PROSEGUR RESEARCH

Highlights

Visión artificial: lo que una máquina puede ver

PROSEGUR
SECURITY

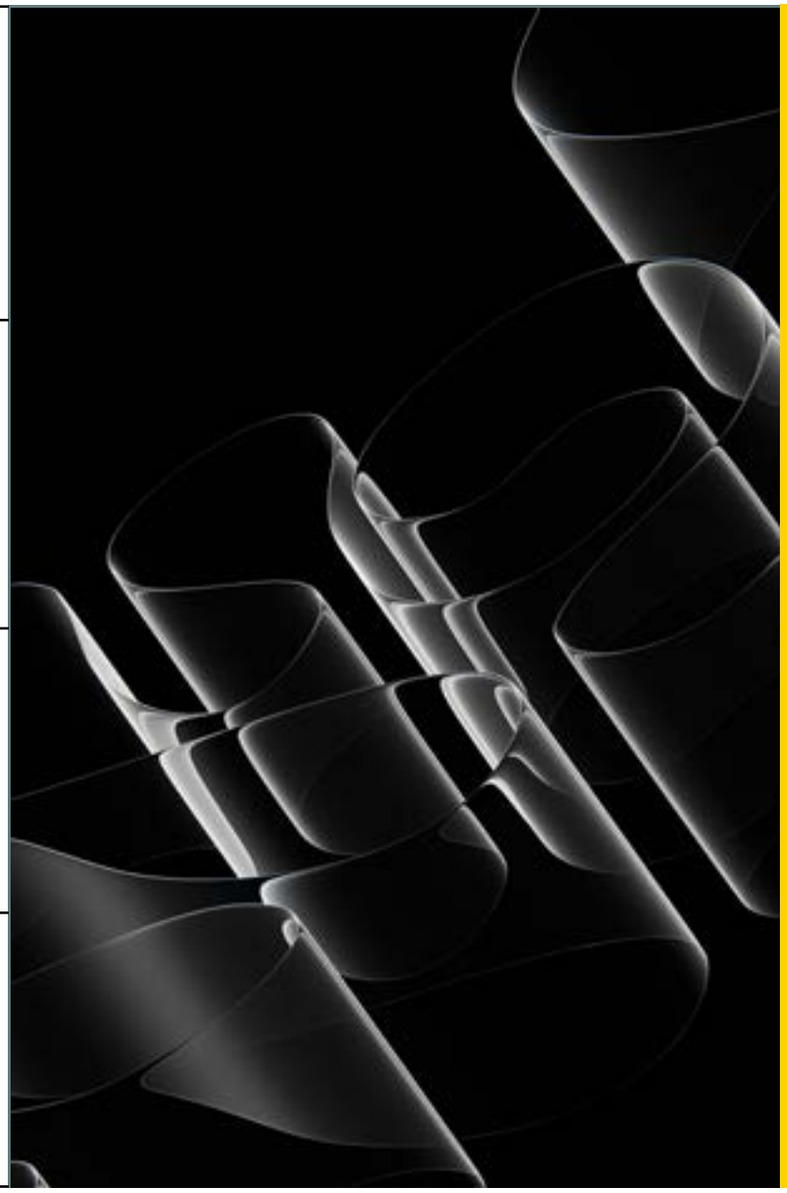
Contenidos

01 **Ver para entender: el origen y la idea de la visión artificial**

02 **¿Para qué puede usarse la visión artificial?**

03 **¿Cómo es posible que una máquina pueda ver?**

04 **Intregración de la visión artificial: progreso tecnológico y sensibilidad humana**



01

Ver para entender: el origen y la idea de la visión artificial

PROSEGUR RESEARCH

Highlights

La visión artificial es la **capacidad de una máquina para interpretar y comprender el mundo visual en base a diferentes modelos algorítmicos para extraer significado.** Aunque hoy suene a tecnología de vanguardia, sus raíces se remontan a las primeras décadas de la informática moderna. En los años 60, investigadores del MIT ya soñaban con dotar a las máquinas de ojos digitales capaces de "ver" el entorno. Pero no fue hasta bien entrada la década de 2010, con la revolución del deep learning, cuando esta disciplina comenzó a desplegarse masivamente en aplicaciones del mundo real.

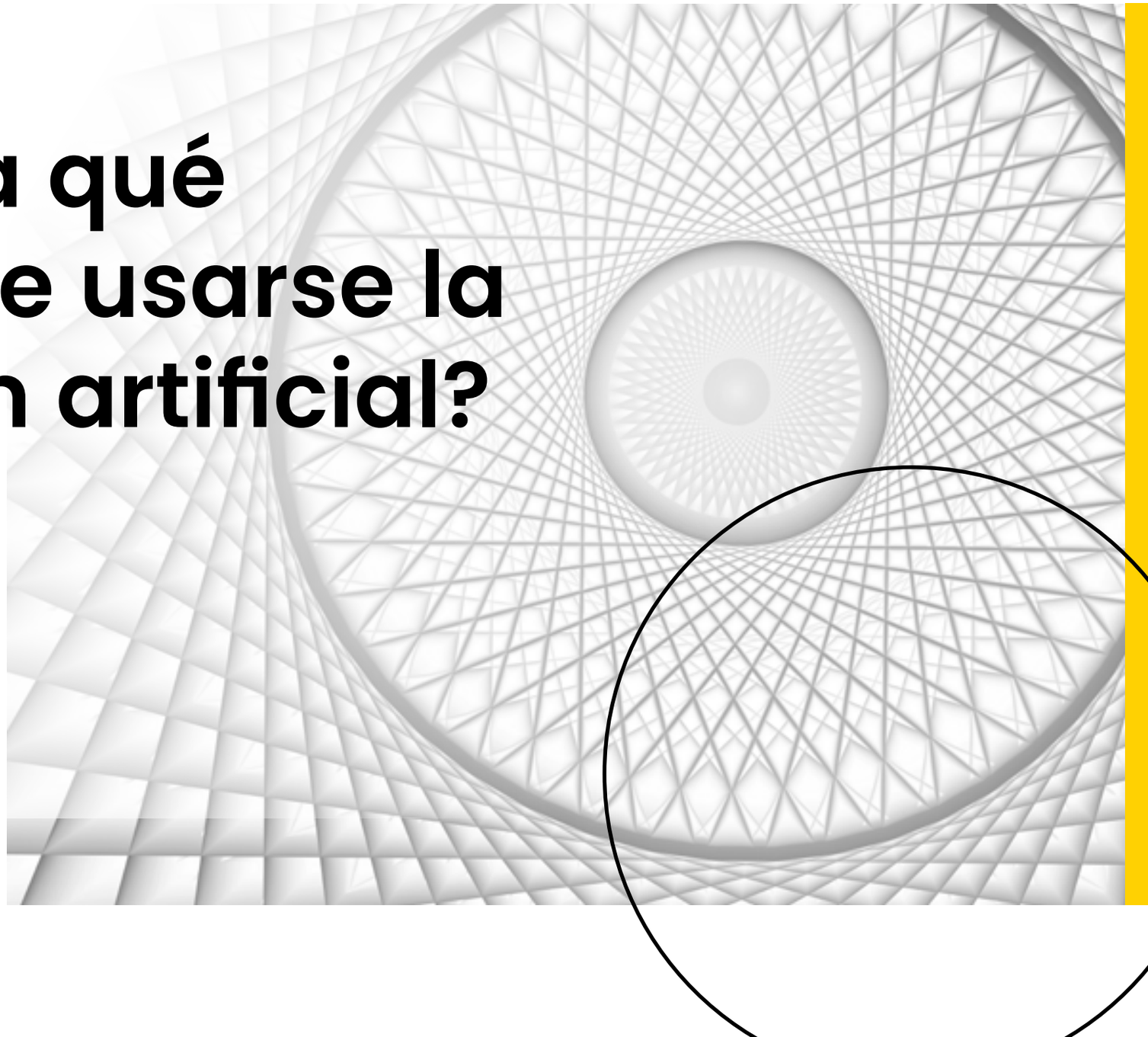
¿Pero qué significa ¨ver¨ para una máquina? La visión de las máquinas

En esencia, la tecnología trata de convertir imágenes –conjuntos de píxeles sin significado inherente– en información comprensible y útil. Para ello, la visión artficial se apoya en una **cadena de procesamiento que imita, a su manera, el sistema visual humano, aunque con procesos radicalmente distintos:** primero capta imágenes a través de sensores (como una cámara), luego las analiza para extraer patrones visuales (bordes, formas, texturas), y finalmente las interpreta, asignándoles significado en función del **contexto.**

Este proceso implica una combinación de óptica, matemáticas, estadística, inteligencia artificial y neurociencia. No basta con reconocer un objeto: hay que **entender qué es, qué está haciendo y cómo se relaciona con el entorno.** Por eso, la visión artificial no solo trata de "ver", sino de **procesar e interpretar lo que se está detectando.** Esta distinción resulta clave para valorar su potencial transformador: no estamos ante una simple evolución de las cámaras de seguridad, sino ante el nacimiento de sistemas capaces de **apoyar y ejecutar decisiones en tiempo real** a partir de estímulos visuales, anticiparse a riesgos, y operar de forma autónoma en entornos complejos.

02

¿Para qué puede usarse la visión artificial?



PROSEGUR RESEARCH
Highlights

La visión artificial es una **tecnología versátil** con un abanico sorprendentemente amplio de aplicaciones prácticas, cuya influencia se extiende transversalmente por distintos sectores de la sociedad. Su potencial reside en la capacidad no solo de automatizar procesos visuales, sino en **mejorar significativamente la precisión y la calidad de las decisiones** tomadas a partir de la interpretación inteligente de imágenes y vídeos[1].

Uno de los ámbitos más inmediatos y de mayor impacto es el de la **seguridad**, donde la visión artificial aporta una **ventaja estratégica determinante al anticipar y detectar amenazas** con una rapidez y precisión superior a las capacidades humanas. Mediante análisis de patrones de comportamiento, esta tecnología permite **identificar en tiempo real situaciones potencialmente peligrosas o sospechosas**, facilitando respuestas ágiles y efectivas. Así, personas, datos y tecnología se integran en un ecosistema de protección eficaz y eficiente en tiempo y recursos.

[1] Para profundizar se recomienda ¨Visión¨ de David Marr.

Adicionalmente, la mejora en la detección de amenazas en las monitorizaciones es un aspecto destacado en la agilidad de los algoritmos específicamente en la detección de movimientos sospechosos, intrusiones o anomalías en el entorno; así se convierten en una herramienta eficaz que complementa las labores del personal de seguridad, añadiendo valor a su imprescindible labor humana. La sinergia entre estas tecnologías que permiten el aprovechamiento de los datos combinados con la relevante experiencia humana se traduce en un nivel de seguridad nunca antes alcanzado, lo que denominamos **seguridad híbrida** en Prosegur.

Un ejemplo de ello sería la utilidad en un evento masivo como un festival de música internacional o una competición deportiva, donde los algoritmos nos permiten garantizar la seguridad de sus asistentes, gracias, por ejemplo, al computer vision. Así, la inteligencia artificial nos permite detectar antes y mejor elementos como armas o comportamientos sospechosos a partir de la información en tiempo real, y desde el profesional a cargo de la cámara se procesa en ese mismo momento, pudiendo informar de manera inmediata al vigilante presencial.

Por eso decimos que empodera, y cada vez con un mayor nivel de sofisticación: por ejemplo, ahora podemos detectar humo mediante imágenes mucho antes que los sensores tradicionales en las condiciones adecuadas de visibilidad.

Otros sectores también beneficiados por este avance tecnológico incluyen el **sector sanitario**, la visión computacional mejora la exactitud diagnóstica y acelera la detección de enfermedades mediante el análisis automatizado de **imágenes médicas** (radiografías, resonancias magnéticas). Estos sistemas facilitan **diagnósticos tempranos y personalizados,** multiplicando la efectividad de los tratamientos y contribuyendo a **salvar vidas.**

En el ámbito del **transporte**, el reconocimiento en tiempo real del entorno impulsa el desarrollo de **vehículos autónomos** capaces de identificar peatones, obstáculos o cambios imprevistos en la carretera, lo que **aumenta la seguridad vial y reduce notablemente la accidentalidad en este tipo de vehículos.**

La **inspección visual automática en la industria** permite detectar al instante **fallos de calidad** o riesgos en procesos de producción críticos. Esta automatización visual no solo **reduce costos**, sino que garantiza **estándares de calidad constantes.**

FInalmente, en la **agricultura**, mediante, por ejemplo, cámaras multiespectrales, la visión artificial monitoriza cultivos con precisión milimétrica, identificando necesidades de agua, fertilización o control de plagas. Asimismo, entre sus aplicaciones más extendidas se encuentra la identificación de plantas y, de manera accesoria a sistemas robóticos, la automatización de las cosechas. En suma, el uso de esta tecnología permite la promoción de un uso más eficiente de recursos naturales, impulsando las capacidades para lograr una **producción sostenible y más segura.**

En definitiva, la visión artificial se eleva más allá de un mero avance tecnológico donde la síntesis entre el ingenio humano y la fuerza del algoritmo crea un tejido inteligente capaz de descifrar el mundo visual con matices sobrehumanos. Es la **alianza entre el ojo digital y la intuición analítica** lo que nos permite leer el pulso de la realidad antes de que la sombra del riesgo se materialice, trazando rutas de protección que vibran con precisión y resiliencia. En este cruce de luz y datos, las máquinas ya no solo ven, sino que también emulan las habilidades humanas de comprensión e intuición, transformando cada fotograma en una oportunidad para anticipar amenazas y reforzar nuestra capacidad de reacción.

03

¿Cómo es posible que una máquina pueda ver?

La visión artificial (o computacional) se beneficia enormemente de la capacidad de las **Redes Neuronales Convolucionales (CNN)** para extraer características visuales complejas y, más recientemente, de la versatilidad de los modelos basados en Transformers (como los Vision Transformers y los **Modelos de Lenguaje-Visión multimodales - VLMs)** para procesar secuencias y relacionar información visual con la semántica del lenguaje natural. **Las CNN son modelos de inteligencia artificial diseñados específicamente para trabajar con imágenes: funcionan aplicando una serie de "filtros"** que captan patrones locales (bordes, texturas) y, a medida que atraviesan sucesivas capas, construyen representaciones cada vez más complejas (formas y objetos). Por su parte, los LLM son sistemas entrenados con enormes volúmenes de texto –y en su variante multimodal, también con datos visuales– **capaces de comprender, relacionar y generar lenguaje natural, aportando contexto y significado a las salidas de las CNN.**

Captura y preprocesado

Las imágenes se obtienen de sensores (cámaras, drones o satélites) y se normalizan para corregir brillo, contraste y eliminar ruido, garantizando insumos homogéneos para el análisis posterior.

Extracción de características

Las CNN despliegan filtros entrenados por capas sucesivas. Las capas iniciales resaltan bordes y texturas; las intermedias combinan estos elementos en figuras simples; y las más profundas ensamblan dichas figuras hasta reconocer objetos completos (por ejemplo, un rostro o un vehículo). Cada filtro ajusta sus parámetros aprendiendo de un gran volumen de imágenes etiquetadas, optimizando así la detección.

Razonamiento contextual

Los LLM multimodales toman los vectores de características generados por la CNN y los enlazan con un vasto corpus de conocimiento textual y visual. Esto permite asignar significado práctico (identificar “contenedor de carga” en lugar de “rectángulo sombreado”), integrar datos externos (ubicación geográfica, registros históricos) para enriquecer la interpretación para, finalmente, emitir conclusiones en lenguaje natural o generado conforme a reglas operativas.

Salida y acción

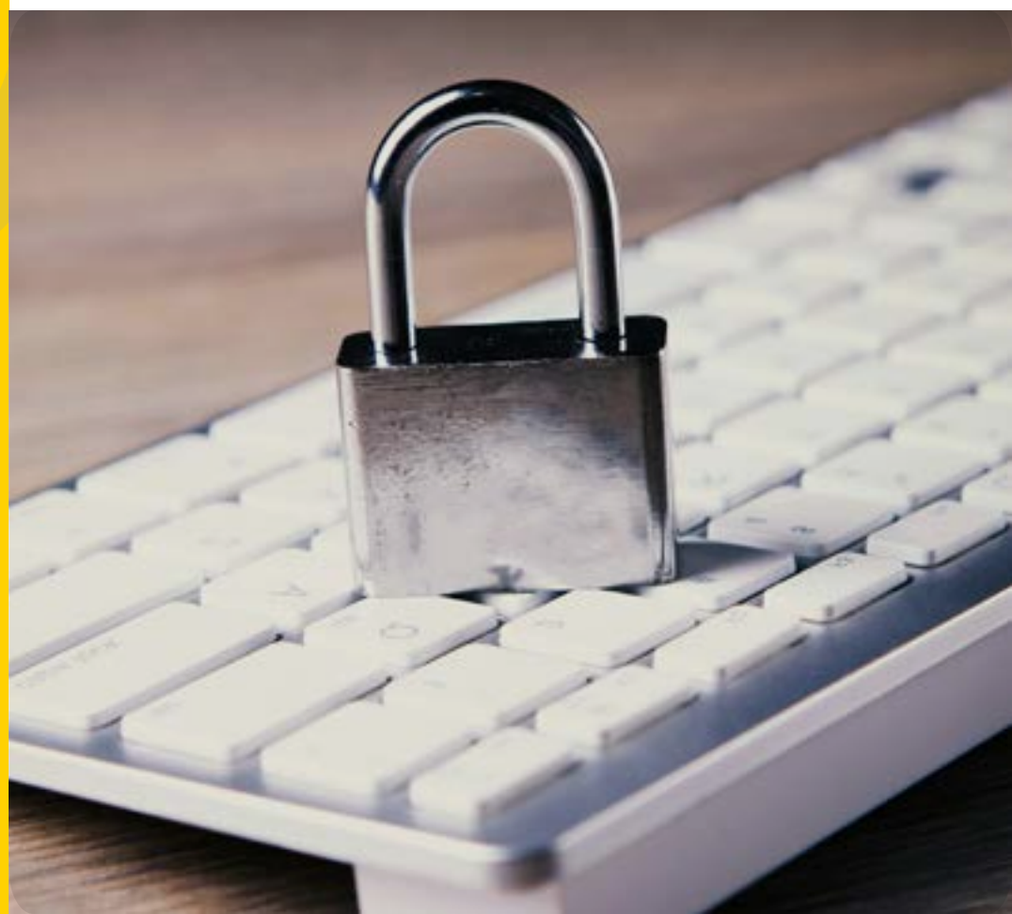
El resultado final combina métricas precisas con descripciones semánticas, habilitando desde alertas de seguridad en tiempo real hasta instrucciones detalladas para sistemas autónomos.

Asimismo, la visión artificial incorpora otras piezas fundamentales que, conjuntamente, forman el **ecosistema** completo de esta tecnología: desde la **segmentación de imágenes**, que separa objetos y fondos; el **procesamiento de vídeo** en tiempo real, que analiza secuencias continuas de fotogramas; la **calibración de sensores**, imprescindible para medir con precisión distancias y ángulos; hasta los **algoritmos de optimización** que mejoran el rendimiento y permiten al sistema aprender de datos etiquetados y descubrir patrones por sí mismo.

De la innovación al protocolo: hacia una seguridad inteligente como norma

La visión artificial ha superado ya la fase de tecnología emergente para consolidarse como una herramienta operativa integrada en el ámbito de la seguridad y más allá. A pesar de la latencia de cierta incertidumbre en torno a los conflictos éticos y legales que subyacen a su utilización, siguiendo la teoría de la "domesticación tecnológica" (Silverstone, Hirsch y Morley, 1992) puede afirmarse que se trata de una **tecnología prácticamente domesticada**, habiendo superado la fase de integración para asentarse como una innovación adaptada a la realidad del mercado en organizaciones y usuarios; quienes comprenden y proyectan su uso como parte de un sistema integrado.

Este fenómeno puede observarse claramente en la **proliferación de sistemas de vigilancia inteligentes**, pues lo que comenzó como experimentos aislados de análisis de vídeo y reconocimiento de patrones hoy forma parte del arsenal estándar de la seguridad –desde la detección automática de intrusos hasta el control biométrico de acceso–, consolidando la visión artificial como un elemento cotidiano en los servicios de seguridad.



La magnitud del cambio y la extensión de la tecnología queda reflejada en las proyecciones de mercado, las cuales estiman que el mercado global de IA para visión computacional **crecerá de 56,4 mil millones de dólares en 2025 a 117,0 mil millones en 2030,** impulsado en gran parte por la creciente adopción de soluciones de “vigilancia inteligente” y por el auge de los vehículos autónomos. Parte del capital tecnológico y cultural de las organizaciones, erigiéndose como un recurso que contribuye a definir nuevos protocolos operativos y a reforzar la resiliencia de los sistemas de protección. En síntesis, la domesticación de la visión artificial señala su madurez: esta tecnología ya no es un experimento marginal, sino un componente establecido en el tejido operativo –especialmente en seguridad– que se manifiesta como un valor competitivo asumido.

04

Integración de la visión artificial progreso tecnológico y sensiblidad humana

El verdadero reto que plantea la visión computacional no radica únicamente en su desarrollo tecnológico, sino en **cómo la sociedad logrará integrarla de manera efectiva, segura y responsable** en herramientas fundamentales ya existentes en nuestro día a día. Como se ha tratado, esta tecnología representa una innovación disruptiva en sectores clave como la seguridad, la sanidad, la movilidad o la industria, lo que requiere **no solo una adaptación tecnológica**, sino también cultural, ética y normativa.

Paralelamente, resulta indispensable promover una **explicación clara y accesible** que permita a todos los actores sociales comprender las implicaciones de esta revolución tecnológica, así como una normativa acorde al progreso técnico que deseamos para la civilización del futuro. Se trata, por tanto, de **educar, divulgar y sensibilizar a la ciudadanía** sobre cómo esta tecnología puede mejorar sustancialmente su calidad de vida.

Una implementación eficaz implica incorporar la visión computacional en **herramientas críticas** que ya forman parte del **tejido social y empresarial**: cámaras de vigilancia, sistemas de diagnóstico médico, vehículos autónomos, controles industriales y sistemas de identificación y autenticación. Este proceso debe ser **progresivo**, asegurando en todo momento que la introducción de esta tecnología **preserve la privacidad y respete los derechos fundamentales**, un equilibrio esencial entre **progreso y protección** en el paradigma de seguridad híbrida que Prosegur impulsa, combinando la fuerza de la inteligencia artificial con la sensibilidad y el criterio humano.

Esta integración requiere una **adaptación progresiva pero decidida de la sociedad**, desde la ciudadanía hasta las organizaciones públicas y privadas. Solo desde el cultivo de **ecosistemas colaborativos entre empresas, gobiernos, academia y sociedad civil** facilitaremos la generación de conocimiento compartido, regulaciones claras y procesos transparentes que conviertan la visión artificial en un motor positivo y seguro del desarrollo humano.

En definitiva, **la clave para afrontar el desafío no es solo tecnológica, sino social**; la visión computacional, integrada responsablemente, promete transformar nuestra realidad cotidiana, siempre que estemos preparados para **acompañar su crecimiento** desde la **ética**, la **transparencia** y la **seguridad**.

PROSEGUR
SECURITY

Garantizamos la seguridad
de las personas, las empresas
y la sociedad en su conjunto.

research@prosegur.com PROSEGUR RESEARCH www.prosegurresearch.com