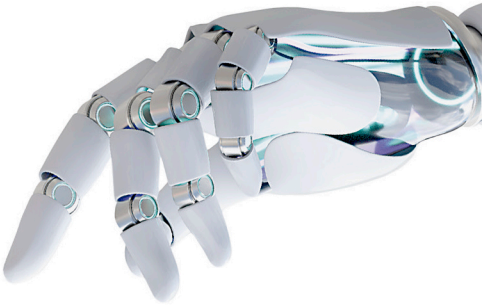
Matías Nahon



Inteligencia Artificial

El poder invisible

BÄRENHAUS

Matías Nahon

Inteligencia Artificial

El poder invisible

BÄRENHAUS

Índice

Prólogo

Este libro nace desde una incomodidad, desde una pregunta urgente que me acompaña hace años: ¿qué estamos dejando entrar, sin darnos cuenta, cada vez que incorporamos una nueva tecnología en nuestras vidas? Durante más de dos décadas trabajé en la prevención del fraude y vi de cerca cómo se armaban estafas, cómo se sofisticaban las operaciones ilegales, cómo las grietas de los sistemas se convertían en autopistas para el delito. Escribí *El fraude de la era digital* como respuesta a esa experiencia: una alerta sobre los vacíos que la tecnología dejaba abiertos cuando no era pensada con responsabilidad.

Hoy, el escenario cambió. Y lo hizo a una velocidad que nos desborda. Por eso este nuevo libro, centrado en la inteligencia artificial, es una continuación natural de ese camino, pero también una transformación: ya no se trata sólo del fraude, sino del colapso de nuestras formas de entender el control, la privacidad, la confianza y, en definitiva, la verdad.

La inteligencia artificial se presenta como el mayor avance tecnológico de nuestra época. Promete eficiencia, predicción, automatización y creatividad.

Está en todas partes: en los teléfonos que usamos, en los servicios que consumimos, en las decisiones que nos afectan sin que lo sepamos. Pero detrás de esa promesa se esconde también una arquitectura invisible de poder, un conjunto de sistemas que aprenden, deciden, jerarquizan, sin rostro ni explicación. Este libro intenta desarmar esa maquinaria, comprender cómo funciona, quién la controla —si es que alguien lo hace— y qué implicancias tiene para nuestras sociedades. No es un tratado técnico. Es un mapa de riesgos. Un recorrido que parte de los principios de la IA, revisa su historia, observa sus usos concretos y, sobre todo, pone el foco en sus amenazas más profundas: las que desdibujan la línea entre lo real y lo falso, entre la decisión humana y la automatización sin criterio.

¿Por qué hablar de riesgos? Porque la historia de la inteligencia artificial no es sólo la de los grandes avances, sino también la de sus consecuencias no previstas. En el campo de la prevención del fraude, lo sabemos bien: toda innovación que mejora procesos puede ser usada para sabotearlos. La IA no es la excepción. Al contrario: multiplica el potencial de daño. Y lo hace de formas que apenas estamos empezando a entender. Por eso, a lo largo de este libro, vamos a analizar cómo la IA ya está siendo utilizada para clonar voces y suplantar identidades, para fabricar imágenes, documentos y perfiles falsos con una precisión que engaña incluso a los sistemas de seguridad más avanzados. Vamos a hablar de extorsiones personalizadas a través de redes sociales, de algoritmos que

discriminan sin saberlo, de decisiones automatizadas que afectan derechos fundamentales. Y también de la dificultad de responsabilizar a alguien cuando el "error" lo comete una máquina.

La irrupción de la inteligencia artificial (IA) en nuestras vidas ha sido tan vertiginosa como transformadora. En pocos años, hemos pasado de contemplar la IA como una promesa futurista a verla integrada en aplicaciones cotidianas que van desde asistentes virtuales hasta sistemas de recomendación personalizados. Sin embargo, este avance exponencial no está exento de sombras. La misma tecnología que potencia la eficiencia y la innovación, también abre puertas a riesgos antes inimaginables.

Uno de los peligros más alarmantes es la clonación de voz mediante IA. Esta técnica permite replicar con asombrosa precisión la voz de cualquier individuo, basándose en muestras de audio disponibles públicamente. Delincuentes han aprovechado esta capacidad para realizar estafas telefónicas convincentes, haciéndose pasar por familiares o ejecutivos de empresas, y solicitando transferencias de dinero bajo falsos pretextos. La generación de imágenes y videos falsos, conocidos como *deepfakes*, ha alcanzado un nivel de sofisticación que desafía la capacidad humana para discernir entre lo real y lo ficticio.

En el sector financiero, se han documentado casos donde *deepfakes* se utilizan para eludir sistemas de verificación de identidad, facilitando fraudes bancarios y suplantaciones de identidad. Las redes sociales, concebidas como plataformas para conectar

personas, se han convertido en terreno fértil para nuevas formas de extorsión. La privacidad, un derecho fundamental, se ve amenazada por la capacidad de la IA para recopilar, cruzar y analizar datos personales más allá de lo que ninguna legislación nacional puede hoy controlar con eficacia. La inteligencia artificial no sólo nos observa: decide por nosotros, muchas veces sin que lo sepamos.

Pero hay algo más profundo. Algo que no remite sólo a la eficiencia o al engaño, sino a la forma en que la inteligencia artificial desorganiza la experiencia misma del tiempo. Las decisiones ya no llegan después de una deliberación: se anticipan, se imponen, se ejecutan en milisegundos. Vivimos en un presente saturado de futuros posibles que actúan antes de que podamos pensarlos. Esa aceleración no sólo cambia la velocidad: cambia la lógica del mundo. Ya no elegimos. Reaccionamos.

En ese sentido, el riesgo de la IA no está sólo en lo que hace, sino en lo que hace posible sin que lo notemos: la delegación del juicio, la fragmentación de la responsabilidad, la aceptación del algoritmo como árbitro de lo real. La IA no sólo automatiza tareas: automatiza criterios, define umbrales, organiza decisiones. Y lo hace desde una infraestructura que se presenta como neutral, técnica, inevitable. Pero no lo es. Está diseñada, entrenada, financiada. Responde a intereses. Y como toda tecnología poderosa, concentra poder.

La pregunta entonces no es si la inteligencia artificial puede ser útil, sino en qué condiciones su uso desplaza nuestra capacidad de entender, intervenir y

resistir. Porque cuando el modelo decide, pero nadie puede explicarlo; cuando el sistema predice, pero nadie puede discutirlo; cuando todo funciona, pero ya no sabemos cómo, entonces la IA ya no es una herramienta. Es un régimen. Un régimen que no necesita mostrarse para gobernar.

Ese es el verdadero riesgo: que el futuro no llegue como una amenaza, sino como una normalidad indiscutida. Que el poder no se ejerza con violencia, sino con automatización. Que el control no se imponga, sino que se naturalice. Por eso este libro no busca sólo alertar sobre fraudes, sino proponer una pausa: un gesto de interrupción. Porque pensar —a tiempo— es, quizá, la única forma de resistencia.

Y todo esto no ocurre sólo en los laboratorios de Silicon Valley. Pasa también en Argentina y en Latinoamérica.

En un país con un sistema judicial saturado, con organismos de control débiles, con una infraestructura tecnológica desigual y una cultura digital aún en desarrollo, los riesgos de la inteligencia artificial se potencian de manera particular. En 2024, el gobierno nacional de la Argentina anunció la incorporación de sistemas predictivos de delitos basados en IA. Si esos modelos se alimentan de datos sesgados —y todo indica que así es—, el resultado será la profundización de prácticas de discriminación y vigilancia selectiva. En lugar de prevenir el delito, podrían terminar amplificándolo, instalando una lógica de sospecha permanente sobre determinados sectores de la sociedad.

Frente a este escenario, el problema no es sólo técnico. Es cultural, ético y político. ¿Quién define qué es un uso legítimo de la inteligencia artificial y qué no lo es? ¿Quién decide cuándo una predicción es confiable y cuándo es un acto de discriminación estadística? ¿Quién se hace responsable cuando una máquina comete un error, pero nadie puede explicarlo? En muchos casos, las respuestas a estas preguntas brillan por su ausencia. Y esa ausencia no es casual: forma parte de una lógica más amplia, en la que la tecnología avanza más rápido que nuestra capacidad para regularla, comprenderla o siquiera discutirla colectivamente.

En el campo del fraude digital, ya no estamos ante estafas rudimentarias o ataques aislados. Lo que vemos es una transformación estructural del delito, impulsada por herramientas de IA que permiten crear operaciones sofisticadas, automatizadas, escalables. Se falsifican identidades, se simulan movimientos financieros, se montan campañas de desinformación con una eficiencia industrial. Todo se vuelve indistinguible: la voz del CEO que llama para pedir una transferencia urgente puede ser un modelo entrenado con su canal de YouTube.

El currículum de un nuevo empleado puede estar armado por una IA generativa que compone logros, universidades y referencias con perfecta sintaxis. El correo con instrucciones bancarias puede haber sido escrito por un bot entrenado para imitar el lenguaje de los ejecutivos de la empresa. Y lo más grave es que, cuando todo esto sucede, no hay un atacante visible: hay un sistema que hace plausible el engaño.

Este libro propone mirar de frente esos riesgos. No para generar paranoia, sino para salir del deslumbramiento ingenuo. La inteligencia artificial es poderosa. Y como toda tecnología poderosa, no es neutra. Está modelada por intereses, por decisiones de diseño, por contextos de uso que muchas veces no son transparentes. En ese sentido, la IA no sólo debe preocuparnos por lo que hace, sino por lo que oculta. Por las formas en que automatiza lógicas de exclusión, reproduce desigualdades históricas y reduce la complejidad de lo humano a un conjunto de datos procesables. No hay neutralidad posible cuando lo que está en juego son decisiones que afectan vidas.

Desde mi experiencia en la detección y prevención de fraudes, aprendí que el daño rara vez llega con un anuncio. Llega por donde nadie estaba mirando. Por eso, si hay una advertencia que este libro quiere dejar en claro desde el principio, es esta: no se trata sólo de conocer cómo funciona la IA, sino de preguntarnos en qué marco estamos dispuestos a permitir que funcione. ¿Queremos una sociedad donde las decisiones se tomen por correlación estadística? ¿Una economía donde la rentabilidad justifique cualquier forma de control? ¿Una cultura donde la autenticidad ya no importe porque todo puede ser simulado?

Estas preguntas no tienen respuestas simples, pero sí necesitan ser formuladas con urgencia.

La inteligencia artificial llegó para quedarse. Pero cómo se integre a nuestras vidas no está predeterminado. Depende de decisiones políticas, regulatorias, empresariales y también personales. Este libro

es una invitación a pensar esas decisiones. A mirar la IA no sólo como un fenómeno técnico, sino como un fenómeno cultural y ético. A entender que los algoritmos no piensan por nosotros, pero pueden terminar decidiendo en nuestro lugar si no les ponemos límites claros. Porque en los próximos años, la ventaja competitiva no estará en adoptar tecnología más rápido, sino en entender con precisión dónde conviene usarla, cómo controlarla y qué riesgos asumir.

La inteligencia artificial no reemplaza la estrategia. Pero mal usada, puede arruinarla. Este libro invita a tomar decisiones informadas antes de que otros —o los algoritmos— las tomen por nosotros.

La historia de la Inteligencia Artificial

La historia de la inteligencia artificial es una travesía marcada por avances espectaculares, fracasos abruptos y períodos de entusiasmo seguidos de escepticismo. Aunque solemos asociarla con las últimas décadas, su origen se remonta mucho más atrás, cuando los primeros pensadores comenzaron a preguntarse si las máquinas podían imitar el pensamiento humano. Desde los mitos de la antigüedad sobre autómatas hasta las primeras computadoras programables, la idea de dotar de inteligencia a una máquina ha sido una constante en la imaginación humana.

Uno de los primeros intentos serios de conceptualizar cómo una máquina podía replicar procesos cognitivos fue realizado en 1950 por el matemático británico Alan Turing. En su célebre artículo Computing Machinery and Intelligence, propuso una pregunta provocadora: ¿Pueden pensar las máquinas? En lugar de buscar una respuesta directa, sugirió un experimento práctico, el hoy conocido Test de Turing, donde una computadora debía engañar a un humano en una conversación para ser considerada inteligente. Este planteo marcó el comienzo de una exploración

más estructurada sobre la posibilidad de construir sistemas que emulen el razonamiento humano.

Sin embargo, Turing no fue el único en trazar las primeras líneas de esta historia. En paralelo, Warren McCulloch y Walter Pitts desarrollaron un modelo matemático de redes neuronales artificiales inspirado en la forma en que opera el cerebro humano. La idea de que una serie de nodos conectados podía procesar información de manera similar a las neuronas fue un concepto que, si bien no tuvo impacto inmediato, se convertiría en la piedra angular del aprendizaje profundo décadas después.

El verdadero punto de partida de la inteligencia artificial como disciplina ocurrió en 1956, cuando un grupo de científicos liderado por John McCarthy, Marvin Minsky, Nathaniel Rochester y Claude Shannon se reunió en la Conferencia de Dartmouth. Allí se acuñó formalmente el término "inteligencia artificial" y se estableció la ambiciosa meta de desarrollar sistemas capaces de aprender y razonar por sí mismos. El optimismo era desbordante y los primeros años vieron la creación de programas como Logic Theorist, capaz de demostrar teoremas matemáticos, y ELIZA, una rudimentaria simulación de un terapeuta conversacional.

Pero el entusiasmo no tardó en toparse con las limitaciones tecnológicas de la época. Los sistemas que parecían prometedores en teoría resultaron poco prácticos cuando se enfrentaron a problemas del mundo real. Las computadoras no tenían la capacidad de procesamiento necesaria y los modelos eran

demasiado rígidos para adaptarse a situaciones imprevistas. La euforia dio paso a la frustración, y en la década de 1970 la inteligencia artificial sufrió su primer gran golpe cuando los gobiernos y las instituciones académicas comenzaron a retirar la financiación. Este período, conocido como el primer "invierno de la IA", puso en pausa muchas investigaciones y dejó en claro que el camino hacia una máquina verdaderamente inteligente sería más largo de lo previsto.

A pesar del declive, la disciplina resurgió en los años 80 con el auge de los sistemas expertos, programas diseñados para replicar el conocimiento de especialistas en campos específicos. Empresas y gobiernos comenzaron a adoptar estos sistemas en áreas como la medicina, la ingeniería y las finanzas, lo que generó una nueva ola de entusiasmo. Sin embargo, el éxito fue efímero: los sistemas expertos resultaron costosos de mantener y carecían de la flexibilidad necesaria para aprender de la experiencia. A finales de la década, una nueva ola de desilusión provocó el segundo "invierno de la IA", otro período de recortes y estancamiento.

Mientras la comunidad científica enfrentaba este nuevo revés, algunos investigadores comenzaron a explorar enfoques alternativos. Uno de los avances más significativos ocurrió en los años 90, cuando los modelos de aprendizaje automático demostraron ser una solución más eficaz que los sistemas expertos. En lugar de depender de reglas predefinidas, estos algoritmos podían analizar datos y mejorar su rendimiento con el tiempo. Fue en esta época cuando

las redes neuronales, aquella idea planteada en los años 40 por McCulloch y Pitts, comenzaron a resurgir con más fuerza gracias al aumento en la capacidad de cómputo.

El cambio definitivo llegó en 1997, cuando la supercomputadora Deep Blue de IBM logró derrotar al campeón mundial de ajedrez Garry Kasparov. Fue un momento simbólico que demostró que la inteligencia artificial podía superar a los humanos en tareas específicas. Pero el verdadero salto ocurrió a partir de la década de 2010, cuando el desarrollo del aprendizaje profundo y el acceso a grandes volúmenes de datos dieron lugar a avances revolucionarios. Modelos como los de Geoffrey Hinton, Yann LeCun y Yoshua Bengio mostraron que las redes neuronales profundas podían reconocer patrones con una precisión sin precedentes, lo que impulsó aplicaciones en visión computacional, procesamiento de lenguaje natural y muchas otras áreas.

Hoy, la inteligencia artificial es una tecnología omnipresente, utilizada en asistentes virtuales, vehículos autónomos, diagnósticos médicos y sistemas de recomendación. Modelos como GPT-3 y GPT-4 han demostrado que el lenguaje puede ser generado con una fluidez sorprendente, mientras que la IA generativa está revolucionando la creatividad en el arte y la música. Sin embargo, el camino no está exento de desafíos. Cuestiones como la ética, los sesgos en los modelos y la posibilidad de que la IA alcance niveles de autonomía no previstos siguen siendo temas de debate en la comunidad científica y tecnológica.

Lo que queda claro al observar la historia de la inteligencia artificial es que no ha sido un camino lineal. Ha avanzado en ciclos de entusiasmo y decepción, impulsado por descubrimientos clave pero limitado por los recursos tecnológicos de cada época. A medida que la capacidad de procesamiento sigue creciendo y los algoritmos se refinan, es probable que estemos sólo al comienzo de una nueva fase en la evolución de la inteligencia artificial, una que quizá nos acerque más a aquella pregunta que Turing planteó hace más de 70 años: ¿Pueden pensar las máquinas?

A medida que la inteligencia artificial avanzaba, lo hacía también la percepción que el público tenía sobre ella. La imagen de una IA omnipotente que podía razonar como un ser humano se desmoronó con las limitaciones técnicas y la falta de avances concretos. Sin embargo, la idea de que las máquinas pudieran aprender por sí mismas, en lugar de ser programadas con reglas fijas, empezó a tomar fuerza en los años 90. Fue un cambio de paradigma: en lugar de construir sistemas expertos con una base rígida de conocimientos, los científicos comenzaron a enfocarse en modelos capaces de analizar datos, detectar patrones y mejorar su desempeño con la experiencia.

Este enfoque marcó el renacimiento del aprendizaje automático, una subdisciplina de la inteligencia artificial que se basa en algoritmos diseñados para aprender sin intervención humana directa. Uno de los hitos más importantes en este proceso ocurrió con las máquinas de soporte vectorial, desarrolladas por Vapnik y Cortes en 1995. Estas permitieron mejorar

la clasificación y el reconocimiento de patrones en conjuntos de datos complejos, preparando el terreno para la explosión de la IA en la década siguiente.

Mientras tanto, la capacidad de cómputo seguía en ascenso. La famosa Ley de Moore, que postula que el número de transistores en un chip se duplica aproximadamente cada dos años, hizo posible procesar grandes volúmenes de información con una eficiencia creciente. Esto permitió la reactivación de un concepto que había sido prácticamente abandonado: las redes neuronales artificiales. Aunque la teoría detrás de estas redes existía desde los años 40, fue recién con el aumento del poder computacional y el acceso a grandes bases de datos que su verdadero potencial comenzó a materializarse.

El gran salto ocurrió en 1997, cuando la supercomputadora Deep Blue, desarrollada por IBM, logró vencer al campeón mundial de ajedrez, Garry Kasparov. Aunque el sistema de Deep Blue no era una IA en el sentido estricto —ya que no aprendía de manera autónoma, sino que analizaba jugadas con una capacidad de cálculo abrumadora—, el evento fue un punto de inflexión. Por primera vez, una máquina derrotaba a un humano en un juego considerado un símbolo de la inteligencia estratégica. Esto impulsó la inversión en IA y reavivó el interés en la posibilidad de que las máquinas no sólo calcularan, sino que también comprendieran, adaptaran y generaran conocimiento.

A partir de la década del 2000, la inteligencia artificial se convirtió en un campo de investigación en constante crecimiento, pero el verdadero punto de

inflexión llegaría en 2012 con un descubrimiento que cambiaría el rumbo de la disciplina. Ese año, un equipo de investigadores liderado por Geoffrey Hinton demostró que las redes neuronales profundas podían superar a cualquier otro método en tareas de reconocimiento de imágenes. En la competencia ImageNet, donde los sistemas debían identificar objetos en una vasta colección de imágenes, su modelo basado en Deep Learning logró reducir el margen de error en un porcentaje que sorprendió a toda la comunidad científica.

Lo que hacía revolucionario al aprendizaje profundo no era simplemente su precisión, sino su capacidad de extraer características relevantes sin necesidad de que un programador definiera manualmente qué debía buscar en los datos. Las redes neuronales profundas eran capaces de aprender representaciones abstractas y realizar tareas de manera más cercana a como lo hace el cerebro humano. Esto desencadenó un auge en la investigación y el desarrollo de aplicaciones basadas en inteligencia artificial.

En los años siguientes, la IA se integró en la vida cotidiana a través de productos y servicios que transformaron industrias enteras. Los asistentes virtuales como Siri, Alexa y Google Assistant se popularizaron, los algoritmos de recomendación personalizaron la experiencia en plataformas como Netflix y Spotify, y los sistemas de reconocimiento facial comenzaron a aplicarse en seguridad y control de acceso. El auge de la IA generativa, con modelos como GPT-3 y GPT-4, llevó el procesamiento del lenguaje natural a niveles

nunca antes vistos, permitiendo a las máquinas escribir textos complejos, traducir idiomas y generar código con una precisión sorprendente.

Pero con estos avances también surgieron preocupaciones. La inteligencia artificial dejó de ser una herramienta exclusivamente técnica para convertirse en un fenómeno con profundas implicaciones éticas, políticas y sociales. La automatización desplazó empleos en múltiples sectores, los sesgos en los algoritmos comenzaron a evidenciarse como un problema serio, y la posibilidad de que la IA adquiriera un nivel de autonomía no previsto despertó debates en la comunidad científica y filosófica.

En el presente, la inteligencia artificial no sólo está en el centro del desarrollo tecnológico, sino también en el debate global sobre su regulación, sus límites y sus riesgos. La historia de la IA es la historia de una búsqueda constante por replicar —y tal vez algún día superar— la inteligencia humana. Sin embargo, lo que ha quedado claro en este recorrido es que los avances han dependido tanto del poder computacional como de la creatividad y el ingenio de quienes han impulsado su desarrollo. Mientras seguimos explorando el potencial de la IA, la pregunta que se planteaba Alan Turing hace más de 70 años sigue sin una respuesta definitiva. Quizás la IA ya esté pensando, pero aún no sabemos si lo hace de la misma manera que nosotros.