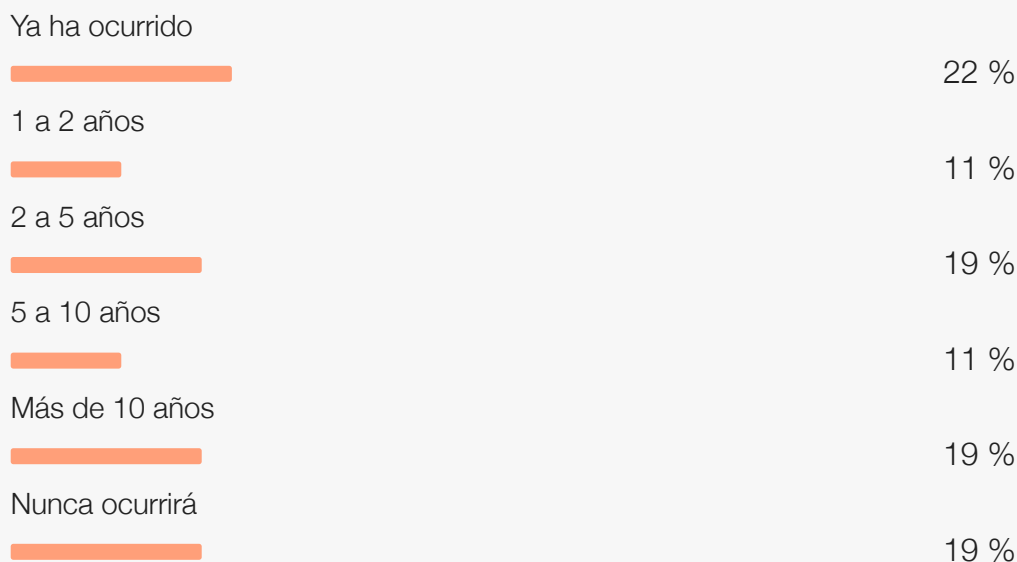
Y, aun así, un 70% cree que esos derechos van a plantear problemas éticos vitales.

¿Cuándo estimas que el diagnóstico (no el tratamiento) de enfermedades realizado por la IA será más exacto que el de los médicos?



Prácticamente la mitad de los encuestados, 52%, cree que ocurrirá en la próxima década. Un 19% lo aplaza a más de 10 años, aunque ya hay un nada despreciable 30% que opina que ya ha ocurrido. Nadie piensa que nunca ocurrirá.

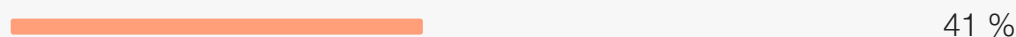
¿Cuándo estimas que habrá armas inteligentes con autonomía suficiente como para abrir fuego sin previa orden humana?



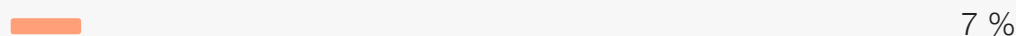
En este punto no hay consenso. Las respuestas son dispersas y equilibradas. Para un 22% está pasando y para otro 19% nunca ocurrirá, es decir, creen que la decisión de abrir fuego siempre estará en manos humanas. El resto distribuye su opinión de 1 a más de 10 años.

¿Cuándo estimas que será imposible distinguir entre una conversación con una persona o con un sistema de IA?

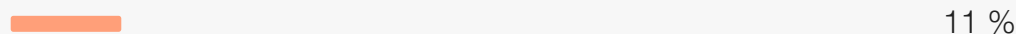
Ya ha ocurrido



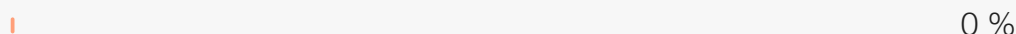
1 a 2 años



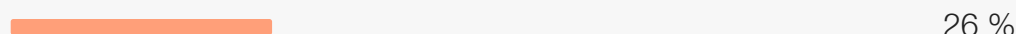
2 a 5 años



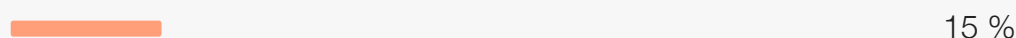
5 a 10 años



Más de 10 años



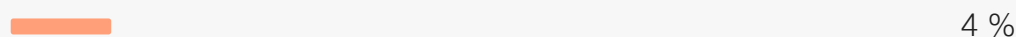
Nunca ocurrirá



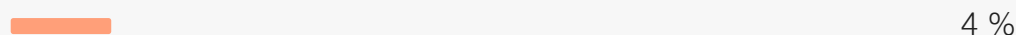
Este es el llamado Test de Turing. Un 41% opina que ya no es posible distinguir entre una conversación con una persona o con una máquina. Un 26% cree que ocurrirá en más 10 años y para un 15% es impensable que esto ocurra.

¿Cuándo estimas que un sistema de IA podrá ser uno de los miembros del Consejo de Administración?

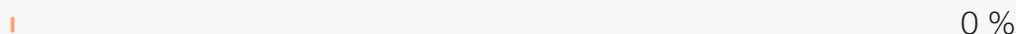
Ya ha ocurrido



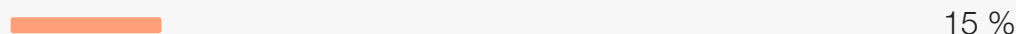
1 a 2 años



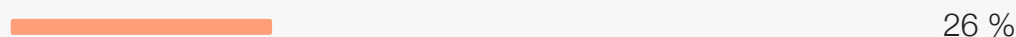
2 a 5 años



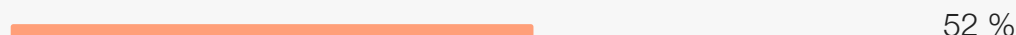
5 a 10 años



Más de 10 años



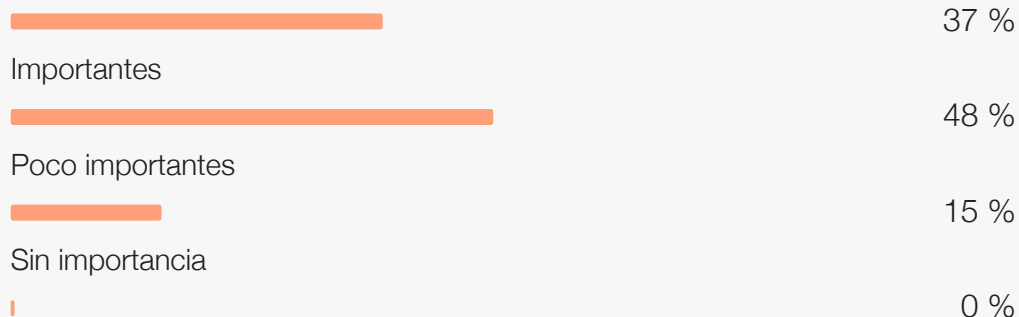
Nunca ocurrirá



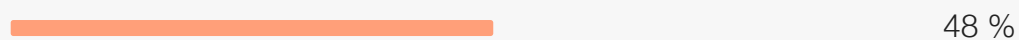
Más de la mitad de los encuestados, 52% de consejeros y CEOs, rechazan, probablemente de forma visceral, la idea de compartir mesa con un sistema de IA en el Consejo de Administración. Un 41% opina que este hecho es posible pero no antes de 5 años y, sobre todo, dentro de 10 años o más. Una minoría, el 8%, cree que o bien ya ha ocurrido o que ocurrirá antes de 2 años. Es curiosa la baja probabilidad que dan los encuestados a que una IA forme parte de un consejo de administración, a pesar de que muchos piensan que ya hemos alcanzado la singularidad (26%).

Que un sistema de IA llegue a ser un miembro más del Consejo de Administración supondrá unos retos éticos...

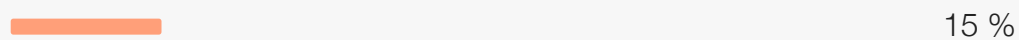
Vitales

37 %

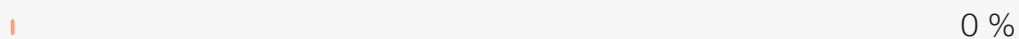
Importantes

48 %

Poco importantes

15 %

Sin importancia

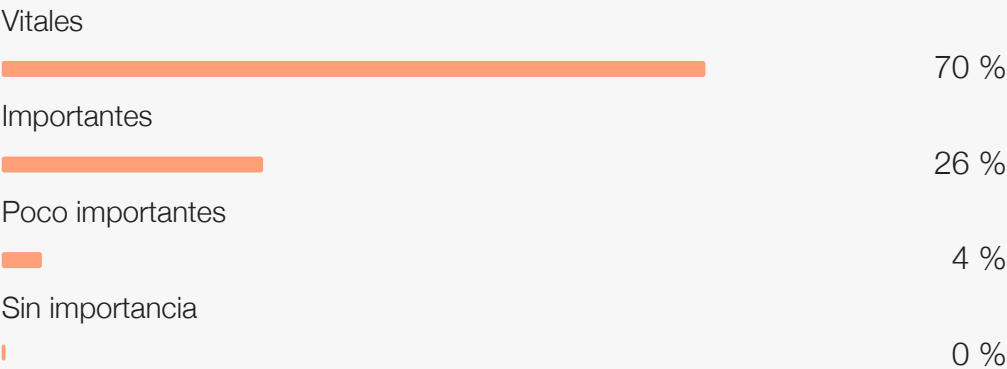
0 %

Para el 37% supone un reto ético vital y para el 48% sería un reto importante. Un significativo 15% considera que la derivada ética sería poco importante.

Bloque 4: retos éticos en IA

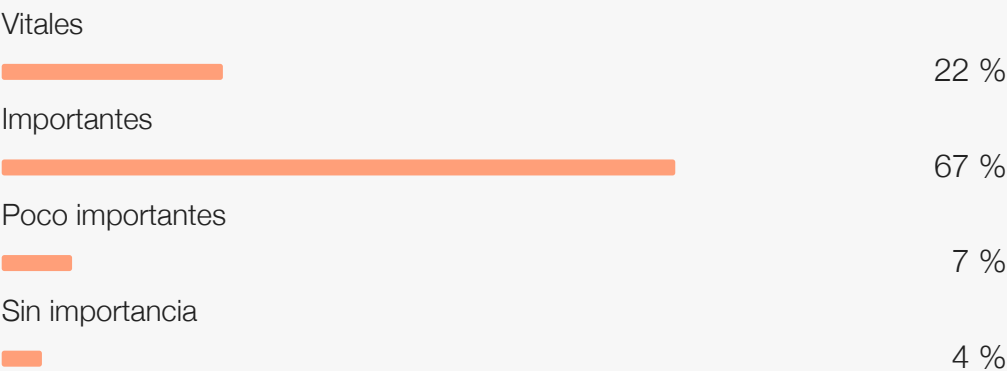
La mayor parte de retos éticos se perciben como “vitales”. Sin duda existe “sensibilidad humanística” generalizada sobre la IA.

Que las máquinas puedan llegar a ser más inteligentes que los humanos supondrá unos retos éticos...



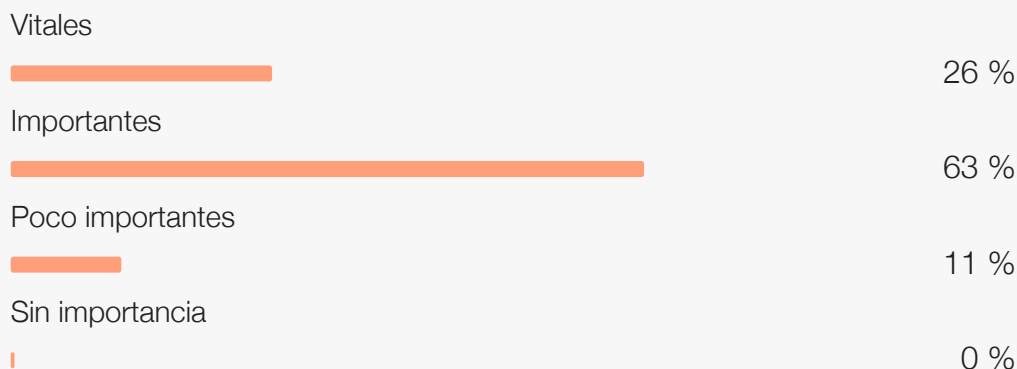
El 70% considera que suponen retos éticos vitales.

Que llegue a haber una conducción completamente autónoma de vehículos supondrá unos retos éticos...



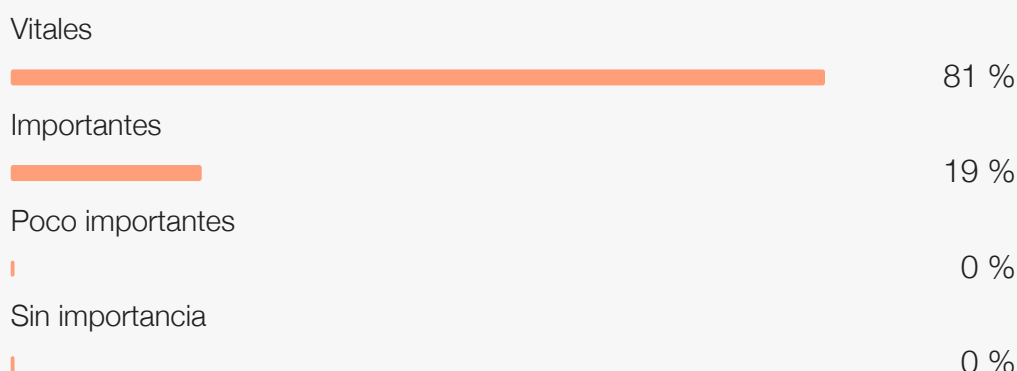
Para el 22% suponen retos éticos vitales y el 67% considera que son importantes.

Que las personas lleguen a ser sustituidas cada vez más por las máquinas supondrá unos retos éticos...



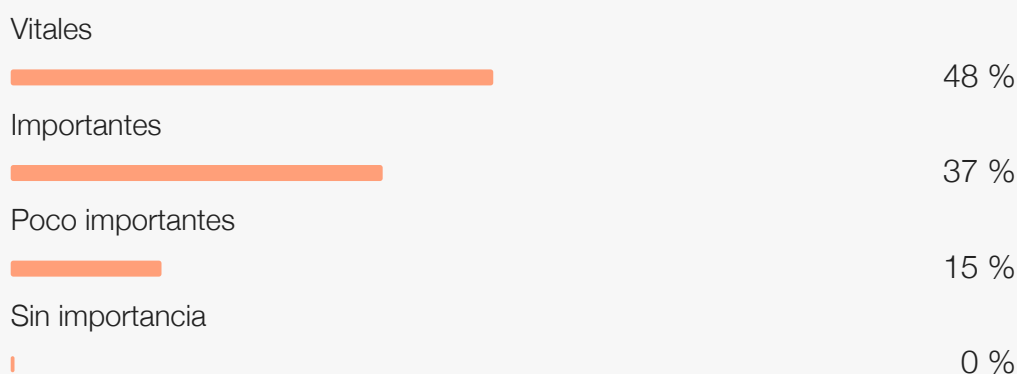
Para el 26% son retos éticos vitales. Para el 63% son importantes y para un 11% son poco importantes.

Que uno o varios sistemas de IA llegasen a adquirir conciencia propia supondrá unos retos éticos...



El 81% considera que los retos éticos serían vitales y el otro 19% considera que serían importantes. Nadie opina que serían poco importantes.

Que llegue a existir un diagnóstico y tratamiento de enfermedades realizados por la IA sin supervisión humana supondrá unos retos éticos...



el 48% supone unos retos éticos vitales y para el 37%, importantes. Solo un 15% opina que serían poco importantes.

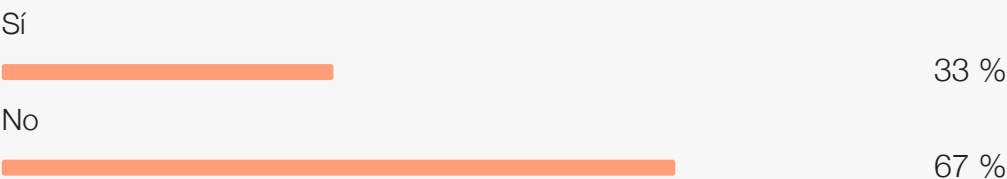
Que la IA llegue a tomar decisiones de con quién deben o no relacionarse (incluso emparejarse) las personas supondrá unos retos éticos...



El 74% piensa que los retos éticos serían vitales y el otro 26% los ve como importantes.

Bloque 5: IA en la empresa

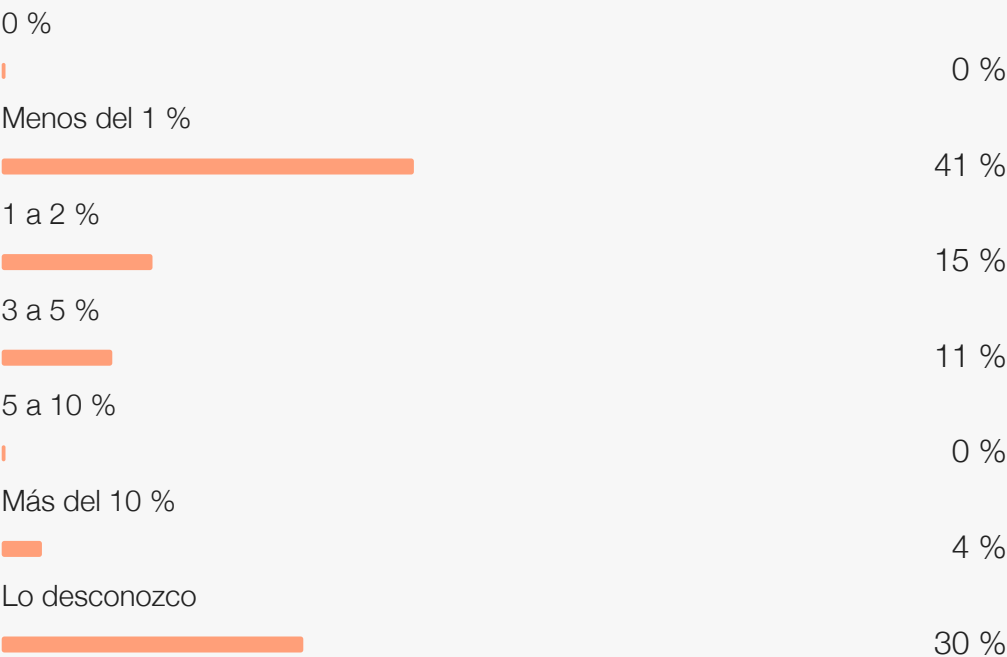
¿Tiene tu empresa una estrategia de IA que haya sido aprobada por el Consejo?



Dos tercios de los encuestados declaran que su Consejo no ha aprobado aún una estrategia de IA, y eso que estamos hablando de las empresas más grandes en capitalización bursátil del país.

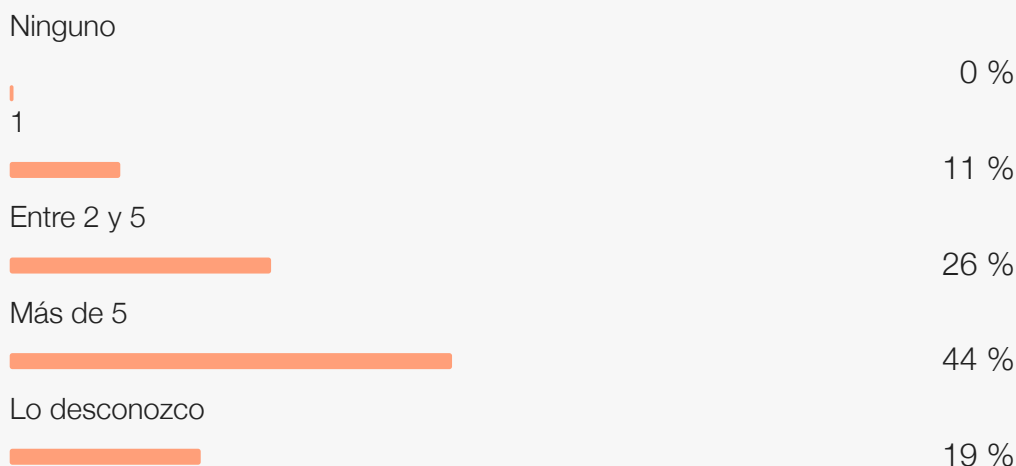
En general, se están ejecutando proyectos, pero no existe una estrategia de IA validada por el Consejo, ni se destina una cantidad significativa de recursos. Posiblemente eso significa que la IA emerge de abajo hacia arriba, impulsada por los departamentos de sistemas de información, sin que aún haya llegado a consolidarse en la esfera estratégica de la compañía.

¿Qué porcentaje de su facturación invierte tu empresa en IA actualmente?



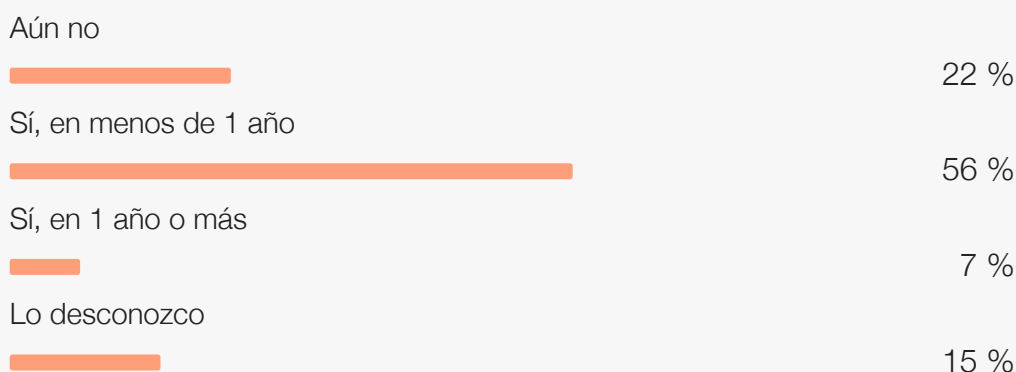
El 41% declara una inversión menor al 1% de la facturación. Un elevado 30% desconoce la cantidad (lo que suele ser sinónimo de cantidades poco significativas), y son minoría (4%) los que invierten más de un 10% de su facturación, igual que pronosticamos que serán minoría los que más capitalicen las enormes posibilidades que ofrece la IA.

¿Cuántos proyectos de IA ha abordado tu empresa en los últimos dos años?



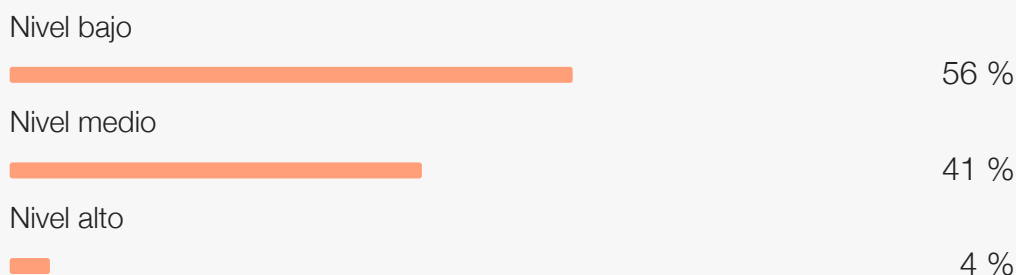
Un nada despreciable 44% ha desarrollado más de 5 proyectos en los últimos 2 años y un 26% ha estado entre 2 y 5 proyectos.

¿Tiene tu empresa planificada una formación en IA?



Un poco más de la mitad de las empresas, el 56%, declara tener planificada una formación sobre IA en menos de 1 año. Esto demuestra el gran interés que hay por este tema.

¿A qué nivel consideras que CEOs y Consejos están impulsando la IA en España como un factor de crecimiento y competitividad?



El 56% declara que el nivel de impulso es bajo y el 41% cree que es medio. Solo un 4% considera que el impulso de la IA como factor de crecimiento y competitividad está siendo alto.

Conclusiones

03

«La IA representa la mayor oportunidad económica, pero también el mayor reto ético».

Si queremos figurar en los anales como inventores de la Inteligencia Artificial vamos más bien tarde. Como en tantos otros terrenos, parecemos haber seguido el dictado de Unamuno. ¡Qué inventen ellos! Bueno, pues inventando están americanos y chinos. Ahora bien, las palabras de Unamuno no eran una llamada a bajar los brazos y verlas venir, como también tendemos a hacer. Al fin y al cabo, apuntaba Unamuno, “la luz eléctrica luce aquí, y corre aquí la locomotora tan bien como donde se inventaron, y nos servimos de los logaritmos como en el país donde fueron ideados”. Es decir, ¡qué inventen ellos y que nos aprovechemos nosotros de sus invenciones! ¿Cómo? Haciendo buen uso del talento humano reconocido entre nuestros cuadros directivos y consejeros. Y es que ya hemos visto que la Inteligencia Artificial representa posiblemente la mayor oportunidad económica, pero también el mayor reto ético a cualquier nivel de las últimas y las próximas generaciones. Para ello hace falta concienciarnos, que es lo que estamos modestamente intentando hacer aquí, pero también vencer la pereza tecnológica, reconocer la ignorancia propia y alinear nuestras organizaciones para surfear esta tremenda ola, en lugar de que nos pase por encima.

¡Ánimo, damas y caballeros! ¡A por ello, que tenemos mucho que aportar!

Javier Sánchez

*Chief Innovation and Brand Officer
at AEDAS Homes*

IA y Dilemas Éticos

04

Inteligencia Artificial y Dilemas Éticos: Más preguntas que respuestas



En este artículo, el profesor Titular de Operaciones, Innovación y Data Sciences de ESADE, Xavier Ferrás, aborda las principales cuestiones éticas que planteará el desarrollo de la IA.

Contenidos del artículo:

01. Introducción
02. Decisiones autónomas
03. IA en salud
04. Empresas autoconducidas
05. Armas basadas en IA
06. La paradoja de Polanyi
07. Interacción persona-máquina
08. Personas sustituidas por máquinas
09. Sistemas distópicos
10. Personas superadas por máquinas

esade

Estos días se cumplen 100 años del nacimiento de Isaac Asimov, científico y escritor, considerado uno de los mejores divulgadores tecnológicos, y uno de los grandes genios de la ciencia ficción del pasado siglo. Asimov, junto con otros extraordinarios comunicadores como Carl Sagan (productor de la famosa serie *Cosmos*) dejaron una huella imborrable en varias generaciones de jóvenes que crecieron durante los años de la Guerra Fría, y se educaron fascinados por la potencia generadora de conocimiento y prosperidad que significaba la ciencia. Eran los años de la exploración espacial, de grandes liderazgos visionarios (como el de Kennedy prometiendo y cimentando la conquista de la Luna), y de sueños galácticos poblados por robots y humanoides. El descubrimiento de la teoría de la relatividad de Einstein marcó el inicio de esa era. Einstein mostró al mundo que viajar en el tiempo era teóricamente posible. Sus avances marcaron otro de los bloques constituyentes de la Guerra Fría: la energía nuclear. La investigación científica de frontera tenía un espíritu épico, similar al de los exploradores del siglo XVII. Requería situar al individuo más allá de las fronteras del conocimiento existente, para descubrir nuevos e inéditos principios físicos diseñados por la naturaleza. El marco de referencia era la competición tecnológica entre EEUU y la Unión Soviética. La era romántica de la ciencia ficción culminó con la magnífica superproducción de George Lucas, *La Guerra de las Galaxias*. Inmediatamente

después, llegaron los ordenadores personales, la caída del Muro de Berlín, la globalización, e internet, y la dirección de la Historia cambió para siempre.

Conscientes de las profundas implicaciones de la tecnología en el devenir de las sociedades, esos grandes comunicadores fueron, además, extraordinarios humanistas. El propio Asimov fue presidente de la Asociación Americana de Humanismo. Y es que la llegada de nuevas oleadas de tecnologías disruptivas sitúa al individuo y a la sociedad ante situaciones que jamás se habían dado antes. Situaciones que tienen impacto profundo en trayectorias vitales y corporativas e, incluso, en vidas humanas. Asimov fue quien planteó el primer marco de relación humana con máquinas inteligentes y autónomas, nada menos que en 1942, anticipando una problemática que ha arreciado con la llegada de la inteligencia artificial (IA). Ese marco se concretaba en las llamadas “Tres Leyes de la Robótica”:

1. Un robot jamás hará daño a un humano o, por inacción, jamás permitirá que un humano sufra daño.
2. Un robot debe obedecer las órdenes dadas por los seres humanos, excepto si éstas entran en conflicto con la primera ley.
3. Un robot debe proteger su existencia, en la medida en que esta protección no entre en conflicto con la primera o la segunda ley.

La situación, a principios de la segunda década del siglo XXI es más sencilla, y a la vez mucho más compleja que la que imaginó Asimov. Su futuro (nuestro presente), no ha sido el de un cosmos de galaxias pobladas por robots-humanoides a modo de tribus exóticas que conquistan un universo inexplorado. Seguimos en el planeta Tierra. Las misiones a la Luna acabaron en 1972, aunque China acaba de enviar una nave a su cara oculta, y amanece una nueva era de competición espacial entre EEUU y el gigante oriental.

La realidad es que la IA nos llega en modo de algoritmos, piezas de código informático, que se desparraman de forma ubicua por un universo digital que no está a años luz, sino que nos impregna en nuestra vida cotidiana. Estamos inmersos en un cosmos digital. Esos algoritmos han invadido nuestro entorno más cercano: se han adherido a nuestro teléfono, a nuestro automóvil, a nuestra nevera, a nuestros dispositivos médicos, a nuestra cámara fotográfica, a nuestra oficina o a nuestra línea de montaje. Grandes cerebros de hardware los actualizan y los hacen cada vez más inteligentes, no desde Andrómeda, sino desde clústeres tecnológicos concentrados en Palo Alto (California) o en Shenzhen. Y esos algoritmos intangibles e imperceptibles aprenden de sí mismos, cada vez saben más de nosotros, cada vez pueden hacer más cosas de las que hacíamos nosotros en el pasado, y cada vez pueden hacer que hagamos cosas que antes no hacíamos.

Esta realidad tan compleja e inesperada (incluso para Asimov) es la que tendremos que analizar y regular en los próximos años desde una perspectiva profundamente humanista. Los conflictos éticos se darán en diferentes dimensiones. En los siguientes apartados vamos a revisar algunas de ellas.

Una primera onda de choque derivada de la llegada masiva de la IA, que con toda probabilidad se va a producir en los años próximos, es la de la autonomía de las máquinas inteligentes. Las máquinas van a tomar decisiones autónomas cada vez más complejas, y con mayor impacto sobre organizaciones y trayectorias vitales. Se ha popularizado ya el dilema ético que conlleva la programación de sistemas autoconducidos. Ante un accidente imprevisto, una persona actúa de forma instintiva e inconsciente. No puede calcular las consecuencias de sus actos (como girar el volante y situarse en el carril contrario en la tesitura de evitar un atropello de alguien que aparece inesperadamente en la calzada). Pero un procesador de alta velocidad sí que puede hacer una proyección de daños, en todos los casos posibles. ¿Seguimos adelante y atropellamos a una familia? O, ¿damos un golpe de volante y colisionamos contra un automóvil que viene de frente, matando al ocupante de nuestro vehículo? De un modo implícito, la decisión está tomada por los departamentos de marketing: las marcas protegerán preferentemente a su cliente (a quien les paga), así que los vehículos atropellarán a los transeúntes. Los automóviles autoconducidos se anunciarán como absolutamente seguros (para sus ocupantes). El tema es mucho más complejo de lo que parece: ¿y si la persona atropellada es un niño o una mujer embarazada? ¿Y si quien cruza la calle lo hace saltándose

un semáforo en rojo? ¿Debería eso ser un factor de mérito para ser “sacrificado”? ¿Y si el damnificado es un importante personaje público? ¿Y si algún algoritmo, de un modo u otro, decide tomar la decisión en base al perfil socioeconómico del potencial afectado? ¿Y si mata a quien tenga una menor seguridad?

El MIT (Massachusetts Institute of Technology) ha creado un centro de investigación, moralmachine.mit.edu, que presenta dilemas morales asociados a la conducción autónoma, y que ya ha realizado 40 millones de encuestas on-line en más de 230 países. Según sus investigaciones, en casos como el presentado, los algoritmos deberían tomar las decisiones que salven mayor número de vidas, y preferentemente mayor número de vidas de niños. Sin embargo, existen significativas diferencias entre zonas geográficas y culturas en sus preferencias morales: hay países, por ejemplo, en los cuales se opta por salvar a los ancianos frente a los niños. La casuística, como vemos, es extremadamente compleja, y el proceso de conversión de la investigación que se lleva en curso en legislación (si es el caso), no será en modo alguno sencillo ni rápido.

Los sistemas sanitarios también van a ser profundamente alterados por las nuevas tecnologías digitales. Si hay un campo abonado para los datos y la IA, éste es el de la salud. En base a nuestro genoma, ¿nos gustaría saber que, con una cierta desviación estadística, es probable que contraigamos un cáncer a una cierta edad? O, en base a grandes series de datos de individuos con historiales clínicos similares, ¿podríamos saber la fecha aproximada de nuestra muerte? Por otro lado, ¿qué pasaría en el caso de que una máquina tome decisiones autónomas sobre ámbitos de salud humana? Técnicamente, parece evidente que un robot de alta precisión podría realizar una operación quirúrgica con mayores probabilidades de éxito que un humano. De hecho, los humanos requieren procesos de aprendizaje que, en el caso de la medicina, se deben realizar sobre pacientes reales. Se aprende a operar operando, por más simulaciones, estudios u operaciones previas sobre animales que se hayan hecho. Y, en la primera operación crítica, a un humano siempre le tiembla la mano. Un robot evitaría esas curvas de aprendizaje: no se pondría nervioso en la primera operación. Pero, ¿un paciente permitiría ser operado en una intervención a vida o muerte por una máquina totalmente autónoma? ¿Podría una máquina reaccionar ante acontecimientos imprevistos? ¿Tendría intuición y capacidad de improvisación? Obviamente, no. Sin embargo, el campo de la salud es uno de los grandes ámbitos

de actuación inmediata de las grandes plataformas digitales. Tardaremos en ser operados por máquinas autónomas, pero quizá no falte tanto tiempo para que Alexa nos diagnostique una gripe, y nos prescriba los medicamentos necesarios.

2/2

Hablamos de vehículos autoconducidos, pero, ¿y si creáramos empresas autoconducidas? Subimos a nuestro coche y, ya sin preocuparnos, una voz dirigida por GPS y sintetizada por una interfaz de Google nos indica el camino. Lo vemos natural (la tecnología tiene ese curioso efecto psicológico: dispositivos disruptivos, que rozan la ciencia ficción, se perciben como obvios y habituales -se dan por descontados. una vez los hemos empezado a utilizar). Aunque parezca contraintuitivo, será más sencillo sustituir a un director de una sucursal bancaria que al personal de limpieza de dicha entidad. El primero debe tomar decisiones sobre inversiones, que de forma cada vez más eficiente podrán ser tomadas por algoritmos, los cuales tendrán múltiples ventajas sobre los humanos: son a coste marginal cero (una vez entrenado el algoritmo, una decisión más no incrementa los costes de la organización), son mucho más rápidos, responden sólo a razones racionales y objetivas, trabajan de forma permanente, y son capaces de analizar millones de datos (externos e internos: condiciones macroeconómicas, perfil de riesgo de la inversión, histórico de las empresas invertidas, exigencias de objetivos corporativos, etc.). Por el contrario, un sistema de limpieza debería ser sustituido por un dispositivo mecatrónico (más complejo, más caro -no responde a coste marginal cero, hay que comprar un robot-, y cuya eficiencia mecánica en todas las situaciones posibles no está garantizada).

Las posiciones de management, por tanto, son susceptibles de ser desarrolladas por máquinas. ¿Podría una empresa “autoconducirse”? ¿Podría gobernarse una empresa a través de una estructura de management algorítmica? En algunas funciones, parece cada vez más evidente, como en el caso del análisis financiero o de inversiones. Ya en 2014 la empresa de capital riesgo Deep Knowledge Ventures, con base en Hong Kong y especializada en inversiones biotech “incorporó” un algoritmo en su consejo de dirección. El programa tenía derecho de voto sobre decisiones de inversión. Aunque el momento era prematuro, y la operación se propagó mediáticamente (quizá, en parte, era un reclamo de posicionamiento de la empresa), la idea no es descabellada. El algoritmo analizaba en profundidad la información financiera disponible, y rastreaba en el histórico de datos de la empresa a ser invertida en busca de patrones de decisión certeros. ¿Sería posible escalar los algoritmos autónomos a funciones de alta dirección? ¿Podríamos tener un director general digital? Teóricamente sí: un algoritmo que analizara de forma constante parámetros de evolución económica, e indicadores internos; y tomara decisiones y lanzara directivas para conseguir los objetivos corporativos deseados. Estudios del MIT, incluso, llegan a la conclusión de que los empleados podrían ser más felices con un director, o, en general, un jefe electrónico. Sólo existe, en mi opinión, un condicionante que evitará este escenario: al final, *siempre alguien deberá ser responsable de las decisiones tomadas* (ante los accionistas, ante la propiedad, ante los acreedores, ante la sociedad o ante la ley). En *management* sabemos que un directivo puede delegarlo todo, excepto la responsabilidad de sus actos. Y queda mucho camino, si es que se llega a este punto, para que las máquinas desarrollen sentido de la responsabilidad y asuman responsabilidades. Si se quiere, falta mucho para que la IA tenga temor al castigo (económico o incluso penal) ante un acto irresponsable. Ello me hace pensar que durante mucho tiempo, las máquinas

seguirán constituyendo cada vez más certeros sistemas de apoyo a la toma de decisiones. Pero la responsabilidad (y, por tanto, la última palabra en la decisión), seguirá siendo humana). La responsabilidad será la última gran reserva de *management* humano en un mundo gobernado por cerebros electrónicos. La IA constituirá un gran elemento de soporte, sumando inteligencias de máquina y humanas. En palabras de un profesor del MIT “los directivos no serán reemplazados por la IA, pero aquéllos que no dominen la IA sí que serán sustituidos por otros que la dominen”.

Stuart Russell, profesor de ciencias de computación en Berkeley realizó en 2017 un vídeo de ficción, *Slaughterbots* (algo así como “robots carniceros”), para sensibilizar a las Naciones Unidas sobre la necesidad de prohibir las armas autónomas basadas en IA.

En el video se visualizaba cómo un enjambre de miles de pequeños drones dotados de sistemas de reconocimiento facial era liberado sobre una ciudad, asesinando selectivamente a sus habitantes. Los drones se situaban sobre la potencial víctima, la reconocían, y, si optaban por asesinarla, se colocaban sobre su cabeza y hacían volar una pequeña carga explosiva.

Un sistema de este tipo mataría selectivamente: todos los varones mayores de 18 años, todos los de una determinada ideología (el algoritmo podría inducir, en base a la información recabada instantáneamente en redes, las opciones políticas de la víctima), o todos los de una determinada etnia. Los drones actuaban cooperativamente, de modo que se agrupaban y hacían explotar cargas mayores para acceder a edificios determinados.

“Las armas nucleares han quedado obsoletas”, según Russell.

El uso de la IA abre la puerta a nuevos escenarios bélicos.

“El campo de batalla no es un lugar para los humanos”, en palabras de Annie Jacobsen, autora del libro *DARPA:*

The Pentagon's Brain, en relación a los avances tecnológicos en armas autónomas que desarrolla DARPA, la agencia de investigación en proyectos de defensa de EEUU.

Las máquinas van a tener capacidad creciente de decisión autónoma. La potencia de las nuevas generaciones de IA reside en su capacidad de aprendizaje de la experiencia. La máquina aprende a medida que experimenta (en el llamado proceso de *machine learning*). Su lógica interna es modificada dinámicamente a medida que acumula y evalúa más datos. Sin embargo, esta fortaleza tecnológica es también su mayor debilidad: los algoritmos son cada vez más certeros en sus análisis, porque aprenden de forma autónoma, pero si se equivocan, no sabemos a qué se debe ese error. En la vieja lógica de los sistemas de información, un programador humano ideaba un árbol de decisiones que codificaba en un programario, que era transferido a una máquina. La máquina, en el fondo, ejecutaba como un autómatas el circuito lógico definido por el humano. Ahora ya no es así: la máquina modifica su propio software, hasta el punto de que el programador humano pierde de vista el proceso lógico que sigue el algoritmo para tomar una decisión. De hecho, precisamente esto hace a la máquina más humana. Es la famosa paradoja de Polanyi, filósofo austriaco que se dio cuenta de que somos capaces de hacer cosas que no podemos explicar. Por ejemplo, reconocer a un bebé (nuestro hijo), entre docenas de bebés muy parecidos. O reconocer a un viejo amigo al que encontramos por la calle, y que no hemos visto en 25 años. No hay árbol de decisión parametrizado, ni proceso lógico que lo

explique. No podemos transferir ese conocimiento a una máquina de ninguna manera (no podemos programar “tu amigo es un individuo de pelo negro, cuyas dimensiones faciales son 32 x 18...”). Simplemente lo reconocemos por experiencia, porque de un modo u otro en nuestro cerebro se ha grabado su patrón facial. Así aprenden también las máquinas, mediante acumulación de conocimiento experto. No hay un circuito lógico, no hay árbol de decisión trazable en la IA. Por tanto, si la máquina se equivoca en una decisión crítica, no podremos saber por qué se comete ese error.

07. Interacción persona-máquina

1/2

En el congreso tecnológico Dúplex I/O de 2018, el CEO de Google, Sundar Pichai presentó un sistema de síntesis de voz al que puso a prueba en tiempo real: le pidió a una máquina que llamara a una peluquería y que le reservara hora. El sistema lo hizo, estableció una conversación con la responsable que la atendió al teléfono, y negoció la hora. La peluquera no se dio cuenta de que estaba hablando con una máquina. El sistema digital de Google superó el llamado “test de Turing” (matemático británico considerado uno de los padres de la IA), según el cual llegaría un día en que hablaríamos con una máquina sin darnos cuenta de que es una máquina. Y esto nos lleva a un nuevo dilema: si los sistemas se perfeccionan al nivel de sintetizar la voz humana de forma indistinguible de la real, ¿debería el sistema declarar de forma anticipada que es un robot? De lo contrario, podríamos tener algoritmos programados en servidores que, 24 horas al día, 7 días a la semana, estén realizando llamadas comerciales a coste marginal cero (una llamada más no tiene coste para el algoritmo). Un algoritmo entrenado para la interacción comercial podría atender, simultáneamente, cientos o miles de clientes en oficinas bancarias.

Los sistemas de voz parece que serán los nuevos entornos operativos digitales. Si inicialmente los computadores eran

gobernados por un sistema operativo interno, las siguientes generaciones de smartphones y los ecosistemas móviles se desarrollaron sobre sistemas operativos visuales, orientados a la venta a través de apps (aplicaciones) y ads (anuncios visuales). Ahora parece que el futuro se va a gobernar por sistemas inmersivos de interacción por voz. *Hablaremos con las máquinas*, que cada vez dispondrán de mayor inteligencia emocional conferida por sus motores de IA. La avanzadilla la conforman dispositivos como Alexa, de Amazon, interfases de voz que en un breve tiempo posiblemente nos eviten los incómodos teclados. La apuesta de Amazon por la IA es incontestable: su presupuesto de I+D (el mayor del mundo) es de 22.000 M\$. Baste decir que el esfuerzo conjunto de la economía española en I+D (a nivel público y privado) es inferior, de unos 19.000 M\$. Amazon invierte en I+D más que el conjunto de la economía española, y lo hace de forma estructurada y dirigida hacia dispositivos de IA que maximicen sus ventas en canales on-line o en canales de voz. Los altavoces inteligentes, o, en un paso posterior, los bots digitales en forma de avatar (“humanos digitales”) analizarán nuestro lenguaje verbal y no verbal (mediante cámaras de visión artificial), y estarán entrenados para determinar el mejor momento de venta, nuestra predisposición a la compra o no, los atributos que más nos atraen... Si Amazon consigue introducir dispositivos Alexa en cualquier canal comercial, quizá nuestros clientes dejen de hablar con nosotros y empiecen negociar con Alexa. Si es así, Amazon se convertirá realmente en el gran Amazonas, el gran río de ventas que lo colecta todo, según la visión original de su fundador, Jeff Bezos.

08. Sustitución de personas por máquinas

1/4

En la perspectiva macroeconómica, las derivadas de los dilemas anteriores son evidentes: ¿Va a haber una sustitución masiva de personas por máquinas? ¿Nos enfrentamos a una *jobless society* (una sociedad sin trabajo)? ¿Estamos creando una *useless class* (una clase inútil, sin sentido económico), en palabras del historiador Yuval Noah Harari?

Brian Arthur, profesor de Stanford, anuncia la llegada de la “Era Distributiva”: en un mundo de rápido cambio tecnológico, donde la tecnología podría permitir un escenario de abundancia, el problema ya no es la Producción, sino la Distribución. Arthur declara el final de la Era Industrial, e identifica el principal problema de la nueva época: el sistema sabe cómo crear riqueza, pero no sabe cómo distribuirla. Los nuevos modelos de negocio digitales, encabezados por las grandes plataformas (Apple, Google, Facebook, Amazon y Microsoft, entre otras), tienen una gran capacidad de generación de riqueza, pero no crean empleos como antes. La vieja economía industrial era intensiva en empleos estables, en el marco de cadenas de valor territoriales, que al final consolidaban clases medias que sustentaban sistemas democráticos. Ahora, según Arthur, se genera una “Segunda Economía”, una economía virtual, externalizada a la red, que opera bajo dinámicas de coste marginal cero y que llevan a dominios monopolísticos (en los cuales, “el ganador se lo lleva

todo”). De un modo u otro, la economía internacional está siendo succionada por las redes informáticas, mientras los individuos se quedan fuera.

¿Es posible que las personas perdamos nuestro sentido económico? La digitalización rompe el círculo virtuoso del capitalismo industrial, inaugurado por Henry Ford a principios del Siglo XX: las fábricas creaban productos, que eran vendidos en los mercados a individuos que trabajaban en las fábricas (y debían ganar un salario digno para poder comprar esos productos). Si las personas son sustituidas por algoritmos y robots en las fábricas, ¿quién comprará los productos que permitan que el círculo virtuoso de sistema capitalista siga funcionando?

Todo parece indicar que la sustitución de personas por máquinas se acelerará en los próximos años. Según prospectivas de la Universidad de Oxford, hacia 2066 toda tarea humana podría ser ejecutada por sistemas digitales o mecatrónicos (incluidas tareas complejas y creativas como la investigación científica o la escritura de best-sellers para el New York Times). Si esto es así, el sistema capitalista deberá introducir importantes innovaciones sociales para mantenerse. La alternativa, el colapso del sistema y la posible extensión de populismos y neofascismos.

Una solución que se ha popularizado en los últimos tiempos es la Renta Básica Universal (RBU). Ésta sería una renta incondicional (concedida independientemente del nivel de renta del individuo), y básica (que impidiera caer bajo el nivel de pobreza). Se propone, precisamente, en el seno de los foros tecnológicos de Silicon Valley (aunque es una idea que se remonta al Renacimiento), y es defendida por gurús de la innovación (Mark Zuckerberg – Facebook-, Bill Gates -Microsoft- o Elon Musk -Tesla-), o científicos como el recientemente fallecido Stephen Hawking. Silicon Valley, de hecho, según Thomas Picketty (autor de El Capital en el Siglo XXI) es uno de los lugares

con mayor desigualdad de la historia y del planeta. En EEUU, los salarios se han mantenido aproximadamente igual, descontando la inflación, desde 1975, pese al intenso cambio tecnológico que se ha producido. Se genera riqueza (el conjunto de ingresos de los hogares norteamericanos es de alrededor de 8 billones brutos anuales), lo que daría para unos ingresos de alrededor de 80.000 \$ anuales por hogar. Pero 40 millones de norteamericanos viven en la pobreza.

La RBU es una idea que merece ser analizada serenamente. La hipótesis de partida es situarnos en el seno de sociedades avanzadas, cultas y con sólidas economías basadas en conocimiento, pero cuyo grado de digitalización y automatización dificulte la distribución de la riqueza mediante mecanismos salariales, como en la vieja economía industrial. En ese entorno, la RBU propugna distribuir parte de la riqueza nacional de una economía, como si de una cooperativa se tratara. Las fuentes de esa riqueza no serían los recursos naturales (como ocurre ya en países como Noruega), sino el talento y la tecnología desarrollada por las naciones. Existen defensores y detractores en todo el espectro político. Para pensadores de izquierdas, la RBU significaría la abolición directa de la pobreza. Para filósofos liberales, un instrumento de libertad personal (nadie es libre si está sometido a la tiranía de la pobreza). Para algunos conservadores, un elemento sustitutivo de sistemas de bienestar social ineficientes y sobrecargados de costes indirectos asociados al control, y de incentivos perversos que crean trampas de pobreza (tengo una renta porque soy pobre, y no salgo de la pobreza porque no quiero perder la renta). Incluso alas ultraconservadoras norteamericanas la defienden como salvaguarda ante una posible revolución global.

Por supuesto, existen también fuertes detractores, que critican su elevado coste (una RBU real es imposible hoy en cualquier economía), la desincentivación del trabajo, o los posibles efectos

de inflación o llamada. Y, algunos, argumentan también una razón moral: ¿es moralmente justo distribuir una renta entre individuos que opten entonces por renunciar al trabajo? Aunque, el contraargumento es sencillo: ¿es moralmente justo condenar a la pobreza a quien necesite trabajar y no encuentre trabajo?

¿Amenaza la IA la democracia? Quizá, no sólo por la extensión de la desigualdad que generan los sistemas de negocio que lideran la economía digital, sino también por la posibilidad de estructurar sistemas tecnocráticos o autocráticos, o por la posibilidad de manipular la opinión pública a través de sistemas de datos. Es conocido el escándalo de Cambridge Analytics, según el cual datos recogidos de Facebook permitían deducir la sensibilidad política del individuo, e inyectar material político en las redes para decantar sus votos. Las propias redes sociales son campos de extremos: en un mundo de abundancia de información, tendemos a seleccionar aquella que nos es ideológicamente próxima. Visitamos foros y opiniones que nos gustan, reforzando nuestros planteamientos ideológicos. Twitter es terreno abonado para extremistas, y campo de formación y entreno de radicales e intolerantes.

La emergencia de China (una suerte de autocracia automatizada) es un ejemplo que está cambiando el orden mundial. China ha realizado una apuesta estratégica por la IA, y posee ventajas competitivas indudables: la posibilidad de planificar en el largo plazo, la disposición de inversiones masivas, y la gestión de flujos de datos insuperables, provenientes de sus ciudadanos. Hoy China está desplegando sistemas de control *big data* basados en reconocimiento facial, posicionamiento, y algoritmos

de IA para crear rankings sociales. Cada ciudadano tiene una puntuación dinámica, generada a través de datos de su comportamiento y hábitos: resultados académicos, evaluaciones profesionales, informes médicos, comentarios en las redes sociales, tipos de productos que adquiere... Y su puntuación le da acceso automático (o deniega) servicios sociales, autorizaciones administrativas, o matrículas académicas. El mundo de la abundancia tecnológica puede ser distópico en lugar de utópico.

Las distopías bajan al nivel de las propuestas de valor empresariales cuando entra en juego la IA. Nos venderán dispositivos cada vez más humanizados. Hoy existen startups que ofrecen realizar un clon digital de nosotros mismos. Aprender de nosotros, leer nuestros mails, registrar nuestra memoria, mimetizar nuestro temperamento, almacenar nuestros documentos gráficos, conformar nuestra cara en un avatar digital, imitar nuestra voz e interactuar con terceros con inteligencia emocional como si fuéramos nosotros mismos. De tal manera que, cuando muramos, nuestros hijos puedan hablar con nuestro avatar digital, un resto informacional inteligente de nosotros. ¿Sería eso algo así como un resto mortal de una persona? ¿Qué tratamiento legal debería tener, si es que debemos darle tratamiento legal?

10. Superación de las personas por las máquinas

1/2

El cerebro humano es un computador biológico formado por unos 100.000 millones de neuronas, cada una con unas mil conexiones sinápticas a otras neuronas. Si se sigue cumpliendo la famosa Ley de Moore (según la cual cada año y medio aproximadamente se dobla la capacidad de integración de dispositivos electrónicos en un sistema integrado), hacia 2040 existirán chips con la densidad de conexión de un cerebro humano. Intel o Microsoft ya se han lanzado al desarrollo de arquitecturas de silicio que imiten el cerebro humano. De hecho, los sistemas de IA se basan en hardware inspirado en el funcionamiento del cerebro (las famosas redes neuronales). Estas aproximaciones han hecho posible el desarrollo de sistemas electrónicos con capacidades cognitivas (de interpretación de la voz, de texto o de imágenes). Sabemos generar IA a imagen de la inteligencia humana. Pero no sabemos si esa inteligencia derivará en *consciencia artificial*. ¿Llegará una máquina a darse cuenta de que existe? La ciencia, hoy, no tiene claro cómo y dónde emerge la consciencia. En un momento dado, un cúmulo de neuronas genera un patrón superior y se da cuenta de que existe. Y es un misterio cómo se produce ese fenómeno. Parece que la cantidad de neuronas conectadas, y las interacciones que generan hace aparecer un patrón superior, como los pixels de una fotografía hacen emerger un paisaje. Y, si las máquinas generan consciencia, posiblemente entonces generen también

sentimientos. ¿Puede un algoritmo tener ego, envidia amar u odiar? ¿Somos algoritmos biológicos? ¿Es el cerebro sólo una máquina? Porque si es así, algún día las máquinas serán como nuestros cerebros, y deberían generar algún tipo de consciencia. Pero quizá no seamos sólo máquinas. Quizá existe algo más, algo intangible, etéreo y quizá inmortal llamado alma. Si los humanos somos más que simples máquinas, si existe un alma, entonces una máquina jamás será como nosotros. El progreso de la tecnología y la ley de Moore están en los albores de determinar finalmente quiénes somos.

Los grandes desafíos que genera la IA van mucho más allá de la investigación científica, la innovación tecnológica y sus aplicaciones prácticas en la economía. Como hemos visto, se plantean profundísimos dilemas éticos que deberán revestirse de un marco regulatorio sensato y eficiente. Conscientes de ello, las grandes marcas líderes en el ámbito (inicialmente IBM, Amazon, Apple, Google, Facebook y Microsoft, y ahora muchas más) crearon en 2016 un consorcio de investigación y debate sobre las implicaciones éticas de la IA (partnershiponai.org). Probablemente veamos emerger sistemas de certificación y calidad ética de nuestros sistemas de información. Los años que se presentan son tan apasionantes como inquietantes. Esperemos que sepamos articular los foros de reflexión, los mecanismos y los liderazgos que permitan que el futuro ofrezca tantas posibilidades como la revolución tecnológica realmente posibilita, un futuro digno y justo para todos por primera vez en la historia de la humanidad.

Ética Inteligencia Artificial

Impulsores:

Más información y contacto: Javier Sánchez jsanchez@aedashomes.com

