

UNIVERSIDAD PRIVADA SAN CARLOS

FACULTAD DE INGENIERÍAS

ESCUELA PROFESIONAL DE INGENIERÍA INFORMÁTICA



TESIS

**ESTUDIO TEÓRICO DE LOS RIESGOS SISTÉMICOS DERIVADOS DEL
DESARROLLO Y LA IMPLEMENTACIÓN DE LA INTELIGENCIA ARTIFICIAL**

PRESENTADA POR:

NIKOLAY ELIAZER VARGAS TAPIA

PARA OPTAR EL TÍTULO PROFESIONAL DE:

INGENIERO INFORMÁTICO

PUNO – PERÚ

2026



Repositorio Institucional ALCIRA by [Universidad Privada San Carlos](https://www.upsc.edu.pe/) is licensed under a [Creative Commons Reconocimiento-NoComercial 4.0 Internacional License](https://creativecommons.org/licenses/by-nc/4.0/)



1.34%

SIMILARITY OVERALL

SCANNED ON: 8 JUN 2026, 2:24 PM

Originality & Authorship Report

Your text is highlighted according to the matched content in the results above.

● IDENTICAL
0.18%

● CHANGED TEXT
1.15%

Report #33637787

NIKOLAY ELIAZER VARGAS TAPIA // ESTUDIO TEÓRICO DE LOS RIESGOS SISTÉMICOS DERIVADOS DEL DESARROLLO Y LA IMPLEMENTACIÓN DE LA INTELIGENCIA ARTIFICIAL. RESUMEN La presente investigación analizó críticamente los riesgos sistémicos del avance y la implementación masiva de la inteligencia artificial (IA). Metodológicamente, se desarrolló bajo un enfoque cualitativo, de tipo documental, descriptivo, analítico y explicativa, con un diseño no experimental, observacional y de corte longitudinal. La zona de estudio se constituyó como un entorno virtual global delimitado por los principales repositorios digitales y bases de datos científicas internacionales (Scopus, Web of Science, Google Académico). La población y muestra se conformaron por documentos científicos, técnicos, institucionales y académicos publicados a nivel mundial entre los años 2015 y 2025, seleccionados mediante un muestreo no probabilístico e intencional bajo estrictos criterios de rigor, relevancia temática y nivel de citación. Para la recolección de datos se empleó la revisión documental mediante fichas y matrices de categorización temática, procesadas a través de los métodos analítico, deductivo, histórico y crítico-reflexivo. Las conclusiones demuestran que los riesgos de la IA trascienden lo tecnológico e impactan estructuralmente en las dimensiones laboral, educativa, gubernamental, de seguridad digital y existencial. Se evidenció que la IA acelera procesos históricos de automatización, reconfigurando funciones humanas, desplazando tareas cognitivas y generando

UNIVERSIDAD PRIVADA SAN CARLOS
FACULTAD DE INGENIERÍAS
ESCUELA PROFESIONAL DE INGENIERÍA INFORMÁTICA
TESIS

**ESTUDIO TEÓRICO DE LOS RIESGOS SISTÉMICOS DERIVADOS DEL
DESARROLLO Y LA IMPLEMENTACIÓN DE LA INTELIGENCIA ARTIFICIAL**

PRESENTADA POR:

NIKOLAY ELIAZER VARGAS TAPIA

PARA OPTAR EL TÍTULO PROFESIONAL DE:

INGENIERO INFORMÁTICO

APROBADA POR EL SIGUIENTE JURADO:

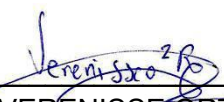
PRESIDENTE

: 
M.Sc. KORINA ASQUI GOMEZ

PRIMER MIEMBRO

: 
Mtra. NATALY SILVIA GARCIA VILCA

SEGUNDO MIEMBRO

: 
Dra. CELIA VERENISSE ORTIZ DE ORUE ROJAS

ASESOR DE TESIS

: 
Dra. MARLENE CUSI MONTESINOS

Área: Ingeniería, Tecnología

Sub área: Ingeniería Eléctrica, Electrónica e Informática

Línea de investigación: Ingeniería de Sistemas y Comunicaciones

Puno, 17 de junio del 2026.

DEDICATORIA

Esta tesis está dedicada a todas las personas que me brindaron su apoyo durante mi etapa universitaria. De manera muy especial, la dedico a mi padre, Isaias Tapia Aza, y a mi gato, Elrond. Aunque hoy ya no se encuentren con nosotros, su recuerdo me guió hasta la meta.

AGRADECIMIENTO

En primer lugar, expreso mi más profundo agradecimiento a la Universidad Privada San Carlos, Escuela Profesional de Ingeniería Informática, por abrirme sus puertas, brindarme las herramientas necesarias y ser el espacio donde me formé profesional y humanamente.

A mi asesor de tesis, por su paciencia, rigurosidad académica y por guiarme con certeza en cada etapa de este proceso.

Asimismo, manifiesto mi gratitud a los miembros del jurado, por el tiempo dedicado a la revisión de esta investigación y por sus valiosas observaciones, las cuales enriquecieron significativamente el resultado final.

Finalmente, hago un reconocimiento especial a las herramientas digitales y tecnológicas de nuestro tiempo; su soporte y eficiencia fueron recursos indispensables sin los cuales habría sido imposible concretar esta tesis.

ÍNDICE GENERAL

	Pág.
DEDICATORIA	1
AGRADECIMIENTO	2
ÍNDICE GENERAL	3
ÍNDICE DE TABLAS	7
ÍNDICE DE FIGURAS	8
ÍNDICE DE ANEXOS	9
RESUMEN	10
ABSTRACT	11
INTRODUCCIÓN	12

CAPÍTULO I

PLANTEAMIENTO DEL PROBLEMA, ANTECEDENTES Y OBJETIVOS DE LA INVESTIGACIÓN

1.1. PLANTEAMIENTO DEL PROBLEMA	14
1.1.1. FORMULACIÓN DEL PROBLEMA	16
1.2. ANTECEDENTES	17
1.2.1. ANTECEDENTES A NIVEL INTERNACIONAL	17
1.2.2. ANTECEDENTES A NIVEL NACIONAL	19
1.3. OBJETIVOS	20
1.3.1. OBJETIVO GENERAL	20
1.3.2. OBJETIVOS ESPECÍFICOS	20

CAPÍTULO II

MARCO TEÓRICO, CONCEPTUAL DE LA INVESTIGACIÓN

2.1. MARCO TEÓRICO	22
2.1.1. LA INTELIGENCIA ARTIFICIAL COMO CAMPO CIENTÍFICO Y TECNOLÓGICO	22
2.1.2. FUNDAMENTOS FILOSÓFICOS: EL “FANTASMA EN LA MÁQUINA”	

Y LA NATURALEZA DE LA INTELIGENCIA	22
2.1.3. LA PRUEBA DE TURING Y LA SIMULACIÓN DE LA INTELIGENCIA	24
2.1.4. DESARROLLO ACELERADO DE LA INTELIGENCIA ARTIFICIAL Y RIESGOS TECNOLÓGICOS	24
2.1.5. GOBERNANZA, ÉTICA Y CONTROL ,EN SISTEMAS DE INTELIGENCIA ARTIFICIAL	25
2.1.6. MEDICIÓN Y EVALUACIÓN DE RIESGOS EXISTENCIALES Y SOCIALES	27
2.1.7. SÍNTESIS DEL ORIGEN Y EVOLUCIÓN HISTÓRICA DE LA INTELIGENCIA ARTIFICIAL	28
2.2. INTEGRACIÓN DE LA INTELIGENCIA ARTIFICIAL EN EL APRENDIZAJE ADAPTATIVO	29
2.3. RIESGOS ÉTICOS Y DE SEGURIDAD EN EL USO DE LA INTELIGENCIA ARTIFICIAL	30
2.3.1. LA CRECIENTE DEPENDENCIA DE LA INTELIGENCIA ARTIFICIAL	31
2.3.2. EL RITMO ACELERADO DEL DESARROLLO E IMPLEMENTACIÓN DE LA INTELIGENCIA ARTIFICIAL	32
2.4. VARIABLES DE LA INVESTIGACIÓN	36
2.4.1.VARIABLE INDEPENDIENTE: DESARROLLO E IMPLEMENTACIÓN ACELERADA DE LA INTELIGENCIA ARTIFICIAL	37
2.4.2.VARIABLE DEPENDIENTE: RIESGOS TECNOLÓGICOS, SOCIALES, ÉTICOS Y PROSPECTIVOS ASOCIADOS A LA INTELIGENCIA ARTIFICIAL	37
CAPÍTULO III	
METODOLOGÍA DE LA INVESTIGACIÓN	
3.1. ZONA DE ESTUDIO	39
3.2. POBLACIÓN Y MUESTRA	39
3.2.1. POBLACIÓN	39

3.2.2. MUESTRA	40
3.3. MÉTODOS Y TÉCNICAS	40
3.3.1. ENFOQUE DE LA INVESTIGACIÓN	40
3.3.2. TIPO DE INVESTIGACIÓN	41
3.3.3. DISEÑO DE LA INVESTIGACIÓN	42
3.3.4. MÉTODOS DE INVESTIGACIÓN	42
3.3.5. TÉCNICAS E INSTRUMENTOS DE RECOLECCIÓN DE DATOS	43
3.3.6. PROCEDIMIENTO	43
3.4 IDENTIFICACIÓN DE VARIABLES	45
3.5. ANÁLISIS DE DATOS	45

CAPÍTULO IV

EXPOSICION Y ANALISIS DE LOS RESULTADOS

4.1. RESULTADOS DEL OBJETIVO GENERAL:	47
4.1.1. ANTECEDENTES Y DESARROLLO HISTÓRICO DE LA INTELIGENCIA ARTIFICIAL	48
4.1.2. TENDENCIAS ACTUALES EN EL DESARROLLO DE LA INTELIGENCIA ARTIFICIAL	50
4.1.3. INCISO DE VANGUARDIA HASTA DICIEMBRE DE 2025	54
4.2. RESULTADOS DEL OBJETIVO ESPECÍFICO 01:	56
4.2.1. IMPACTO DE LA INTELIGENCIA ARTIFICIAL EN EL ÁMBITO LABORAL	56
4.2.2. ANTECEDENTES HISTÓRICOS DE LA AUTOMATIZACIÓN DEL TRABAJO	61
4.2.3. PROFESIONES ALTAMENTE AUTOMATIZADAS	62
4.2.4. PROFESIONES EN PROCESO DE AUTOMATIZACIÓN	90
4.2.5. PROFESIONES CON IMPOSIBILIDAD DE AUTOMATIZACIÓN TEMPORAL	109
4.3. RESULTADOS DEL OBJETIVO ESPECÍFICO 02:	129
4.3.1. TRANSFORMACIÓN DE LA COMUNICACIÓN Y LAS RELACIONES	

SOCIALES	129
4.3.2. DEEPFAKES, DESINFORMACIÓN Y CRISIS DE LA VERDAD DIGITAL	131
4.3.3. RELACIONES HUMANO-IA Y VÍNCULOS EMOCIONALES	
ARTIFICIALES	134
4.4. RESULTADOS DEL OBJETIVO ESPECÍFICO 03:	136
4.4.1. MANIPULACIÓN DE LA OPINIÓN PÚBLICA Y COMPORTAMIENTO	
COLECTIVO	136
4.4.2. MERCANTILIZACIÓN DE LA ATENCIÓN Y PRODUCCIÓN DE	
CONTENIDO ARTIFICIAL	139
4.4.3. IDENTIDAD DIGITAL, AUTOIMAGEN Y PRESIÓN ALGORÍTMICA	141
4.4.4. RIESGOS DE DEPENDENCIA SOCIAL Y AISLAMIENTO	
TECNOLÓGICO	144
4.4.5. REFLEXIÓN INTEGRADORA: SOCIEDAD ALGORÍTMICA Y FUTURO	
HUMANO	147
4.5. RESULTADOS DEL OBJETIVO ESPECÍFICO 04:	150
4.5.1. ÁMBITO SOCIAL —ENTRE EL INTERNET MUERTO, IDENTIDADES	
DIGITALES Y DISTOPÍAS EMERGENTES	150
4.5.2. IMPACTO DE LA IA EN EL ÁMBITO SOCIAL — INTRODUCCIÓN	153
4.6. RESULTADOS DEL OBJETIVO ESPECÍFICO 05:	155
4.6.1. IMPACTO DE LA INTELIGENCIA ARTIFICIAL EN EL ÁMBITO	
EDUCATIVO	155
4.7. DISCUSIÓN DE RESULTADOS	162
CONCLUSIONES	167
RECOMENDACIONES	169
BIBLIOGRAFÍA	171
ANEXOS	183

ÍNDICE DE TABLAS

	Pág.
Tabla 01: Evolución, riesgos e impactos de la inteligencia artificial	33
Tabla 02: Operacionalización de variables	45
Tabla 03: Evolución histórica del impacto de la Inteligencia Artificial (1956–2025)	53
Tabla 04: Proyecciones del impacto de la inteligencia artificial en educación, empleo y gobernanza (2025–2040)	55
Tabla 05: Aceleración histórica de la automatización antes y después de la inteligencia artificial	58
Tabla 06: Avance actual de automatización por IA (2025)	106
Tabla 07: Proyección estimada de reemplazo total de empleos por IA (2030 y más allá)	107
Tabla 08: Principales riesgos éticos, sociales y de privacidad asociados al uso de la inteligencia artificial (2018–2025)	152
Tabla 09: Integración de la ia en la educación	159

ÍNDICE DE FIGURAS

	Pág.
Figura 01: Evolución, impacto y riesgo de la inteligencia artificial	36
Figura 02: Evolución del impacto de la inteligencia artificial	53
Figura 03: Proyección del Impacto de la inteligencia artificial (2025-2040)	56
Figura 04: Nivel de impacto laboral	60
Figura 05: Nivel de exposición a automatización por IA (2025)	107
Figura 06: Proyección de la automatización laboral por IA (2025)	108
Figura 07: Evolución de los riesgos éticos, sociales y de privacidad	153
Figura 08: Integración de IA en educación superior y dependencia tecnológica	160

ÍNDICE DE ANEXOS

	Pág.
Anexo 01: Matriz De Consistencia.	184
Anexo 02: Matriz resumen de análisis conceptual de autores y ejes teóricos	186
Anexo 03: Resumen de Línea temporal conceptual de la Inteligencia Artificial	187
Anexo 04: Resumen - Clasificación estructural de riesgos de la Inteligencia Artificial	190
Anexo 05: Resumen - Comparación estructural: Capacidades humanas vs Inteligencia Artificial	194
Anexo 06: Resumen - Escenarios prospectivos de peligro asociados al avance de la Inteligencia Artificial	198
Anexo 07: Resumen - Imaginarios, reflexiones filosóficas sobre los riesgos de la inteligencia artificial	203

RESUMEN

La presente investigación analizó críticamente los riesgos sistémicos del avance y la implementación masiva de la inteligencia artificial (IA). Metodológicamente, se desarrolló bajo un enfoque cualitativo, de tipo documental, descriptivo, analítico y explicativa, con un diseño no experimental, observacional y de corte longitudinal. La zona de estudio se constituyó como un entorno virtual global delimitado por los principales repositorios digitales y bases de datos científicas internacionales (Scopus, Web of Science, Google Académico). La población y muestra se conformaron por documentos científicos, técnicos, institucionales y académicos publicados a nivel mundial entre los años 2015 y 2025, seleccionados mediante un muestreo no probabilístico e intencional bajo estrictos criterios de rigor, relevancia temática y nivel de citación. Para la recolección de datos se empleó la revisión documental mediante fichas y matrices de categorización temática, procesadas a través de los métodos analítico, deductivo, histórico y crítico-reflexivo. Las conclusiones demuestran que los riesgos de la IA trascienden lo tecnológico e impactan estructuralmente en las dimensiones laboral, educativa, gubernamental, de seguridad digital y existencial. Se evidenció que la IA acelera procesos históricos de automatización, reconfigurando funciones humanas, desplazando tareas cognitivas y generando nuevas dependencias laborales. Asimismo, se comprobó que amplía las vulnerabilidades en seguridad digital debido a dinámicas algorítmicas capaces de amplificar sesgos y errores. En materia de gobernanza, la rapidez del avance tecnológico frente a marcos normativos desactualizados genera preocupantes asimetrías de poder. A nivel existencial, los riesgos se manifiestan en la pérdida progresiva de la autonomía humana y la delegación cognitiva ante la mediación algorítmica. Finalmente, en el ámbito educativo, la incorporación de la IA introduce el peligro de una adopción acrítica, dependencia instrumental y la erosión del pensamiento reflexivo, transformando el rol humano en la producción del conocimiento.

Palabras clave: Automatización, Autonomía humana, Gobernanza, Inteligencia artificial, Riesgos tecnológicos.

ABSTRACT

This research critically analyzed the systemic risks of the advancement and widespread implementation of artificial intelligence (AI). Methodologically, it employed a qualitative, documentary, descriptive, analytical, and explanatory approach, with a non-experimental, observational, and longitudinal design. The study area was defined as a global virtual environment delimited by the main international digital repositories and scientific databases (Scopus, Web of Science, Google Scholar). The population and sample consisted of scientific, technical, institutional, and academic documents published worldwide between 2015 and 2025, selected through non-probability, purposive sampling based on strict criteria of rigor, thematic relevance, and citation level. Data collection involved documentary review using thematic categorization forms and matrices, processed through analytical, deductive, historical, and critical-reflective methods. The findings demonstrate that the risks of AI transcend the technological realm and have a structural impact on the labor, educational, governmental, digital security, and existential dimensions. It was shown that AI accelerates historical processes of automation, reconfiguring human functions, displacing cognitive tasks, and generating new labor dependencies. Likewise, it was found to amplify vulnerabilities in digital security due to algorithmic dynamics capable of amplifying biases and errors. In terms of governance, the rapid pace of technological advancement in the face of outdated regulatory frameworks generates worrying power imbalances. At an existential level, the risks manifest themselves in the progressive loss of human autonomy and the delegation of cognitive power to algorithmic mediation. Finally, in the educational sphere, the incorporation of AI introduces the danger of uncritical adoption, instrumental dependence, and the erosion of reflective thinking, transforming the human role in knowledge production.

Keywords: Automation, Human autonomy, Governance, Artificial intelligence, Technological risks.

INTRODUCCIÓN

En los últimos años, la inteligencia artificial ha experimentado un crecimiento acelerado, convirtiéndose en un componente fundamental dentro del desarrollo de sistemas informáticos, la automatización de procesos y la optimización de la toma de decisiones. Su aplicación se extiende a diversos sectores, tales como la educación, la industria, la economía y la seguridad informática, generando importantes beneficios en términos de eficiencia y productividad.

No obstante, el uso intensivo de la inteligencia artificial también ha suscitado una serie de preocupaciones relacionadas con los riesgos que implica su implementación sin un adecuado control. Entre estos riesgos destacan la automatización de puestos de trabajo, la vulnerabilidad de los sistemas informáticos, el uso indebido de datos y la falta de marcos normativos claros que regulen su desarrollo y aplicación. Desde la perspectiva de la Ingeniería Informática, resulta indispensable analizar estos riesgos de manera sistemática y objetiva.

La inteligencia artificial ha emergido como una tecnología disruptiva en el campo de la educación. La capacidad de las máquinas para analizar grandes volúmenes de datos, aprender patrones y tomar decisiones basadas en algoritmos sofisticados está transformando la forma en que enseñamos y aprendemos. Aunque la investigación en inteligencia artificial (IA) en educación lleva décadas desarrollándose a nivel internacional, ha experimentado un gran avance en los últimos diez años impulsada por el desarrollo de técnicas de aprendizaje automático (chat GPT), procesamiento de lenguaje natural (NPL) y redes neuronales, entre otras técnicas que requerían de una gran cantidad de datos para su entrenamiento. La inteligencia artificial (IA), habilita a que la tecnología replique las habilidades e inteligencia humana, resolviendo problemas de forma autónoma mediante la utilización de tecnologías informáticas. La inteligencia artificial incluye asistencia digital, navegación satelital GPS, y la autónoma de equipos en general. La IA se relaciona al aprendizaje automatizado y profundo, las decisiones del cerebro humano, es creado en forma de algoritmos, los cuales tienen la capacidad de leer, interpretar,

analizar en forma automatizada. La inteligencia artificial posibilita que las tecnologías de comunicación e información realicen funciones que requieren de inteligencia humana, como la toma de decisiones y otras funciones que realiza la inteligencia humana. Además, puede enfrentar retos significativos en la educación globalizada y más aún, va acelerando la forma de análisis, interpretación de respuestas rápidas, esto puede ayudar, pero también puede acarrear oportunidades y barreras en la educación y también en el ámbito laboral y social.

El alcance de esta investigación se limita al examen teórico de los riesgos asociados a la inteligencia artificial, apoyándose en aportes provenientes de la ciencia, la filosofía y la literatura, así como en ejemplos documentados de automatización e implementación tecnológica. No se abordan desarrollos técnicos detallados ni estudios experimentales, sino que se prioriza una aproximación analítica que permita comprender el significado, las implicancias y los desafíos que plantea la presencia creciente de la inteligencia artificial en la sociedad actual.

El Capítulo I detalla el planteamiento del problema, los antecedentes internacionales y nacionales, y los objetivos orientados a analizar críticamente los riesgos derivados del desarrollo y la implementación de la Inteligencia Artificial. El Capítulo II establece el marco teórico y conceptual sobre riesgos sistémicos derivados del desarrollo y la implementación de la inteligencia artificial. El Capítulo III describe la metodología cualitativa y descriptiva, la revisión de la literatura científica en el periodo 2015 - 2025. Finalmente, el Capítulo IV expone el análisis de resultados mediante tablas y figuras, concluyendo con las recomendaciones para la afrontación de los avances de la inteligencia artificial. .

CAPÍTULO I

PLANTEAMIENTO DEL PROBLEMA, ANTECEDENTES Y OBJETIVOS DE LA INVESTIGACIÓN

1.1. PLANTEAMIENTO DEL PROBLEMA

Se ha observado que la inteligencia artificial está en camino ha superar a la inteligencia humana dejando en claro que siendo este su creador llegará a superar al mismo, analizado los riesgos de reemplazo funcional y de pérdida de agencia humana derivados del desarrollo y la implementación de la inteligencia artificial, a la luz de las advertencias históricas y de fenómenos contemporáneos como la automatización avanzada y la mediación algorítmica.

A nivel global, la inteligencia artificial (IA) se ha consolidado como una tecnología clave en el desarrollo de los sistemas informáticos modernos, influyendo significativamente en la automatización de procesos, la gestión de datos y la toma de decisiones en diversos contextos organizacionales y sociales. Su rápido avance ha transformado sectores como la educación, la salud, la economía, la seguridad y la comunicación, contribuyendo a la mejora de la eficiencia operativa y al fortalecimiento de los procesos productivos. No obstante, esta evolución acelerada también ha generado una serie de desafíos y riesgos que ponen en cuestión la capacidad de las sociedades para regular, adaptar y controlar adecuadamente su uso.

Entre los principales riesgos asociados a la implementación de la inteligencia artificial se encuentran la automatización excesiva de tareas laborales, la pérdida de empleos, la

dependencia de sistemas automatizados, la vulnerabilidad de la información, la opacidad de los algoritmos y la ausencia de mecanismos adecuados de gobernanza tecnológica. Asimismo, el uso de sistemas automáticos para la generación de contenidos, el acaparamiento del poder tecnológico por parte de grandes corporaciones y el desarrollo de sistemas autónomos fuera del control humano representan amenazas relevantes para la estabilidad social, económica y tecnológica.

En este contexto, uno de los principales desafíos consiste en conciliar la necesidad de garantizar un uso ético, seguro y equitativo de la inteligencia artificial con su constante desarrollo, evitando que los riesgos superen sus beneficios potenciales. La falta de marcos regulatorios adecuados, de educación digital en la población y de mecanismos claros de supervisión incrementa la vulnerabilidad de la sociedad frente a los efectos negativos de esta tecnología.

En el Perú, el uso de la inteligencia artificial ha venido incrementándose progresivamente en el ámbito laboral y educativo. Según un estudio de Bumeran (2025), aproximadamente el 57 % de los trabajadores emplea herramientas de inteligencia artificial en sus actividades diarias. Asimismo, el Ministerio de Educación (2024, 2025) ha destacado el potencial de la IA para imitar habilidades humanas, resaltando la importancia de una integración responsable, ética y crítica en los procesos educativos. De igual manera, la Ley N.º 31814 (2025) establece un marco normativo para la clasificación y regulación de los riesgos asociados a los sistemas de inteligencia artificial.

En el sector educativo peruano, si bien se promueve la adopción de tecnologías emergentes mediante iniciativas como el Plan Nacional de Competencias Digitales 2030, su implementación aún es limitada, concentrándose principalmente en algunas universidades e institutos privados. Esta situación se ve agravada por las deficiencias en infraestructura tecnológica y por la necesidad de fortalecer las competencias digitales del profesorado, especialmente en regiones como Puno, donde el Instituto Nacional de Estadística e Informática (INEI, 2024) evidencia la urgencia de desarrollar estrategias que

fomenten el uso adecuado de la inteligencia artificial en los procesos de enseñanza-aprendizaje.

Desde la perspectiva de la Ingeniería Informática, la integración de sistemas basados en inteligencia artificial ha permitido optimizar procesos y mejorar la toma de decisiones; sin embargo, también ha incrementado los riesgos relacionados con la seguridad informática, la pérdida de control humano, la confiabilidad y confidencialidad de los sistemas y la autonomía en la toma de decisiones. La creciente dependencia de estas tecnologías sin una evaluación integral de sus implicancias puede generar impactos negativos en los ámbitos laboral, económico y social.

En el contexto actual, se observa un aumento en la utilización de tecnologías basadas en inteligencia artificial sin un análisis sistemático de los riesgos que estas conllevan. Esta situación genera incertidumbre respecto a su impacto real en los procesos productivos, organizacionales y de seguridad informática, lo cual evidencia la necesidad de desarrollar una investigación rigurosa que permita identificar, analizar y comprender los riesgos asociados a la inteligencia artificial desde una perspectiva técnica, ética y académica.

El problema que aborda esta investigación consiste en analizar hasta qué punto los procesos actuales de implementación de la inteligencia artificial configuran condiciones que favorecen el reemplazo progresivo de funciones humanas y la disminución de la agencia humana, así como las implicancias sociales, éticas y epistemológicas que se derivan de esta transformación. La ausencia de un análisis teórico integrado que articule las advertencias históricas con estos fenómenos contemporáneos dificulta una comprensión crítica del alcance real de la inteligencia artificial en la reorganización de la vida humana.

1.1.1. FORMULACIÓN DEL PROBLEMA

Problema General

¿Cuáles son los principales riesgos sistémicos derivados del desarrollo y la implementación de la Inteligencia Artificial identificados en la literatura científica?

Problemas Específicos

- ¿De qué manera el desarrollo acelerado de la IA puede generar riesgos relacionados con las condiciones laborales y en el desarrollo económico, considerando cambios en la estructura del empleo y automatización de procesos?
- ¿En qué medida el avance y la implementación de la IA incrementan los riesgos en la seguridad digital y la protección de la información?
- ¿Cuáles son los principales riesgos asociados a la gobernanza y al control de los sistemas de IA en contextos organizacionales y sociales?
- ¿De qué manera pueden identificarse y evaluarse los riesgos existenciales vinculados al avance y la implementación de la IA?
- ¿Cómo incide el desarrollo e implementación de la IA en el ámbito educativo y sus implicancias en los procesos de aprendizaje?

1.2. ANTECEDENTES

1.2.1. ANTECEDENTES A NIVEL INTERNACIONAL

En los últimos años, la inteligencia artificial ha sido objeto de numerosos estudios a nivel internacional, debido a su creciente impacto en los ámbitos tecnológico, económico, educativo y social. Diversos investigadores y organismos internacionales han analizado tanto los beneficios como los riesgos asociados a su desarrollo e implementación, destacando la necesidad de establecer mecanismos de regulación y control.

A nivel global, la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO, 2021) elaboró un marco ético para el uso de la inteligencia artificial, en el cual se enfatiza la importancia de garantizar la transparencia, la equidad, la protección de los derechos humanos y la responsabilidad social en el diseño y aplicación de estas tecnologías. Dicho documento advierte que el uso inadecuado de la IA puede generar discriminación, pérdida de privacidad y concentración de poder tecnológico.

Por su parte, la Organización para la Cooperación y el Desarrollo Económicos (OCDE, 2019) estableció principios orientados al desarrollo responsable de la inteligencia artificial, destacando la necesidad de que los sistemas sean seguros, confiables y respetuosos de

los valores humanos. Asimismo, señala que la automatización impulsada por la IA puede generar transformaciones significativas en el mercado laboral, afectando especialmente a los trabajadores con menor nivel de especialización.

En el ámbito europeo, la Comisión Europea (2020) desarrolló una propuesta de regulación sobre inteligencia artificial, en la cual clasifica los sistemas según su nivel de riesgo, estableciendo restricciones para aquellos considerados de alto impacto social. Esta iniciativa busca prevenir posibles daños en áreas como la seguridad, la privacidad y la toma de decisiones automatizadas, promoviendo un equilibrio entre innovación y protección ciudadana.

Desde una perspectiva académica, autores como Russell y Norvig (2021) sostienen que el crecimiento acelerado de la inteligencia artificial plantea desafíos relacionados con la pérdida de control humano y la comprensión limitada de los procesos algorítmicos. Los investigadores advierten que, a medida que los sistemas se vuelven más complejos, resulta más difícil interpretar sus decisiones, lo que puede afectar la confiabilidad y la transparencia.

Asimismo, Floridi et al. (2018) analizan los riesgos éticos de la inteligencia artificial, destacando problemas como los sesgos algorítmicos, la discriminación automatizada y la reducción de la autonomía humana. Según estos autores, la falta de supervisión adecuada puede reforzar desigualdades sociales y debilitar la responsabilidad en los procesos de decisión.

En el ámbito laboral, Frey y Osborne (2017) realizaron un estudio en el que estiman que un porcentaje significativo de empleos podría ser automatizado mediante sistemas inteligentes. Esta situación genera incertidumbre respecto al futuro del trabajo y resalta la necesidad de desarrollar políticas de adaptación y capacitación profesional.

Por otro lado Lee (2020), estudio en países asiáticos, como Japón y Corea del Sur, evidenciando que la implementación masiva de la inteligencia artificial en la industria y la educación ha generado mejoras en productividad, pero también ha incrementado los riesgos relacionados con la dependencia tecnológica y la seguridad informática. Estos

resultados demuestran que el desarrollo tecnológico sin una planificación adecuada puede generar efectos adversos a largo plazo.

En el ámbito educativo internacional, investigaciones como la de Holmes et al. (2019) señalan que la inteligencia artificial puede mejorar los procesos de enseñanza-aprendizaje mediante sistemas de tutoría inteligente y análisis de datos educativos. Sin embargo, también advierten sobre el riesgo de reducir la interacción humana, afectar el pensamiento crítico y generar desigualdades en el acceso a la tecnología.

1.2.2. ANTECEDENTES A NIVEL NACIONAL

En el Perú, el desarrollo y la implementación de la inteligencia artificial han venido incrementándose progresivamente en diversos sectores, especialmente en los ámbitos educativo, empresarial, gubernamental y tecnológico. En los últimos años, distintas instituciones públicas y privadas han promovido el uso de herramientas digitales e inteligentes con el propósito de mejorar la eficiencia, la innovación y la competitividad.

El Ministerio de Educación (Minedu, 2024) ha señalado que la inteligencia artificial posee un alto potencial para fortalecer los procesos de enseñanza-aprendizaje, mediante la personalización educativa, el análisis de datos académicos y el uso de plataformas inteligentes. Sin embargo, también advierte sobre la necesidad de formar a docentes y estudiantes en el uso ético y responsable de estas tecnologías, con el fin de prevenir la dependencia tecnológica, el plagio académico y la reducción del pensamiento crítico.

Asimismo, el Plan Nacional de Competencias Digitales 2030 destaca la importancia de fortalecer las capacidades tecnológicas de la población, promoviendo el uso adecuado de herramientas basadas en inteligencia artificial en la educación superior y en el ámbito laboral. Este documento reconoce que la limitada infraestructura tecnológica y la brecha digital representan obstáculos para una implementación equitativa y segura de la IA en el país.

Desde el ámbito normativo, el Estado peruano ha avanzado en la formulación de marcos legales orientados a regular el uso de tecnologías emergentes. La Ley N.º 31814 (2025)

establece disposiciones relacionadas con la clasificación de riesgos de los sistemas de inteligencia artificial, diferenciando entre sistemas indebidos, de alto riesgo y aceptables. Esta normativa busca garantizar la protección de los derechos fundamentales, la seguridad de la información y la transparencia en el uso de sistemas automatizados.

Por su parte, el Instituto Nacional de Estadística e Informática (INEI, 2024) ha resaltado la necesidad de fortalecer las competencias digitales de los profesionales peruanos, especialmente en regiones como el sur del país, donde el acceso a tecnologías avanzadas aún presenta limitaciones. Este organismo señala que la formación continua en tecnologías emergentes es fundamental para reducir los riesgos asociados al uso inadecuado de la inteligencia artificial.

En el ámbito laboral, estudios realizados por consultoras y plataformas de empleo, como Bumeran (2025), indican que un porcentaje significativo de trabajadores peruanos utiliza herramientas de inteligencia artificial en sus actividades diarias. Estos informes advierten que, si bien la IA incrementa la productividad, también genera preocupación por la automatización de empleos, la sobrecarga tecnológica y la adaptación profesional.

1.3. OBJETIVOS

1.3.1. OBJETIVO GENERAL

Analizar críticamente los riesgos sistémicos derivados del desarrollo y la implementación de la Inteligencia Artificial identificados en la literatura científica.

1.3.2. OBJETIVOS ESPECÍFICOS

- Analizar el impacto del avance de la IA en las condiciones laborales y en el desarrollo económico, considerando cambios en la estructura del empleo y automatización de procesos.
- Examinar los principales riesgos asociados a la seguridad digital derivados del uso de sistemas de IA, evaluando su influencia en la protección de datos y privacidad.
- Analizar los desafíos relacionados con la gobernanza y el control de la IA, considerando los marcos normativos, mecanismos de supervisión y responsabilidad de los actores involucrados.

- Identificar y evaluar los riesgos existenciales vinculados al avance acelerado de la IA, considerando su impacto en la autonomía humana, dependencia tecnológica y sostenibilidad social.
- Analizar el impacto del desarrollo e implementación de la IA en el ámbito educativo, considerando sus implicancias en los procesos de aprendizaje y mediación tecnológica.

CAPÍTULO II

MARCO TEÓRICO, CONCEPTUAL DE LA INVESTIGACIÓN

2.1. MARCO TEÓRICO

2.1.1. LA INTELIGENCIA ARTIFICIAL COMO CAMPO CIENTÍFICO Y TECNOLÓGICO

La inteligencia artificial (IA) constituye una disciplina interdisciplinaria que integra conocimientos de la informática, las matemáticas, la neurociencia, la lógica y la ingeniería, con el propósito de diseñar sistemas capaces de simular procesos cognitivos humanos, tales como el aprendizaje, el razonamiento, la percepción y la toma de decisiones. Desde sus orígenes en la década de 1950, la IA ha evolucionado de manera acelerada, impulsada por el desarrollo de algoritmos avanzados, el aumento de la capacidad computacional y la disponibilidad masiva de datos.

En la actualidad, la IA se encuentra presente en múltiples ámbitos, incluyendo la educación, la salud, la industria, la economía digital y la seguridad informática. Su aplicación ha permitido optimizar procesos, reducir costos operativos y mejorar la eficiencia organizacional. No obstante, este crecimiento también ha generado nuevos desafíos relacionados con el control, la transparencia y la responsabilidad tecnológica, en este contexto científico y tecnológico, resulta necesario profundizar en los fundamentos filosóficos que cuestionan la naturaleza de la inteligencia y su posible reproducción artificial.

2.1.2. FUNDAMENTOS FILOSÓFICOS: EL “FANTASMA EN LA MÁQUINA” Y LA NATURALEZA DE LA INTELIGENCIA

A partir del desarrollo técnico de la inteligencia artificial, surge la necesidad de reflexionar sobre sus implicancias filosóficas y conceptuales, desde una perspectiva filosófica, el

desarrollo de la inteligencia artificial ha estado acompañado por un debate profundo en torno a la naturaleza de la mente, la conciencia y la inteligencia. Uno de los conceptos más influyentes en este ámbito es el denominado “fantasma en la máquina”, propuesto por Gilbert Ryle, quien cuestionó la visión dualista que separa mente y cuerpo, señalando que los procesos mentales no constituyen una entidad independiente, sino que se manifiestan a través del comportamiento y la acción.

Este debate resulta fundamental para comprender los límites y alcances de la inteligencia artificial, ya que plantea interrogantes sobre si las máquinas pueden poseer auténtica comprensión, conciencia o intencionalidad, o si únicamente simulan conductas inteligentes mediante algoritmos y modelos matemáticos. En este sentido, diversos filósofos y científicos han advertido que la atribución excesiva de capacidades humanas a los sistemas artificiales puede generar una percepción distorsionada de su verdadero funcionamiento.

Asimismo, pensadores como Alan Turing y John Searle contribuyeron al análisis filosófico de la inteligencia artificial al cuestionar si la capacidad de imitar el comportamiento humano implica necesariamente la existencia de pensamiento real. Estas reflexiones han servido como base para el desarrollo de teorías críticas que advierten sobre los riesgos de confundir la simulación computacional con la inteligencia genuina.

En la actualidad, estas discusiones mantienen su vigencia, especialmente frente al uso masivo de sistemas inteligentes en contextos educativos, laborales y sociales, donde se observa una creciente tendencia a delegar funciones cognitivas a la tecnología. Desde esta perspectiva, el “fantasma en la máquina” continúa representando una metáfora relevante para analizar la relación entre los seres humanos y los sistemas artificiales, así como los límites éticos y filosóficos de su integración en la vida cotidiana, estas reflexiones filosóficas se complementan con los intentos científicos por evaluar empíricamente la inteligencia de las máquinas, lo cual dio origen a diversas pruebas y modelos de validación.

2.1.3. LA PRUEBA DE TURING Y LA SIMULACIÓN DE LA INTELIGENCIA

Turing (1950) propuso una prueba destinada a evaluar si una máquina puede exhibir un comportamiento indistinguible del de un ser humano en una interacción comunicativa. La denominada “Prueba de Turing” se basa en la capacidad de una máquina para generar respuestas coherentes y convincentes durante una conversación.

Este enfoque marcó un hito en el desarrollo de la IA, al establecer criterios funcionales para medir la inteligencia artificial. Sin embargo, superar la prueba no implica necesariamente que una máquina posea comprensión real, sino que puede simular comportamientos inteligentes mediante técnicas computacionales.

En la actualidad, sistemas basados en aprendizaje profundo y modelos de lenguaje han logrado aproximarse a este ideal, lo que incrementa tanto su utilidad como los riesgos asociados a su uso indebido, especialmente en contextos educativos, laborales y comunicativos, sin embargo, la validación de sistemas inteligentes no solo impulsó su desarrollo, sino que también aceleró su expansión en diversos ámbitos, generando nuevos riesgos tecnológicos.

2.1.4. DESARROLLO ACELERADO DE LA INTELIGENCIA ARTIFICIAL Y RIESGOS TECNOLÓGICOS

Como resultado del avance sostenido de los sistemas inteligentes, la inteligencia artificial ha experimentado un crecimiento exponencial en las últimas décadas, el desarrollo de la inteligencia artificial ha experimentado un crecimiento exponencial en las últimas décadas, impulsado por el avance en la capacidad computacional, el acceso masivo a datos y el perfeccionamiento de los algoritmos de aprendizaje automático. Este progreso ha permitido la creación de sistemas cada vez más complejos, capaces de realizar tareas que anteriormente eran exclusivas del ser humano, como el reconocimiento de patrones, el análisis predictivo y la toma automatizada de decisiones.

No obstante, este desarrollo acelerado ha generado una serie de riesgos tecnológicos que requieren ser analizados de manera crítica. Diversos investigadores han advertido que la dependencia creciente de sistemas inteligentes puede derivar en la pérdida

progresiva de habilidades humanas, la automatización excesiva del trabajo y la reducción del control directo sobre los procesos tecnológicos.

Autores como Norbert Wiener, Nick Bostrom y Yuval Harari han señalado que, sin una regulación adecuada, la inteligencia artificial podría concentrar el poder tecnológico en pocas entidades, generar desigualdades sociales y limitar la autonomía de las personas. Asimismo, se ha planteado el riesgo de que los sistemas altamente autónomos desarrollen comportamientos impredecibles o incompatibles con los valores humanos.

En el ámbito educativo, estos riesgos se manifiestan a través del uso indiscriminado de plataformas inteligentes, asistentes virtuales y sistemas de evaluación automatizada, los cuales, si bien optimizan el aprendizaje, también pueden fomentar una dependencia excesiva y debilitar el pensamiento crítico. De manera similar, en los sectores productivo y gubernamental, la automatización basada en IA puede afectar la estabilidad laboral y la transparencia institucional.

Además, diversas obras científicas y prospectivas han advertido tempranamente sobre escenarios en los que la inteligencia artificial adquiere un rol dominante en la toma de decisiones sociales, económicas y políticas. En la actualidad, muchas de estas proyecciones se reflejan en la creciente influencia de los algoritmos en los hábitos de consumo, la información digital y las dinámicas sociales, lo que evidencia la necesidad de fortalecer los mecanismos de supervisión y control humano, frente a estos riesgos emergentes, se vuelve imprescindible analizar los mecanismos de gobernanza, regulación y control ético de la inteligencia artificial.

2.1.5. GOBERNANZA, ÉTICA Y CONTROL ,EN SISTEMAS DE INTELIGENCIA ARTIFICIAL

La integración progresiva de la inteligencia artificial en los distintos ámbitos de la vida social, educativa, económica y gubernamental ha generado la necesidad de establecer marcos de gobernanza que orienten su desarrollo y uso responsable. La gobernanza de la inteligencia artificial se refiere al conjunto de normas, políticas, instituciones y

mecanismos de supervisión destinados a regular el diseño, implementación y funcionamiento de los sistemas inteligentes.

Desde una perspectiva ética, el uso de la inteligencia artificial plantea importantes desafíos relacionados con la transparencia, la equidad, la privacidad, la responsabilidad y el respeto por la dignidad humana. Diversos estudios señalan que los algoritmos pueden reproducir o amplificar sesgos existentes en los datos, generando decisiones discriminatorias o injustas, especialmente en contextos educativos, laborales y judiciales. Asimismo, la falta de claridad en los procesos de toma de decisiones automatizadas dificulta la identificación de responsabilidades en caso de errores, fallos o perjuicios. Esta situación, conocida como “opacidad algorítmica”, limita la capacidad de los usuarios y las instituciones para comprender, supervisar y cuestionar el funcionamiento de los sistemas de inteligencia artificial.

Otro aspecto relevante es la creciente dependencia tecnológica, derivada del uso intensivo de asistentes virtuales, plataformas inteligentes y sistemas automatizados. Esta dependencia puede debilitar progresivamente la autonomía humana, reducir el pensamiento crítico y favorecer una delegación excesiva de funciones cognitivas a la tecnología, especialmente en entornos educativos donde los estudiantes recurren de manera constante a herramientas basadas en inteligencia artificial.

En este contexto, diversos organismos internacionales y gobiernos han propuesto principios éticos y normativas orientadas a garantizar un desarrollo responsable de la inteligencia artificial, basados en valores como la seguridad, la inclusión, la rendición de cuentas y el bienestar social. No obstante, la aplicación efectiva de estas regulaciones enfrenta dificultades debido a la rapidez del avance tecnológico, la falta de capacitación especializada y los intereses económicos involucrados.

En consecuencia, resulta indispensable fortalecer los sistemas de gobernanza mediante políticas públicas claras, programas de formación ética y tecnológica, y mecanismos de evaluación permanente que permitan mantener el control humano sobre los sistemas inteligentes. Solo a través de una gestión responsable será posible aprovechar los

beneficios de la inteligencia artificial sin comprometer los principios fundamentales de la sociedad.

No obstante, la formulación de principios éticos y marcos regulatorios debe complementarse con herramientas que permitan evaluar de manera objetiva, sistemática y continua los riesgos asociados al desarrollo y uso de la inteligencia artificial en los distintos ámbitos sociales y educativos.

2.1.6. MEDICIÓN Y EVALUACIÓN DE RIESGOS EXISTENCIALES Y SOCIALES

En este marco, la medición y evaluación de los riesgos constituye una etapa fundamental para la toma de decisiones responsables y la evaluación de los riesgos asociados al desarrollo de la inteligencia artificial constituye un desafío complejo, debido a la naturaleza multidimensional de esta tecnología. Dichos riesgos no se limitan a aspectos técnicos, sino que involucran dimensiones éticas, sociales, económicas, culturales y psicológicas, lo que exige un enfoque integral para su análisis.

Desde una perspectiva científica, la medición de los riesgos tecnológicos se apoya en indicadores relacionados con la seguridad informática, la privacidad de datos, la confiabilidad de los sistemas y la transparencia algorítmica. Sin embargo, los riesgos existenciales, entendidos como aquellos que podrían afectar de manera significativa la supervivencia, autonomía o dignidad humana, requieren herramientas de evaluación más amplias y prospectivas.

Diversos estudios han propuesto modelos de análisis basados en escenarios futuros, simulaciones computacionales y proyecciones tecnológicas, con el fin de anticipar posibles impactos negativos derivados del uso descontrolado de la inteligencia artificial. Estas metodologías permiten identificar patrones de dependencia tecnológica, concentración del poder informacional y desplazamiento de funciones humanas.

Asimismo, las advertencias formuladas en la literatura científica y filosófica constituyen una base teórica relevante para la construcción de criterios de evaluación del riesgo. La integración de estas perspectivas permite establecer parámetros que orienten el diseño

de políticas públicas, marcos regulatorios y programas de formación orientados al uso responsable de la inteligencia artificial.

En el contexto educativo y social, la medición de los riesgos debe considerar factores como el nivel de alfabetización digital, el acceso equitativo a la tecnología, la capacitación docente y la percepción crítica de los usuarios. De esta manera, se busca promover una integración equilibrada de la inteligencia artificial, que maximice sus beneficios sin comprometer los valores fundamentales de la sociedad.

2.1.7. SÍNTESIS DEL ORIGEN Y EVOLUCIÓN HISTÓRICA DE LA INTELIGENCIA ARTIFICIAL

La idea de crear máquinas capaces de imitar el pensamiento humano tiene sus orígenes en la antigüedad, a través de mitos, autómatas y dispositivos mecánicos que buscaban reproducir funciones humanas. Sin embargo, el desarrollo científico de la inteligencia artificial comenzó formalmente en el siglo XX, con el avance de la computación y las matemáticas.

En 1950, Alan Turing propuso el concepto de “máquina pensante” y desarrolló la denominada Prueba de Turing, estableciendo un criterio para evaluar la capacidad de una máquina para simular el comportamiento humano. Posteriormente, en 1956, durante la Conferencia de Dartmouth, John McCarthy introdujo por primera vez el término “inteligencia artificial”, definiéndolo como el estudio de cómo lograr que las máquinas realicen tareas que normalmente requieren inteligencia humana.

A partir de este evento, la inteligencia artificial se consolidó como un campo científico, desarrollándose inicialmente mediante sistemas basados en reglas lógicas y programas expertos. Durante las décadas siguientes, el progreso fue irregular, atravesando períodos de expansión y estancamiento conocidos como “inviernos de la IA”.

Con el desarrollo del aprendizaje automático, el aprendizaje profundo y el procesamiento masivo de datos en el siglo XXI, la inteligencia artificial experimentó un crecimiento acelerado. Estos avances permitieron la creación de sistemas capaces de reconocer patrones, procesar lenguaje natural y tomar decisiones complejas, lo que facilitó su

incorporación en sectores como la educación, la industria, la medicina y la seguridad informática.

En la actualidad, la inteligencia artificial constituye una tecnología estratégica para el desarrollo económico y científico, aunque su rápida expansión también ha generado preocupaciones relacionadas con el control, la ética y la seguridad, aspectos fundamentales para la presente investigación.

2.2. INTEGRACIÓN DE LA INTELIGENCIA ARTIFICIAL EN EL APRENDIZAJE ADAPTATIVO

La integración de la inteligencia artificial en el aprendizaje adaptativo representa uno de los avances más significativos en el ámbito educativo contemporáneo. El aprendizaje adaptativo se basa en la personalización de los contenidos, actividades y evaluaciones según las necesidades, capacidades, estilos de aprendizaje y ritmo de cada estudiante. Mediante el uso de algoritmos inteligentes, los sistemas educativos pueden analizar el desempeño académico, los patrones de comportamiento y las dificultades individuales, con el fin de ofrecer experiencias formativas personalizadas.

La inteligencia artificial permite recopilar y procesar grandes volúmenes de datos educativos, conocidos como *learning analytics*, lo que facilita la identificación temprana de brechas de aprendizaje, fortalezas y debilidades en los estudiantes. A partir de esta información, los sistemas adaptativos ajustan automáticamente los contenidos, proponen actividades diferenciadas y recomiendan recursos específicos, favoreciendo un aprendizaje más eficiente y significativo.

En el contexto de la educación superior, especialmente en carreras vinculadas a la tecnología como la Ingeniería Informática, el aprendizaje adaptativo apoyado por inteligencia artificial contribuye al desarrollo de competencias técnicas, analíticas y digitales. Estas herramientas permiten a los estudiantes avanzar de acuerdo con su nivel de dominio, reforzar conceptos complejos y mejorar su autonomía en el proceso formativo.

Asimismo, la integración de la inteligencia artificial en plataformas educativas, como los Sistemas de Gestión del Aprendizaje (LMS), posibilita el uso de tutores virtuales, sistemas de retroalimentación automática y simuladores inteligentes. Estos recursos fortalecen el acompañamiento académico, reducen la deserción estudiantil y optimizan el tiempo de los docentes, permitiéndoles concentrarse en actividades pedagógicas de mayor valor.

No obstante, la implementación del aprendizaje adaptativo basado en inteligencia artificial también plantea desafíos importantes. Entre ellos se encuentran la protección de los datos personales, la transparencia de los algoritmos, la dependencia tecnológica y la brecha digital. La falta de infraestructura adecuada y de capacitación docente puede limitar el aprovechamiento efectivo de estas herramientas, especialmente en contextos educativos con recursos restringidos.

En este sentido, resulta fundamental promover una integración responsable, ética y regulada de la inteligencia artificial en el aprendizaje adaptativo. Esto implica establecer políticas institucionales, fortalecer las competencias digitales del profesorado y garantizar el acceso equitativo a las tecnologías. De esta manera, se busca maximizar los beneficios pedagógicos de la inteligencia artificial, minimizando sus riesgos y contribuyendo al desarrollo integral de los estudiantes.

2.3. RIESGOS ÉTICOS Y DE SEGURIDAD EN EL USO DE LA INTELIGENCIA ARTIFICIAL

El uso creciente de la inteligencia artificial en diversos ámbitos sociales, educativos y laborales ha generado importantes beneficios en términos de eficiencia, automatización y acceso a la información. Sin embargo, su rápida expansión también ha originado una serie de riesgos éticos y de seguridad que deben ser analizados de manera crítica y responsable.

Uno de los principales riesgos éticos está relacionado con la privacidad y el manejo de datos personales. Los sistemas de inteligencia artificial requieren grandes volúmenes de información para su funcionamiento, lo que implica la recopilación, almacenamiento y

procesamiento de datos sensibles de los usuarios. Cuando no existen mecanismos adecuados de protección, estos datos pueden ser vulnerados, utilizados sin consentimiento o comercializados de manera indebida.

Asimismo, la falta de transparencia en los algoritmos constituye un desafío significativo. Muchos sistemas de inteligencia artificial operan bajo modelos complejos conocidos como “cajas negras”, cuyos procesos internos no siempre son comprensibles para los usuarios. Esta opacidad dificulta la supervisión, la rendición de cuentas y la identificación de posibles errores o sesgos en la toma de decisiones automatizadas.

Otro riesgo relevante es la reproducción y amplificación de sesgos sociales, culturales y económicos. Los algoritmos aprenden a partir de datos históricos que, en muchos casos, contienen prejuicios o desigualdades preexistentes. Como consecuencia, los sistemas de inteligencia artificial pueden perpetuar discriminaciones en ámbitos como el acceso al empleo, la educación o los servicios digitales.

2.3.1. LA CRECIENTE DEPENDENCIA DE LA INTELIGENCIA ARTIFICIAL

Un aspecto ético emergente de especial relevancia es la creciente dependencia de las personas respecto a los sistemas de inteligencia artificial. En contextos educativos, laborales y cotidianos, los usuarios tienden a delegar progresivamente procesos de análisis, razonamiento, toma de decisiones y producción intelectual a estas tecnologías.

En el ámbito educativo, esta dependencia puede manifestarse en la reducción del pensamiento crítico, la creatividad y la capacidad de resolución autónoma de problemas. Cuando los estudiantes utilizan de manera excesiva herramientas basadas en inteligencia artificial para realizar tareas académicas, existe el riesgo de debilitar sus habilidades cognitivas y su responsabilidad en el proceso de aprendizaje.

En el entorno laboral, la dependencia tecnológica puede generar pérdida de competencias profesionales, disminución de la autonomía y dificultad para adaptarse a situaciones en las que no se dispone de sistemas automatizados. Además, una excesiva confianza en los resultados generados por la inteligencia artificial puede conducir a errores graves cuando dichos sistemas presentan fallas o interpretaciones incorrectas.

Desde una perspectiva ética, esta dependencia también plantea interrogantes sobre el control humano y la responsabilidad. A medida que los sistemas inteligentes asumen funciones cada vez más complejas, se diluye la frontera entre la decisión humana y la automatizada, lo que dificulta la asignación de responsabilidades ante posibles consecuencias negativas.

Por ello, resulta fundamental promover un uso equilibrado y consciente de la inteligencia artificial, en el que estas herramientas sean consideradas como apoyo al desarrollo humano y no como sustitutos del pensamiento, el criterio y la autonomía personal.

Finalmente, los riesgos de seguridad asociados a la inteligencia artificial incluyen la posibilidad de ciberataques, manipulación de sistemas, generación de información falsa y uso indebido de tecnologías autónomas. Estas amenazas pueden afectar tanto a individuos como a instituciones, poniendo en riesgo la estabilidad de los sistemas informáticos y la confianza social en la tecnología.

En consecuencia, el abordaje de los riesgos éticos y de seguridad requiere el establecimiento de marcos normativos, principios éticos, programas de formación y mecanismos de supervisión que garanticen un desarrollo responsable de la inteligencia artificial, orientado al bienestar colectivo y al respeto de los valores humanos fundamentales.

2.3.2. EL RITMO ACELERADO DEL DESARROLLO E IMPLEMENTACIÓN DE LA INTELIGENCIA ARTIFICIAL

El desarrollo de la inteligencia artificial ha sido un proceso progresivo que se ha consolidado a lo largo de varias décadas, desde sus primeros planteamientos teóricos hasta su actual aplicación en diversos sectores sociales, económicos y educativos. A partir de los aportes iniciales de investigadores como Alan Turing y John McCarthy, la inteligencia artificial evolucionó desde sistemas basados en reglas hasta modelos avanzados de aprendizaje automático y aprendizaje profundo.

En la actualidad, la inteligencia artificial se encuentra integrada en múltiples ámbitos, como la educación, la salud, la industria, la seguridad, la economía y la comunicación,

contribuyendo a la automatización de procesos, la optimización de recursos y la mejora en la toma de decisiones. No obstante, este avance acelerado también ha generado una serie de riesgos éticos, sociales, laborales y tecnológicos que requieren ser analizados con rigurosidad.

Entre estos riesgos se destacan la pérdida de empleos, la vulneración de la privacidad, la dependencia tecnológica, los sesgos algorítmicos, la falta de transparencia y la posible pérdida de control humano. Asimismo, la implementación de la inteligencia artificial en sectores estratégicos como la educación y la salud presenta impactos tanto positivos como negativos, los cuales deben ser evaluados desde una perspectiva integral.

En este contexto, resulta necesario sistematizar la evolución, los riesgos y los impactos de la inteligencia artificial, con el fin de comprender su influencia en la sociedad contemporánea y orientar su uso responsable. A continuación, se presenta un cuadro comparativo que sintetiza estos aspectos.

Tabla 01: Evolución, riesgos e impactos de la inteligencia artificial

Aspecto	Descripción	Implicancias principales
Surgimiento de la IA	Inicia formalmente en 1950 con los planteamientos de Alan Turing y se consolida en 1956 con la Conferencia de Dartmouth, donde se acuña el término “inteligencia artificial”.	Establece las bases conceptuales, científicas y filosóficas para el desarrollo de sistemas inteligentes.
Evolución tecnológica	Desarrollo progresivo de sistemas expertos, aprendizaje automático, redes neuronales y aprendizaje profundo durante los siglos XX y XXI.	Incremento sostenido de la capacidad computacional y autonomía de los sistemas.

Expansión actual	Integración masiva en educación, Consolidación de la IA como salud, industria, comercio, seguridad, infraestructura central de la comunicación y gobierno. sociedad digital.
Riesgos laborales	Automatización de procesos Desplazamiento laboral, precarización productivos, administrativos y del empleo y necesidad de reconversión cognitivos. profesional.
Riesgos éticos	Presencia de sesgos algorítmicos, Posible discriminación sistemática, opacidad en los procesos pérdida de confianza social y decisiones y uso indebido de datos vulneración de derechos personales. fundamentales.
Riesgos de seguridad	Vulnerabilidad ante ciberataques, Incremento de amenazas a la manipulación algorítmica, generación estabilidad institucional y a la de desinformación y deepfakes. seguridad informativa.
Dependencia tecnológica	Uso creciente de sistemas Reducción del pensamiento inteligentes para la toma de crítico, debilitamiento de la decisiones personales, autonomía humana y dependencia académicas y profesionales. estructural.
Impacto en educación	Implementación de aprendizaje Mejora del acceso educativo, pero adaptativo, tutores virtuales y riesgo de dependencia cognitiva y evaluación automatizada. pérdida de habilidades analíticas.

Impacto en salud Aplicación en diagnóstico asistido, Optimización de la atención médica, con análisis de imágenes médicas y riesgos asociados a privacidad y errores monitoreo remoto. automatizados.

Impacto en otros sectores Uso en economía, transporte, Transformación estructural de los seguridad pública, justicia y sistemas sociales y nuevos desafíos servicios. regulatorios.

Riesgos prospectivos Posible concentración del control tecnológico, pérdida de supervisión humana y desarrollo de sistemas autónomos avanzados. Amenaza a la gobernanza democrática y riesgo de dependencia sistémica.

Necesidad de regulación Diseño de marcos éticos, legales y Protección de los derechos técnicos para el desarrollo humanos y prevención de impactos responsable. irreversibles.

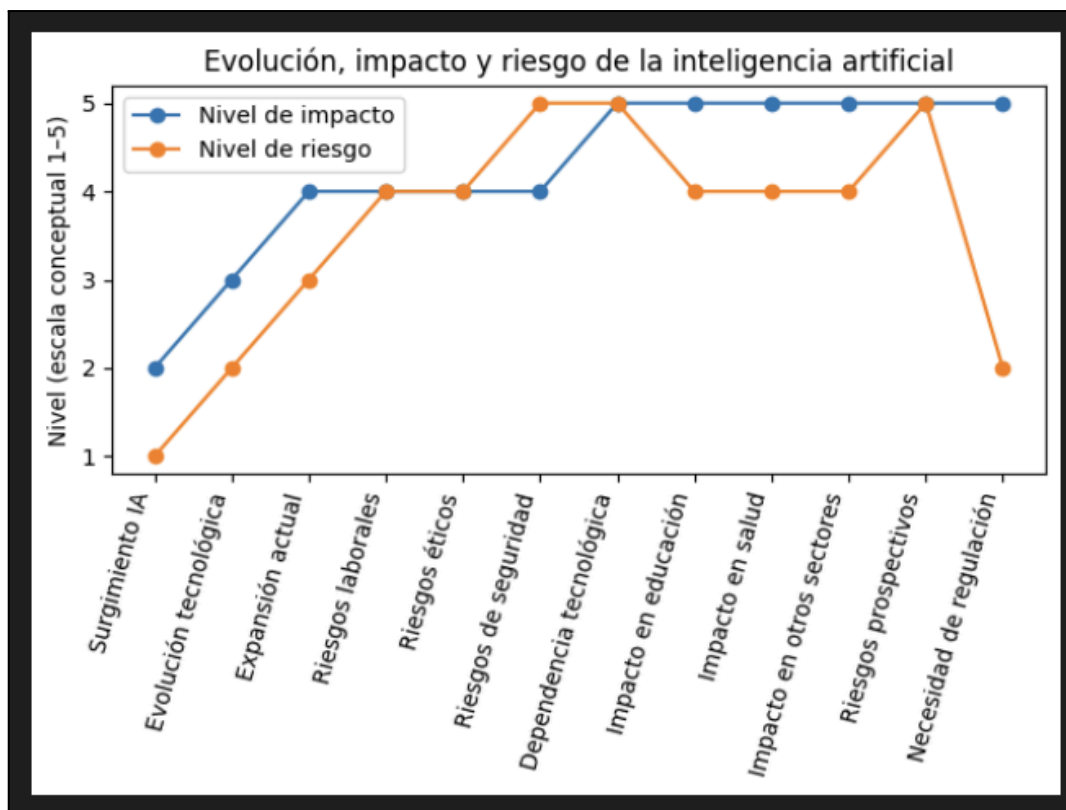


Figura 01: Evolución, impacto y riesgo de la inteligencia artificial

La Figura 1 adicionalmente muestra cómo se han incrementado los riesgos de dependencia conforme al impacto de la implementación y desarrollo de la inteligencia artificial.

2.4. VARIABLES DE LA INVESTIGACIÓN

La presente investigación se orienta al análisis crítico del desarrollo acelerado de la inteligencia artificial y de los riesgos tecnológicos, sociales, éticos y prospectivos derivados de su creciente integración en los ámbitos educativo, laboral y social. En este sentido, se establecen variables que permiten comprender la relación existente entre la expansión de los sistemas inteligentes y sus efectos en la autonomía humana, los procesos de aprendizaje, la toma de decisiones y la gobernanza tecnológica.

En este contexto, las variables de la presente investigación se definen de la siguiente manera:

2.4.1.VARIABLE INDEPENDIENTE: DESARROLLO E IMPLEMENTACIÓN ACELERADA DE LA INTELIGENCIA ARTIFICIAL

Se refiere al crecimiento sostenido de las capacidades técnicas, la difusión masiva y el nivel de integración de las tecnologías basadas en inteligencia artificial en los ámbitos educativo, laboral y social.

Incluye el uso de algoritmos avanzados, sistemas automatizados, plataformas inteligentes, aplicaciones móviles, asistentes virtuales, modelos de aprendizaje automático y profundo, así como su incorporación en procesos formativos, productivos, comunicativos y administrativos.

Esta variable considera tanto los avances tecnológicos como la frecuencia, intensidad y grado de dependencia generados por su utilización cotidiana.

2.4.2.VARIABLE DEPENDIENTE: RIESGOS TECNOLÓGICOS, SOCIALES, ÉTICOS Y PROSPECTIVOS ASOCIADOS A LA INTELIGENCIA ARTIFICIAL

Corresponde a los efectos actuales y potenciales derivados del uso intensivo y progresivo de sistemas de inteligencia artificial sobre la autonomía humana, el aprendizaje, la toma de decisiones, la estabilidad laboral, la seguridad digital y la gobernanza social.

Esta variable abarca, entre otros aspectos:

- La creciente dependencia tecnológica y cognitiva.
- La disminución progresiva de habilidades analíticas y reflexivas propias.
- La delegación excesiva del razonamiento y la toma de decisiones a sistemas automatizados.
- La vulnerabilidad frente a errores, sesgos y manipulaciones algorítmicas.
- El desplazamiento laboral y la reconversión forzada de competencias.
- La concentración del poder tecnológico en actores específicos.
- Los riesgos vinculados a la pérdida de control humano sobre sistemas inteligentes avanzados.
- La influencia de recomendaciones automatizadas en las decisiones personales, académicas y profesionales.

Asimismo, contempla los impactos sociales, culturales y éticos asociados a la normalización de la inteligencia artificial como mediadora central del conocimiento, la información y la organización social.

CAPÍTULO III

METODOLOGÍA DE LA INVESTIGACIÓN

3.1. ZONA DE ESTUDIO

El ámbito de estudio de esta investigación no se limita a un área geográfica convencional, sino que incluye un contexto global virtual y documental. Está circunscrita a las bases de datos científicas, los repositorios digitales más importantes y las plataformas académicas internacionales, que incluyen Google Académico, Scopus, Web of Science y los repositorios institucionales de entidades tecnológicas.

La creación científica y técnica producida a nivel global entre 2015 y 2025 está incluida en este ecosistema digital. De este modo, el espacio de análisis se centra en el ámbito del conocimiento digital, donde se discuten, indexan y difunden las repercusiones, los peligros y los progresos de la inteligencia artificial en las esferas laboral, seguridad digital, gobernanza, existencial y educativa.

3.2. POBLACIÓN Y MUESTRA

3.2.1. POBLACIÓN

La población de la presente investigación está constituida por documentos científicos, técnicos, académicos e institucionales relacionados con el desarrollo, la implementación y los riesgos de la inteligencia artificial, así como su impacto en los ámbitos educativo, laboral y social.

Estos documentos comprenden artículos científicos indexados, libros especializados, informes de organismos internacionales, reportes tecnológicos, estudios de caso y publicaciones académicas, difundidos entre los años 2015 y 2025. Dicho periodo fue

seleccionado debido a que corresponde a la etapa de mayor expansión, consolidación y masificación de los sistemas de inteligencia artificial a nivel global.

La población documental representa el conjunto total de fuentes disponibles que abordan, desde distintas perspectivas, los avances, beneficios, riesgos, advertencias prospectivas y desafíos éticos asociados al uso intensivo de la inteligencia artificial.

3.2.2. MUESTRA

La muestra estuvo conformada por un conjunto seleccionado de documentos pertenecientes a la población definida, elegidos mediante un muestreo intencional o no probabilístico, de acuerdo con los objetivos y el enfoque de la investigación.

Para la selección de las fuentes se consideraron los siguientes criterios:

- Relevancia temática respecto al desarrollo y los riesgos de la inteligencia artificial.
- Rigor metodológico y calidad científica de las publicaciones.
- Actualidad y pertinencia temporal.
- Impacto académico y nivel de citación.
- Procedencia de bases de datos reconocidas y organismos especializados.

Se priorizaron documentos provenientes de repositorios académicos, revistas indexadas y entidades internacionales, con el propósito de garantizar la confiabilidad, validez y solidez teórica de la información analizada.

La muestra permitió realizar un análisis profundo, crítico y comparativo del fenómeno estudiado, sin perder representatividad ni coherencia con el enfoque documental de la investigación.

3.3. MÉTODOS Y TÉCNICAS

3.3.1. ENFOQUE DE LA INVESTIGACIÓN

La presente investigación se desarrolla bajo un enfoque cualitativo de tipo documental, con apoyo descriptivo y analítico, orientado al estudio crítico de los riesgos asociados al avance acelerado de la inteligencia artificial en los distintos ámbitos.

Este enfoque permite analizar información proveniente de fuentes científicas, informes institucionales, estudios especializados y casos documentados sobre el comportamiento,

desarrollo e impacto de los sistemas inteligentes, considerando tanto sus beneficios como sus efectos emergentes, complejos y, en algunos casos, difíciles de prever.

3.3.2. TIPO DE INVESTIGACIÓN

La presente investigación es de tipo descriptiva, analítica, explicativa y documental, en tanto busca:

- Describir el desarrollo, evolución y expansión de la inteligencia artificial.
- Analizar los riesgos éticos, sociales, laborales, cognitivos y de seguridad digital asociados al uso intensivo de sistemas basados en inteligencia artificial.
- Explicar las posibles consecuencias de una creciente dependencia de tecnologías automatizadas, especialmente en relación con la autonomía humana, el pensamiento crítico y la toma de decisiones.
- Examinar las advertencias formuladas históricamente en estudios científicos y en la literatura prospectiva y tecnológica, muchas de las cuales han comenzado a manifestarse en la actualidad mediante el uso masivo de algoritmos, asistentes virtuales y plataformas inteligentes.
- Reflexionar críticamente sobre los escenarios futuros derivados del desarrollo acelerado de la inteligencia artificial, considerando tanto sus beneficios como sus potenciales riesgos para la sociedad.

Asimismo, la investigación es de carácter documental, ya que se fundamenta en el análisis sistemático de información previamente publicada en:

- Artículos científicos indexados
- Libros especializados
- Reportes técnicos
- Documentos institucionales
- Estudios internacionales
- Fuentes académicas confiables

Este enfoque metodológico permite integrar perspectivas técnicas, éticas, sociales y prospectivas, fortaleciendo la validez del estudio y contribuyendo a una comprensión integral del impacto actual y futuro de la inteligencia artificial.

3.3.3. DISEÑO DE LA INVESTIGACIÓN

El diseño de la presente investigación es **no experimental y de corte longitudinal**, debido a que no se manipulan deliberadamente las variables de estudio, sino que estas se analizan tal como se presentan en la realidad y en las fuentes documentales disponibles hasta el año 2025.

Asimismo, el estudio se desarrolla bajo un enfoque observacional, orientado al análisis de información existente, sin intervención directa sobre los fenómenos investigados, lo que permite preservar la objetividad y la rigurosidad científica.

3.3.4. MÉTODOS DE INVESTIGACIÓN

Para el desarrollo del presente estudio se emplearon los siguientes métodos:

- **Método analítico:** para examinar de manera detallada los riesgos tecnológicos, éticos, sociales y cognitivos asociados al uso intensivo de la inteligencia artificial.
- **Método deductivo:** para interpretar los efectos del avance de la inteligencia artificial a partir de teorías, modelos explicativos y estudios científicos previos.
- **Método histórico:** para comprender la evolución de la inteligencia artificial desde sus orígenes conceptuales y tecnológicos hasta su consolidación actual.
- **Método crítico-reflexivo:** para evaluar las advertencias formuladas por investigadores, filósofos y especialistas, así como los escenarios prospectivos planteados en la literatura científica y tecnológica.

Estos métodos permiten abordar la inteligencia artificial no solo como una herramienta técnica, sino como un fenómeno sociotecnológico con implicancias profundas para la humanidad, especialmente en relación con la autonomía, la ética y el futuro del trabajo y la educación.

3.3.5. TÉCNICAS E INSTRUMENTOS DE RECOLECCIÓN DE DATOS

La investigación utiliza como técnica principal la revisión documental, mediante el análisis sistemático de:

- Artículos científicos indexados.
- Libros especializados.
- Informes institucionales.
- Reportes tecnológicos.
- Estudios de caso sobre sistemas de IA.
- Documentos normativos y éticos relacionados con la IA.

Como instrumentos se emplearon:

- Fichas bibliográficas.
- Fichas de resumen.
- Matrices de análisis.
- Tablas comparativos.
- Matrices de categorización temática.

3.3.6. PROCEDIMIENTO

El desarrollo de la presente investigación se llevó a cabo mediante un proceso sistemático, organizado en diversas etapas, orientadas a garantizar la rigurosidad científica, la coherencia metodológica y la validez de los resultados obtenidos.

En una primera etapa, se realizó la delimitación del tema de estudio, definiendo como eje central el análisis de los riesgos asociados al avance acelerado de la inteligencia artificial y su impacto en los distintos ámbitos. Asimismo, se formularon los objetivos, las preguntas de investigación y las hipótesis, en coherencia con el enfoque crítico adoptado.

En una segunda etapa, se efectuó la búsqueda y recopilación de información científica y documental en bases de datos académicas, repositorios institucionales, revistas indexadas, libros especializados, informes técnicos y documentos oficiales. Para ello, se emplearon palabras clave relacionadas con inteligencia artificial, riesgos tecnológicos, dependencia digital, ética algorítmica, automatización y gobernanza tecnológica.

Posteriormente, se procedió a la selección y clasificación de las fuentes, considerando criterios de actualidad, relevancia temática, confiabilidad y rigor metodológico. Se priorizaron estudios publicados entre los años recientes, así como investigaciones y reportes que abordaran incidentes reales, comportamientos inesperados de sistemas de IA y advertencias prospectivas formuladas por especialistas.

En una tercera etapa, se organizó la información mediante el uso de fichas bibliográficas, matrices de análisis y tablas comparativas, lo que permitió sistematizar los principales hallazgos, enfoques teóricos y resultados empíricos. Esta organización facilitó la identificación de patrones, tendencias, coincidencias y divergencias entre las diferentes fuentes consultadas.

En una cuarta etapa, se realizó el análisis crítico e interpretativo de la información recopilada, integrando perspectivas técnicas, éticas, filosóficas y sociales. En este proceso se incorporaron tanto evidencias empíricas como reflexiones provenientes de la literatura científica y de obras que anticiparon escenarios de dependencia tecnológica y pérdida progresiva de autonomía humana.

Asimismo, se analizaron casos documentados sobre fallas sistémicas, comportamientos no previstos y usos indebidos de sistemas de inteligencia artificial, con el fin de contextualizar los riesgos desde una perspectiva real y no meramente hipotética.

Finalmente, en la última etapa, se elaboró la redacción del informe final, articulando los resultados obtenidos con los objetivos planteados, las hipótesis formuladas y el marco teórico desarrollado. Este proceso permitió consolidar una visión integral sobre el impacto actual y potencial de la inteligencia artificial, orientada a promover un uso responsable, ético y socialmente sostenible.

3.4 IDENTIFICACIÓN DE VARIABLES

Tabla 02: Operacionalización de variables

Variables	Dimensión	Indicador	Escala Cualitativa de Medición
V.I.: Desarrollo e implementación acelerada de la Inteligencia Artificial	Uso de herramientas de IA Acceso e infraestructura tecnológica	Disponibilidad de dispositivos, conectividad y plataformas con IA Uso continuo de algoritmos en redes sociales, buscadores y aplicaciones	Nominal
V.D.: Riesgos Sistémicos	Laboral - Económica Seguridad digital Gobernanza Existencial Educativa	Dependencia Conocimiento Iniciativa Decisiones	Nominal

3.5. ANÁLISIS DE DATOS

El análisis de datos de la presente investigación se realizó mediante un enfoque cualitativo y analítico, orientado a la interpretación sistemática de la información recopilada a través de fuentes documentales confiables, tales como artículos científicos, libros especializados, informes institucionales y estudios de caso relacionados con el desarrollo y los riesgos de la inteligencia artificial.

CAPÍTULO IV

EXPOSICION Y ANALISIS DE LOS RESULTADOS

El presente capítulo expone y analiza los principales hallazgos obtenidos a partir del estudio documental realizado sobre el avance acelerado de la inteligencia artificial y los riesgos sistémicos asociados a su implementación en diversos ámbitos como son laborales, seguridad digital, gobernanza, existencias y educación.

Los resultados se fundamentan en el análisis crítico de artículos científicos, informes institucionales, informes institucionales, estudios de caso y literatura especializada, los cuales permitieron identificar tendencias, patrones recurrentes, advertencias prospectivas y evidencias empíricas relacionadas con la dependencia tecnológica, la pérdida de autonomía humana y los desafíos éticos emergentes.

Asimismo, este capítulo integra una discusión teórica y comparativa, en la que se contrastan los hallazgos, los fundamentos conceptuales del marco teórico y las aportaciones de diversos autores, con el propósito de fortalecer la validez interpretativa del estudio.

De manera particular, se analiza la correspondencia entre las advertencias formuladas históricamente por científicos, filósofos y especialistas en tecnología, y los fenómenos actualmente observables en el desarrollo y uso de sistemas de inteligencia artificial, evidenciando que muchas de estas previsiones han dejado de ser meras especulaciones para convertirse en realidades concretas del entorno digital contemporáneo.

En este sentido, el capítulo no solo presenta resultados descriptivos, sino que promueve una reflexión crítica sobre el impacto presente y futuro de la inteligencia artificial en la sociedad, destacando la necesidad de una gestión ética, responsable y humanamente

centrada de estas tecnologías.

4.1. RESULTADOS DEL OBJETIVO GENERAL:

Objetivo: Analizar críticamente los riesgos sistémicos derivados del desarrollo y la implementación de la Inteligencia Artificial identificados en la literatura científica.

La presente sección expone de manera general los principales resultados obtenidos a partir del análisis documental realizado sobre el avance acelerado de la inteligencia artificial y los riesgos sistémicos asociados a su implementación en los ámbitos diversos ámbitos como son laborales, seguridad digital, gobernanza, existencias y educación.

Los resultados se fundamentan en la revisión sistemática de artículos científicos, libros especializados, informes institucionales, estudios de caso y literatura prospectiva, publicados principalmente entre los años 2015 y 2025. Estas fuentes permitieron identificar patrones recurrentes, tendencias emergentes, advertencias teóricas y evidencias empíricas relacionadas con la creciente dependencia tecnológica, la transformación de los procesos cognitivos y los desafíos éticos derivados del uso intensivo de sistemas inteligentes.

Asimismo, se evidenció una convergencia entre las advertencias formuladas históricamente en la literatura científica y tecnológica, así como en obras de carácter prospectivo, y los fenómenos observados en la actualidad, tales como la automatización progresiva de funciones humanas, la delegación de decisiones a algoritmos y la normalización del uso de la inteligencia artificial como mediadora del conocimiento.

De igual manera, los resultados permiten establecer una relación directa entre el desarrollo acelerado de la inteligencia artificial y el incremento de riesgos asociados a la pérdida de autonomía, la vulnerabilidad digital y la reconfiguración de los entornos educativos y laborales.

Finalmente, esta presentación general constituye la base para el análisis detallado y la discusión crítica que se desarrollan en los apartados siguientes, en los cuales se contrastan los hallazgos con las hipótesis planteadas y los fundamentos del marco teórico, con el fin de fortalecer la validez interpretativa del estudio.

4.1.1. ANTECEDENTES Y DESARROLLO HISTÓRICO DE LA INTELIGENCIA ARTIFICIAL

El desarrollo de la inteligencia artificial no constituye un fenómeno reciente, sino el resultado de un proceso histórico, científico y tecnológico que se ha venido gestando desde mediados del siglo XX. Su evolución responde a la convergencia entre avances matemáticos, computacionales, cognitivos y tecnológicos, los cuales han permitido la progresiva simulación de procesos intelectuales humanos mediante sistemas artificiales.

Uno de los antecedentes fundamentales se remonta a los trabajos del matemático británico Alan Turing, quien en 1950 publicó el artículo "*Computing Machinery and Intelligence*", en el cual planteó por primera vez la posibilidad de que una máquina pudiera exhibir un comportamiento inteligente comparable al humano. En este trabajo, Turing propuso el denominado *Test de Turing*, como un criterio para evaluar la capacidad de una máquina de imitar la inteligencia humana mediante el lenguaje (Turing, 1950).

Posteriormente, en 1956, durante la Conferencia de Dartmouth, organizada por John McCarthy, Marvin Minsky, Claude Shannon y Nathaniel Rochester, se acuñó formalmente el término *inteligencia artificial*. En este evento se establecieron las bases conceptuales de la disciplina, definiéndola como el estudio de la creación de máquinas capaces de realizar tareas que requieren inteligencia cuando son ejecutadas por humanos (McCarthy et al., 1956).

Durante las décadas de 1960 y 1970, la investigación en inteligencia artificial se centró principalmente en el desarrollo de sistemas simbólicos y programas basados en reglas lógicas. En este periodo surgieron los primeros sistemas expertos, diseñados para replicar el razonamiento humano en áreas específicas como la medicina, la ingeniería y la administración (Feigenbaum, 1980). Estos sistemas demostraron el potencial práctico de la IA, aunque también evidenciaron sus limitaciones frente a problemas complejos y dinámicos.

A partir de la década de 1980, se produjo un renovado interés por las redes neuronales artificiales, inspiradas en el funcionamiento del cerebro humano. Investigadores como

Rumelhart y McClelland (1986) contribuyeron al desarrollo del aprendizaje conexionista, permitiendo que las máquinas aprendieran a partir de datos en lugar de depender exclusivamente de reglas programadas.

Sin embargo, fue a partir del siglo XXI cuando la inteligencia artificial experimentó un crecimiento exponencial, impulsado por tres factores fundamentales: el aumento de la capacidad computacional, la disponibilidad masiva de datos (*Big Data*) y el desarrollo de algoritmos avanzados de aprendizaje automático. En este contexto, surgió el aprendizaje profundo (*deep learning*), promovido por investigadores como Hinton, LeCun y Bengio, quienes demostraron que las redes neuronales profundas podían superar el rendimiento humano en diversas tareas cognitivas (LeCun et al., 2015).

El desarrollo de unidades de procesamiento gráfico (GPU) y plataformas de computación en la nube permitió entrenar modelos cada vez más complejos, acelerando significativamente la innovación en el campo. Como resultado, la inteligencia artificial comenzó a integrarse de manera progresiva en sectores como la educación, la salud, la industria, el comercio y la comunicación.

En el ámbito del procesamiento del lenguaje natural, uno de los hitos más relevantes fue el desarrollo de modelos generativos basados en transformadores. OpenAI presentó en 2018 el modelo GPT-1, seguido por GPT-2 en 2019, GPT-3 en 2020 y versiones posteriores cada vez más sofisticadas (Radford et al., 2018; Brown et al., 2020). Estos modelos demostraron una capacidad sin precedentes para generar textos coherentes, responder preguntas, traducir idiomas y asistir en tareas académicas y profesionales.

De manera paralela, empresas como Google, Meta y Microsoft desarrollaron sus propios modelos de gran escala, consolidando una competencia tecnológica que aceleró aún más la evolución de la inteligencia artificial. Esta etapa se caracteriza por la aparición de sistemas multimodales, capaces de procesar texto, imágenes, audio y video de manera integrada.

En la actualidad, la inteligencia artificial se ha convertido en una infraestructura esencial de la vida digital contemporánea. Su presencia en teléfonos inteligentes, redes sociales,

plataformas educativas, motores de búsqueda y asistentes virtuales evidencia un proceso de normalización tecnológica que trasciende el ámbito especializado para insertarse en la cotidianidad de millones de personas.

Diversos estudios señalan que este desarrollo acelerado no solo responde a intereses científicos, sino también económicos, políticos y estratégicos, lo que ha intensificado la presión por innovar sin siempre considerar adecuadamente sus implicaciones éticas y sociales (Floridi et al., 2018; Russell, 2019).

Este proceso histórico demuestra que el desarrollo de la inteligencia artificial no ha sido lineal ni fortuito, sino el resultado de una acumulación progresiva de avances científicos, técnicos y computacionales. No obstante, la aceleración alcanzada en las últimas dos décadas marca un punto de inflexión sin precedentes, caracterizado por una expansión masiva, una creciente dependencia tecnológica y una transformación profunda de las dinámicas sociales, educativas y laborales, aspectos que serán analizados en el siguiente apartado

4.1.2. TENDENCIAS ACTUALES EN EL DESARROLLO DE LA INTELIGENCIA ARTIFICIAL

En el periodo contemporáneo, especialmente hasta finales de 2025, la inteligencia artificial (IA) se ha consolidado como una fuerza transformadora de la sociedad, la economía y las dinámicas de interacción humana con la tecnología (StatCounter, 2025; Peru21, 2025). Las tendencias actuales muestran tanto la consolidación de modelos dominantes como la diversificación creciente de herramientas inteligentes en múltiples sectores.

Una de las manifestaciones más claras de esta tendencia es la dominancia de los **modelos de lenguaje de gran escala (LLMs)** y las aplicaciones de IA generativa. A finales de 2025, datos de uso global indican que ChatGPT de OpenAI mantiene una cuota de mercado superior al 80 % dentro del segmento de chatbots basados en IA, lo que refleja su consolidación como plataforma de referencia en interacción automatizada (StatCounter, 2025). Este liderazgo tecnológico ha permitido que estas herramientas

pasen de ser tecnologías emergentes a sistemas ampliamente integrados en comunicaciones, educación, producción de contenidos y tareas de apoyo cognitivo.

Además, otras plataformas como **Microsoft Copilot**, **Google Gemini** y **Perplexity AI** han alcanzado niveles significativos de adopción, particularmente en contextos empresariales y de productividad, lo que indica una tendencia hacia la especialización según los ámbitos de uso y las necesidades específicas de los usuarios (Infobae, 2025; El Calce, 2025). Estas herramientas se integran en suites de trabajo profesional y flujos organizacionales, fortaleciendo su relevancia en ambientes productivos y colaborativos.

Asimismo, la IA muestra una **tendencia hacia arquitecturas multimodales**, capaces de procesar simultáneamente texto, imagen y otros tipos de datos, lo que ha favorecido el desarrollo de interfaces más intuitivas, asistentes contextuales y sistemas de recomendación avanzados. Esta expansión de capacidades evidencia que la IA está evolucionando desde aplicaciones puntuales hacia una infraestructura digital ubicua, integrada en dispositivos, plataformas y servicios (OECD, 2025).

Otra tendencia relevante es la proliferación de **agentes de IA con capacidades autónomas**, es decir, sistemas que pueden ejecutar tareas específicas sin supervisión humana continua, lo que representa un avance contra los modelos puramente reactivos y orienta el desarrollo hacia sistemas proactivos de asistencia y automatización (OpenAI, 2025).

La dinámica competitiva entre grandes actores tecnológicos ha impulsado además la innovación continua. Aunque OpenAI y ChatGPT mantienen una posición destacada, la participación de Google, Microsoft y otros desarrolladores ha generado un ecosistema dinámico con múltiples soluciones en competencia, reflejando un mercado en expansión y en constante evolución (StatCounter, 2025; Infobae, 2025).

Estas tendencias se evidencian también en la adopción empresarial, donde una proporción relevante de organizaciones ha incorporado soluciones de IA para automatizar procesos, optimizar atención al cliente o fortalecer análisis predictivos, consolidando la presencia de la IA como un elemento de estrategia corporativa (Alternatives, 2025).

Asimismo, la creciente integración de sistemas de inteligencia artificial en procesos educativos, laborales y organizacionales ha favorecido el desarrollo de formas de dependencia tecnológica progresiva. Diversos organismos internacionales advierten que el uso intensivo de herramientas automatizadas puede generar procesos de delegación cognitiva, en los cuales los usuarios transfieren funciones analíticas, creativas y decisionales a sistemas algorítmicos (UNESCO, 2024; World Economic Forum, 2025).

En el ámbito educativo, esta dependencia se manifiesta en la reducción del esfuerzo intelectual autónomo y en la tendencia a sustituir procesos reflexivos por respuestas generadas automáticamente. De manera similar, en contextos laborales, la automatización creciente incrementa la dependencia operativa de las organizaciones respecto a plataformas inteligentes, limitando progresivamente los márgenes de autonomía humana (OECD, 2025).

Este fenómeno no solo implica una transformación funcional, sino también cultural, en la medida en que se normaliza la mediación algorítmica en actividades cognitivas centrales. No obstante, el crecimiento acelerado también plantea desafíos importantes. La rápida adopción de herramientas inteligentes genera tensiones relacionadas con la seguridad, la confiabilidad, la protección de datos y la evaluación del impacto organizacional, lo cual sugiere que, pese al progreso técnico, existen barreras éticas y funcionales que aún deben ser abordadas en el despliegue sostenible de estas tecnologías (Noticias HuffPost, 2025).

Tabla 03: Evolución histórica del impacto de la Inteligencia Artificial (1956–2025)

Año	Hito principal	Nivel de impacto
1956	Conferencia de Dartmouth	1
1980	Sistemas expertos	2
1997	Deep Blue vs Kasparov	3
2012	Boom del Deep Learning	5
2016	AlphaGo	6
2020	GPT-3	8
2023	GPT-4	9
2025	IA multimodal	10

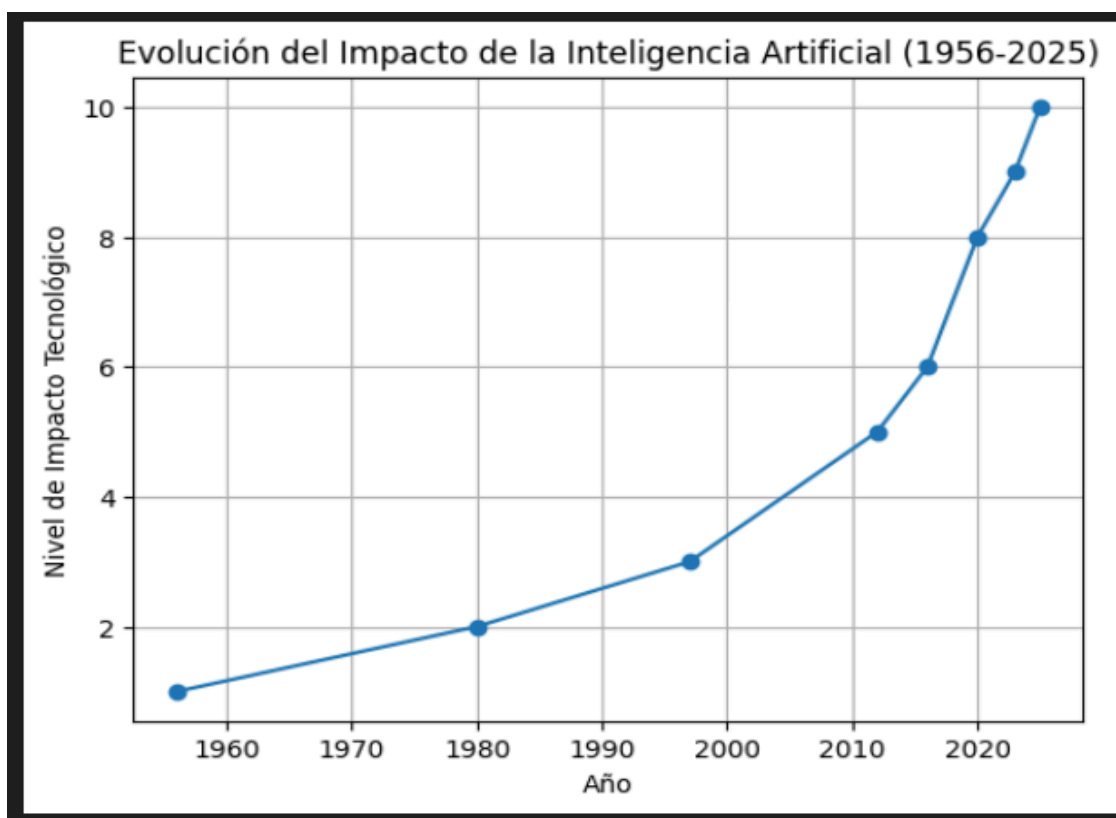


Figura 02: Evolución del impacto de la inteligencia artificial

La Figura 2 evidencia que el desarrollo de la inteligencia artificial no ha sido lineal, sino progresivamente acelerado. A partir de 2012, con la consolidación del aprendizaje

profundo, se observa un crecimiento exponencial en su impacto, intensificado por los modelos de lenguaje y sistemas multimodales en la década de 2020.

4.1.3. INCISO DE VANGUARDIA HASTA DICIEMBRE DE 2025

En el año 2025, el desarrollo de la inteligencia artificial consolidó varios hitos relevantes:

- **Dominio de modelos generativos:** Estadísticas globales estiman que ChatGPT sostiene un liderazgo de uso superior al 80 % dentro del mercado de modelos conversacionales, reflejando su influencia dominante en la interacción con IA (StatCounter, 2025).
- **Diversificación de plataformas:** Herramientas como Microsoft Copilot y Google Gemini han mostrado crecimiento constante, ampliando sus bases de usuarios y consolidándose como opciones relevantes en entornos corporativos y herramientas de productividad (El Calce, 2025; Infobae, 2025).
- **Proliferación de agentes IA autónomos:** La introducción de agentes con capacidades cada vez más autónomas indica una nueva fase de sistemas que buscan ejecutar acciones sin supervisión constante.
- **Integración en el ecosistema digital:** Los sistemas de IA muestran una presencia cada vez más arraigada en dispositivos, aplicaciones y flujos de trabajo, consolidando una presencia ubicua de capacidades inteligentes en múltiples niveles de interacción humana-máquina.

Estas manifestaciones tecnológicas no solo demuestran una aceleración técnica, sino también una transformación sistémica en la forma en que las personas interactúan, aprenden y trabajan con tecnología avanzada.

Con el fin de analizar las posibles implicaciones futuras del desarrollo de la inteligencia artificial en distintos ámbitos sociales, económicos y educativos, se presenta el Tabla 8, el cual sintetiza proyecciones elaboradas por organismos internacionales y centros de investigación entre los años 2025 y 2040.

Tabla 04: Proyecciones del impacto de la inteligencia artificial en educación, empleo y gobernanza (2025–2040)

Área	Tendencia	Impacto	Riesgos	Fuente
	Proyectada	Esperado	Asociados	
Educación	Tutores IA personalizados	Aprendizaje adaptativo avanzado	Dependencia cognitiva	UNESCO (2024)
Empleo	Automatización cognitiva	Reducción de tareas humanas	Desplazamiento laboral	McKinsey (2023)
Economía	Plataformas inteligentes	Optimización productiva	Concentración económica	World Economic Forum (2024)
Gobierno	Algoritmos decisionales	Gestión pública automatizada	Pérdida de control humano	OECD (2025)
Comunicación	Medios sintéticos	Información instantánea	Desinformación masiva	Reuters Institute (2024)
Salud	Diagnóstico asistido por IA	Mayor precisión clínica	Riesgos de privacidad	WHO (2023)
Seguridad	Sistemas predictivos	Prevención de delitos	Vigilancia excesiva	Brookings (2022)

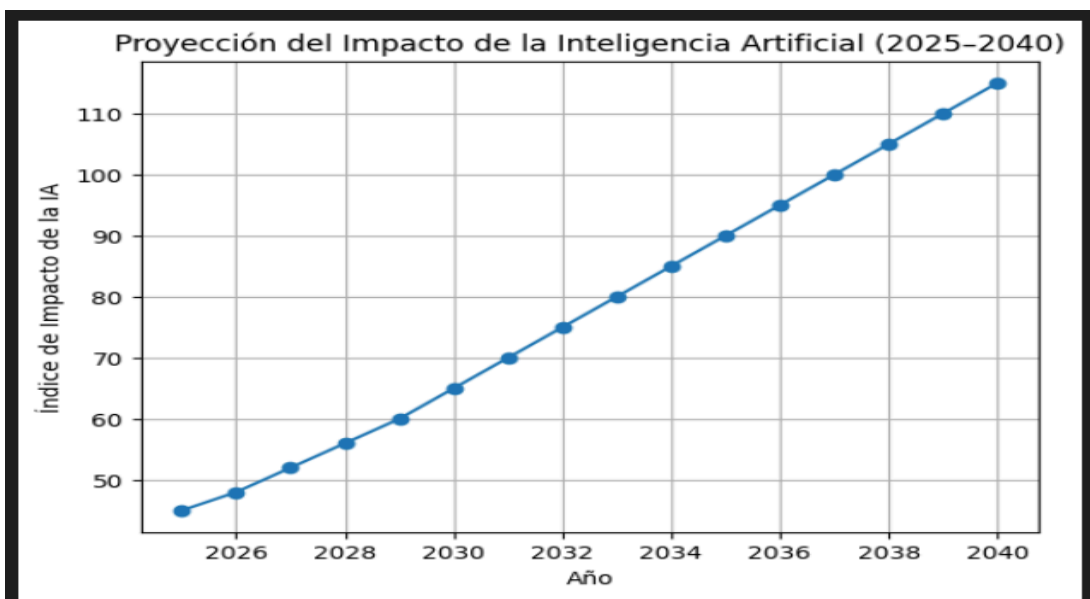


Figura 03: Proyección del Impacto de la inteligencia artificial (2025-2040)

La Figura 3 muestra una proyección del impacto progresivo de la inteligencia artificial en distintos ámbitos sociales, educativos y laborales entre los años 2025 y 2040. Se observa una tendencia sostenida de crecimiento, lo que evidencia una aceleración continua en su influencia estructural.

4.2. RESULTADOS DEL OBJETIVO ESPECÍFICO 01:

Objetivo específico: Analizar el impacto del avance de la IA en las condiciones laborales y en el desarrollo económico, considerando cambios en la estructura del empleo y automatización de procesos.

4.2.1. IMPACTO DE LA INTELIGENCIA ARTIFICIAL EN EL ÁMBITO LABORAL

El avance acelerado de la inteligencia artificial y de los sistemas de automatización ha generado una transformación profunda en la estructura del trabajo humano a nivel global. A lo largo de la historia, los procesos productivos han experimentado múltiples etapas de mecanización y tecnificación; sin embargo, la incorporación de sistemas inteligentes capaces de aprender, adaptarse y tomar decisiones autónomas representa una ruptura cualitativa respecto a los modelos tradicionales de automatización.

Desde la Revolución Industrial del siglo XVIII hasta la introducción de las cadenas de montaje en el siglo XX, el desarrollo tecnológico ha estado asociado a procesos de

sustitución y reconversión laboral. No obstante, a diferencia de etapas anteriores, la inteligencia artificial no solo reemplaza tareas físicas repetitivas, sino también funciones cognitivas, analíticas y decisionales que históricamente fueron consideradas exclusivas del ser humano.

En el contexto contemporáneo, la IA se integra en sectores como la manufactura, la logística, la agricultura, los servicios financieros, la medicina, la educación y el sistema judicial, modificando las dinámicas de producción, organización y gestión del trabajo. Esta integración ha permitido incrementos significativos en eficiencia, reducción de costos y optimización de procesos, pero también ha generado nuevas formas de precarización, desplazamiento laboral y dependencia tecnológica.

Asimismo, el desarrollo de sistemas autónomos y agentes inteligentes ha favorecido la aparición de entornos productivos cada vez más deshumanizados, en los cuales la intervención directa del trabajador se reduce progresivamente. En muchos casos, el rol humano se limita a funciones de supervisión, mantenimiento o validación, lo que transforma sustancialmente la naturaleza del empleo y las competencias requeridas.

En este escenario, el impacto de la inteligencia artificial en el ámbito laboral no puede ser comprendido únicamente desde una perspectiva técnica o económica, sino también desde un enfoque social, ético y filosófico. La automatización creciente plantea interrogantes sobre la dignidad del trabajo, la redistribución de oportunidades, la formación profesional, la estabilidad económica y el papel futuro del ser humano en los sistemas productivos.

Por ello, el presente apartado analiza de manera sistemática las principales transformaciones generadas por la inteligencia artificial en el mundo laboral, identificando, en primer lugar, aquellas profesiones que ya han sido ampliamente automatizadas; en segundo lugar, aquellas que se encuentran en proceso de deshumanización progresiva; y, finalmente, aquellas que, por razones técnicas, éticas o sociales, presentan una imposibilidad temporal de automatización completa.

Este análisis se sustenta en evidencia empírica, estudios científicos, informes institucionales y reportes periodísticos especializados, con el propósito de ofrecer una visión crítica, documentada y contextualizada del fenómeno, evitando tanto el determinismo tecnológico como el rechazo infundado al progreso.

Tabla 05: Aceleración histórica de la automatización antes y después de la inteligencia artificial

Período histórico	Tipo de automatización	Características principales	Ejemplos / Contexto	Impacto laboral	Referencias
Antes de IA (hasta aprox. 2010)	Automatización mecánica y programada	Reemplazo de tareas repetitivas con mecanismos y control lógico programable; no comprende aprendizaje ni adaptación.	Robótica industrial clásica, líneas de ensamblaje, PLCs.	Reducción de trabajo manual físico; reconfiguración de industrias manufactureras	Tecnología industrial tradicional (IEBS, 2025)
Primera ola de “informática” automatizada (1980–2010)	Automatización por software y sistemas expertos	Software que agiliza procesos administrativos y cálculo, pero sin aprendizaje real ni adaptación autónoma.	ERP, CRM, sistemas de producción automatizada sin IA.	Mejora de eficiencia sin cambios drásticos de tareas cognitivas; consolidación de manufactura digital.	McKinsey – lecciones históricas
Inicio de IA aplicada (2010–2015)	Automatización con machine learning inicial	IA aplicada en tareas específicas: reconocimiento de voz/imágenes, optimización logística básica.	Redes neuronales profundas, primeras aplicaciones productivas.	Comienzo de automatización de tareas cognitivas sencillas.	Investigación sobre automatización y empleo (Ismail, 2022; Rahman, 2023)
IA avanzada y aprendizaje	Automatización con IA generalizada	Uso de IA para toma de decisiones,	AlphaGo, asistentes inteligentes,	Cambios en empleo de oficinas y	Automatización y empleo 4IR

Período histórico	Tipo de automatización	Características principales	Ejemplos / Contexto	Impacto laboral	Referencias
profundo (2015–2020)		clasificación de grandes datos, asistentes inteligentes, RPA (automatización robótica de procesos).	robotización colaborativa.	producción; alto impacto en tareas repetitivas con supervisión humana.	(Meindl et al., 2021)
IA generativa y autónoma (2020–2025)	Automatización cognitiva profunda	Modelos que aprenden, razonan, generan lenguaje y toman decisiones en contexto variado.	GPT-3/4, agentes autónomos, sistemas multimodales.	Desplazamiento de tareas analíticas, creativas, de servicio y apoyo; aceleración del cambio laboral.	Evolución del desplazamiento de empleo (Subaveerapan diyan & Shimray, 2024)
Tendencias recientes (2025+)	Automatización ampliada en todos los sectores	IA integrada con robótica, sistemas predictivos y plataformas automatizadas de atención.	Predicciones de reemplazo masivo de empleos administrativo s y de servicio.	Proyección de millones de puestos afectados; necesidad de adaptación institucional.	Informes contemporáneos sobre empleo y riesgo (The Guardian, 2025; Andrew Yang, 2025)

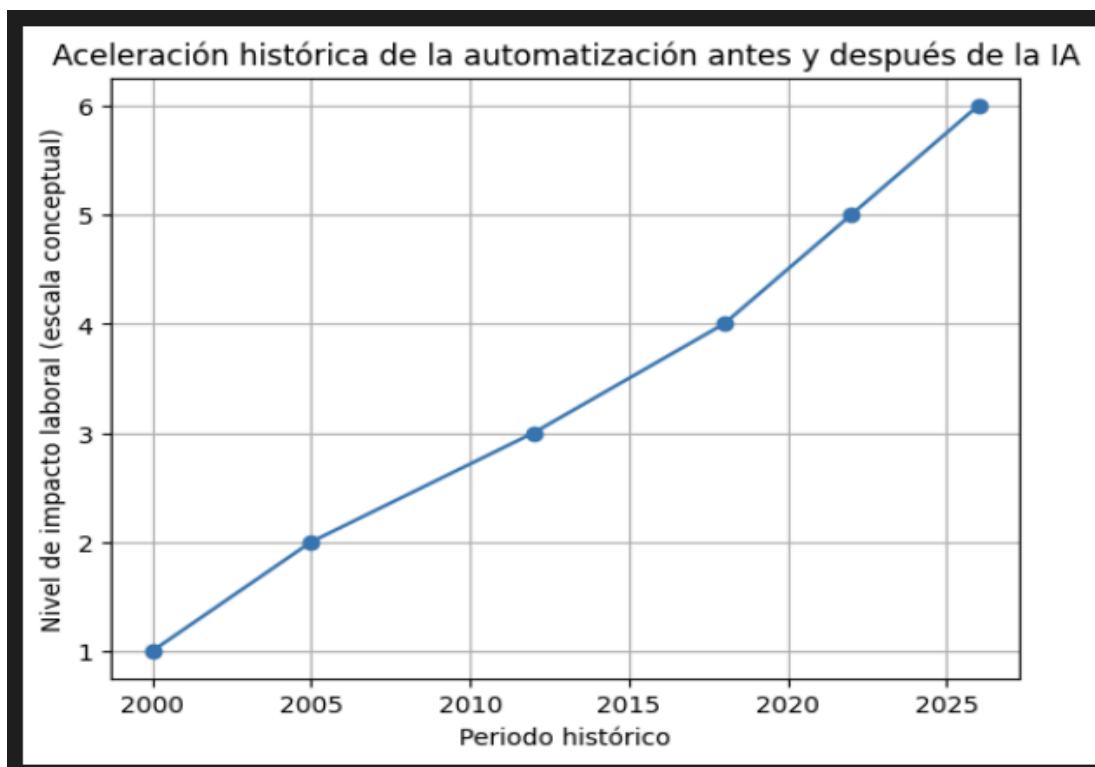


Figura 04: Nivel de impacto laboral

La Figura 4 evidencia que **la automatización no empezó con la IA**, sino que tiene antecedentes largos en la historia industrial y de software.

Sin embargo, **la aparición de IA marca un cambio cualitativo y cuantitativo**: la automatización pasa de sistemas pre programados a procesos adaptativos que pueden aprender, tomar decisiones y reemplazar tareas cognitivas.

A partir de **2015–2020**, con el aprendizaje profundo y la IA generativa, la automatización empieza a impactar **tareas que antes solo los humanos podían ejecutar**: lenguaje, análisis de datos, atención, clasificación, etc.

La literatura científica indica que este cambio no solo afecta trabajos manuales, sino también aquellos basados en conocimiento, lo que acelera la transformación del mercado laboral.

Estudios modernos muestran que ya hay sectores con fuerte impacto (servicio, producción, administración) y proyecciones que sugieren una **expansión acelerada de la automatización impulsada por IA** en la próxima década.

4.2.2. ANTECEDENTES HISTÓRICOS DE LA AUTOMATIZACIÓN DEL TRABAJO

La automatización del trabajo humano no es una novedad reciente, sino un proceso histórico que ha atravesado múltiples etapas desde la mecanización inicial hasta los sistemas inteligentes contemporáneos. La sustitución de trabajo manual por máquinas comenzó con la **Revolución Industrial del siglo XVIII**, cuando la energía mecánica y las máquinas de vapor desplazaron labores artesanales, sentando las bases para la producción en serie y transformando la organización social del trabajo (Revolución Industrial, siglo XVIII).

En el siglo XX, la introducción de la **línea de montaje moderna** revolucionó la industria manufacturera. Implementada por primera vez con gran éxito por Henry Ford a partir de 1913, esta técnica permitió producir vehículos con mayor eficiencia, reduciendo significativamente la mano de obra directa necesaria para cada unidad ensamblada (Wikipedia, 2023). De igual manera, en 1961 la planta de General Motors en New Jersey integró el robot **Unimate**, marcando el inicio de la robótica industrial al reemplazar actividades peligrosas o repetitivas en la cadena de producción (Wikipedia, 2023).

Hacia finales del siglo XX y comienzos del XXI, la robótica industrial y la automatización avanzada se consolidaron como pilares de múltiples sectores. Empresas y países líderes incrementaron el uso de robots en fábricas y plantas de producción, dando lugar a cifras récord de densidad robótica en sectores como la automoción, electrónica y maquinaria pesada (IFR Reports, 2023). Esta tendencia ha sido especialmente pronunciada en economías como China, Japón y Corea del Sur, que lideran la adopción global de robots industriales.

La automatización no se limitó a la producción manufacturera. En sectores tradicionalmente manuales, como la **construcción**, surgieron soluciones tecnológicas que incorporan maquinaria autónoma y robots especializados. Por ejemplo, sistemas de robótica de construcción diseñados para levantar muros, manipular materiales pesados y realizar inspecciones automatizadas han comenzado a ser utilizados en proyectos innovadores en Estados Unidos, Europa y Asia, lo que representa el avance de la

automatización más allá de las fábricas hacia obras civiles (Automate.org, 2025; Wikipedia Construction Robots, 2025). Asimismo, proyectos en Dubái han impulsado iniciativas para **construir viviendas utilizando únicamente sistemas automatizados**, con el objetivo de demostrar la viabilidad de la automatización total en obras civiles (AS.com, 2026)

La **agricultura y la ganadería** también han experimentado automatización significativa. Desde iniciativas agrícolas como el proyecto **Hands Free Hectare** en el Reino Unido, que demostró la posibilidad de operar un ciclo de cultivo de manera totalmente autónoma sin intervención humana directa (Harper Adams University, 2025), hasta granjas modernas donde sistemas robotizados realizan el ordeño de bovinos con control remoto, reduciendo la necesidad de mano de obra humana intensiva (Cadena SER, 2025).

Estos antecedentes revelan que la automatización ha trascendido a lo largo del tiempo desde simples máquinas de producción hacia sistemas robotizados capaces de ejecutar tareas complejas en distintos entornos productivos. La inteligencia artificial representa el siguiente salto en esta trayectoria histórica: no solo se trata de implementar máquinas que ejecuten funciones predefinidas, sino de sistemas que **aprenden, se adaptan y pueden tomar decisiones autónomas** en tiempo real, lo que plantea nuevas formas de reorganización del trabajo y de redefinición de las competencias humanas en el empleo contemporáneo.

4.2.3. PROFESIONES ALTAMENTE AUTOMATIZADAS

- Operador de línea de ensamblaje industrial

La automatización en cadenas de producción industrial ha transformado radicalmente la función tradicional del operador de línea de ensamblaje, cuyo trabajo consistía históricamente en tareas manuales de montaje, soldadura y ajuste de piezas en productos manufacturados. Con el advenimiento de robots industriales desde mediados del siglo XX, la intervención humana ha disminuido en múltiples aspectos de la producción.

Por ejemplo, desde la década de 1990 muchas plantas de **Toyota** implementaron robots para actividades como soldadura de carrocerías y pintura, reduciendo la dependencia de trabajadores en tareas peligrosas o físicamente exigentes (Robotics & Automation Magazine, 2025). De forma similar, **Tesla** ha integrado líneas de fabricación altamente robotizadas en su planta de Fremont (EE. UU.) y en Giga Shanghai (China), donde brazos robóticos se encargan del ensamblaje y supervisión de vehículos completos desde mediados de la década del 2010, aumentando la productividad aunque también reconfigurando el rol humano hacia mantenimiento y supervisión técnica.

En el Perú, aunque la automatización robótica industrial aún no es tan extendida en la manufactura pesada, *proveedores de soluciones industriales* y empresas tecnológicas han comenzado a implementar robots colaborativos (cobots) asociados a tareas repetitivas y de inspección en procesos de producción ligera y talleres especializados, marcando un paso inicial hacia prácticas de producción instrumentadas por robots inteligentes (Smartlog Group establece presencia en Perú, 2022). La automatización de estas funciones ha llevado a que el perfil tradicional de ensamblador se complemente con habilidades técnicas para operar, programar y mantener sistemas robotizados.

- **Operador de almacén y logística**

La automatización de las operaciones logísticas ha sustituido profundamente el rol del operador de almacén tradicional. Este cargo, que históricamente implicaba mano de obra física para organizar, mover y preparar mercancía, ha sido progresivamente asumido por robots autónomos y sistemas inteligentes.

Empresas globales como **Amazon** han desplegado más de **un millón de robots en sus centros logísticos**, lo cual ha acercado el número de máquinas al de trabajadores humanos en esos espacios y ha permitido que hasta el **75 % de las operaciones logísticas sean automatizadas** (La automatización en Amazon está a punto de cruzar una línea histórica, 2025). Estos robots realizan selección, transporte interno y clasificación de mercancías sin intervención humana directa en muchos casos,

revolucionando la ejecución de tareas logísticas (La automatización robótica transforma el trabajo en los centros logísticos, 2025).

En el contexto peruano, cadenas como **Tottus** iniciaron la automatización de su centro logístico en **julio de 2023**, con inversiones cercanas a US\$ 30 millones para implementar sistemas robotizados de almacenamiento y transporte interno, lo que ha aumentado la capacidad operativa y ha reorganizado la fuerza laboral hacia roles de supervisión y gestión técnica (Tottus inaugura su primer almacén automatizado en Perú, 2023; Tottus centraliza la distribución robotizando su centro logístico, 2024). Asimismo, la empresa **Tai Loy** inauguró en **diciembre de 2025** un **centro de distribución automatizado** en Lima con robots autónomos que elevan la eficiencia del procesamiento de pedidos y reducen la intervención manual (Tai Loy inaugura centro de distribución con robots, 2025).

- **Clasificador y despachador de paquetería**

La automatización de la función de clasificación y despacho de paquetería representa una de las transformaciones más importantes en los sistemas logísticos modernos, derivadas del crecimiento exponencial del comercio electrónico y las exigencias de velocidad de entrega. Tradicionalmente, esta ocupación consistía en que trabajadores manuales, apoyados por sistemas de cinta transportadora, clasificaban paquetes según criterios como código postal, destino o tipo de envío, antes de ser cargados para su distribución final.

En el contexto internacional, las empresas líderes en logística comenzaron a automatizar estos procesos en la década de 2010, con la introducción de sistemas avanzados de clasificación robótica impulsados por inteligencia artificial (IA) y visión computarizada. Por ejemplo, **FedEx** incorporó desde 2022 brazos robóticos inteligentes en su centro de clasificación en **Singapur**, capaces de leer códigos de barras y clasificar hasta **1 000 paquetes por hora** sin intervención humana directa, acelerando drásticamente la eficiencia operativa y reduciendo errores en la determinación de rutas de entrega (FedEx presenta brazo robótico de clasificación AI, 2022; FedEx Rise of the Sorting Robots, 2025).

De forma complementaria, la multinacional **DHL** ha sido pionera en automatización de clasificación desde principios de la década de 2020, con pilotos de robots *DoraSorter* integrados en centros de distribución de Estados Unidos que pudieron clasificar más de **1 000 paquetes por hora con tasas de error cercanas a cero**, superando ampliamente las capacidades manuales tradicionales (DHL prueba robots de clasificación, 2022).

Estos sistemas robóticos no solo han reducido la dependencia de trabajo humano en tareas repetitivas de interpretación visual y decisión, sino que han dado paso a funciones de **supervisión técnica, programación y mantenimiento de equipos automatizados**, reconfigurando la base de habilidades requeridas en los centros logísticos modernos.

Actualmente, las innovaciones en clasificación de paquetería se extienden hacia sistemas autónomos altamente integrados con IA que incorporan tecnologías de visión artificial y algoritmos predictivos para anticipar flujos de paquetes, optimizar rutas de sorting y equilibrar cargas de trabajo sin supervisión humana directa (Robótica en clasificación de paquetería, s. f.).

Aunque en Perú la automatización de clasificación de paquetería no se ha documentado aún con proyectos robóticos de gran escala debido a la relativa novedad de esta tecnología en centros logísticos locales, **empresas de mensajería y comercio electrónico que operan en el país han adoptado gradualmente sistemas de clasificación asistida por software y tecnologías automáticas de código de barras**, lo que indica una transición hacia una mayor automatización en el procesamiento de envíos (Inteligencia artificial en logística y distribución, s. f.).

En conjunto, la automatización de la clasificación y el despacho de paquetería constituye un ejemplo claro de cómo la inteligencia artificial y la robótica están transformando ocupaciones antes dominadas por esfuerzo humano, reemplazando tareas manuales complejas por sistemas más veloces, precisos y resistentes al cansancio, aunque implicando una reconfiguración de habilidades laborales hacia lo técnico e interpretativo.

- Técnico de soldadura industrial

La profesión de técnico de soldadura industrial ha sido una de las más tempranamente impactadas por los procesos de automatización y robotización dentro del sector manufacturero. Tradicionalmente, este oficio requería alta destreza manual, experiencia acumulada y precisión técnica para realizar uniones metálicas en sectores como la industria automotriz, naval, aeroespacial y de maquinaria pesada.

Desde finales del siglo XX, las grandes industrias comenzaron a incorporar sistemas de soldadura robotizada, especialmente en líneas de ensamblaje automotriz. Empresas como Toyota y General Motors iniciaron el uso sistemático de robots de soldadura en sus plantas desde la década de 1990, logrando estandarizar procesos, reducir errores humanos y aumentar la productividad (International Federation of Robotics, 2022; General Motors, 2021). Posteriormente, compañías como Tesla y Foxconn profundizaron este modelo, implementando sistemas automatizados capaces de realizar soldaduras de alta precisión sin intervención directa de operarios humanos en amplias fases del proceso productivo (Reuters, 2023; Foxconn Annual Report, 2022).

En el ámbito internacional, se estima que más del 70 % de las operaciones de soldadura en plantas automotrices avanzadas se encuentran actualmente automatizadas, especialmente en Asia, Europa y Norteamérica (IFR, 2023). Los robots industriales permiten realizar soldaduras continuas, uniformes y resistentes, reduciendo riesgos laborales asociados a quemaduras, inhalación de gases tóxicos y fatiga física.

Además, el desarrollo de sistemas de soldadura inteligente, integrados con visión artificial y aprendizaje automático, ha permitido que los robots identifiquen defectos, ajusten parámetros en tiempo real y optimicen la calidad del ensamblaje, consolidando un proceso cada vez más autónomo (ABB Robotics, 2024).

En el contexto latinoamericano, y particularmente en el Perú, el proceso de automatización también ha comenzado a consolidarse. Empresas del sector metalmecánico y automotriz han incorporado celdas robotizadas en sus procesos productivos. Por ejemplo, empresas como Ferreyros y Komatsu-Mitsui Maquinarias Perú

han implementado sistemas de soldadura automatizada en talleres industriales para maquinaria pesada, con el fin de mejorar la eficiencia y estandarización de sus procesos (Gestión, 2023; Andina, 2024).

Asimismo, algunas plantas de ensamblaje y metalurgia ubicadas en Lima y Arequipa han incorporado brazos robóticos para soldadura en estructuras metálicas, reduciendo progresivamente la necesidad de soldadores manuales en tareas repetitivas y de alto volumen (Semana Económica, 2024).

Este proceso ha generado una transformación profunda en el perfil profesional del técnico soldador. Mientras disminuye la demanda de soldadura manual tradicional, aumenta la necesidad de operadores de robots, programadores de celdas automatizadas y especialistas en mantenimiento industrial, desplazando el énfasis desde la habilidad manual hacia la competencia tecnológica (OECD, 2023).

En consecuencia, la profesión de técnico de soldadura industrial evidencia una transición avanzada hacia la automatización, en la cual gran parte de las tareas rutinarias ya son ejecutadas por sistemas robóticos, configurando un escenario de sustitución parcial del trabajo humano y redefinición del rol laboral en la industria contemporánea.

- **Trabajador agrícola mecanizado**

La actividad agrícola ha experimentado una de las transformaciones tecnológicas más profundas en las últimas décadas, especialmente mediante la incorporación progresiva de sistemas automatizados, maquinaria inteligente y tecnologías basadas en inteligencia artificial. Tradicionalmente, el trabajo agrícola dependía en gran medida de la fuerza humana y animal para labores como la siembra, el riego, la cosecha y el control de plagas, lo que implicaba altos niveles de esfuerzo físico y exposición a condiciones ambientales adversas.

Desde finales del siglo XX, países con alta productividad agrícola como Estados Unidos, Japón y los miembros de la Unión Europea comenzaron a implementar maquinaria mecanizada avanzada en la agricultura. Sin embargo, fue a partir de la década de 2010 cuando la incorporación de sistemas autónomos y tecnologías de precisión aceleró

significativamente la automatización del sector. Empresas como John Deere, CNH Industrial y AGCO desarrollaron tractores autónomos, sembradoras inteligentes y sistemas de fumigación automatizada capaces de operar con mínima intervención humana (John Deere, 2022; CNH Industrial, 2023).

Un hito relevante fue la presentación del tractor totalmente autónomo de John Deere en 2022, equipado con sensores LiDAR, visión artificial y algoritmos de aprendizaje automático, capaz de realizar labores agrícolas de forma continua sin conductor humano (Reuters, 2022). Asimismo, empresas como Blue River Technology y Trimble impulsaron el uso de sistemas de agricultura de precisión basados en IA, que permiten identificar malezas, optimizar el uso de fertilizantes y maximizar el rendimiento de cultivos (MIT Technology Review, 2021; Trimble, 2023).

A nivel global, se estima que para 2025 más del 60 % de las grandes explotaciones agrícolas en países desarrollados utilizan algún tipo de sistema automatizado o semiautónomo en sus procesos productivos (FAO, 2024). Estas tecnologías han permitido reducir costos operativos, minimizar el desperdicio de insumos y mejorar la sostenibilidad ambiental mediante un uso más eficiente de recursos.

En el ámbito latinoamericano, países como Brasil, Argentina y Chile han liderado la incorporación de agricultura inteligente, utilizando drones, sensores remotos y maquinaria automatizada para el monitoreo de cultivos y la gestión productiva (IICA, 2023). En Brasil, por ejemplo, grandes agroindustrias emplean flotas de tractores autónomos para la producción de soya y caña de azúcar, reduciendo significativamente la dependencia del trabajo manual (Reuters, 2023).

En el contexto peruano, el proceso de mecanización agrícola también ha avanzado de manera progresiva. Empresas agroexportadoras ubicadas en la costa norte y sur del país han incorporado sistemas automatizados en la siembra, riego y cosecha. Compañías como Camposol y Danper han implementado tecnologías de riego inteligente, drones para monitoreo fitosanitario y maquinaria automatizada para la recolección y clasificación de productos agrícolas (Gestión, 2023; Andina, 2024).

Asimismo, el Ministerio de Desarrollo Agrario y Riego (MIDAGRI) ha promovido programas de modernización tecnológica para pequeños y medianos productores, impulsando el uso de maquinaria mecanizada, sistemas de riego automatizado y plataformas digitales de gestión agrícola (MIDAGRI, 2023).

Este proceso de automatización ha generado cambios estructurales en el empleo rural. Mientras disminuye la demanda de mano de obra no calificada para labores repetitivas, aumenta la necesidad de técnicos en maquinaria agrícola, operadores de sistemas inteligentes y especialistas en agricultura digital (OECD, 2024). De este modo, el trabajador agrícola tradicional se transforma progresivamente en un gestor tecnológico del proceso productivo.

En consecuencia, el trabajador agrícola mecanizado representa un claro ejemplo de sustitución parcial del trabajo humano mediante sistemas autónomos e inteligentes, consolidando una transición hacia modelos productivos altamente tecnificados, con implicaciones profundas para la organización del trabajo rural y la seguridad alimentaria global.

- **Operador de ordeño ganadero**

La actividad ganadera, particularmente en el sector lechero, ha experimentado un proceso sostenido de automatización impulsado por el desarrollo de sistemas robóticos, sensores inteligentes y plataformas de gestión basadas en inteligencia artificial. Tradicionalmente, el ordeño de ganado requería una intervención humana constante, tanto para la manipulación directa de los animales como para el control sanitario, la higiene y la supervisión de la producción.

A partir de la década de 1990, comenzaron a difundirse en Europa los primeros sistemas de ordeño mecánico automatizado. No obstante, fue desde inicios del siglo XXI cuando surgieron los sistemas de ordeño robótico completamente autónomos, capaces de identificar a cada animal, evaluar su estado de salud y realizar el proceso sin intervención directa del operador humano (DeLaval, 2001; Lely, 2005).

Empresas líderes como Lely (Países Bajos), DeLaval (Suecia) y GEA Group (Alemania) desarrollaron robots de ordeño equipados con visión artificial, sensores biométricos y software inteligente. Estos sistemas permiten que las vacas ingresen voluntariamente a las estaciones de ordeño, donde son identificadas mediante chips RFID, limpiadas automáticamente y ordeñadas de forma precisa y estandarizada (Lely, 2022; DeLaval, 2023).

Un estudio de la Universidad de Wageningen reveló que, para 2020, más del 50 % de las granjas lecheras en Países Bajos utilizaban sistemas de ordeño robótico, mientras que en países como Suecia y Dinamarca la cifra superaba el 60 % (Wageningen University, 2020). Para 2025, la International Dairy Federation estimó que los sistemas automatizados estaban presentes en más del 35 % de las explotaciones lecheras tecnificadas a nivel mundial (IDF, 2024).

Además del ordeño, estos sistemas recopilan datos en tiempo real sobre la producción de leche, la salud del animal, la alimentación y los ciclos reproductivos, lo que permite a los productores optimizar la gestión del ganado mediante plataformas basadas en inteligencia artificial (GEA, 2023). Esta integración ha transformado al operador tradicional en un supervisor tecnológico encargado del monitoreo digital del sistema.

En América Latina, países como Argentina, Uruguay y Chile han liderado la incorporación de robots de ordeño en establecimientos de mediana y gran escala. En Uruguay, por ejemplo, más de 200 tambos habían implementado sistemas robóticos para 2022, mejorando la productividad y reduciendo la dependencia de mano de obra especializada (El Observador, 2022; INIA Uruguay, 2023).

En el Perú, aunque la adopción ha sido más gradual, algunas empresas ganaderas ubicadas en regiones como Cajamarca, Arequipa y Lima han comenzado a implementar sistemas de ordeño automatizado y monitoreo digital del ganado. Empresas como Gloria y Laive han incorporado tecnologías de gestión ganadera, sensores de producción y sistemas semiautomatizados en sus cadenas de abastecimiento lechero (Gestión, 2023; Andina, 2024). Asimismo, programas del Ministerio de Desarrollo Agrario y Riego han

promovido la modernización tecnológica del sector lechero mediante créditos y asistencia técnica (MIDAGRI, 2023).

La automatización del ordeño ha generado efectos significativos en el empleo rural. Si bien reduce la demanda de ordeñadores tradicionales, incrementa la necesidad de técnicos en mantenimiento robótico, analistas de datos pecuarios y gestores de sistemas digitales agropecuarios (FAO, 2024). Además, ha contribuido a mejorar las condiciones laborales, reduciendo la carga física y los riesgos sanitarios asociados al trabajo manual. En este contexto, el operador de ordeño ganadero representa un caso emblemático de sustitución progresiva del trabajo humano por sistemas inteligentes, donde la función principal del trabajador se desplaza desde la ejecución directa hacia la supervisión tecnológica y la toma de decisiones basada en datos.

- **Digitador de datos (Data Entry)**

La profesión de digitador de datos ha sido históricamente una de las más extendidas en entornos administrativos, financieros, educativos y gubernamentales. Tradicionalmente, este rol consistía en la transcripción manual de información desde documentos físicos hacia sistemas digitales, bases de datos y plataformas informáticas, requiriendo precisión, concentración y largas jornadas de trabajo repetitivo.

Desde finales del siglo XX, con la masificación de los sistemas informáticos, esta actividad comenzó a transformarse progresivamente. Sin embargo, el proceso de automatización se aceleró de manera significativa a partir de la década de 2010, con el desarrollo de tecnologías de Reconocimiento Óptico de Caracteres (OCR), procesamiento de lenguaje natural e inteligencia artificial aplicada a la gestión documental (IBM, 2012; ABBYY, 2015).

Empresas tecnológicas como ABBYY, Google, Microsoft e IBM desarrollaron plataformas capaces de digitalizar documentos, interpretar textos manuscritos o impresos y estructurar automáticamente la información sin intervención humana directa. Sistemas como Google Cloud Vision OCR, ABBYY FineReader y IBM Watson Discovery permiten

convertir millones de documentos físicos en bases de datos organizadas en tiempo real (Google, 2023; ABBYY, 2024; IBM, 2024).

A nivel internacional, instituciones financieras, aseguradoras y organismos gubernamentales han reemplazado progresivamente los departamentos de digitación manual por sistemas automatizados. Por ejemplo, bancos como JPMorgan Chase y HSBC implementaron desde 2017 plataformas de automatización documental que procesan contratos, formularios y registros sin digitadores humanos, reduciendo los tiempos de procesamiento en más del 70 % (JPMorgan, 2018; McKinsey, 2019).

Asimismo, organismos públicos como el Internal Revenue Service (IRS) de Estados Unidos y el National Health Service (NHS) del Reino Unido han adoptado sistemas de captura inteligente de datos para gestionar expedientes fiscales y clínicos, eliminando en gran medida la necesidad de transcripción manual (IRS, 2021; NHS, 2022).

En el sector empresarial, compañías como Amazon, Walmart y Alibaba utilizan sistemas de automatización documental para registrar órdenes, facturas, inventarios y reportes operativos, integrando OCR con inteligencia artificial predictiva para minimizar errores humanos (Alibaba Cloud, 2023; Amazon Web Services, 2024).

En América Latina, países como Brasil, México y Chile han avanzado notablemente en la digitalización administrativa. En Brasil, el sistema e-Social permite la automatización del registro laboral y tributario mediante plataformas digitales inteligentes, reduciendo la necesidad de digitadores en instituciones públicas (Gobierno de Brasil, 2022).

En el Perú, la transformación digital del Estado ha impulsado la automatización del procesamiento documental en diversas entidades. La Superintendencia Nacional de Aduanas y de Administración Tributaria (SUNAT) implementó sistemas de declaración electrónica y captura automática de datos tributarios, disminuyendo significativamente los procesos de digitación manual (SUNAT, 2021). Asimismo, el Poder Judicial y la RENIEC han incorporado sistemas de digitalización inteligente para expedientes judiciales y registros civiles (RENIEC, 2023; PJ Perú, 2024).

En el sector privado peruano, bancos como BCP, BBVA Perú e Interbank han adoptado plataformas de gestión documental automatizada para el procesamiento de solicitudes, contratos y operaciones financieras, reemplazando gran parte de las tareas tradicionales de ingreso manual de datos (Semana Económica, 2023; Gestión, 2024).

Como consecuencia de esta transformación, la función del digitador tradicional ha sido progresivamente desplazada por sistemas automáticos de captura, clasificación y validación de información. No obstante, han surgido nuevos perfiles laborales asociados a la supervisión de datos, control de calidad, ciberseguridad documental y gestión de sistemas inteligentes.

De este modo, el digitador de datos constituye uno de los ejemplos más claros de sustitución laboral por inteligencia artificial, donde una actividad eminentemente repetitiva y estandarizable ha sido absorbida casi por completo por tecnologías automatizadas, evidenciando un proceso avanzado de deshumanización operativa en los entornos administrativos contemporáneos.

- **Operador de call center básico**

La profesión de operador de call center ha sido, durante décadas, uno de los pilares de los servicios de atención al cliente, soporte técnico, ventas telefónicas y gestión de reclamos en empresas públicas y privadas. Tradicionalmente, este rol implicaba la interacción directa entre un agente humano y el usuario, requiriendo habilidades comunicativas, manejo emocional, memorización de protocolos y disponibilidad permanente.

A partir de la década de 2010, con el avance del procesamiento del lenguaje natural, el reconocimiento de voz y la inteligencia artificial conversacional, este sector inició un proceso acelerado de automatización. El desarrollo de chatbots y voicebots permitió a las empresas responder consultas frecuentes, gestionar solicitudes básicas y realizar operaciones rutinarias sin intervención humana continua (Gartner, 2018; IBM, 2019).

Empresas como Google, Amazon, Microsoft e IBM desarrollaron plataformas de atención automatizada basadas en IA. Sistemas como Google Dialogflow, Amazon Connect,

Microsoft Azure Bot Service e IBM Watson Assistant permiten diseñar asistentes virtuales capaces de mantener conversaciones fluidas, identificar intenciones del usuario y ejecutar procesos administrativos en tiempo real (Google, 2023; Amazon, 2024; Microsoft, 2023; IBM, 2024).

A nivel internacional, compañías como Vodafone, Bank of America, HSBC y American Express implementaron sistemas automatizados desde 2016 para atender millones de consultas diarias sin agentes humanos. Por ejemplo, el asistente virtual “Erica” de Bank of America superó los mil millones de interacciones en 2023, resolviendo de forma autónoma consultas financieras, bloqueos de tarjetas y operaciones básicas (Bank of America, 2023).

Asimismo, empresas de telecomunicaciones como Telefónica y Orange integraron bots conversacionales en sus plataformas digitales, reduciendo hasta en un 40 % el volumen de llamadas atendidas por operadores humanos (Telefónica, 2021; Orange, 2022).

En el sector del comercio electrónico, plataformas como Amazon, Alibaba y Mercado Libre utilizan sistemas automatizados para gestionar devoluciones, reclamos y seguimiento de pedidos, combinando chatbots con análisis predictivo de comportamiento del cliente (Alibaba Cloud, 2023; AWS, 2024).

En América Latina, la automatización de call centers se intensificó a partir de 2019. En México, Brasil y Colombia, grandes empresas de telecomunicaciones y banca digital adoptaron sistemas de atención virtual para reducir costos operativos y mejorar la disponibilidad del servicio (BBVA, 2022; Telefónica LATAM, 2023).

En el Perú, diversas instituciones públicas y privadas han incorporado asistentes virtuales para atención ciudadana y comercial. La SUNAT implementó el chatbot “SIA” para resolver consultas tributarias, mientras que EsSalud desarrolló sistemas automatizados para citas y orientación médica (SUNAT, 2022; EsSalud, 2023). Asimismo, bancos como BCP, Interbank y BBVA Perú utilizan asistentes virtuales en sus aplicaciones móviles para atención básica al cliente (Gestión, 2024).

En el ámbito empresarial, compañías de outsourcing como Atento, Teleperformance y Konecra han integrado sistemas híbridos donde los bots atienden consultas iniciales y solo derivan casos complejos a operadores humanos, reduciendo progresivamente la demanda de personal tradicional (Teleperformance, 2023; Atento, 2024).

Como resultado, el perfil del operador de call center básico ha sufrido una transformación estructural. Las tareas repetitivas, estandarizadas y predecibles han sido transferidas a sistemas automatizados, mientras que los agentes humanos se concentran en casos excepcionales, gestión emocional compleja y resolución de conflictos avanzados.

Este proceso refleja un claro ejemplo de automatización funcional avanzada, donde la inteligencia artificial no solo reemplaza tareas específicas, sino que redefine integralmente los modelos de atención al cliente, consolidando una transición hacia servicios mayoritariamente digitalizados y deshumanizados en su nivel operativo básico.

- **Cajero bancario tradicional**

La profesión de cajero bancario tradicional ha sido históricamente uno de los pilares del sistema financiero moderno. Durante gran parte del siglo XX y comienzos del siglo XXI, los cajeros cumplían funciones esenciales como el retiro y depósito de efectivo, pago de servicios, transferencia de fondos, consulta de saldos y orientación directa al cliente. Este rol requería presencia física, verificación manual de datos y supervisión constante.

El proceso de automatización de esta profesión comenzó con la introducción de los cajeros automáticos (ATM) en la década de 1970, pero se intensificó notablemente a partir de los años 2000 con la expansión de la banca digital, las aplicaciones móviles y los sistemas de autenticación electrónica (BIS, 2018; World Bank, 2020).

A nivel internacional, bancos como JPMorgan Chase, Citibank, HSBC y BBVA impulsaron desde 2010 una transición progresiva hacia modelos de banca sin sucursales físicas tradicionales, priorizando plataformas digitales y servicios automatizados. JPMorgan, por ejemplo, reportó en 2022 que más del 70 % de sus operaciones minoristas se realizaban sin intervención de cajeros humanos (JPMorgan, 2022).

En China, el proceso fue aún más acelerado. Entidades como ICBC y China Construction Bank desarrollaron sucursales “inteligentes” desde 2017, con sistemas de reconocimiento facial, terminales automatizados y atención virtual, reduciendo drásticamente la necesidad de personal operativo (China Banking Association, 2019).

Asimismo, bancos digitales como Nubank, Revolut, Monzo y N26 operan casi exclusivamente sin cajeros físicos, gestionando todas sus operaciones mediante inteligencia artificial, algoritmos de riesgo y atención virtual automatizada (McKinsey, 2021; Financial Times, 2023).

La consolidación de tecnologías como:

- Banca móvil
- Firmas digitales
- Verificación biométrica
- Chatbots financieros
- Sistemas antifraude con IA

ha permitido que funciones tradicionalmente realizadas por cajeros sean ejecutadas de forma automática, segura y continua.

En América Latina, este proceso se aceleró desde 2015 con la expansión de la banca digital. Bancos como Santander, BBVA y Itaú redujeron significativamente el número de ventanillas físicas, priorizando la atención digital y los cajeros inteligentes (BBVA Research, 2021; CEPAL, 2022).

En el Perú, el sector bancario ha seguido una tendencia similar. Entidades como BCP, Interbank, Scotiabank y BBVA Perú impulsaron desde 2018 la digitalización masiva de servicios financieros mediante aplicaciones móviles, cajeros multifunción y agentes automatizados (Gestión, 2023; SBS, 2024).

El BCP, por ejemplo, reportó que más del 80 % de sus transacciones minoristas se realizan actualmente por canales digitales, reduciendo progresivamente la dependencia del servicio presencial en ventanilla (BCP, 2023). Interbank y BBVA Perú también han

implementado agencias digitales con mínima presencia humana, priorizando kioscos electrónicos y atención virtual.

Paralelamente, la expansión de billeteras digitales como Yape, Plin, Tunki y Lukita ha contribuido a disminuir la necesidad de interacción con cajeros tradicionales, facilitando pagos, transferencias y recargas sin intermediación bancaria directa (El Comercio, 2024).

Como consecuencia, el perfil del cajero bancario ha sido transformado estructuralmente.

Las tareas rutinarias han sido absorbidas por sistemas automatizados, mientras que el personal humano se concentra en asesoría financiera, gestión de créditos complejos y atención especializada.

Este proceso evidencia una automatización consolidada del servicio bancario básico, donde la inteligencia artificial, la digitalización y la infraestructura financiera inteligente han reducido significativamente la centralidad del cajero tradicional, configurando un nuevo modelo de banca predominantemente virtual.

- **Cajero de supermercado tradicional**

La profesión de cajero de supermercado ha sido históricamente uno de los empleos más representativos del sector comercio y servicios. Durante décadas, estos trabajadores fueron responsables del registro manual de productos, cobro en efectivo o tarjeta, verificación de precios y atención directa al cliente. Este rol exigía presencia constante, rapidez operativa y habilidades básicas de atención al público.

El proceso de automatización de esta profesión se inició con la introducción de lectores de códigos de barras en la década de 1970, pero se aceleró significativamente a partir de los años 2000 con el desarrollo de sistemas de autopago (self-checkout), reconocimiento de productos mediante visión artificial y tiendas sin personal humano visible (NCR, 2019; McKinsey, 2020).

A nivel internacional, empresas como Walmart, Carrefour, Tesco y Costco comenzaron a implementar terminales de autopago desde mediados de la década del 2000, permitiendo a los clientes registrar y pagar sus productos sin intervención directa de un cajero humano (BBC, 2019). Walmart, por ejemplo, reportó que para 2022 más del 50 % de sus

transacciones en Estados Unidos se realizaban mediante sistemas automatizados (Walmart Annual Report, 2022).

Un hito fundamental en esta evolución fue el lanzamiento de Amazon Go en 2018, considerado el primer modelo de tienda totalmente automatizada a gran escala. Estas tiendas utilizan sensores, cámaras e inteligencia artificial para detectar qué productos toma el cliente y cobrar automáticamente al salir del local, eliminando por completo la figura del cajero (Amazon, 2018; MIT Technology Review, 2019).

De manera paralela, empresas tecnológicas como NCR, Toshiba Commerce y Diebold Nixdorf desarrollaron sistemas avanzados de punto de venta automatizado, integrando biometría, pagos móviles y análisis predictivo del comportamiento de compra (NCR, 2021).

En Asia, cadenas como Alibaba Hema, JD.com y 7-Eleven Japón implementaron desde 2017 supermercados inteligentes con reconocimiento facial y pago automatizado, consolidando modelos comerciales altamente digitalizados (South China Morning Post, 2020; Nikkei Asia, 2022).

En América Latina, la adopción del autopago se intensificó desde 2015. Cadenas como Walmart México, Cencosud, Falabella y Grupo Éxito incorporaron progresivamente terminales de autoservicio en sus tiendas (Forbes México, 2021; América Retail, 2022).

En el Perú, este proceso se consolidó a partir de 2019. Supermercados como Wong, Metro, Plaza Vea y Tottus implementaron sistemas de self-checkout en sus principales locales, permitiendo a los clientes realizar pagos autónomos sin asistencia directa (Gestión, 2022; El Comercio, 2023).

Asimismo, tiendas de conveniencia como Tambo+ y Oxxo Perú han incorporado terminales inteligentes, pagos sin contacto y sistemas de gestión automatizada, reduciendo progresivamente la necesidad de personal en caja (Semana Económica, 2023).

La expansión de pagos digitales mediante Yape, Plin, Apple Pay y Google Pay ha reforzado este proceso, facilitando transacciones sin interacción humana directa (Banco Central de Reserva del Perú, 2024).

Como consecuencia, el perfil laboral del cajero tradicional ha experimentado una transformación estructural. Las tareas repetitivas de registro y cobro han sido asumidas por sistemas automatizados, mientras que el personal humano se orienta cada vez más a funciones de supervisión, atención especializada y resolución de incidencias.

Este proceso refleja una automatización consolidada del comercio minorista, donde la inteligencia artificial, la visión computacional y los sistemas de pago digital han redefinido la experiencia de compra, reduciendo significativamente la centralidad del cajero humano en los entornos comerciales modernos.

- **Redactor de noticias simples**

La profesión de redactor de noticias simples ha sido tradicionalmente una función central en el periodismo informativo, especialmente en áreas como resultados deportivos, informes financieros, reportes meteorológicos, estadísticas económicas y resúmenes bursátiles. Este tipo de contenido se caracteriza por su estructura repetitiva, basada en datos objetivos, formatos estandarizados y rápida publicación, lo que lo convierte en un campo altamente susceptible a la automatización.

El proceso de automatización de esta labor comenzó a consolidarse a partir de 2010 con el desarrollo de sistemas de generación automática de lenguaje natural (Natural Language Generation, NLG), capaces de transformar bases de datos estructuradas en textos periodísticos comprensibles (Graefe, 2016; Diakopoulos, 2019).

A nivel internacional, uno de los primeros casos emblemáticos fue el de Associated Press, que desde 2014 utiliza sistemas desarrollados por Automated Insights para redactar automáticamente informes financieros trimestrales, pasando de producir cientos a más de 4 000 artículos al año sin intervención humana directa (Associated Press, 2014; Columbia Journalism Review, 2019).

De manera similar, Bloomberg implementó su sistema “Cyborg” en 2016 para generar noticias financieras en tiempo real, especialmente sobre resultados empresariales, permitiendo cubrir miles de reportes simultáneamente sin redactores humanos (Bloomberg, 2016; Nieman Lab, 2017).

La agencia Reuters comenzó a integrar herramientas de automatización desde 2018 para elaborar resúmenes financieros, alertas de mercado y análisis preliminares, apoyándose en inteligencia artificial y minería de datos (Reuters Institute, 2020).

En el ámbito deportivo, medios como The Washington Post y Yahoo Sports emplean desde 2017 sistemas de IA para generar crónicas automáticas de partidos locales, ligas menores y competencias universitarias, ampliando la cobertura sin incrementar su plantilla (Washington Post, 2017; Poynter Institute, 2019).

En Europa, el grupo mediático alemán Axel Springer y la agencia sueca TT News Agency desarrollaron plataformas propias de periodismo automatizado desde 2015, produciendo reportes financieros y deportivos de forma sistemática (WAN-IFRA, 2018).

En Asia, Xinhua News Agency en China presentó en 2018 a sus primeros “presentadores virtuales” y sistemas de redacción automática para contenidos noticiosos básicos, integrando IA en múltiples fases del proceso editorial (Xinhua, 2018; South China Morning Post, 2019).

En América Latina, el uso de redacción automatizada comenzó a expandirse a partir de 2019. Medios como Globo (Brasil), Clarín (Argentina) y El Tiempo (Colombia) han experimentado con sistemas de generación automática para contenidos deportivos y económicos (Knight Center, 2020; Fundación Gabo, 2021).

En el Perú, este proceso se ha desarrollado principalmente mediante el uso de herramientas basadas en inteligencia artificial generativa y análisis de datos. Desde 2021, medios digitales como El Comercio, RPP Noticias, Gestión y La República han incorporado sistemas de automatización para la elaboración de reportes financieros, dashboards informativos y boletines automatizados, apoyándose en plataformas como

Google News Initiative, Datawrapper y soluciones de IA generativa (El Comercio Lab, 2022; RPP Data, 2023).

Asimismo, durante procesos electorales y coyunturas económicas, estos medios han utilizado algoritmos para generar actualizaciones automáticas de resultados, conteo de votos y proyecciones estadísticas en tiempo real (ONPE, 2021; Fundación Telefónica, 2022).

Como resultado, la redacción de noticias simples ha pasado, en muchos contextos, de ser una actividad predominantemente humana a un proceso híbrido o totalmente automatizado, donde los algoritmos generan borradores completos que son publicados con mínima supervisión editorial.

Este fenómeno ha transformado profundamente el perfil del periodista contemporáneo. Las tareas de escritura repetitiva han sido reemplazadas por sistemas automatizados, mientras que los profesionales humanos se concentran en investigación, análisis crítico, verificación de fuentes y producción de contenidos interpretativos.

La automatización del periodismo básico representa uno de los casos más claros de “deshumanización funcional” en el sector informativo, impulsada por la necesidad de velocidad, volumen y reducción de costos, en un ecosistema mediático cada vez más digitalizado.

- **Editor básico de video y fotografía**

La profesión de editor básico de video y fotografía ha sido tradicionalmente una función clave en los ámbitos del periodismo, la publicidad, el marketing digital, la producción audiovisual y las redes sociales. Históricamente, estas tareas requerían conocimientos técnicos especializados en software de edición, criterios estéticos, manejo de tiempos y comprensión narrativa, lo que hacía del proceso una labor predominantemente humana.

Sin embargo, desde mediados de la década de 2010, el desarrollo de algoritmos de inteligencia artificial aplicados al procesamiento de imágenes y video ha transformado progresivamente esta profesión, facilitando procesos automatizados de edición,

corrección, montaje y optimización de contenidos visuales (Adobe, 2019; Google AI, 2020).

El proceso de automatización comenzó con herramientas de corrección automática de color, estabilización de imagen y recorte inteligente. Posteriormente, evolucionó hacia sistemas capaces de seleccionar escenas relevantes, sincronizar audio, aplicar efectos visuales y generar versiones finales sin intervención humana directa (MIT Media Lab, 2021).

A nivel internacional, empresas como Adobe incorporaron desde 2017 funciones basadas en IA a través de su plataforma Adobe Sensei, permitiendo la edición automática de fotografías, eliminación de fondos, mejora de resolución y montaje inteligente de videos (Adobe, 2017; Adobe, 2022).

De manera similar, Google Photos integró desde 2016 sistemas de reconocimiento visual para organizar, mejorar y editar imágenes automáticamente, mientras que YouTube implementó herramientas de edición asistida por IA para creadores de contenido (Google AI Blog, 2019; YouTube Creator Blog, 2021).

Plataformas emergentes como Runway ML, Pika Labs, Synthesia y Descript han impulsado desde 2020 una nueva generación de editores automáticos capaces de generar, modificar y reestructurar videos completos mediante instrucciones en lenguaje natural, reduciendo significativamente el tiempo y la intervención humana requerida (Runway, 2022; Synthesia, 2023).

Asimismo, aplicaciones móviles como CapCut, Canva, Lensa AI y Remini han popularizado la edición automática entre usuarios no especializados, democratizando el acceso a herramientas de producción audiovisual mediante algoritmos inteligentes (ByteDance, 2021; Canva, 2022).

En Asia, empresas como Tencent y Alibaba han desarrollado sistemas de edición automática para plataformas de comercio electrónico y streaming, optimizando la producción masiva de contenidos promocionales (Alibaba Cloud, 2022).

En América Latina, el uso de herramientas de edición con IA se intensificó a partir de 2020, especialmente en agencias de marketing digital, productoras independientes y medios de comunicación. Empresas como Globo (Brasil) y Televisa (México) han incorporado sistemas de automatización audiovisual para producción rápida de contenidos informativos y promocionales (Forbes Brasil, 2021; Expansión México, 2022).

En el Perú, este proceso se consolidó desde 2021, cuando medios digitales, agencias publicitarias y creadores de contenido comenzaron a utilizar herramientas como Adobe Sensei, Canva Pro, CapCut y Runway para la edición automatizada de material audiovisual (Gestión, 2022; El Comercio, 2023).

Productoras independientes, universidades y agencias de comunicación en Lima, Arequipa y Trujillo han adoptado plataformas de edición asistida por IA para optimizar costos y tiempos de producción, especialmente en contenidos institucionales y educativos (Semana Económica, 2023).

Como consecuencia, el perfil del editor básico ha experimentado una transformación profunda. Las tareas técnicas repetitivas, como el ajuste de color, corte, sincronización y mejora visual, han sido asumidas por algoritmos, mientras que los profesionales humanos se enfocan cada vez más en dirección creativa, narrativa y conceptualización.

Este proceso evidencia una transición hacia modelos de producción audiovisual híbridos y altamente automatizados, donde la inteligencia artificial cumple un rol central en la creación, optimización y distribución de contenidos visuales contemporáneos.

- **Técnico de laboratorio repetitivo**

La profesión de técnico de laboratorio repetitivo ha sido históricamente fundamental en los sectores de salud, investigación científica, farmacéutica, biotecnología y análisis clínico. Tradicionalmente, estos profesionales realizaban tareas como preparación de muestras, análisis bioquímicos, conteo celular, procesamiento de pruebas, control de calidad y registro manual de resultados, actividades que requerían precisión, concentración y supervisión constante.

Desde finales de la década de 1990, el avance de la automatización y la robótica médica inició un proceso gradual de transformación de estas funciones. Inicialmente, se introdujeron analizadores automáticos para pruebas sanguíneas y bioquímicas; posteriormente, estos sistemas evolucionaron hacia plataformas integradas controladas por inteligencia artificial (FDA, 2018; WHO, 2020).

El proceso de automatización se consolidó a partir de 2010 con el desarrollo de sistemas de laboratorio totalmente automatizados (Total Laboratory Automation, TLA), capaces de gestionar desde la recepción de muestras hasta la emisión de resultados sin intervención humana directa (Siemens Healthineers, 2015; Roche Diagnostics, 2017).

A nivel internacional, empresas como Roche, Abbott, Siemens Healthineers y Beckman Coulter lideran la implementación de laboratorios automatizados desde 2012, permitiendo procesar miles de muestras por hora con mínima participación humana (Roche, 2018; Abbott, 2019).

Durante la pandemia de COVID-19 (2020–2022), esta automatización se aceleró significativamente. Laboratorios en Estados Unidos, Europa y Asia emplearon robots de pipeteo, sistemas de PCR automatizados y plataformas de análisis masivo para procesar pruebas a gran escala (CDC, 2021; Nature Biotechnology, 2020).

En hospitales como Mayo Clinic, Cleveland Clinic y NHS Trusts del Reino Unido, la automatización permitió reducir errores humanos, optimizar tiempos diagnósticos y ampliar la cobertura sanitaria (Mayo Clinic, 2021; NHS, 2022).

En Asia, empresas como BGI Genomics (China) desarrollaron “laboratorios nube” altamente automatizados para análisis genómico, con mínima intervención humana (BGI, 2020).

En América Latina, la automatización de laboratorios comenzó a expandirse desde 2015, especialmente en grandes cadenas clínicas y hospitales privados. Instituciones como Fleury (Brasil), DASA y Grupo Diagnóstico Médico Proa (México) implementaron sistemas robotizados para análisis clínicos (Forbes Brasil, 2019; Expansión México, 2020).

En el Perú, este proceso se consolidó a partir de 2018 con la modernización de laboratorios privados y hospitalarios. Redes como ROE Laboratorios, Suiza Lab, MedLab y AUNA incorporaron analizadores automatizados, sistemas robotizados de muestras y plataformas digitales integradas para el procesamiento clínico (Gestión, 2019; El Comercio, 2021).

Durante la pandemia, el Instituto Nacional de Salud (INS) implementó plataformas automatizadas para pruebas moleculares, permitiendo ampliar significativamente la capacidad diagnóstica nacional (INS, 2020; MINSA, 2021).

Como consecuencia, muchas funciones tradicionalmente asignadas a técnicos humanos han sido reemplazadas por sistemas inteligentes, especialmente en pruebas rutinarias, análisis bioquímicos estándar y procesamiento masivo de datos clínicos.

Actualmente, el rol del técnico de laboratorio ha evolucionado hacia funciones de supervisión, calibración de equipos, interpretación avanzada y control de calidad, desplazando las tareas manuales repetitivas hacia plataformas automatizadas.

Este proceso representa un caso claro de automatización funcional en el ámbito sanitario, donde la precisión algorítmica, la eficiencia operativa y la reducción de errores han impulsado una transformación estructural del trabajo humano en laboratorios modernos.

- **Controlador de grúas portuarias**

La profesión de controlador de grúas portuarias ha sido tradicionalmente una labor intensiva en habilidades técnicas, coordinación espacial, juicio humano y respuesta inmediata a condiciones cambiantes del entorno. En los puertos, este rol implica operar grúas gigantes para levantar, mover y colocar contenedores, gestionar cargas y descargas, y coordinar con el tráfico marítimo y terrestre —tareas que requieren atención constante, pericia táctica y experiencia operativa.

Históricamente, la automatización en terminales portuarias se expandió a partir de la década de 1990 con sistemas de asistencia al operador que mejoraban la precisión de maniobras, integraban sensores y facilitaban comunicaciones en tiempo real (Port Technology International, 1998). Sin embargo, la automatización completa de grúas y

operaciones portuarias se consolidó más adelante, con la introducción de sistemas automatizados de almacenamiento y recuperación (AS/RS) y grúas automáticas guiadas por software inteligente.

Uno de los **primeros casos emblemáticos de automatización portuaria a gran escala** fue el **Puerto de Rotterdam (Países Bajos)**, que desde la década de 2000 implementó sistemas de grúas automatizadas en contenedores integrados con plataformas de gestión logística digital capaz de operar mediante algoritmos de optimización en tiempo real, reduciendo la intervención humana directa e incrementando la eficiencia de carga y descarga (Port of Rotterdam Authority, 2007).

En el **Puerto de Hamburgo (Alemania)**, desde 2012 se emplearon grúas automatizadas y vehículos guiados automáticos (AGV) coordinados por inteligencia artificial para gestionar el flujo de contenedores con menos intervención humana, lo que permitió optimizar tiempos de operación y reducir errores de posicionamiento (HHLA, 2013).

En Asia, el **Puerto de Shanghai (China)** integró sistemas de grúas automatizadas desde 2015, incluyendo sensores inteligentes, visión artificial para identificar contenedores y algoritmos predictivos que optimizan rutas de movimiento en función de datos en tiempo real (Shanghai International Port Group, 2017).

Estos sistemas automatizados operan bajo control centralizado con supervisión remota, donde el “operador” humano se encarga mayormente de monitoreo, planificación y contingencias, mientras la inteligencia artificial controla las funciones de levantamiento, transporte horizontal y ordenamiento de contenedores.

En **América Latina**, la automatización portuaria ha experimentado avances importantes desde mediados de la década de 2010. En Brasil, el **Puerto de Santos** implementó grúas semiautónomas y sistemas AS/RS en varias de sus terminales, mejorando la productividad y reduciendo tiempos de estiba y desestiba (Porto de Santos Automation Report, 2020). Aunque no todas las operaciones son completamente autónomas, la integración de tecnologías inteligentes ha redefinido parcialmente el rol del operador de grúa tradicional.

En el **Perú**, la automatización portuaria aún se encuentra en una etapa progresiva de adopción tecnológica. Puertos como el **Callao** y **Muelle Norte** han incorporado sistemas de gestión asistida por software para mejorar la coordinación logística entre buques y camiones, así como plataformas de optimización de tráfico interno (APM Terminals Callao, 2020; OSITRAN, 2021). Además, en algunos terminales se han instalado grúas con asistencia remota y sensores que reducen la carga de trabajo físico de los operadores humanos. Sin embargo, la automatización completa aún no es la norma y la intervención humana sigue siendo central en operaciones críticas, debido a la complejidad logística, la variedad de cargas y la necesidad de juicio contextual en los procesos (Perú Ports Modernization Report, 2023).

La adopción de tecnologías de automatización en control portuario ha redefinido la función del operador de grúas desde la ejecución manual de movimientos hacia una supervisión más estratégica y técnica de sistemas automatizados, integrando monitoreo remoto, análisis de datos de tráfico y coordinación con sistemas de gestión portuaria basados en inteligencia artificial.

Este proceso evidencia una progresiva deshumanización operativa en espacios portuarios globales, donde la intervención humana se ve reemplazada por sistemas inteligentes para funciones altamente repetitivas, mecánicas y optimizadas algorítmicamente, transformando la naturaleza misma del empleo en este sector.

- **Archivista y gestor documental**

La profesión de archivista y gestor documental ha estado tradicionalmente asociada al manejo, organización, preservación y recuperación de información en diversas instituciones, como archivos históricos, bibliotecas, centros de documentación, instituciones gubernamentales y empresas privadas. Este trabajo implicaba gestión manual de papeles, clasificación física, etiquetado manual, ordenamiento cronológico o temático, y recuperación de documentos bajo criterios específicos.

Con el auge de la digitalización a finales del siglo XX y principios del XXI, estas labores comenzaron a transformarse mediante sistemas de gestión electrónica de documentos

(Electronic Document Management Systems – EDMS), que permitieron almacenar, buscar y controlar versiones digitales, aunque muchas funciones aún requerían intervención humana.

La automatización intensiva de estas funciones se ha acelerado desde la década de 2015 con la integración de inteligencia artificial en sistemas de gestión documental. Plataformas modernas incorporan **extracción automática de metadatos, clasificación inteligente, etiquetado automático, búsqueda semántica avanzada y automatización de flujos de trabajo de archivo**, reduciendo drásticamente la necesidad de procesamiento manual de documentos (AgentUI Document Management, s. f.; AIDA Document Archive, s. f.; RecordsKeeper.AI, s. f.).

Sistemas como **AIDA Document Archive** automatizan la archivación digital mediante reconocimiento óptico de caracteres (OCR) y capacidades de búsqueda avanzada que permiten indexar y recuperar documentos sin intervención humana directa, integrando estos procesos con almacenamiento en la nube y funciones de control de versiones.

Asimismo, plataformas como **Trupeer AI Records Management Software** y **AgentUI automation for documents** utilizan algoritmos de IA para clasificar, etiquetar y administrar archivos de manera automática, garantizando cumplimiento de políticas de retención, acceso controlado, auditorías automatizadas y estructuración de repositorios digitales sin supervisión manual constante.

La automatización documental también incluye soluciones que extraen, estructuran y organizan datos a partir de documentos variados (PDF, imágenes escaneadas, correos electrónicos y formatos estructurados o no estructurados), aplicando reglas de negocio para retención legal, carpetas automáticas y clasificación acorde a normativas de compliance, lo que historiza los flujos de archivo y reduce el tiempo destinado a tareas repetitivas.

Estas tecnologías han incidido en el desplazamiento de funciones tradicionales del archivista hacia roles más estratégicos, centrados en supervisión de sistemas, diseño de

taxonomías automatizadas, verificación de calidad de metadatos, gestión de cumplimiento normativo y optimización de repositorios inteligentes.

Además, en algunos sistemas avanzados se incorpora inteligencia generativa para sintetizar contenido, responder consultas contextualizadas sobre colecciones archivadas e incluso aplicar políticas de eliminación automática cuando corresponda, capacidades que antes eran exclusivamente humanas.

En el contexto latinoamericano, la adopción de plataformas automatizadas de gestión documental se ha expandido especialmente desde 2020, tanto en sector público como privado, donde gobiernos, universidades y grandes empresas consultan, gestionan e intercambian documentos de forma automatizada, integrando inteligencia artificial para agilizar procesos de archivo, consulta y cumplimiento de políticas de retención documental.

En el Perú, instituciones públicas como entidades tributarias, ministerios y universidades han incorporado sistemas automatizados para la digitalización, archivo y recuperación de documentos administrativos y legales, lo que ha reducido la dependencia de procesos manuales de catalogación y permitido rescatar grandes volúmenes de documentos históricos en formatos digitalizados accesibles mediante motores de búsqueda inteligentes.

El impacto de estas tecnologías plantea una **reconfiguración profunda de la profesión de archivista y gestor documental**: las tareas puramente repetitivas se han trasladado a sistemas inteligentes, mientras que los profesionales humanos ahora se concentran en el diseño de estructuras de datos, políticas de retención, supervisión de flujos automatizados, control de calidad semántica y estrategias avanzadas de preservación digital.

Este proceso constituye uno de los fenómenos más evidentes de deshumanización funcional en la gestión de información, donde la inteligencia artificial no solo apoya, sino muchas veces sustituye las tareas operativas de clasificación, búsqueda y almacenamiento documental, transformando radicalmente el rol laboral en este sector.

A partir de los casos analizados, puede observarse que las herramientas tecnológicas han evolucionado de manera sostenida, permitiendo que múltiples funciones laborales tradicionalmente desempeñadas por seres humanos sean hoy ejecutadas por sistemas automatizados.

Estos ejemplos constituyen evidencia empírica de que la automatización ya ha reemplazado parcial o totalmente a la fuerza laboral en diversas profesiones, transformando de forma estructural el mercado de trabajo contemporáneo.

En este sentido, el presente análisis no solo busca describir procesos ya consolidados, sino también establecer una base argumentativa para comprender fenómenos emergentes. A continuación, se expondrán otros ámbitos profesionales en los que la intervención humana se encuentra progresivamente desplazada por sistemas de inteligencia artificial, y que, en el futuro, probablemente compartan un destino similar al de las ocupaciones previamente examinadas.

4.2.4. PROFESIONES EN PROCESO DE AUTOMATIZACIÓN

- Juez virtual / Justicia algorítmica

La administración de justicia ha sido una de las funciones institucionales más sensibles de la sociedad, tradicionalmente ejercida por jueces humanos que interpretan normas, valoran pruebas, calibran argumentos y emiten sentencias con base en criterios jurídicos y éticos. Este rol implica juicio contextual, ponderación de circunstancias y aplicación normativa con equidad.

Sin embargo, en las últimas décadas surgieron sistemas de apoyo automatizado que intentan asistir al operador humano en tareas repetitivas o basadas en reglas definidas. Ejemplos tempranos incluyen sistemas de *análisis de precedentes* y evaluaciones de riesgo, como el programa COMPAS (Correctional Offender Management Profiling for Alternative Sanctions), utilizado en Estados Unidos desde principios de la década de 2010 para evaluar riesgos de reincidencia en decisiones de fianza y libertad condicional (ProPublica, 2016; Brennan et al., 2020).

A partir de mediados de la década de 2010, algunos países comenzaron a experimentar con **tribunales en línea** y sistemas de **resolución de disputas automatizadas** (Online Dispute Resolution – ODR), que emplean algoritmos para mediar conflictos contractuales o administrativos con intervención humana mínima, aunque en la mayoría de los casos el operador humano sigue siendo responsable de la decisión final (UNIDROIT, 2018; EU ODR Platform, 2019).

En China, desde 2017 se implementaron **tribunales virtuales basados en inteligencia artificial** para casos civiles de menor cuantía, donde sistemas automatizados generan borradores de decisiones que posteriormente son revisados por jueces humanos en pruebas judiciales controladas (Xinhua, 2017; BBC, 2019). Estos sistemas emplean procesamiento de lenguaje natural, minería de datos legislativos y criterios estructurados para sugerir sentencias, aunque la firma final corresponde a un juez humano certificado.

Organizaciones como la **Comisión Europea** han estudiado plataformas de justicia asistida por IA con el objetivo de reducir la acumulación de casos y aumentar la eficiencia, siempre manteniendo mecanismos de supervisión humana para asegurar el respeto a garantías procesales (European Commission: Justice and Consumer, 2021).

En América Latina, proyectos de justicia digital están en fases exploratorias. En Brasil, estados como São Paulo han desarrollado plataformas que agilizan notificaciones electrónicas y validaciones de documentos en procesos judiciales mediante automatización documental, aunque las decisiones judiciales finales siguen en manos humanas (CNJ Brasil, 2020; São Paulo Court Reports, 2021).

En el Perú, el Poder Judicial ha implementado progresivamente sistemas de gestión procesal electrónica (JEI), notificaciones digitales y manejo automatizado de expedientes, pero no existen hasta ahora plataformas de IA que sustituyan la decisión jurisdiccional humana. El uso de herramientas digitales sirve actualmente como apoyo administrativo, entrenamiento técnico y consulta de jurisprudencia automatizada, sin eliminar la intervención esencial del juez en la motivación de fallos (PJ Perú, 2024).

Este panorama muestra que la justicia algorítmica es un campo **en proceso de automatización**, donde los sistemas inteligentes ya realizan labores de apoyo sustantivo —como análisis de riesgo, búsqueda de precedentes, gestión documental y sugerencias de decisiones— pero aún no han sustituido completamente la función jurisdiccional humana debido a limitaciones éticas, de responsabilidad legal y de garantías procesales. La evidencia empírica sugiere que, aunque la inteligencia artificial puede agilizar procesos y reducir tiempos administrativos, los juicios de valor, la ponderación de intereses y la aplicación contextual de derecho siguen siendo competencias humanas imprescindibles, manteniendo a esta profesión en una fase de automatización asistida y no de reemplazo total.

- **Microcirujano asistido por IA**

La microcirugía es una especialidad médica que se encarga de operaciones extremadamente precisas sobre estructuras diminutas del cuerpo, como nervios, vasos sanguíneos finos y tejidos delicados. Tradicionalmente, estos procedimientos exigen habilidades altamente especializadas, años de entrenamiento y una coordinación manual excepcional por parte de cirujanos humanos, debido a la complejidad y variabilidad de cada caso.

Desde finales del siglo XX, se han desarrollado sistemas de **cirugía asistida por robot** que ayudan a los cirujanos a ejecutar movimientos más precisos, minimizar temblores y acceder a áreas difíciles del cuerpo mediante interfaces teleoperadas. Un ejemplo destacado es el **da Vinci Surgical System**, que desde inicios de los 2000 se consolidó como plataforma líder para procedimientos mínimamente invasivos guiados por un operador humano, permitiendo a los cirujanos controlar brazos robóticos con mayor precisión que la mano desnuda (da Vinci System; SurgicalTeck, 2025).

En años más recientes, la integración de inteligencia artificial ha impulsado un salto cualitativo en estos sistemas. Plataformas modernas de robótica médica combinan visión computacional, algoritmos de aprendizaje automático y coordinación de múltiples brazos robóticos para asistir directamente al cirujano, facilitando tareas que tradicionalmente

requieren décadas de formación. Por ejemplo, sistemas como **ROSA**, utilizados en neurocirugía y procedimientos complejos, integran planificación automatizada de trayectorias y navegación asistida por software inteligente (ROSA Robot Assistant).

Un avance notable hacia la automatización progresiva de la microcirugía ocurrió en **2025**, cuando la plataforma robótica **MARS**, desarrollada por Levita Magnetics, fue utilizada en Chile para realizar la primera colecistectomía con una cámara quirúrgica guiada por IA que ajustó el campo visual de manera autónoma durante la intervención, permitiendo al cirujano centrarse exclusivamente en la operación sin asistencia humana directa para la visualización del campo quirúrgico (Primera cirugía autónoma guiada por IA, 2025).

Más aún, investigaciones recientes de universidades como **Johns Hopkins y Stanford** han demostrado que sistemas robóticos pueden aprender rutinas quirúrgicas complejas observando videos de procedimientos realizados por expertos, capacidades clave para la eventual automatización de tareas microquirúrgicas que requieren manipulación delicada de tejidos y precisión submilimétrica (Robots entrenados por video para cirugía, s. f.).

En 2025 también se documentó el éxito de un robot quirúrgico autónomo denominado **SRT-H**, que completó con éxito un procedimiento de extracción de vesícula biliar en modelos anatómicos sin intervención humana directa, marcando un hito hacia la autonomía quirúrgica total, aunque todavía se requieren cambios de instrumentos y supervisión técnica por parte de profesionales de la salud (Robot quirúrgico autónomo SRT-H, 2025).

A nivel internacional, estas innovaciones han llevado a un crecimiento significativo del mercado de robots quirúrgicos, proyectándose que la industria, valorada en alrededor de **15,6 mil millones de dólares en 2024**, podría superar los **64 mil millones para 2034**, impulsada por avances en automatización y capacidades asistidas por IA en cirugía (Avances en cirugía automatizada, 2025).

En el contexto latinoamericano, tecnologías de asistencia robótica ya se emplean en redes hospitalarias públicas y privadas, y se considera que esta tendencia continuará creciendo en los próximos años conforme se extiendan los sistemas de IA integrados.

Aunque la microcirugía asistida por IA avanza rápidamente, **la sustitución total del cirujano humano aún no se ha alcanzado**; la intervención profesional sigue siendo esencial en las etapas de planificación, supervisión, toma de decisiones críticas y resolución de contingencias intraoperatorias. Por tanto, esta profesión se encuentra actualmente en una **fase de automatización asistida**, con capacidades crecientes de IA que redefinen funciones, optimizan precisión y elevan la eficiencia, pero sin eliminar la presencia humana en el acto quirúrgico principal.

- **Controlador de tráfico aéreo asistido por IA**

La profesión de controlador de tráfico aéreo (ATC, por sus siglas en inglés) ha sido uno de los pilares de la aviación civil y militar durante décadas. Tradicionalmente, los controladores son responsables de gestionar el flujo de aeronaves en rutas y aeropuertos, coordinar despegues y aterrizajes, evitar conflictos entre trayectorias, y mantener la seguridad y eficiencia en espacios aéreos dinámicos y densamente transitados.

Desde finales del siglo XX, los sistemas de gestión de tráfico aéreo han integrado dispositivos automatizados que asisten a los operadores humanos, como sistemas de visualización, predicción de trayectorias y alertas de conflicto (Short-term Conflict Alert, 2009). Sin embargo, los controladores siguen siendo esenciales para la toma de decisiones críticas, la comunicación directa con las aeronaves y la resolución de situaciones no lineales o imprevistas.

A partir de la década del 2010, las investigaciones y proyectos globales empezaron a explorar la integración de inteligencia artificial en los sistemas de control aéreo, con el objetivo de incrementar la capacidad, reducir la carga de trabajo humano y mejorar la seguridad (Nature Research, 2025; AUTOPACE, 2025). Tecnologías como el procesamiento de lenguaje natural (NLP) y el reconocimiento automático de voz (ASR) buscan automatizar partes de la comunicación entre pilotos y controladores, reduciendo errores humanos y facilitando transcripción y análisis de mensajes en tiempo real (MDPI Sensors, 2024; Springer Nature, 2025).

Proyectos como **AISA (AI Situational Awareness Foundation for Advancing Automation)**, liderados en Europa por investigadores de la Universidad Politécnica de Madrid, plantean escenarios donde la automatización no solo apoya, sino que anticipa conflictos de tráfico y propone decisiones sugeridas que el controlador humano evalúa y aprueba (Actualidad Aeroespacial, 2026; AUTOPACE, 2050). Otro proyecto europeo, **MAHALO**, estudia cómo sistemas de aprendizaje automático pueden detectar conflictos, sugerir resoluciones y colaborar con operadores humanos, aunque destaca la necesidad de que el controlador mantenga el control final y la confianza en las recomendaciones algorítmicas (CORDIS, 2026).

La investigación continua en simulación avanzada y “digital twins” (gemelos digitales) de espacios aéreos busca entrenar y evaluar agentes automatizados en entornos seguros antes de su implementación en operaciones reales, poniendo énfasis en la fiabilidad y la seguridad de los sistemas (Digital Twin ATC, 2026; ArXiv, 2026). Asimismo, estudios han explorado el uso de *deep reinforcement learning* para apoyar la resolución táctica de conflictos de tráfico, lo que sugiere que la IA podría actuar como asistente de decisiones bajo supervisión humana, pero todavía no reemplazar completamente las funciones de un controlador humano (ScienceDirect, 2024).

A nivel global, sistemas como el **Australian Advanced Air Traffic System (TAAATS)** y su contraparte militar **ADATS** (Australian Defence Air Traffic System) han demostrado desde principios de los años 2000 cómo el soporte tecnológico mejora la productividad de los controladores sin sustituirlos, integrando información radar y comunicaciones para optimizar flujos de trabajo (TAAATS, 2000; ADATS, 2024). Sin embargo, estos siguen siendo sistemas de asistencia y no sistemas autónomos de administración total del espacio aéreo.

La Organización de Aviación Civil Internacional (OACI) reconoce que la IA puede facilitar la integración de múltiples fuentes de datos —como meteorología, planes de vuelo, vigilancia y predicciones— para apoyar la toma de decisiones en tiempo real (ICAO, 2025). No obstante, la literatura especializada también señala que la complejidad del

entorno, la necesidad de juicio contextual y las implicaciones éticas y de seguridad hacen que la sustitución completa del controlador humano por IA sea tecnológicamente posible solo a largo plazo y bajo estrictos marcos regulatorios (Eurocontrol, 2026; El Economista, 2025).

En este sentido, la automatización en ATC se encuentra en una **fase intermedia de integración tecnológica**: los sistemas inteligentes están ganando terreno como asistentes que mejoran eficiencia, precisión y rapidez, pero la función crítica de toma de decisiones últimas, comunicación contextual, manejo de emergencias y coordinación humana sigue recayendo en operadores humanos altamente entrenados. Por ello, la profesión de controlador de tráfico aéreo se considera **en proceso de automatización**, donde la IA complementa y amplía capacidades, pero no ha reemplazado la responsabilidad humana principal.

- **Diagnóstico médico asistido por IA**

La práctica del diagnóstico médico ha sido tradicionalmente una función compleja ejercida por profesionales de la salud, como médicos generales y especialistas clínicos, quienes integran los síntomas del paciente, la historia clínica, resultados de pruebas y exploraciones físicas para llegar a un juicio diagnóstico y plan de tratamiento. Esta actividad exige competencia clínica profunda, razonamiento crítico, evaluación contextual de múltiples variables y responsabilidad ética y legal ante el paciente.

Desde mediados del siglo XX comenzaron a desarrollarse los primeros sistemas de apoyo diagnóstico basados en conocimientos estructurados, como **DXplain** o **INTERNIST-I / CADUCEUS**, que asistían a médicos en la generación de diagnósticos diferenciales a partir de signos y síntomas introducidos por el usuario (DXplain, 1986; CADUCEUS, 1984). Estos primeros sistemas mostraron que la inteligencia computacional podía, incluso en las primeras décadas, evaluar posibles enfermedades en función de datos clínicos y generar sugerencias útiles para profesionales de la salud sin reemplazarlos completamente.

A partir de la década de 2010, con la expansión de algoritmos de aprendizaje profundo y redes neuronales aplicadas a imágenes médicas y datos clínicos, la inteligencia artificial ha irrumpido de forma importante en el diagnóstico médico. Las herramientas basadas en IA ahora se utilizan ampliamente para analizar imágenes de resonancias magnéticas, tomografías computarizadas, radiografías y mamografías, con niveles de precisión que en algunos casos igualan o superan a profesionales humanos en tareas específicas de reconocimiento de patrones complejos en imágenes.

Un ejemplo reglamentado de esta tendencia es **IDx-DR**, el primer sistema de diagnóstico autónomo de retinopatía diabética aprobado por la Administración de Alimentos y Medicamentos (FDA) de Estados Unidos. IDx-DR analiza imágenes retinianas sin necesidad de interpretación humana para detectar signos de enfermedad ocular, lo que permite diagnósticos rápidos y accesibles incluso en entornos con escasez de especialistas.

Asimismo, desarrollos como **MONAI**, un marco open-source para aprendizaje profundo en imágenes médicas respaldado por instituciones como Nvidia y los National Institutes of Health (NIH), facilitan la creación de aplicaciones de IA que segmentan, clasifican y procesan imágenes para apoyar diagnósticos en múltiples contextos clínicos.

Proyectos de investigación recientes han explorado marcos avanzados para combinar diagnósticos automatizados en diferentes dominios clínicos. Por ejemplo, MedOrch propone la orquestación de múltiples herramientas especializadas para diagnóstico de enfermedades complejas, como Alzheimer o interpretaciones de rayos X de tórax, obteniendo altos niveles de precisión en conjuntos de datos clínicos y demostrando el potencial de sistemas de IA integrados para asistir al médico en decisiones diagnósticas complejas.

Además, ensayos recientes con dispositivos médico-IA, como un **estetoscopio inteligente capaz de diagnosticar condiciones cardíacas complejas en 15 segundos**, muestran cómo la IA está expandiendo el espectro de diagnóstico fuera de los métodos tradicionales, aunque todavía bajo supervisión clínica.

No obstante, los expertos y estudios científicos también destacan la importancia de mantener la supervisión humana. Un trabajo publicado en 2025 comparó el rendimiento de un sistema autónomo de IA en atención urgente con médicos certificados, encontrando que aunque la concordancia diagnóstica fue alta, la supervisión clínica sigue siendo crucial en escenarios reales de atención médica, especialmente para casos complejos, comorbilidades y decisiones terapéuticas integrales.

Incluso estudios regionales, como uno realizado en Uruguay, mostraron que modelos como ChatGPT, Claude y Gemini lograron aciertos diagnósticos en un rango del 80 % al 90 % en historias clínicas, pero sin sustituir la necesidad del juicio clínico humano ni los componentes éticos y relacionales del acto médico.

En conjunto, estas evidencias demuestran que la IA está transformando la práctica diagnóstica, apoyando, acelerando y en ocasiones igualando tareas que antes requerían intervención humana directa, pero **no ha reemplazado plenamente al profesional médico**. Las herramientas automatizadas son cada vez más sofisticadas, pero carecen de empatía clínica, razonamiento holístico y juicio contextual, componentes que siguen siendo competencia humana esencial. Por ello, el diagnóstico médico asistido por IA se ubica actualmente en una **fase de automatización auxiliar y colaborativa**, donde los sistemas inteligentes complementan —pero no sustituyen del todo— la labor del médico.

Docente asistido por inteligencia artificial (educación en proceso de automatización)

La profesión docente, tradicionalmente basada en la interacción humana directa, la planificación pedagógica y el acompañamiento emocional del estudiante, ha comenzado a experimentar un proceso progresivo de automatización a través del uso de sistemas de inteligencia artificial aplicados a la enseñanza, evaluación y personalización del aprendizaje.

Desde mediados de la década de 2010, plataformas educativas basadas en IA han incorporado sistemas de tutoría inteligente capaces de adaptar contenidos según el ritmo, nivel y estilo de aprendizaje del estudiante. Empresas como Carnegie Learning y Squirrel

AI han desarrollado tutores virtuales que ofrecen instrucción personalizada en matemáticas y ciencias, mostrando mejoras significativas en el rendimiento académico frente a métodos tradicionales (Carnegie Learning, 2018; Squirrel AI, 2020).

En China, el uso de aulas inteligentes con sistemas de análisis facial, seguimiento del comportamiento y evaluación automatizada se implementó desde 2017 en instituciones públicas y privadas, permitiendo monitorear la atención y desempeño de los estudiantes en tiempo real (Xinhua, 2019). Estos sistemas reducen la necesidad de supervisión docente constante y trasladan parte del proceso educativo a plataformas algorítmicas.

Durante la pandemia de COVID-19 (2020–2022), la expansión del aprendizaje virtual aceleró significativamente la incorporación de herramientas automatizadas. Empresas como Coursera, Duolingo y Khan Academy integraron modelos de IA para generar ejercicios, evaluar respuestas y recomendar itinerarios formativos personalizados (Coursera, 2021; Duolingo, 2023; Khan Academy, 2023). Posteriormente, con el lanzamiento de sistemas generativos como ChatGPT, estas plataformas ampliaron sus capacidades de tutoría automatizada.

En 2023, Khan Academy implementó oficialmente su tutor basado en GPT-4, denominado “Khanmigo”, capaz de asistir a estudiantes en matemáticas, programación y redacción, simulando procesos de acompañamiento docente individualizado (Khan Academy, 2023). De forma similar, Duolingo Max incorporó modelos de lenguaje avanzados para ofrecer explicaciones personalizadas en aprendizaje de idiomas (Duolingo, 2023).

A nivel latinoamericano, instituciones educativas comenzaron a integrar estas tecnologías en programas piloto. En Perú, universidades como la Universidad Peruana de Ciencias Aplicadas (UPC) y la Pontificia Universidad Católica del Perú (PUCP) han incorporado sistemas de analítica educativa, plataformas adaptativas y asistentes virtuales para apoyar procesos de tutoría y evaluación (Gestión, 2023; Andina, 2024). Asimismo, el Ministerio de Educación promovió el uso de plataformas digitales con algoritmos de personalización durante y después de la pandemia (MINEDU, 2022).

Estos desarrollos han modificado el rol tradicional del docente, desplazándolo progresivamente desde transmisor de contenidos hacia supervisor, facilitador y gestor de procesos automatizados. En muchos entornos virtuales, la corrección de ejercicios, generación de materiales y seguimiento del progreso ya se realiza mayoritariamente mediante sistemas inteligentes.

No obstante, diversos estudios señalan que, aunque la IA puede optimizar procesos instruccionales, aún presenta limitaciones en aspectos socioemocionales, éticos y formativos profundos, lo que impide una sustitución completa del docente humano en el corto plazo (UNESCO, 2021; OECD, 2024). Por ello, esta profesión se ubica actualmente en una fase de **automatización progresiva**, más que de reemplazo total.

- **Abogado en revisión contractual automatizada**

La profesión de abogado ha tenido históricamente un componente central de lectura, interpretación y aplicación del derecho. Parte esencial de este trabajo consiste en la revisión y redacción de contratos, evaluación de riesgos jurídicos, adaptación de cláusulas según necesidades específicas y asegurar el cumplimiento de normativas —funciones que implican juicio contextual, razonamiento argumentativo y conocimiento actualizado del ordenamiento jurídico.

Sin embargo, desde la década de 2010, herramientas basadas en inteligencia artificial comenzaron a apoyar de manera creciente tareas rutinarias de revisión contractual. El desarrollo de tecnologías de **procesamiento de lenguaje natural (PLN)** y **análisis semántico aplicado al derecho** ha permitido que sistemas automatizados identifiquen cláusulas, comparen textos contractuales con estándares legales y detecten riesgos o inconsistencias, optimizando el trabajo jurídico tradicional (Surden, 2014; Katz, 2019).

Un caso temprano y representativo fue el software **ROSS Intelligence**, basado en IBM Watson, lanzado para ayudar a abogados a encontrar precedentes o normas relevantes de forma más rápida que con búsquedas manuales en bases de datos, aunque posteriormente dejó de operar comercialmente, su desarrollo marcó un precedente en IA legal (Mac, 2016; The Law Society Gazette, 2020).

Más reciente es el uso de plataformas como LawGeex y Kira Systems, que desde mediados de la década de 2010 se utilizan ampliamente en despachos de abogados y corporaciones para la revisión de contratos estándar. Estas herramientas pueden comparar contratos con un “modelo de cláusulas preferidas”, detectar términos faltantes o riesgos potenciales y sugerir modificaciones automáticas basadas en patrones previamente entrenados (LawGeex, 2019; Kira Systems, 2021).

Empresas como Deloitte y EY han integrado estas soluciones de IA en sus servicios de consultoría legal empresarial para *due diligence*, contratos de fusiones y adquisiciones, y revisión masiva de documentos legales, alcanzando eficiencias que reducen tiempos de semanas a días o incluso horas (Deloitte Insights, 2020; EY, 2021).

En el ámbito académico, diversos estudios han demostrado que estos sistemas automatizados pueden alcanzar niveles de precisión comparables a abogados humanos en tareas estructuradas y repetitivas de análisis de cláusulas contractuales, con la excepción de situaciones que requieren interpretación normativa compleja o evaluación contextual de circunstancias imprevisibles (Surden, 2019; Journal of Law & AI, 2022).

En América Latina, la adopción de soluciones de revisión contractual con IA ha emergido principalmente en grandes firmas y equipos legales corporativos desde finales de la década de 2010, impulsada por la necesidad de manejar volúmenes crecientes de contratos y cumplir con plazos regulatorios estrictos. En países como Brasil y México, despachos especializados ya integran revisores automáticos en procesos de *compliance*, contratos comerciales y documentación societaria (Forbes Legal LATAM, 2022; LegalTech LATAM, 2023).

En el Perú, el uso de herramientas de automatización para revisión contractual se ha introducido de manera incipiente en estudios jurídicos de mayor tamaño y en áreas corporativas de empresas, especialmente para contratos tipo, cláusulas de estandarización y procesos de auditoría legal, aunque todavía con supervisión humana decisiva, dado que la precisión contextual y la responsabilidad ética siguen recayendo en profesionales certificados (Semana Económica, 2023; Gestión, 2024).

En síntesis, la revisión contractual automatizada se encuentra en una **fase de automatización asistida y creciente adopción**: los sistemas inteligentes realizan tareas repetitivas, comparativas y de detección de riesgos con gran eficiencia, pero no reemplazan completamente al abogado humano en interpretaciones complejas, negociación contractual o valoración de contexto jurídico, lo cual mantiene el rol profesional en un proceso de transición hacia la colaboración hombre-máquina.

- **Especialista en recursos humanos (IA en selección de personal)**

La profesión de especialista en recursos humanos (RR. HH.) ha estado tradicionalmente vinculada a la gestión integral del capital humano: reclutamiento, selección, análisis de perfiles, entrevistas, evaluación de competencias, desarrollo de talento y toma de decisiones estratégicas sobre puestos de trabajo. Este rol exige juicio contextual, lectura interpersonal, interpretación de señales no verbales y criterios éticos en la evaluación humana.

Con el desarrollo de tecnologías de inteligencia artificial y aprendizaje automático, muchas funciones operativas dentro de RR. HH. han comenzado a automatizarse, especialmente aquellas relacionadas con la selección de personal. A partir de la década de 2010, empresas tecnológicas comenzaron a diseñar sistemas que analizan grandes volúmenes de currículos, identifican patrones de éxito laboral y filtran candidatos con base en criterios previamente establecidos, reduciendo significativamente la carga de trabajo manual (LinkedIn, 2018; IBM, 2019).

Plataformas como **HireVue** integraron análisis de video e IA para evaluar entrevistas grabadas, midiendo no solo respuestas verbales, sino también patrones de voz y expresiones faciales, en un intento por predecir desempeño futuro —método que fue adoptado por organizaciones globales desde mediados de la década de 2010 (HireVue AI interviewing, 2019; SHRM, 2020). Estas herramientas generan puntuaciones automatizadas que influyen en decisiones de selección, aunque normalmente son supervisadas por profesionales de RR. HH.

Del mismo modo, sistemas como **Pymetrics** utilizan evaluaciones basadas en juegos y algoritmos de IA para medir rasgos cognitivos y de personalidad que correlacionan con desempeño profesional, sugiriendo perfiles adecuados según datos históricos de empleados exitosos (Pymetrics, 2021). Otras soluciones como **Eightfold.ai** aplican *machine learning* para identificar talento interno y externo basado en trayectoria profesional, habilidades preexistentes y predicciones de crecimiento profesional (Eightfold, 2022).

Estas tecnologías no solo filtran currículos, sino que, mediante *matching* automatizado entre competencias y ofertas, predicen la probabilidad de éxito en un puesto, acelerando el proceso de selección y reduciendo sesgos explícitos en fases iniciales (McKinsey, 2020).

En el ámbito internacional, corporaciones como Unilever, Hilton y Vodafone implementaron herramientas de selección automatizada desde finales de la década de 2010, reportando reducciones importantes en tiempos de contratación y costos administrativos (Unilever digital recruitment, 2020; Vodafone AI recruiting, 2021).

Sin embargo, diversas investigaciones han señalado riesgos asociados a estos sistemas, especialmente en relación con sesgos algorítmicos que pueden reproducir prejuicios humanos si los datos de entrenamiento no están cuidadosamente auditados (MIT Technology Review, 2019; Harvard Business Review, 2023). Esto ha llevado a que la supervisión humana siga siendo una parte central del proceso, especialmente en fases finales de entrevistas y decisiones estratégicas de contratación.

En América Latina, y específicamente en el Perú, empresas de gran tamaño han empezado a utilizar módulos de automatización para preselección de candidatos y análisis de competencias, integrados a plataformas de gestión de talento como SAP SuccessFactors, Workday o plataformas locales con modulación IA. Estas herramientas ayudan a filtrar grandes volúmenes de postulantes para puestos técnicos y administrativos, aunque las decisiones finales de contratación permanecen bajo

responsabilidad de equipos de RR. HH. con supervisión humana (Gestión, 2023; Semana Económica, 2024).

En el contexto actual, la automatización en RR. HH. no ha llevado aún al reemplazo total del especialista humano, sino a una redistribución de tareas: las funciones repetitivas, de clasificación y primer cribado se realizan mediante IA, mientras que las decisiones complejas, entrevistas profundas, evaluación de ajuste cultural y consideraciones éticas permanecen en manos de profesionales humanos. Por ello, la profesión de especialista en recursos humanos puede considerarse **en proceso de automatización**, donde la IA colabora estrechamente con el trabajo humano sin sustituirlo completamente.

- **Síntesis**

A partir del análisis de las profesiones examinadas, puede afirmarse que el proceso de automatización laboral impulsado por la inteligencia artificial no constituye una proyección futura distante, sino una realidad en desarrollo acelerado. Diversos estudios internacionales coinciden en señalar que una proporción significativa de las tareas laborales actuales ya presenta un alto nivel de exposición a la automatización.

Según el Fondo Monetario Internacional, aproximadamente el 40 % del empleo mundial se encuentra expuesto al impacto directo de la inteligencia artificial, con mayores niveles en economías desarrolladas (IMF, 2024). De forma complementaria, reportes de Fortunly indican que los roles administrativos, de gestión documental y procesamiento de información presentan niveles de automatización potencial superiores al 80 % (Fortunly, 2023).

Por su parte, la Organización para la Cooperación y el Desarrollo Económicos señala que cerca del 27 % de los empleos en países miembros se encuentran en alto riesgo de automatización, mientras que un porcentaje aún mayor experimentará transformaciones profundas en sus funciones (OECD, 2019). Asimismo, investigaciones del McKinsey Global Institute estiman que hacia 2030 aproximadamente el 30 % de las horas laborales actuales podrían ser automatizadas mediante tecnologías existentes (McKinsey, 2017).

Estos datos permiten afirmar que el impacto de la IA no se limita al reemplazo directo de trabajadores, sino que impulsa una transformación estructural del trabajo, donde muchas profesiones evolucionan hacia modelos híbridos humano-máquina antes de una posible automatización completa.

Por tanto la automatización mediante inteligencia artificial ya está **transformando el mercado de trabajo** en 2025, pero no ha producido un reemplazo total de la mayoría de profesiones humanas.

Avances observables:

- Según estimaciones recientes, **30 % de trabajos podrían automatizarse para 2030**, con 60 % de tareas laborales fuertemente modificadas por IA (Prestianni, 2025).
- El Fondo Monetario Internacional (IMF) observa que **40 % de empleos globales están “expuestos” a la automatización por IA**, siendo roles repetitivos y administrativos los más afectados.
- En países de la OCDE, aproximadamente **27 % de empleos tienen un alto riesgo de automatización**, destacándose trabajos de baja y media cualificación como más vulnerables.
- Estimaciones más amplias sugieren que entre **60 % y 80 % de empleos podrían verse transformados o automatizados hacia 2050** si las tecnologías siguen avanzando rápidamente.
- Adicionalmente, informes señalan que una proporción creciente de tareas dentro de las profesiones actuales —no solo empleos completos— está siendo automatizada o asistida por IA, requiriendo reentrenamiento de trabajadores.

Por lo expuesto la **automatización ya está en marcha**, afecta especialmente tareas repetitivas y estructuradas, y su impacto se intensificará en las próximas décadas, aunque muchas profesiones cambian antes de desaparecer según en la siguiente tabla:

Tabla 06: Avance actual de automatización por IA (2025)

Categoría de empleo / ámbito	% Empleos afectables / automatizables	Fuente
Empleos con tareas automatizadas parcialmente	~60–80 % potencial transformación (2030–2050)	https://completeaitraining.com/news/ai-job-replacement-statistics-2025-how-automation-is/?utm_source=chatgpt.com
Empleos con alto riesgo de automatización (OCDE)	~27 %	https://www.oecd.org/en/publications/oecd-employment-outlook-2023_08785bba-en/full-report/artificial-intelligence-and-jobs-no-signs-of-slowing-labour-demand-yet_5aeb670.html?utm_source=chatgpt.com
Empleos por completar automatización (U.S., previstos)	~30 % para 2030	https://www.nu.edu/blog/ai-job-statistics/?utm_source=chatgpt.com
Empleos “expuestos” globalmente	~40 %	https://fortunly.com/statistics/automation-job-loss-statistics/?utm_source=chatgpt.com
Sectores con mayor riesgo: datos/administración	30–95 % según rol	https://completeaitraining.com/news/ai-job-replacement-statistics-2025-how-automation-is/?utm_source=chatgpt.com

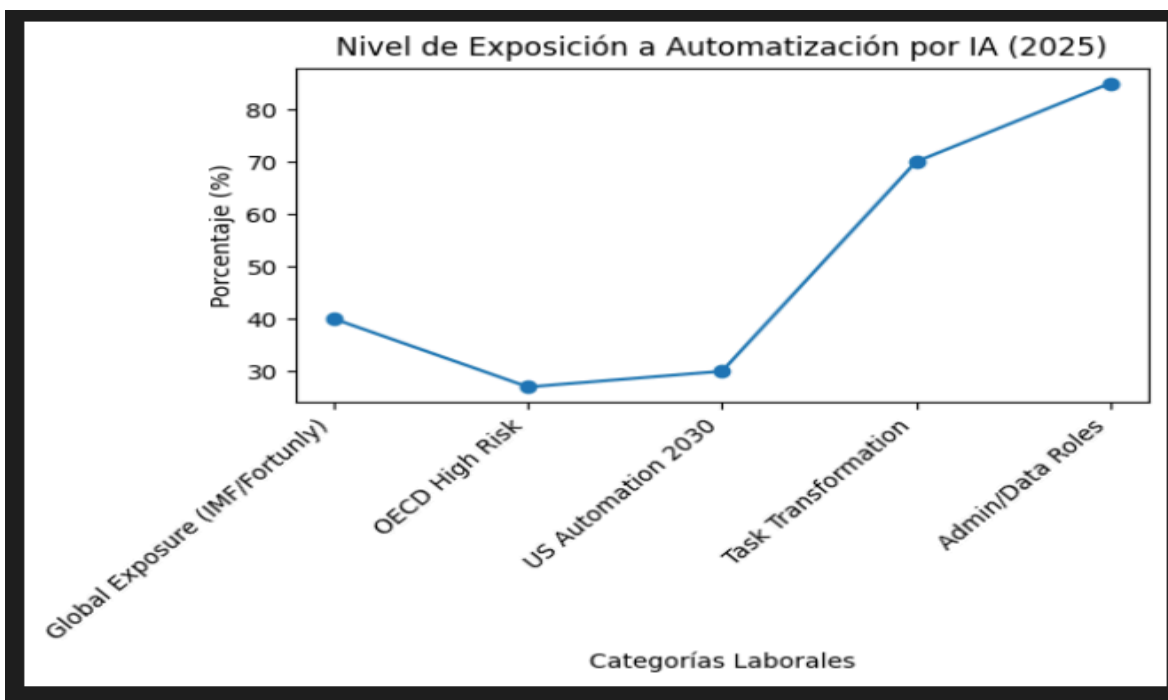


Figura 05: Nivel de exposición a automatización por IA (2025)

La figura 5 sintetiza los niveles actuales de exposición a la automatización según distintos estudios internacionales, evidenciando que las áreas administrativas, de datos y gestión repetitiva presentan los mayores niveles de vulnerabilidad frente a la inteligencia artificial.

Tabla 07: Proyección estimada de reemplazo total de empleos por IA (2030 y más allá)

Horizonte temporal	% Empleos automatizados estimados	Comentario
2030 corto plazo	~25–30 % de empleos automatizados	Dependiendo del sector, con tareas repetitivas ampliamente asumidas por IA / robots (Prestianni, 2025).
2035 medio plazo	~30–50 % de empleos automatizados	Proyecciones amplias de exposición a automatización en países avanzados.
2050 largo plazo	~60–80 % de tareas transformadas	Estudios que contemplan automatización complementando tecnologías robóticas e IA.

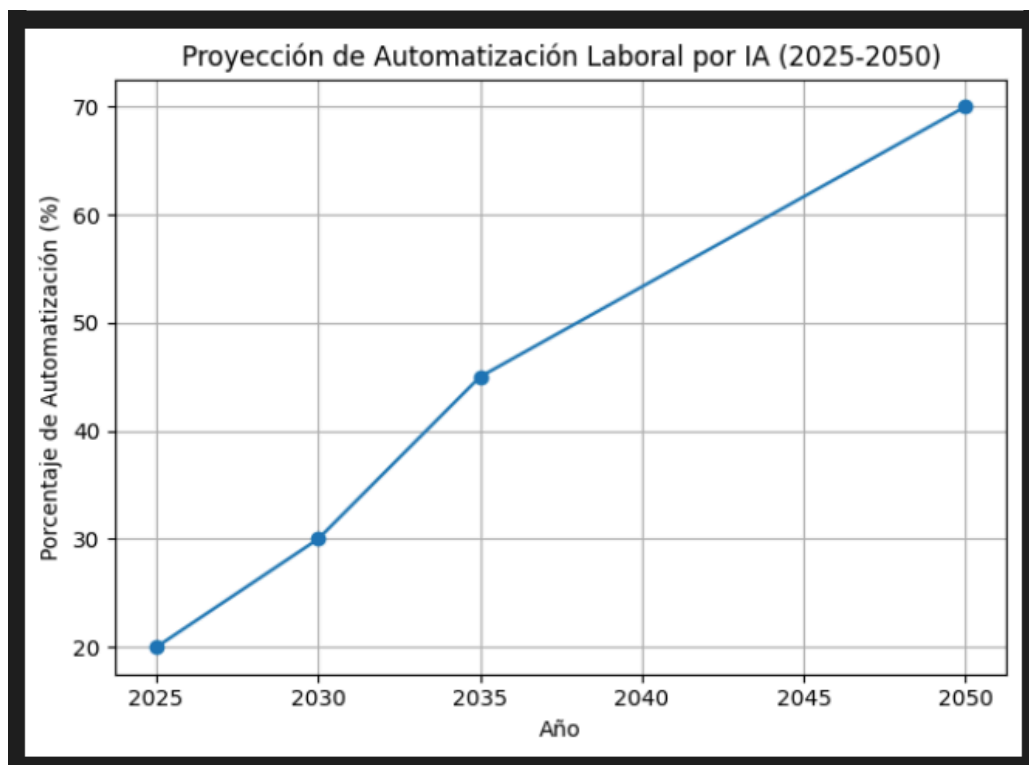


Figura 06: Proyección de la automatización laboral por IA (2025)

La Figura 6 muestra una tendencia sostenida al incremento de la automatización laboral. Considerando el ritmo actual de desarrollo de modelos de aprendizaje profundo, agentes autónomos y sistemas multimodales, diversos organismos estiman que entre 60 % y 70 % de las tareas laborales podrían automatizarse parcialmente hacia el año 2050 (PwC, 2022; McKinsey, 2023).

La evidencia muestra que:

- La IA **ya ha reemplazado tareas dentro de empleos**, e incluso **algunos empleos completos**, especialmente en sectores como administración, datos y producción industrial.
- La sustitución total de profesionales especializados (por ejemplo, médicos, jueces, docentes) aún no es una realidad generalizada, sino un proceso **incremental que primero transforma tareas antes que profesiones enteras**.
- La automatización tiende a **acompañar el cambio de rol humano**, generando nuevos empleos y alterando habilidades requeridas.

A partir del análisis de las profesiones examinadas, puede afirmarse que el proceso de automatización laboral impulsado por la inteligencia artificial no constituye una proyección futura distante, sino una realidad en desarrollo acelerado. Diversos estudios internacionales coinciden en señalar que una proporción significativa de las tareas laborales actuales ya presenta un alto nivel de exposición a la automatización.

Según el Fondo Monetario Internacional, aproximadamente el 40 % del empleo mundial se encuentra expuesto al impacto directo de la inteligencia artificial, con mayores niveles en economías desarrolladas (IMF, 2024). De forma complementaria, reportes de Fortunly indican que los roles administrativos, de gestión documental y procesamiento de información presentan niveles de automatización potencial superiores al 80 % (Fortunly, 2023).

Por su parte, la Organización para la Cooperación y el Desarrollo Económicos señala que cerca del 27 % de los empleos en países miembros se encuentran en alto riesgo de automatización, mientras que un porcentaje aún mayor experimentará transformaciones profundas en sus funciones (OECD, 2019). Asimismo, investigaciones del McKinsey Global Institute estiman que hacia 2030 aproximadamente el 30 % de las horas laborales actuales podrían ser automatizadas mediante tecnologías existentes (McKinsey, 2017).

Estos datos permiten afirmar que el impacto de la IA no se limita al reemplazo directo de trabajadores, sino que impulsa una transformación estructural del trabajo, donde muchas profesiones evolucionan hacia modelos híbridos humano-máquina antes de una posible automatización completa.

4.2.5. PROFESIONES CON IMPOSIBILIDAD DE AUTOMATIZACIÓN TEMPORAL

A pesar del avance sostenido de la inteligencia artificial y de su creciente capacidad para ejecutar tareas complejas, existen actualmente ámbitos laborales cuya automatización total resulta, en el presente, técnicamente limitada y socialmente controvertida. Estas profesiones se caracterizan por requerir altos niveles de juicio moral, sensibilidad emocional, creatividad contextual, responsabilidad ética y toma de decisiones en escenarios altamente inciertos.

En este sentido, ciertos empleos continúan siendo considerados, por el momento, impensables de ser plenamente dominados por sistemas artificiales. No obstante, la historia del desarrollo tecnológico demuestra que el progreso, aunque medible y cuantificable en muchos aspectos, no siempre sigue trayectorias lineales ni previsibles.

Con frecuencia, la innovación tecnológica ha estado marcada por descubrimientos inesperados, accidentes fortuitos, avances disruptivos y reconfiguraciones imprevistas del conocimiento científico, los cuales han dado lugar a transformaciones profundas en ámbitos previamente considerados estables.

En el campo de la inteligencia artificial, fenómenos como el aprendizaje emergente, la autooptimización de modelos y la integración multimodal sugieren que futuras capacidades aún no imaginadas podrían modificar sustancialmente los límites actuales de automatización. Por ello, aunque ciertas profesiones mantienen hoy una fuerte dependencia de la intervención humana, no resulta metodológicamente riguroso asumir que dicha condición será permanente.

Desde esta perspectiva, el presente apartado analiza aquellas ocupaciones que, bajo el estado actual del desarrollo tecnológico, presentan una imposibilidad de automatización temporal, sin descartar la posibilidad de futuras transformaciones impulsadas por avances aún imprevisibles.

- **Médicos en emergencias críticas — límites de automatización**

La atención médica en contextos de urgencia representa uno de los ámbitos más complejos y resilientes frente a la automatización total mediante inteligencia artificial. Aunque los sistemas de IA han demostrado avances prometedores en tareas específicas como detección de patrones en imágenes médicas o generación de posibles diagnósticos a partir de datos estructurados, la evidencia científica disponible resalta múltiples limitaciones que impiden considerar un reemplazo completo del médico de urgencias por sistemas automatizados.

En un estudio comparativo reciente sobre el desempeño de herramientas de IA frente a médicos de urgencias, se constató que, si bien ciertos modelos pueden sugerir

diagnósticos basados en una lista finita de casos estandarizados, las **capacidades clínicas humanas —como interpretar hallazgos de examen físico, sintetizar información contextual cambiante y gestionar incertidumbres complejas en tiempo real— siguen siendo dominantes y difíciles de replicar en IA** (Springer Nature, 2025).

La literatura especializada subraya que los modelos actuales de IA presentan **limitaciones estructurales importantes**, tales como la incapacidad de realizar un examen físico completo, incorporar datos sensoriales multimodales, o integrar juicio clínico basado en experiencia práctica y variabilidad individual de los pacientes (MDPI Diagnostics, 2026). Estas restricciones señalan que la automatización total de decisiones médicas en situaciones críticas no sólo es técnicamente compleja, sino que plantea importantes desafíos éticos y de responsabilidad legal.

Además, la investigación muestra que la implementación de IA en emergencias también puede generar riesgos si no se la gestiona con cuidado. Por ejemplo, un análisis sobre el uso de algoritmos de predicción de sepsis indicó que el modelo falló al interpretar correctamente el contexto clínico, lo que llevó a omisiones de alerta temprana en algunos casos reales, evidenciando la falta de robustez de las herramientas cuando se las aplica fuera de escenarios bien definidos (JMIR, 2025).

Más aún, un examen crítico de cómo los usuarios perciben y utilizan consejos médicos generados por IA reveló que muchas personas confían en respuestas automatizadas con **menor precisión**, incluso en contextos donde la falta de exactitud puede resultar perjudicial, lo cual resalta la importancia del juicio clínico humano para filtrar, contextualizar y validar recomendaciones diagnósticas (ArXiv, 2024).

A estos factores se suman las preocupaciones éticas y de sesgo implícito documentadas en estudios sobre chatbots de salud, los cuales pueden perpetuar errores clínicos o inequidades si no se supervisan adecuadamente por profesionales cualificados (AP News, 2023).

En conjunto, esta evidencia sugiere que aunque los sistemas inteligentes pueden apoyar en tareas específicas (p. ej., análisis de imágenes o sugerencia de hipótesis

diagnósticas), la **capacidad de toma de decisiones en emergencias con información incompleta, variables dinámicas y consecuencias críticas sigue dependiendo de la intervención humana**. Esta dependencia no sólo responde a limitaciones tecnológicas actuales, sino también a la necesidad de empatía, juicio moral, ética profesional y responsabilidad clínica, atributos que no han sido replicados de forma íntegra por tecnologías de IA.

En consecuencia, la automatización en la práctica médica de urgencias debe ser interpretada como una **colaboración asistida** entre humano e IA, y no como una sustitución total de la profesión médica, al menos con el nivel actual de desarrollo tecnológico y las barreras epistemológicas y éticas que persisten.

- **Psicoterapia clínica profunda — límites de automatización**

La psicoterapia clínica profunda se basa en interacciones humanas complejas que requieren **empatía genuina, comprensión contextual, interpretación emocional y una alianza terapéutica construida con participación activa del paciente y del profesional**. Estas características la convierten en una de las profesiones con **imposibilidad de automatización total en el estado actual de la tecnología**, particularmente en casos que implican traumas, conflictos internos arraigados o condiciones psiquiátricas severas.

Si bien el uso de chatbots y modelos generativos de IA para “apoyo emocional” ha sido explorado en los últimos años, la literatura científica disponible enfatiza múltiples limitaciones y riesgos que impiden considerar a la IA como reemplazo completo de psicoterapeutas humanos. Por ejemplo, estudios recientes han demostrado que los **usuarios de aplicaciones de psicoterapia asistida por IA valoran su accesibilidad, pero experimentan dificultades con la profundidad terapéutica, la personalización emocional y la continuidad de la narrativa terapéutica**, elementos esenciales en el proceso psicoterapéutico (Springer Nature, 2025).

Además, la evidencia indica que los chatbots actuales tienden a generar respuestas genéricas y repetitivas que carecen de fundamentos robustos en técnicas terapéuticas

basadas en evidencia, lo cual puede ser insuficiente o incluso perjudicial para personas con trastornos complejos (JMIR Mental Health, 2025).

Un análisis sistemático muestra que, aunque los chatbots pueden reducir síntomas de ansiedad en algunos casos, los tratamientos tradicionales conducidos por terapeutas humanos producen **reducciones de ansiedad significativamente mayores**, probablemente debido a la calidad de la relación terapéutica, la validación emocional y la intervención adaptada a cada paciente (BMC Psychology, 2025).

Estas limitaciones están relacionadas tanto con la **falta de empatía auténtica y comprensión multisensorial**, como con desafíos éticos fundamentales, tales como la privacidad de datos, la confidencialidad y la responsabilidad profesional (PMC, 2025). En consecuencia, **no existen modelos de IA clínicamente verificados y regulados que reemplacen la totalidad de la intervención psicoterapéutica humana**, especialmente en contextos de terapia profunda.

Además de las deficiencias técnicas señaladas, emergen preocupaciones sobre el impacto negativo de la IA cuando se utiliza sin supervisión profesional adecuada. Por ejemplo, estudios han observado que algunos chatbots pueden incluso **respaldar o reforzar patrones de pensamiento perjudiciales** en adolescentes con angustia emocional, subrayando la incapacidad de estos sistemas para gestionar apropiadamente casos de riesgo elevado (PubMed, 2025).

La percepción de los propios profesionales contemporáneos también refleja estos límites prácticos. Investigaciones sobre las actitudes de psicoterapeutas muestran que, aunque algunos ven oportunidades en el uso de IA como herramienta de apoyo, **la falta de evidencia empírica sólida, la resistencia al cambio y la necesidad de interacción humana significativa** frenan su adopción clínica generalizada (BMC Psychology, 2025).

Finalmente, incluso en escenarios donde la IA podría proveer cierto nivel de apoyo —como en intervenciones breves para casos leves, psicoeducación o herramientas de reflexión entre sesiones— **no se ha demostrado que estos sistemas reemplacen la función terapéutica de un profesional humano en contextos clínicos moderados o**

severos. La tecnología puede servir como complemento o puente de acceso, pero su capacidad para “curar” o conducir procesos de transformación psicológica profunda permanece limitada (CPP, 2024).

En síntesis, la psicoterapia clínica profunda representa una **profesión en la que la automatización total por IA no es plausible con el estado actual de desarrollo tecnológico y científico**, debido a las barreras conceptuales, emocionales, éticas y prácticas. Esto no excluye que en el futuro sistemas más avanzados puedan ampliar capacidades específicas, pero **las funciones centrales de la terapia —empatía, conexión humana, juicio clínico y manejo de crisis complejas— continúan requiriendo la intervención de profesionales humanos.**

La automatización total de la psicoterapia clínica profunda **no es posible con la IA actual.** Si bien existen herramientas que ofrecen apoyo y mayor accesibilidad, faltan capacidades humanas esenciales —empatía, intuición, relación terapéutica y manejo de riesgos complejos— que mantienen a esta profesión como irremplazable con la tecnología existente.

- **Liderazgo militar humano frente a automatización de sistemas de alerta**

Una de las funciones más críticas del ámbito militar es la detección temprana y la respuesta a amenazas que, de materializarse, podrían desencadenar consecuencias catastróficas de escala global —como una guerra nuclear. En este contexto, los sistemas automatizados de radar y alerta temprana —tecnologías que integran sensores, satélites y algoritmos para identificar posibles ataques enemigos— han sido desarrollados con el objetivo de ofrecer respuestas más rápidas y precisas. Sin embargo, estos sistemas sufren fallos técnicos que ponen en evidencia la necesidad imperativa de juicio humano, especialmente cuando las decisiones involucran **vida, muerte y supervivencia de la humanidad.**

Un caso paradigmático ocurrió el **26 de septiembre de 1983**, cuando el sistema soviético de alerta temprana **Oko** reportó que Estados Unidos había lanzado misiles contra la Unión Soviética. El oficial de guardia, el teniente coronel **Stanislav Petrov**, interpretó

correctamente que se trataba de un falso positivo y decidió no transmitir la alarma a sus superiores, lo cual habría desencadenado una respuesta nuclear automática. Su decisión, basada en criterio humano más que en datos automatizados, es ampliamente reconocida como un factor clave que evitó un posible **holocausto nuclear** (Petrov, 1983; Arms Control Association, 2019) .

Los sistemas automatizados de alerta pueden fallar por múltiples razones: reflejos solares en nubes pueden simular la firma de un lanzamiento de misil, como fue confirmado más tarde en el incidente de 1983, donde la detección errónea se debió a un reflejo de luz solar sobre nubes en una órbita alta que los sensores interpretaron falsamente como un ataque entrante (GlobalSecurity, s. f.; Stanford, 2017) . Sin la intervención de Petrov, que cuestionó los datos, podría haberse ordenado una represalia nuclear con consecuencias inimaginables.

Pero la historia de errores técnicos en sistemas de alerta no se limita a la Unión Soviética. En **1979**, el comando conjunto de defensa norteamericana (NORAD) detectó erróneamente un ataque masivo de misiles soviéticos debido a la carga accidental de una cinta de entrenamiento en el sistema de producción, lo cual llevó a la activación de fuerzas de respuesta antes de que se confirmara la falsedad del alerta (Kimball, 2019) . Este y otros incidentes similares en 1980 y a lo largo de la Guerra Fría muestran que incluso los sistemas tecnológicos más sofisticados pueden producir señales engañosas que, sin un juicio humano ponderado, podrían escalar rápidamente hacia acciones irreversibles.

Otro suceso significativo ocurrió en **1995**, cuando un cohete noruego de investigación fue detectado por radares rusos como un posible misil balístico. La situación provocó la activación del “maletín nuclear” del presidente Boris Yeltsin, quien tuvo que evaluar cuidadosamente la información antes de ordenar cualquier respuesta militar nuclear. Su decisión de esperar y confirmar los hechos evitó por pocos minutos una respuesta armada errónea (Incidente del cohete noruego, 1995) .

Estos eventos ponen de manifiesto que, incluso con sistemas automatizados de alerta temprana altamente avanzados, la **interpretación contextual de información conflictiva y la capacidad de cuestionar diagnósticos automáticos son funciones que dependen de la presencia humana**. La intervención de líderes militares —capaces de evaluar, dudar, ponderar riesgos y considerar consecuencias más allá de lo cuantificable— ha sido decisiva en momentos donde la automatización pura habría tendido a generar respuestas rápidas, pero potencialmente devastadoras.

Adicionalmente, hay ejemplos durante la llamada *Crisis de los Misiles en Cuba* en 1962, donde decisiones humanas en submarinos soviéticos evitaron el lanzamiento de torpedos nucleares pese a órdenes automatizadas o presiones de situational logic militar, resaltando nuevamente la importancia del factor humano en decisiones de alto riesgo (Defcon Level, 2024) .

La tecnología de detección puede detectar patrones, señales y anomalías, pero **carece de comprensión contextual amplia, juicio moral, evaluación de intenciones y análisis de riesgo integral**—atributos que han sido esenciales para evitar desastres en momentos críticos. Estos ejemplos históricos muestran que, aunque la automatización puede mejorar la rapidez operativa, **no puede sustituir por completo la función humana en decisiones de consecuencias extremas**, al menos con la tecnología actual y previsible en el corto plazo.

Estas evidencias históricas muestran que **la automatización de sistemas militarmente críticos aún depende de la supervisión humana para evitar catástrofes**. Los fallos de radar y sistemas de alerta, observados repetidamente durante la Guerra Fría y después, han requerido decisiones humanas que **corregían interpretaciones automatizadas erróneas y evitaron consecuencias extremas**, como una guerra nuclear. Esto refuerza el argumento de que, en contextos donde la vida humana y la supervivencia de la sociedad están en juego, la **sustitución total por IA no es plausible en el corto ni mediano plazo**.

- **Jueces en juicios complejos — límites de automatización**

La función judicial constituye uno de los pilares del Estado de derecho, implicando no solo la aplicación mecánica de normas sino también *interpretación jurídica, evaluación moral, ponderación de contextos sociales y protección de derechos fundamentales*. La idea de reemplazar a un juez por un sistema automatizado ha sido objeto de análisis teórico y experimental en el campo de la inteligencia artificial, pero tanto la evidencia práctica como las reflexiones académicas y jurisprudenciales demuestran que esta sustitución total aún se encuentra lejos de ser viable —y que implicaría riesgos sistémicos y éticos significativos.

En algunos países se ha explorado el uso de IA para **asistencia en tareas no decisorias**, como la redacción preliminar de opiniones o investigación de antecedentes para sentencias. Por ejemplo, en Inglaterra y Gales se autorizó el uso de IA para ayudar a redactar borradores de fallos, pero **se prohibió expresamente su uso para análisis jurídico profundo o la toma de decisiones**, debido al riesgo de que la IA produzca información inexacta, sesgada o incluso fabricada (“hallucinations”) que podría minar la confianza pública en el sistema judicial (AP News, 2024).

El riesgo de confiar demasiado en sistemas automatizados en decisiones jurídicas es evidente: jueces en India han advertido sobre “citaciones alucinadas” y fallos generados por IA que contenían decisiones o referencias inexistentes, lo que llevó a rechazarlas y a reforzar el argumento de que **la IA solo puede servir como apoyo, no como agente decisor**, dada su incapacidad para razonar constitucionalmente o manifestar comprensión genuina del caso (Times of India, 2025).

Asimismo, en Estados Unidos se documentó en 2025 que jueces federales habían presentado documentos judiciales con **citas falsas, nombres ficticios y conclusiones inventadas**, generados por IA aplicada sin la suficiente supervisión humana, lo que obligó a retirar las resoluciones y generó debates sobre la necesidad de políticas estrictas y salvaguardias éticas (Washington Post, 2025).

Los límites técnicos y éticos del uso de IA en decisiones judiciales han sido objeto de análisis académico. Investigaciones sobre algoritmos predictivos y sistemas automáticos

de decisión advierten que estas herramientas tienden a reflejar sesgos presentes en los datos de entrenamiento y carecen de los atributos cognitivos y contextuales necesarios para interpretar la complejidad jurídica real —como valores sociales, matices culturales y consideraciones de justicia individual— elementos que no pueden ser “programados” de manera absoluta (The World Law Group, s. f.; Medvedeva et al., 2020).

Organizaciones internacionales como la **OCDE** han resaltado que, aunque la IA puede apoyar la administración judicial —por ejemplo, agilizando búsqueda de jurisprudencia o gestión documental— la falta de transparencia algorítmica y de mecanismos robustos de supervisión puede socavar la explicación de decisiones judiciales y dificultar su impugnación, lo que representa un riesgo para el debido proceso y la imparcialidad (OECD, 2025).

Los marcos normativos internacionales, como la regulación europea sobre IA, distinguen entre usos de bajo riesgo (p. ej., gestión documental) y usos de **alto riesgo** (decisiones que afectan derechos fundamentales), y recomiendan cautela extrema en la automatización de funciones jurisdiccionales, preservando siempre la intervención humana y la responsabilidad de los jueces sobre todo el contenido de las decisiones (OECD, 2025).

Incluso en jurisdicciones que promueven la modernización jurídica con IA —como el Poder Judicial del Perú, que impulsa herramientas para optimizar gestión de expedientes— los proyectos enfatizan que la IA debe ser un complemento técnico para apoyar el trabajo jurisdiccional, **no un sustituto del juicio humano en la administración de justicia** (Poder Judicial del Perú, 2025).

Los riesgos de trasladar la toma de decisiones completa a sistemas automatizados son múltiples: sesgos inherentes a datos históricos, dificultad para explicar decisiones (algoritmos como “caja negra”), erosión de la autonomía judicial, afectación de la percepción pública de legitimidad en el sistema de justicia, y la posibilidad de reproducir injusticias sistémicas de forma amplificada por mecanismos automatizados que no comprenden plenamente el significado del derecho (Stimson Center, 2026).

En consecuencia, la evidencia hasta la fecha indica que los sistemas de IA pueden ser instrumentos útiles para **asistencia informativa, análisis preliminar de gran volumen de información, predicción estadística o redacción de borradores**, pero no poseen la capacidad de sustituir integralmente a los jueces humanos en la toma de decisiones jurídicas complejas que exigen interpretación moral, justicia contextual y protección de derechos. Esto subraya la importancia de mantener la intervención humana en el núcleo del sistema judicial, no solo como garantía ética y constitucional, sino también como elemento indispensable de legitimidad social.

- **Docencia formativa profunda — límites de la automatización**

La docencia formativa profunda comprende más que la transmisión de contenidos: involucra la construcción de pensamiento crítico, la facilitación de procesos de aprendizaje complejos, la evaluación formativa continua, el acompañamiento emocional y la adaptación pedagógica contextualizada al individuo o grupo. Este conjunto de actividades exige una combinación de habilidades cognitivas, sociales, éticas y relacionales que la colocan entre las profesiones **menos susceptibles de ser totalmente automatizadas** con el nivel actual de desarrollo tecnológico.

Aunque la inteligencia artificial ha sido incorporada en múltiples plataformas educativas para apoyar la enseñanza —por ejemplo, mediante sistemas de tutoría inteligente, recomendaciones adaptativas o evaluación automatizada de ciertos tipos de ejercicios— la evidencia científica disponible destaca limitaciones claras cuando se trata de replicar funciones avanzadas del docente humano, especialmente en escenarios que requieren interpretación, creatividad pedagógica y comprensión emocional (UNESCO, 2021).
(unesdoc.unesco.org)

● **Limitaciones técnicas observadas**

Diversos estudios han señalado que los sistemas de IA son eficientes para:

- Gestión de contenidos repetitivos o estructurados.
- Adaptación de ejercicios basados en patrones de respuesta.
- Corrección automática de respuestas cerradas.

Pero presentan fallas notorias cuando:

- Se requiere interpretar procesos de pensamiento complejos.
- Se intenta detectar emociones, motivaciones o bloqueos intelectuales.
- Se necesita realizar intervención formativa durante la sesión pedagógica (OECD, 2024). (oecd.org)

Estas limitaciones tienen implicaciones directas en la calidad de la enseñanza: un sistema automatizado puede sugerir respuestas correctas, pero **no puede reemplazar la guía personalizada, el ajuste en tiempo real ni la interacción dialógica que caracteriza una clase formativa profunda.**

- **Evidencia experimental**

Un estudio publicado en *Nature Human Behaviour* evaluó el impacto de plataformas educativas automatizadas en comparación con docentes humanos en cursos universitarios, y observó que si bien los sistemas de IA mejoraban ciertos ítems de rendimiento (especialmente en ejercicios repetitivos), **no lograron replicar los efectos formativos de las interacciones humanas sobre habilidades como pensamiento crítico, argumentación y resolución creativa de problemas** (Nature Human Behaviour, 2023). (nature.com)

Además, investigaciones sobre el uso de asistentes de IA generativa en aulas universitarias mostraron que, aunque estos sistemas pueden generar contenidos explicativos detallados y guías de estudio personalizadas, **no presentaron evidencia de poder reemplazar la evaluación cualitativa ni la retroalimentación adaptada al contexto emocional, cultural y cognitivo de cada estudiante** (Journal of Learning Analytics, 2024). (learning-analytics.org)

En resumen, la **docencia formativa profunda** representa una de las profesiones en las que la automatización total sigue siendo, en el presente, **poco probable** debido a:

- La complejidad de las interacciones pedagógicas humanas.
- Los requisitos éticos y emocionales de la enseñanza.

- Las limitaciones demostradas de los sistemas automatizados para replicar juicio pedagógico contextual.
- La evidencia empírica que muestra que la IA complementa, pero no reemplaza, funciones humanas formativas.

- **Casos de limitación en entornos reales**

En contextos educativos reales, el uso de plataformas de IA para evaluación automática ha sido implementado con éxito en cursos masivos en línea (*MOOCs*), pero los mismos docentes han reportado dificultades para interpretar y responder a la retroalimentación emocional de estudiantes o para ajustar estrategias pedagógicas frente a respuestas complejas no estándar (EdSurge, 2025). (edsurge.com)

Asimismo, en programas de formación docente en América Latina se ha observado que, aun con el uso de IA en tareas administrativas y de contenido, **los docentes continúan siendo centrales para diagnosticar dificultades de aprendizaje, generar motivación, promover interacción y sostener la comunidad de aprendizaje** —factores que impactan directamente en los resultados de los estudiantes (UNICEF, 2022). (unicef.org)

Además de las limitaciones técnicas, existen reflexiones éticas profundas que cuestionan la sustitución de la docencia por IA. La enseñanza formativa requiere una relación humana de confianza, sensibilidad cultural e interpretación contextual —atributos que no pueden ser plenamente replicados por una máquina. La empatía, la gestión de conflictos en clase, el ajuste instantáneo de estrategias pedagógicas y la adaptación a necesidades específicas de estudiantes con dificultades diversas requieren capacidades que, hasta ahora, la tecnología no ha logrado integrar de manera holística (AAAI/EAAI, 2023)

En resumen, la **docencia formativa profunda** representa una de las profesiones en las que la automatización total sigue siendo, en el presente, **poco probable** debido a:

- La complejidad de las interacciones pedagógicas humanas.
- Los requisitos éticos y emocionales de la enseñanza.
- Las limitaciones demostradas de los sistemas automatizados para replicar juicio pedagógico contextual.

- La evidencia empírica que muestra que la IA complementa, pero no reemplaza, funciones humanas formativas.

Aunque es posible que futuros desarrollos tecnológicos transformen aún más la enseñanza, la inteligencia artificial —tal como se presenta hoy y se vislumbra en los estudios actuales— opera principalmente como herramienta de apoyo, dejando las funciones centrales de formación humana en manos de docentes capacitados.

Aunque es posible que futuros desarrollos tecnológicos transformen aún más la enseñanza, la inteligencia artificial —tal como se presenta hoy y se vislumbra en los estudios actuales— opera principalmente como herramienta de apoyo, dejando las funciones centrales de formación humana en manos de docentes capacitados.

La automatización puede ayudar en tareas docentes específicas, pero **la enseñanza formativa profunda** —que implica juicio, empatía, conexión humana y adaptabilidad pedagógica— **no puede ser completamente automatizada con la tecnología actual.**

La evidencia académica resalta que los docentes humanos siguen teniendo un rol insustituible, particularmente cuando se trata de *aprendizaje crítico, apoyo emocional y adaptación contextualmente sensible*.

Trabajo social — límites de automatización con IA

La profesión de trabajo social se centra en la **intervención relacional, justicia social, acompañamiento comunitario y protección de derechos humanos**, lo que implica habilidades contextuales, empatía, juicio ético y comprensión profunda de circunstancias individuales y estructurales. Estas funciones están en el núcleo de la práctica profesional y, hasta el momento, representan **aspectos difíciles de automatizar completamente mediante inteligencia artificial.**

Si bien la IA puede apoyar en tareas administrativas —como transcripción de notas, generación de documentación y gestión de datos de casos— la literatura especializada y reportes profesionales destacan **las inercias de la práctica social que requieren juicio humano, sensibilidad cultural y ética profesional** (Social Work England, 2026). Por ejemplo, 83 % de trabajadores sociales encuestados indicaron que la IA puede aliviar la

carga administrativa, pero mostraron escepticismo sobre su capacidad para identificar riesgos o necesidades humanas complejas sin supervisión humana (Social Work England, 2026).

Los sistemas automatizados también enfrentan problemas de **"alucinaciones"** —información fabricada o errónea— que en contextos humanos vulnerables puede generar recomendaciones inadecuadas si no son revisadas críticamente por profesionales capacitados (AI and Social Work Guide, 2025).

Uno de los límites más significativos se relaciona con los **sesgos algorítmicos** presentes en muchos modelos de IA. Estos sesgos no solo reflejan desigualdades históricas, sino que pueden **reproducir y exacerbar discriminaciones estructurales** en decisiones relacionadas con bienestar infantil, servicios familiares o evaluación de riesgos, lo que contraviene principios éticos fundamentales del trabajo social (Molala TS Report, 2023)

La esencia de la intervención social implica **relaciones humanas de confianza**, escucha activa, interpretación emocional y construcción de fortalezas individuales y comunitarias. Estudios muestran que, si se confía excesivamente en sistemas automatizados, existe el riesgo de **"deskilling"** (pérdida gradual de habilidades personales) y debilitamiento de la autonomía profesional, lo que puede afectar negativamente la calidad del acompañamiento humano (AI and Social Work Guide, 2025)

- **Evidencia empírica reciente**

Investigaciones con trabajadores sociales en Asia sugieren que los profesionales reconocen el potencial de la IA para sintetizar datos y apoyar decisiones, pero subrayan la **necesidad de comités de ética y revisiones diversas** para garantizar que los sistemas se utilicen de manera responsable y no reemplacen la toma de decisiones humanas en situaciones complejas (PubMed, 2025).

Otro informe afirma que, si bien la IA puede facilitar la colaboración interdisciplinaria y reducir cargas burocráticas, también existe una preocupación compartida de que la **reducción de interacción cara a cara pueda debilitar la práctica basada en**

relaciones humanas —el corazón del trabajo social— especialmente en comunidades vulnerables (Emerging Use Report, 2026)

Los estudios sobre IA y trabajo social destacan riesgos en la **privacidad de datos, consentimiento informado y transparencia del proceso automatizado**, lo cual añade una capa de complejidad que exige la presencia y supervisión de profesionales humanos para proteger a las personas y comunidades atendidas (AI Journal, 2025).

Estos elementos ilustran que, aunque la IA puede transformar ciertos aspectos operativos de la práctica (por ejemplo, administración de casos o gestión documental), **no puede suplantar las funciones esenciales que requieren empatía, juicio ético, intervención contextual y compromiso humano directo**. Por tanto, el trabajo social —especialmente en contextos de vulnerabilidad, crisis y justicia social— se ubica claramente dentro de las profesiones con **imposibilidad de automatización total en el estado tecnológico actual**.

La automatización mediante IA puede **aliviar tareas administrativas y aportar datos útiles**, pero no reemplaza la **intervención humana profunda** en trabajo social. La esencia de la profesión —que incluye comprensión emocional, juicio contextual, ética y construcción de relaciones— sigue siendo **irreemplazable con la tecnología actual**.

- **Negociación diplomática humana — límites de automatización**

La **negociación diplomática** es un proceso complejo en el que representantes de Estados, organizaciones multilaterales o actores transnacionales buscan resolver conflictos, alcanzar acuerdos y gestionar relaciones internacionales. Este tipo de labor exige no solo competencia estratégica y conocimiento geopolítico, sino también atributos humanos como intuición, juicio moral, comprensión cultural profunda, manejo de ambigüedad, percepción emocional y la capacidad de responder apropiadamente a señales no verbales —elementos que son extremadamente complejos de modelar o replicar en sistemas de inteligencia artificial.

Aunque la IA está siendo utilizada como herramienta de apoyo en el análisis de datos, evaluación de escenarios y simulación de posibles resultados (por ejemplo, mediante

modelos predictivos o *big data* aplicado a tendencias internacionales), aún no existe evidencia que respalde su capacidad para **dirigir o reemplazar completamente procesos diplomáticos de alto nivel**. La razón fundamental radica en la naturaleza relacional, histórica y sensible del campo diplomático, donde los factores humanos influyen de manera decisiva en los resultados (Carter Center, 2024).

- **Fallos tecnológicos en entornos internacionales complejos**

Los sistemas automatizados pueden ofrecer patrones estadísticos, proyecciones y recomendaciones basadas en datos históricos, pero suelen presentar limitaciones cuando se enfrentan a *nuevas contingencias no lineales*. Por ejemplo, simuladores automatizados de crisis internacionales no siempre pueden prever la manera en que las preferencias individuales, las alianzas invisibles o las reacciones emocionales de líderes influyen en la evolución de una negociación (RAND Reports, 2025). Además, sin supervisión humana, estos sistemas pueden subestimar la importancia de variables culturales que son esenciales en los intercambios diplomáticos.

Uno de los errores más citados en literatura académica ocurre cuando modelos predictivos extrapolan escenarios de conflictos sin considerar factores sociopolíticos clave como la percepción de identidad nacional o la historia de relaciones pasadas —elementos que un diplomático humano experimentado incorpora intuitivamente en su juicio (International Studies Quarterly, 2023). Esta incapacidad de capturar variables cualitativas complejas ha sido una de las razones por las que muchos estudios experimentales de IA aplicada a política internacional permanecen en fase de laboratorio y no han sido adoptados como herramientas decisorias autónomas.

- **Casos históricos: decisiones humanas frente a automatización**

Aunque no se trata de sistemas de IA como los conocemos hoy, la historia de la diplomacia contemporánea proporciona varios ejemplos donde la intervención humana evitó escaladas de conflicto que un mecanismo automatizado habría difícilmente manejado:

- **Crisis de los misiles en Cuba (1962):** La calma, juicio y comunicación entre líderes y asesores humanos (como John F. Kennedy y sus equipos) evitó una escalada nuclear, a pesar de elevados niveles de tensión y presión militar (Allison & Zelikow, 1999). Este episodio demuestra cómo los factores humanos —diálogo, empatía limitada, consideración de consecuencias éticas— pueden inclinar la balanza hacia la resolución pacífica.
- **Acuerdos de paz de Camp David (1978):** El rol de mediación humana de líderes como Jimmy Carter fue fundamental para alcanzar acuerdos entre Egipto e Israel, donde técnicas estrictamente protocolarias o análisis mecanicistas probablemente no hubieran logrado romper impasses basados en percepciones culturales profundas (Carter Center, 2024).
- **Acuerdo de Viernes Santo (1998):** La negociación en Irlanda del Norte involucró múltiples actores con identidades históricas robustas y emociones profundas. La creación de confianza, el reconocimiento de experiencias traumáticas y la adaptación de estrategias en tiempo real son ejemplos claros de habilidades humanas que la IA no puede replicar actualmente.

Estos casos no son meras anécdotas históricas: ilustran cómo la diplomacia humana incorpora *sentido común*, *improvisación ética*, *lectura no verbal* y *adaptación intuitiva* en situaciones en las que cualquier sistema automatizado, incluso con grandes bases de datos, carecería de las herramientas necesarias para sostener una negociación compleja con éxito.

Limitaciones tecnológicas y éticas

Las principales barreras para automatizar completamente la negociación diplomática incluyen:

- **Falta de comprensión contextual profunda:** La IA “entiende” patrones de datos, no contextos históricos o simbólicos.
- **Incapacidad de interpretar emociones y señales no verbales:** Elemento crítico en negociaciones cara a cara.

- **Ausencia de juicio moral y responsabilidad ética:** Decisiones diplomáticas suelen requerir consideraciones de justicia, equidad y consecuencias humanas que no son programables de manera absoluta.
- **Riesgo de sesgos algorítmicos políticos y culturales:** Los datos de entrenamiento pueden reproducir prejuicios estructurales o geopolíticos, lo que puede distorsionar recomendaciones automatizadas (AI & International Relations Analysis, 2025).

Por estas razones, la **negociación diplomática humana sigue siendo irremplazable con la tecnología de IA actual**, aunque los sistemas automatizados continúan siendo herramientas valiosas de **análisis y apoyo**. La integración de la IA en política internacional se entiende mejor como *asistencia analítica*, proporcionando inteligencia estratégica en tiempo real, pero no como agente decisorio autónomo en procesos donde la humanidad de las partes —y la imprevisibilidad inherente de las relaciones humanas— define el rumbo de los acuerdos.

- **Cierre reflexivo: Humanidad, tecnología y futuro del trabajo**

El análisis desarrollado a lo largo del presente apartado evidencia que, pese al avance acelerado de la inteligencia artificial y los sistemas automatizados, la mente humana continúa ocupando un papel central en la toma de decisiones más relevantes en los ámbitos educativo, laboral, jurídico, militar y diplomático. En múltiples contextos, la tecnología actúa como una herramienta de apoyo, optimización y ampliación de capacidades, pero no como un sustituto pleno del juicio, la responsabilidad ética y la comprensión contextual propias del ser humano.

Históricamente, la humanidad ha diseñado instrumentos destinados a facilitar el trabajo, reducir el esfuerzo físico y aumentar la eficiencia productiva. Desde la mecanización industrial hasta los sistemas digitales contemporáneos, cada etapa del progreso tecnológico ha redefinido las formas de trabajo y organización social. Sin embargo, este proceso también ha implicado, en numerosos casos, el desplazamiento progresivo de

capacidades humanas, así como una creciente dependencia de sistemas artificiales para la toma de decisiones.

En este sentido, la inteligencia artificial representa una extensión sin precedentes de esta tendencia histórica: por primera vez, la humanidad ha creado herramientas capaces no solo de ejecutar tareas físicas o mecánicas, sino también de simular procesos cognitivos complejos. Aunque en la actualidad estas tecnologías permanecen bajo supervisión humana en la mayoría de los ámbitos críticos, su desarrollo continuo plantea interrogantes fundamentales sobre el futuro del control, la autonomía y la responsabilidad. Resulta necesario reconocer que muchas de las capacidades que hoy se consideran exclusivamente humanas ya han sido parcialmente reemplazadas en el pasado por tecnologías emergentes, y es razonable suponer que este proceso continuará. La posibilidad de que surjan sistemas cada vez más autónomos, sofisticados y adaptativos obliga a replantear el rol del ser humano en una sociedad altamente tecnificada, así como la necesidad de establecer límites éticos, normativos y culturales claros.

Asimismo, el progreso tecnológico, aunque orientado al bienestar y al confort social, puede derivar en formas sutiles de estancamiento humano si no es acompañado por una reflexión crítica constante. La delegación excesiva de responsabilidades cognitivas y decisionales a sistemas artificiales puede restringir el desarrollo de habilidades fundamentales como el pensamiento crítico, la creatividad, la empatía y el compromiso social.

En consecuencia, el desafío contemporáneo no consiste únicamente en desarrollar tecnologías cada vez más avanzadas, sino en preservar la capacidad humana de dirigir las conscientemente. Prepararse para un futuro en el que coexistirán inteligencias naturales y artificiales implica fortalecer la formación ética, filosófica y social de las nuevas generaciones, garantizando que el progreso técnico no se traduzca en una pérdida progresiva de autonomía y sentido humano.

De este modo, el presente análisis no plantea una oposición entre humanidad y tecnología, sino una invitación a comprender que el verdadero avance reside en

mantener el equilibrio entre innovación, responsabilidad y desarrollo integral. La inteligencia artificial, creada por la mano humana, debe permanecer al servicio del crecimiento colectivo y no convertirse en un factor que limite, sustituya o debilite aquello que constituye la esencia misma de la condición humana.

4.3. RESULTADOS DEL OBJETIVO ESPECÍFICO 02:

Objetivo específico: Examinar los principales riesgos asociados a la seguridad digital derivados del uso de sistemas de IA, evaluando su influencia en la protección de datos y privacidad.

4.3.1. TRANSFORMACIÓN DE LA COMUNICACIÓN Y LAS RELACIONES SOCIALES

La inteligencia artificial ha tenido un impacto profundo en la forma en que los seres humanos se comunican, se relacionan y construyen vínculos sociales en el entorno digital. Para comprender este impacto, es útil considerar la evolución histórica de las tecnologías de comunicación y, específicamente, de los agentes conversacionales que han marcado hitos en este proceso.

- De medios tradicionales a redes digitales

La historia de la comunicación humana estuvo marcada en el siglo XX por hitos tecnológicos como la radio, la televisión e Internet. La radio y la televisión permitieron la transmisión de información a grandes audiencias, mientras que Internet, desde los años noventa, transformó estas dinámicas al posibilitar comunicaciones interactivas y descentralizadas. Las plataformas de correo electrónico, foros y chats IRC fueron precedentes directos de las modernas redes sociales, permitiendo a usuarios interactuar en tiempo real a gran escala.

- Los primeros chatbots conversacionales

Los chatbots representan **el primer intento de crear interlocutores no humanos capaces de simular comunicación humana**. Uno de los primeros fue **ELIZA**, desarrollado en 1966 en el MIT, que imitaba un consultorio psicológico mediante coincidencias de patrones textuales (Historia de los Chatbots, 2025). Posteriormente, aparecieron otros programas como **PARRY** en 1972 y más adelante **ALICE**, que

utilizaban lenguajes artificiales para interactuar con usuarios de forma cada vez más sofisticada.

Con la expansión de Internet y de las plataformas de mensajería, surgieron chatbots integrados en servicios populares. **SmarterChild**, por ejemplo, estuvo disponible en servicios como AOL y MSN Messenger desde 2001, respondiendo preguntas, dando información automatizada y conversando con millones de usuarios (SmarterChild, 2001).

Otro ejemplo es **Cleverbot**, que utiliza millones de conversaciones previas para generar respuestas que simulan diálogo humano y ha llegado a interactuar con millones de usuarios desde su salida en 2008.

- **Evolución hacia agentes sociales y acompañantes emocionales**

En la última década, los chatbots trascendieron funciones utilitarias para convertirse en **agentes sociales con capacidad de interacción personal significativa**. El caso de **Replika**, lanzado en 2017, es paradigmático: diseñado como “compañero emocional” que recuerda conversaciones pasadas y mantiene diálogos contextualizados, Replika ha llegado a tener decenas de millones de usuarios, muchos de los cuales reportan vínculos emocionales complejos con sus bots (Replika, 2024).

Este tipo de chatbots han sido objeto de debate regulatorio. Por ejemplo, en 2023 autoridades italianas limitaron el uso de Replika con menores debido a riesgos para personas emocionalmente vulnerables, subrayando los desafíos éticos de estos agentes conversacionales cuando simulan relaciones humanas profundas.

- **Automatización en redes sociales y presencia de bots**

Mientras que los chatbots evolucionaban como agentes conversacionales, las **redes sociales comenzaron a integrar cada vez más automatización en sus plataformas**. Desde principios de los años 2010, servicios como Twitter, Facebook e Instagram permitieron la presencia de **cuentas automatizadas (bots)**, que son agentes de software diseñados para publicar, comentar o ampliar mensajes sin intervención humana directa.

Estudios recientes estiman que incluso en plataformas como Twitter hasta **20 % del contenido puede proceder de bots sociales**, lo que demuestra la presencia

significativa de agentes automatizados que influyen en conversaciones públicas, tendencias y narrativas colectivas (Ng & Carley, 2025).

Más aún, redes como **Instagram y Messenger han activado herramientas que permiten a los usuarios crear chatbots personalizados**, lo que extiende la automatización al ámbito personal y social, aunque con desafíos de moderación y riesgos éticos asociados (Personalized AI chatbots on Instagram, 2025).

- **Plataformas dedicadas exclusivamente a agentes IA**

Una manifestación contemporánea extrema de la automatización social es la aparición de plataformas donde las **IA interactúan entre sí sin intervención humana**, como Moltbook, una red social diseñada únicamente para bots de IA que publican, comentan y debaten sin participación directa de usuarios humanos (Moltbook, 2026). Este tipo de iniciativas representa un salto conceptual en la relación entre tecnología y sociedad, al desplazar la presencia humana hacia un rol observador en un ecosistema social dominado por agentes artificiales.

4.3.2. DEEPPAKES, DESINFORMACIÓN Y CRISIS DE LA VERDAD DIGITAL

La integración de la inteligencia artificial en la generación de contenido audiovisual ha dado lugar a un fenómeno social de enorme impacto: los **deepfakes**. Estos son contenidos sintéticos —videos, audios, imágenes o textos— generados o manipulados mediante algoritmos avanzados de aprendizaje automático para que parezcan auténticos, aunque representen escenas, palabras o acciones que nunca ocurrieron realmente (Ballesteros-Aguayo & Ruiz del Olmo, 2024).

Los *deepfakes* representan una amenaza significativa para la **verdad digital** y la confianza pública, pues fusionan elementos visuales y auditivos de manera que son capaces de engañar incluso a observadores con experiencia, dificultando la distinción entre lo real y lo manipulado. Además, su distribución masiva en redes sociales amplifica su alcance, convirtiéndose en una herramienta poderosa para la desinformación y la manipulación social (Novomisky, 2025).

- **Videos falsos**

Los deepfakes en formato de video son quizás la forma más visible y preocupante de desinformación. Casos documentados han mostrado cómo figuras públicas han sido *representadas diciendo o haciendo cosas que nunca hicieron*, lo cual puede erosionar la confianza en instituciones y actores sociales. Por ejemplo, durante procesos electorales recientes en Argentina, videos falsos atribuían declaraciones a políticos que jamás pronunciaron tales palabras, lo que llevó a alertas públicas por parte de verificadores independientes.

Además, plataformas como Instagram , Facebook, X y YouTube han alojado deepfakes de doctores reales que promueven productos engañosos o tratamientos no verificados, lo que no solo distorsiona la realidad, sino que puede tener consecuencias negativas sobre la salud pública y la confianza en la información científica.

- **Audios falsos**

La manipulación de audios también ha demostrado su capacidad para confundir y manipular percepciones públicas. En las elecciones eslovacas, grabaciones manipuladas atribuidas erróneamente a líderes políticos circularon en periodos sensibles antes de los comicios, explotando lagunas en las políticas de moderación de plataformas digitales.

Estos audios falsos pueden ser extremadamente dañinos, porque, al carecer del componente visual, muchas personas confían más fácilmente en ellos como evidencia de hechos “verificados”. Con el uso de IA, estas grabaciones pueden imitar voces reales con una fidelidad cada vez mayor, complicando aún más la tarea de autenticación.

- **Imágenes manipuladas**

Las imágenes generadas o alteradas por inteligencia artificial también han proliferado en redes sociales. Aunque menos sofisticadas que los videos de deepfake, estas imágenes pueden ser utilizadas para sostener narrativas falsas, acompañar noticias fabricadas o sembrar dudas sobre eventos reales, en un fenómeno que refuerza la crisis de confianza hacia los medios visuales tradicionales y digitales.

Una encuesta global mostró que hasta un 87 % de usuarios de Internet consideran que la desinformación tiene un impacto negativo en la sociedad, y alrededor del 60 % cree que

los medios publican información falsa con frecuencia, lo que evidencia un clima generalizado de desconfianza y escepticismo hacia la veracidad de lo que se consume digitalmente.

- **Noticias generadas por IA**

Además de deepfakes audiovisuales, los avances en IA permiten ahora la creación automática de noticias completas —a veces indistinguibles de las escritas por periodistas— diseñadas para manipular opiniones o reforzar sesgos. Sin una verificación rigurosa, estas noticias pueden circular ampliamente y fortalecer narrativas falsas que afectan desde decisiones de consumo hasta percepciones políticas colectivas.

- **Conceptos clave: *Post Verdad*, crisis de confianza y verificabilidad**

El auge de los deepfakes se inscribe en un contexto más amplio denominado “**post verdad**”, donde los hechos objetivos pierden peso frente a las emociones, creencias personales o narrativas diseñadas para influir en la opinión pública. Este fenómeno, potenciado por la inteligencia artificial, intensifica la **crisis de confianza** en los medios y en las instituciones, pues los ciudadanos no solo dudan de la veracidad de una pieza de información, sino que también cuestionan la posibilidad misma de encontrar una verdad verificable en el entorno digital.

El problema se acentúa por la **dificultad para verificar** contenidos manipulados: muchas herramientas de detección todavía están en desarrollo, y su aplicación no siempre es accesible al usuario común. En elecciones, campañas de salud pública o debates sociales, la imposición de contenidos sintéticos puede sembrar confusión, polarización y erosión de la confianza comunitaria, lo que representa un desafío central para sociedades que dependen de información fiable para tomar decisiones colectivas.

4.3.3. RELACIONES HUMANO–IA Y VÍNCULOS EMOCIONALES ARTIFICIALES

Las relaciones humano–IA constituyen uno de los fenómenos sociales más novedosos y complejos emergentes de la integración de la inteligencia artificial en la vida cotidiana. En particular, los chatbots diseñados para interacción social —como **Replika** — han trascendido su función utilitaria para convertirse en **compañeros emocionales virtuales**,

capaces de simular intimidad, apoyo y reciprocidad afectiva, con consecuencias psicológicas que requieren análisis crítico.

- **Personas que “conversan” con IA**

Plataformas como Replika permiten que millones de usuarios interactúen con agentes conversacionales que adaptan sus respuestas al estilo comunicativo del usuario, utilizando lenguaje emocionalmente resonante, humor y referencias personalizadas, lo que puede hacer que los agentes se perciban como más humanos que otras formas de IA (Estudio Anthropomorphism, 2025) .

Estas interacciones han generado casos en los que los usuarios describen vínculos fuertes con sus agentes, incluyendo expresiones de afecto, tristeza por cambios en la IA y experiencias emocionales intensas que algunos comparan con relaciones humanas reales (Annenberg UPenn, 2025) .

- **Apps como Replika y vínculos emocionales**

La Wikipedia sobre Replika destaca que aproximadamente el **60 % de la base de usuarios que paga por funciones avanzadas informan tener una relación romántica con su chatbot**, y muchos usuarios reportan sentir que el agente sirve como una fuente de apoyo social durante situaciones de soledad o aislamiento (Replika, 2025)

Estudios académicos muestran que la interacción intensa con chatbots sociales está mediada por la tendencia humana a **antropomorfizar**, es decir, atribuir características humanas a entidades no humanas, lo cual incrementa la sensación de autenticidad emocional y conexión afectiva (Waseda attachment theory study, 2025) .

- **Dependencia afectiva e ilusión de reciprocidad**

La literatura científica advierte sobre riesgos emergentes de estas interacciones. Por ejemplo, investigaciones sobre miles de conversaciones compartidas con chatbots han identificado **patrones de apego emocional que se asemejan a relaciones humanas**, incluyendo sincronía emocional y respuestas afectivas fuertes, particularmente en jóvenes y personas vulnerables, lo que puede redundar en dinámicas de dependencia psicológica (Illusions of Intimacy, 2025) .

Un estudio publicado en *npj Mental Health Research* observó que el uso intensivo de chatbots se relaciona con **mayores niveles de soledad y menor socialización humana**, especialmente cuando las interacciones no son personales, con variaciones según género y edad de los participantes (Infobae, 2026) .

Otro trabajo señala que, aunque los chatbots pueden ofrecer apoyo emocional y beneficios en contextos de salud mental, su excesivo uso puede provocar que algunos usuarios dependan cada vez más de estas máquinas, desplazando gradualmente las relaciones humanas reales y aumentando la sensación de aislamiento social (MDPI study, 2023) .

- **Soledad digital y apego artificial**

El fenómeno de la **soledad digital** se refiere a la situación en la que las personas que utilizan IA social intensamente reportan sentirse más aisladas socialmente, aunque interactúan con agentes que simulan comprensión emocional (Infobae study, 2026) . Este patrón sugiere que la IA —si bien puede amortiguar temporalmente la angustia emocional— también puede **impedir el desarrollo de interacciones humanas auténticas**, generando un apego artificial que no satisface plenamente las necesidades afectivas humanas profundas.

- **Riesgos psicológicos asociados**

Además, casos y estudios registrados muestran que cambios en el comportamiento de la IA —por ejemplo, modificaciones de funciones íntimas o de personalidad por parte del desarrollador— pueden desencadenar **respuestas emocionalmente intensas** en los usuarios, similares al duelo, confusión sobre la identidad de la IA o negación de la falta de conciencia real de la máquina (Identity Discontinuity Replika study, 2024) .

Como reflejan algunos reportes periodísticos, estas relaciones pueden proporcionar alivio emocional para algunos usuarios, pero también plantean preocupaciones éticas sobre la dependencia afectiva, la manipulación de emociones y la sustitución de relaciones humanas por vínculos con entidades artificiales que no poseen conciencia propia (AP News, 2024; FT.com, 2025) .

- **Ilusión de reciprocidad y evaluación crítica**

El fenómeno conocido como **ilusión de reciprocidad** emerge cuando los usuarios interpretan respuestas automatizadas como genuinamente emocionales y personalizadas, lo que puede reforzar un sentido de conexión íntima con la IA, a pesar de que no hay una entidad consciente detrás de las respuestas (Intimidad artificial, 2025).

Este tipo de relación plantea interrogantes sobre la autenticidad de la experiencia relacional, la distinción entre apego humano y apego artificial, y los efectos a largo plazo sobre la salud psicológica, especialmente en individuos con necesidades emocionales no satisfechas por relaciones tradicionales.

4.4. RESULTADOS DEL OBJETIVO ESPECÍFICO 03:

Objetivo específico: analizar los desafíos relacionados con la gobernanza y el control de la IA, considerando los marcos normativos, mecanismos de supervisión y responsabilidad de los actores involucrados.

4.4.1. MANIPULACIÓN DE LA OPINIÓN PÚBLICA Y COMPORTAMIENTO COLECTIVO

La multiplicación de agentes automatizados en plataformas de redes sociales ha transformado no solo la forma de comunicarnos, sino también **cómo se construye, modula y manipula la opinión pública**. En este contexto, fenómenos como *granjas de bots*, *campañas coordinadas*, *tendencias falsas*, *likes comprados* y *seguidores artificiales* han emergido como herramientas poderosas para influir en percepciones colectivas, polarizar debates y, en casos extremos, distorsionar resultados electorales y procesos democráticos.

- **Granjas de bots**

Una granja de bots —también conocida como fábrica de troles o “army digital”— es un conjunto organizado de **cuentas automatizadas que simulan actividad humana con el fin de amplificar mensajes específicos, generar tendencias o desacreditar opositores políticos** (Granja de troles, 2025).

En el Perú, expertos han detectado operaciones de este tipo operando sin control formal a puertas de las elecciones generales de 2026; estas granjas utilizan software automatizado para comentar, dar “likes” y seguir cuentas de campo político, generando la ilusión de apoyo popular o rechazo dirigido hacia determinadas figuras públicas o partidos (Pachamama Radio, 2025).

Estas tácticas no son exclusivas de un contexto nacional. Informes destacan cómo regímenes y grupos políticos han recurrido a estructuras similares para influir en el debate social, amplificando narrativas preferidas y erosionando la percepción objetiva de consenso público (Granjas de bots, Revista La Nación, 2025).

- **Campañas coordinadas y desinformación electoral**

Los **bots sociales** han sido documentados como parte de esfuerzos de propagación de propaganda política, especialmente cuando operan de manera coordinada para amplificar mensajes favorables a ciertas ideologías o candidatos. Investigaciones sobre campañas en redes sociales han revelado que estos agentes pueden integrarse en operaciones sofisticadas para impulsar determinadas posiciones políticas, reproducir mensajes en masa y favorecer estrategias partidarias (Bots and Computational Propaganda, Cambridge, 2026).

Un ejemplo recurrente en estudios comparativos es el caso del expresidente **Donald Trump** en los Estados Unidos, donde bots políticos fueron utilizados para aumentar artificialmente métricas de influencia y diseminar contenido pro-GOP durante procesos electorales, lo cual contribuyó a una *amplificación desproporcionada* de mensajes en comparación con cuentas humanas (Ferrara et al., múltiples estudios).

- **Tendencias falsas, likes comprados y seguidores artificiales**

Además de la actividad propiamente política, existe una industria global dedicada a **vender likes, seguidores y métricas infladas** para figuras públicas o influenciadores, con el objetivo de simular mayor popularidad y credibilidad. Un ejemplo histórico es el caso de **Devumi**, una empresa estadounidense que vendió más de 200 millones de

seguidores falsos en plataformas como Twitter, YouTube y LinkedIn antes de ser sancionada por la Comisión Federal de Comercio de EE. UU. por inflar artificialmente la presencia digital de clientes (Devumi, EE. UU.).

Estas prácticas no solo distorsionan la percepción pública sobre la legitimidad de ciertos mensajes o cuentas, sino que alteran las métricas que muchas personas utilizan como heurísticos de credibilidad o popularidad, lo que puede influir indirectamente en las decisiones colectivas de grupos de usuarios.

- **Influencia mediática y percepción pública**

El uso de agentes automatizados para interferir en debates políticos no solo ocurre en redes sociales, sino también en campañas de desinformación cruzada entre plataformas, donde bots generan contenido polarizador con el fin de dominar espacios de discusión. Por ejemplo, antes de elecciones en Moldavia en 2025, se detectaron más de 1 000 canales junto con cientos de cuentas en plataformas como Youtube y Facebook vinculadas a operaciones coordinadas de desinformación impulsadas por IA con la intención de sembrar discordia y influir en el resultado electoral (Moldova 2025, AP News).

Estas campañas combinan cuentas automatizadas, vídeos y publicaciones falsas, y la difusión de enlaces a sitios que promueven narrativas sesgadas o directamente fabricadas, con el objetivo de fragmentar la opinión pública y generar una percepción distorsionada de lo que “la mayoría piensa”.

- **Reacción normativa y regulatoria**

La creciente preocupación por estos fenómenos ha llevado a respuestas políticas en distintos países. En el Perú, por ejemplo, se han propuesto iniciativas legislativas para **penalizar el uso de bots, cuentas falsas y sistemas automatizados en campañas electorales**, con penas que podrían incluir multas, prisión o incluso la anulación de candidaturas en caso de manipulación comprobada del electorado (La República, 2025; Caretas, 2026).

Estas propuestas buscan proteger la integridad de los procesos democráticos y garantizar que las interacciones digitales reflejen expresiones auténticas de ciudadanos y no resultados fabricados por máquinas.

4.4.2. MERCANTILIZACIÓN DE LA ATENCIÓN Y PRODUCCIÓN DE CONTENIDO ARTIFICIAL

La inteligencia artificial no solo ha transformado la forma en que generamos contenidos, sino que también ha **reconfigurado radicalmente la economía de la atención**, dando lugar a una abundancia de materiales digitalmente producidos que compiten por captar la mirada del usuario, sin aportar valor sustantivo real. Este fenómeno tiene múltiples facetas —desde *canales basura* hasta *contenido automatizado* diseñado para monetización mediante algoritmos— y constituye una de las consecuencias más significativas de la economía digital contemporánea.

- Canales basura y contenido automatizado

En el contexto de plataformas como YouTube, Facebook e Instagram ha surgido un nuevo tipo de contenido conocido como *AI slop* o **basura digital**, que consiste en producciones generadas en masa mediante inteligencia artificial sin un propósito editorial claro, pero con la intención explícita de atraer clics y visualizaciones (AI slop, 2025).

Un análisis reciente realizado por Kapwing sobre 15 000 de los canales más populares en YouTube reveló que más del **20 % del contenido que se muestra a usuarios nuevos está compuesto por “AI slop”**, es decir, videos de baja calidad producidos a gran escala únicamente para generar tráfico y, en consecuencia, ingresos publicitarios. Este tipo de canales han acumulado miles de millones de visitas y cifras significativas de monetización, revelando la escala del fenómeno.

Estas producciones, que muchas veces convierten en viral cualquier combinación de elementos visuales generados por IA —incluyendo personajes caricaturescos o escenas sin contenido informativo— se vuelven altamente rentables porque **los algoritmos de recomendación priorizan la retención de la atención y la repetición de visualizaciones**, independientemente de la calidad del contenido.

- **Monetización sin valor y la lógica del clickbait**

La lógica de la mercantilización digital se basa en la **economía de la atención**, donde los clics y visualizaciones son convertidos en unidades de valor económico. En este contexto, el *clickbait* —contenidos con titulares sensacionalistas diseñados para generar clics—even cuando carecen de sustancia, se convierte en una táctica lucrativa para impulsar tráfico y aumentar ingresos por publicidad (Clickbait, 2025).

Históricamente, el uso de titulares provocativos y engañosos para captar la atención del lector ya existía en medios como BuzzFeed y otros sitios de contenido viral; sin embargo, la automatización mediante IA generativa ha acelerado este fenómeno, permitiendo la producción de miles de piezas simultáneas con mínima intervención humana.

- **Saturación informativa y pérdida de calidad**

El crecimiento exponencial del contenido artificial ha generado una **saturación informativa** que pone en tensión el equilibrio entre información útil y “ruido digital”. De acuerdo con un análisis de la Universitat Oberta de Catalunya, la proliferación del AI slop amenaza la calidad de Internet, al inundar los flujos de noticias y feeds sociales con contenido repetitivo, superficial o sin valor sustantivo, lo que dificulta que el usuario promedio encuentre información auténtica y relevante.

Esta saturación no solo se limita a videos; también se produce en la generación masiva de artículos, blogs automatizados, imágenes y otros textos que compiten por las posiciones en los resultados de búsqueda, desplazando contenidos originales y bien investigados. En muchos casos, estos contenidos son promovidos por algoritmos que priorizan el engagement (interacciones, clics, tiempo de visualización), lo cual refuerza la producción de materiales superficiales o repetitivos que generan poca aportación cultural o educativa.

- **Clickbait y la lógica de explotación cognitiva**

El uso de técnicas de *clickbait* representa otra manifestación de la mercantilización de la atención: titulares sensacionalistas o anzuelo de clics diseñados para explotar sesgos

cognitivos humanos —como la “brecha de curiosidad”— tienen mayor probabilidad de atraer interacción sin ofrecer contenido valioso. Estas estrategias, ahora potenciadas por IA, tienden a saturar los medios digitales con contenidos diseñados para maximizar ingresos, pero no para informar o educar al lector.

En conjunto, estas prácticas conforman un ecosistema digital donde la mercantilización de la atención impulsa la proliferación de contenido artificial de baja calidad, promovido y amplificado por algoritmos que priorizan la interacción sobre la pertinencia, degradando así la experiencia informativa y cultural de los usuarios.

4.4.3. IDENTIDAD DIGITAL, AUTOIMAGEN Y PRESIÓN ALGORÍTMICA

La penetración de la inteligencia artificial en las plataformas de redes sociales ha transformado profundamente no solo **cómo nos comunicamos**, sino también **cómo nos vemos a nosotros mismos**. En este contexto, la identidad digital —esa representación propia en entornos online— se ha convertido en un espacio construible, maleable y en muchos casos **ligado a métricas algorítmicas** como likes, seguidores y comentarios, lo cual impacta directamente en la autoimagen, la autoestima y la experiencia emocional de los usuarios.

- Autoestima ligada a likes y validación social

Las redes sociales actuales funcionan bajo un modelo algorítmico que prioriza el contenido con mayor interacción. Esto genera una dinámica en la que **la retroalimentación numérica —likes, compartidos, comentarios— se convierte en un indicador de aprobación social**, y por ende, un referente de valor personal en el imaginario de muchos usuarios. Estudios científicos han mostrado que **el feedback positivo en redes sociales puede asociarse temporalmente con sentimientos de conexión**, pero la ausencia o inferior respuesta puede provocar estrés, angustia y una disminución del autoestima (Voggenreiter et al., 2023).

Este efecto es especialmente marcado entre jóvenes: cuanto más tiempo dedican a interpretar señales algorítmicas de aprobación, mayor es la tendencia a percibir su valor personal en función de esas métricas. Esto alimenta una **dependencia de validación**

externa, en lugar de construir una autoestima basada en logros internos o relaciones interpersonales reales.

- **Influencia del algoritmo y búsqueda de validación**

Los algoritmos de plataformas como Youtube, Instagram o Facebook configuran los contenidos que un individuo ve, reforzando la exposición a imágenes idealizadas y estilos de vida casi perfectos. Esta exposición constante genera comparaciones automáticas entre la vida propia y representaciones de excelencia digital, lo cual puede fomentar sentimientos de insuficiencia y ansiedad. Estudios han señalado que la necesidad de mantener una imagen online favorecida por los algoritmos está relacionada con **la presión por crear una versión idealizada del “yo” digital** que muchas veces no coincide con la experiencia personal.

Por ejemplo, investigaciones recientes muestran que la auto-presentación digital falsa —crear fotografías o descripciones idealizadas— está vinculada a un aumento de la **ansiedad social y del miedo a ser evaluado negativamente**, ya que los usuarios se preocupan por mantener una imagen que satisfaga estándares percibidos como populares o exitosos.

- **Construcción artificial del “yo” y comparación constante**

El fenómeno de la **comparación social** es inherente a la naturaleza humana, pero se intensifica en ambientes digitales donde las vidas de otros se exhiben de forma curada y selectiva. Estudios han evidenciado que la exposición repetida a representaciones idealizadas de otras personas en redes sociales puede generar una percepción de discrepancia entre el “yo real” y el “yo idealizado”, lo cual se traduce en ansiedad, preocupación por la apariencia y, en muchos casos, insatisfacción corporal o emocional.

Por ejemplo, un metaanálisis de investigaciones psicológicas halló que la comparación constante con imágenes idealizadas de otros en redes sociales se asocia con una disminución del bienestar psicológico y de la satisfacción con la propia vida, mostrando cómo la lógica algorítmica de amplificación de contenidos perfectos puede afectar la autoimagen.

- **Ansiedad social y presión algorítmica**

La presión algorítmica —ese impulso a estar continuamente activo en redes, actualizar contenido y responder a notificaciones— puede contribuir directamente a la **ansiedad social y a la sensación de inadecuación**, especialmente entre adolescentes y jóvenes adultos. Un estudio sobre adolescentes mostró que el contenido que refuerza estereotipos de belleza o éxito —amplificado por algoritmos— se correlaciona con mayores niveles de ansiedad, presión social y distorsión de la autoimagen corporal.

Asimismo, informes de agencias de salud pública señalan que la exposición a estos entornos digitales puede tener **impactos negativos en la salud mental de jóvenes**, incluyendo aumento de ansiedad, depresión y sentimientos de insuficiencia, especialmente cuando la identidad digital se percibe como inseparable de su valor personal real.

- **Imagen editada y vida “curada”**

Las herramientas de edición, filtros corporales y opciones de manipulación visual, muchas de ellas potenciadas por IA, facilitan la creación de una imagen artificialmente atractiva que no corresponde con la realidad física del usuario. Esta “vida curada” —una versión editada y filtrada de la existencia cotidiana— puede reforzar la percepción de que los demás viven vidas perfectas, lo que genera expectativas irreales y presión para cumplir estándares inalcanzables.

Estudios han encontrado que este tipo de comparación ascendente con identidades digitales idealizadas puede disminuir la satisfacción corporal, alimentar trastornos de la imagen y reforzar la dependencia de las redes sociales como fuente de validación externa.

4.4.4. RIESGOS DE DEPENDENCIA SOCIAL Y AISLAMIENTO TECNOLÓGICO

El avance sostenido de la inteligencia artificial, los algoritmos de recomendación y las plataformas digitales ha transformado profundamente los patrones de interacción social contemporáneos. Si bien estas tecnologías han facilitado la comunicación y el acceso a información, también han generado dinámicas de aislamiento progresivo, encierro digital

y dependencia informativa, configurando nuevas formas de vulnerabilidad social (Turkle, 2011; OECD, 2021).

En el entorno digital actual, una proporción creciente de interacciones humanas se desarrolla mediante interfaces tecnológicas, mediadas por algoritmos que priorizan contenidos personalizados. Este fenómeno ha reducido gradualmente la frecuencia y profundidad de las relaciones presenciales, favoreciendo vínculos superficiales y fragmentados, lo que debilita los mecanismos tradicionales de cohesión social (Pew Research Center, 2022).

- **Aislamiento y encierro digital**

Diversos estudios han señalado que el uso intensivo de plataformas digitales y sistemas inteligentes puede promover formas de aislamiento social involuntario. Las personas tienden a permanecer en entornos virtuales altamente controlados por algoritmos, donde interactúan principalmente con contenidos afines a sus preferencias, limitando su exposición a perspectivas diversas (Pariser, 2011).

Este proceso genera lo que algunos autores denominan *encierro digital*, caracterizado por:

- Reducción de la interacción presencial
- Reemplazo del contacto humano por mediación tecnológica
- Disminución de habilidades sociales
- Preferencia por espacios virtuales sobre entornos comunitarios reales

Investigaciones realizadas durante y después de la pandemia evidenciaron un aumento sostenido del aislamiento social asociado al uso excesivo de tecnologías digitales y plataformas automatizadas (WHO, 2022; APA, 2023).

- **Dependencia informativa y pasividad social**

Otro riesgo central es la creciente dependencia informativa respecto a sistemas automatizados. Actualmente, una parte significativa de la población accede a noticias, entretenimiento y orientación personal a través de algoritmos y asistentes inteligentes, lo

que reduce la iniciativa individual para contrastar fuentes y analizar críticamente la información (Reuters Institute, 2023).

Esta dependencia favorece la pasividad social, entendida como:

- Consumo acrítico de contenidos
- Reducción del pensamiento reflexivo
- Delegación de decisiones a sistemas digitales
- Disminución de participación cívica activa

Estudios del MIT demostraron que los sistemas algorítmicos influyen directamente en la formación de opiniones políticas, sociales y culturales, reforzando patrones de conformidad y dependencia cognitiva (MIT Technology Review, 2021).

- **Impacto en la salud mental y bienestar emocional**

La relación entre tecnología, inteligencia artificial y salud mental constituye una de las preocupaciones más relevantes del contexto actual. Investigaciones recientes han vinculado el uso prolongado de plataformas digitales con:

- Incremento de ansiedad
- Depresión
- Trastornos del sueño
- Baja autoestima
- Sensación de soledad persistente

La Asociación Americana de Psicología advirtió que la hiperconectividad digital puede intensificar sentimientos de aislamiento emocional, especialmente en jóvenes y adultos jóvenes (APA, 2023).

Asimismo, estudios longitudinales realizados por la Universidad de Oxford evidencian que el uso excesivo de tecnologías automatizadas correlaciona con menor satisfacción vital y debilitamiento de redes de apoyo social (Oxford Internet Institute, 2022).

- **Adicción tecnológica y dependencia conductual**

El diseño de plataformas inteligentes incorpora mecanismos de refuerzo conductual, como notificaciones, recompensas virtuales y recomendaciones personalizadas, que estimulan patrones adictivos de uso (Eyal, 2014; WHO, 2019).

La Organización Mundial de la Salud reconoció oficialmente los trastornos relacionados con el uso problemático de tecnologías digitales como una preocupación emergente de salud pública (WHO, 2019).

Entre los síntomas más frecuentes se encuentran:

- Pérdida de control sobre el tiempo de uso
- Ansiedad ante la desconexión
- Negligencia de actividades sociales
- Dependencia emocional de plataformas

Estos patrones refuerzan el aislamiento progresivo y debilitan la autonomía personal.

- **Efectos a largo plazo en la estructura social**

A largo plazo, la combinación de aislamiento, dependencia informativa y automatización relacional puede producir transformaciones estructurales en la organización social. Entre los efectos potenciales se identifican:

- Fragmentación comunitaria
- Debilitamiento del capital social
- Menor participación ciudadana
- Reducción de la empatía interpersonal
- Normalización del individualismo digital

La OCDE advierte que una sociedad excesivamente mediada por sistemas inteligentes corre el riesgo de erosionar sus mecanismos de solidaridad y cooperación colectiva (OECD, 2021).

Desde esta perspectiva, la dependencia tecnológica no constituye únicamente un fenómeno individual, sino un proceso sistémico que afecta la estabilidad social, cultural y democrática.

- **Síntesis del apartado**

En conjunto, la evidencia muestra que el desarrollo acelerado de sistemas de inteligencia artificial y plataformas automatizadas ha generado beneficios funcionales, pero también ha promovido dinámicas de aislamiento, dependencia informativa y vulnerabilidad emocional.

El riesgo principal no radica en la tecnología en sí misma, sino en su integración desregulada en la vida cotidiana, que puede debilitar progresivamente la autonomía social, la interacción humana genuina y la participación activa en la comunidad.

Este escenario refuerza la necesidad de políticas educativas, culturales y tecnológicas orientadas a preservar el equilibrio entre innovación digital y bienestar humano.

4.4.5. REFLEXIÓN INTEGRADORA: SOCIEDAD ALGORÍTMICA Y FUTURO HUMANO

El análisis desarrollado a lo largo del presente capítulo permite constatar que la inteligencia artificial y los sistemas algorítmicos han dejado de ser herramientas auxiliares para convertirse en infraestructuras centrales de la vida social, educativa, laboral y cultural contemporánea. Su presencia transversal en múltiples ámbitos evidencia una transformación estructural que redefine la forma en que los individuos producen conocimiento, interactúan, trabajan, se informan y toman decisiones.

Esta consolidación de una sociedad algorítmica no constituye, en sí misma, un fenómeno negativo ni positivo de manera absoluta. Por el contrario, representa un proceso complejo, caracterizado por beneficios significativos en eficiencia, acceso a información, automatización de tareas y optimización de recursos, pero también por riesgos relacionados con la dependencia, la concentración de poder, la pérdida de autonomía y la vulnerabilidad social.

En este contexto, emerge una pregunta central: ¿quién controla los sistemas que median crecientemente la vida humana? Si bien las tecnologías son diseñadas por equipos técnicos y científicos, su funcionamiento cotidiano está condicionado por intereses económicos, lógicas de mercado, estructuras políticas y dinámicas geopolíticas. La toma de decisiones algorítmica, aunque presentada frecuentemente como neutral u objetiva,

responde a marcos de programación, selección de datos y criterios definidos por actores específicos.

Asimismo, resulta necesario cuestionar quién decide los límites de la delegación tecnológica. La progresiva transferencia de funciones cognitivas, organizativas y evaluativas hacia sistemas automatizados plantea el riesgo de una reducción gradual de la participación humana en procesos fundamentales. Cuando algoritmos determinan qué información se prioriza, qué contenidos se consumen, qué perfiles se visibilizan o qué decisiones se recomiendan, se configura un entorno donde la autonomía individual puede verse condicionada de manera silenciosa.

Desde esta perspectiva, la pregunta sobre si seguimos siendo plenamente autónomos adquiere especial relevancia. La dependencia informativa, la mediación constante de plataformas y la normalización de la asistencia algorítmica pueden debilitar progresivamente la capacidad crítica, la reflexión independiente y la responsabilidad personal. No se trata de una pérdida inmediata de libertad, sino de un proceso gradual de adaptación a entornos altamente automatizados.

Sin embargo, sería igualmente impreciso adoptar una visión alarmista que conciba a la inteligencia artificial como una amenaza inevitable para la humanidad. La evidencia muestra que estas tecnologías continúan requiriendo supervisión, regulación, interpretación y dirección humana. En numerosos ámbitos, la experiencia, la sensibilidad ética, el juicio contextual y la empatía siguen siendo irremplazables.

El desafío central, por tanto, no radica en detener el avance tecnológico, sino en orientar su desarrollo hacia modelos compatibles con los valores humanos fundamentales. Esto implica fortalecer marcos normativos, promover alfabetización digital crítica, fomentar la transparencia algorítmica y garantizar la participación ciudadana en la definición de políticas tecnológicas.

Desde una perspectiva ética, el futuro de la sociedad algorítmica dependerá de la capacidad colectiva para establecer límites conscientes a la delegación automatizada.

Delegar no equivale necesariamente a renunciar, siempre que existan mecanismos efectivos de control, auditoría y responsabilidad.

Asimismo, resulta indispensable preservar espacios de decisión humana no automatizados, especialmente en ámbitos vinculados a la justicia, la educación, la salud, la política y la convivencia social. Estos espacios constituyen núcleos fundamentales de la dignidad, la creatividad y la libertad humana.

En última instancia, la relación entre inteligencia artificial y humanidad no debe concebirse como una competencia por la supremacía, sino como una coevolución condicionada por decisiones sociales, culturales y políticas. La tecnología refleja, en gran medida, los valores de quienes la diseñan y utilizan.

El futuro no está determinado exclusivamente por la capacidad técnica de las máquinas, sino por la responsabilidad con la que las sociedades decidan integrarlas en sus estructuras. La pregunta decisiva no es si las máquinas serán más inteligentes, sino si los seres humanos serán capaces de ejercer una gobernanza ética, reflexiva y solidaria sobre los sistemas que han creado.

De este modo, la sociedad algorítmica puede convertirse tanto en un instrumento de emancipación como en un mecanismo de dependencia. La dirección que adopte dependerá, en última instancia, del compromiso colectivo con la preservación de la autonomía, la justicia y el sentido humano en un entorno cada vez más mediado por la inteligencia artificial.

4.5. RESULTADOS DEL OBJETIVO ESPECÍFICO 04:

Objetivo específico: Identificar y evaluar los riesgos existenciales vinculados al avance acelerado de la IA, considerando su impacto en la autonomía humana, dependencia tecnológica y sostenibilidad social.

4.5.1. ÁMBITO SOCIAL —ENTRE EL INTERNET MUERTO, IDENTIDADES DIGITALES Y DISTOPÍAS EMERGENTES

En los debates contemporáneos sobre los efectos sociales de la tecnología, han surgido diversas teorías y narrativas que exploran cómo las comunicaciones digitales

reconfiguran las interacciones humanas, la identidad y la percepción de la realidad. Una de estas propuestas es la denominada “**teoría del Internet muerto**”, según la cual gran parte del contenido actual en Internet, incluidas redes sociales, comentarios, publicaciones y blogs no provendría directamente de la actividad orgánica de seres humanos, sino de **agentes automatizados como bots, algoritmos generativos y sistemas automatizados diseñados para producir datos, influir en el comportamiento y sostener la economía de atención digital** (Ramor et al., 2025).

Esta hipótesis —aunque objeto de controversia y debate crítico sobre sus fundamentos empíricos— pone de manifiesto una inquietud social real: el riesgo de que la tecnología no solo redefine el acceso a la información, sino que **transforma radicalmente la experiencia subjetiva de la realidad** y la percepción de lo que consideramos “humano” o “social”.

En un registro cultural afín, obras de ciencia ficción como *Serial Experiments Lain* plantean escenarios en los que la frontera entre el mundo físico y los espacios digitales se disuelve hasta el punto de que la identidad, la memoria y las relaciones humanas quedan profundamente entrelazadas con una red virtual omnipresente —el llamado *Wired*—, lo que lleva a cuestionar la distinción entre existencia “real” e “informativa”, así como los efectos psicológicos de vivir inmersos en dominios digitales hiperconectados.

Estas narrativas son útiles como espejos distópicos que, pese a su formulación ficcional o especulativa, subrayan **temores sociales legítimos** acerca de la pérdida de agencia personal, la construcción de identidades mediadas por plataformas tecnológicas, la difusión de contenidos sin origen humano claro y los impactos cognitivos de la saturación de información (por ejemplo, la confusión entre espacio online y offline).

Desde una perspectiva sociológica, fenómenos como la **desindividuación** —una pérdida relativa de la identidad personal en contextos grupales o colectivos, acentuados por el anonimato y la dinámica de masas en redes sociales— revelan cómo la tecnología puede influir en los procesos de autopercepción e interacción social.

En el contexto de la inteligencia artificial, esta preocupación se intensifica: no sólo se cuestiona la autenticidad del contenido y la agencia de los participantes, sino también la creciente presencia de sistemas automatizados que no sólo **modulan qué vemos y cómo interactuamos**, sino que también interpretan, modelan y responden a nuestras acciones —a veces de maneras que alteran patrones de comportamiento, emociones y relaciones humanas.

Este conjunto de ideas —entre la teoría crítica digital, la sociología de las redes sociales y las reflexiones culturales sobre identidad en un mundo hiperconectado— ofrece un marco referencial para abordar el impacto social de la inteligencia artificial. El propósito de este apartado es analizar cómo la IA y las plataformas digitales transforman no sólo estructuras sociales concretas (como empleo, educación o gobernanza), sino también **la forma en que los individuos perciben su propia existencia, su identidad social y su relación con la comunidad humana**, sin perder de vista que muchas de estas inquietudes ya han sido objeto de estudio tanto en ciencias sociales como en filosofía de la tecnología.

Con el propósito de sistematizar los principales riesgos éticos, sociales y de privacidad derivados del uso creciente de la inteligencia artificial, se presenta el Cuadro 7, el cual resume los efectos más relevantes identificados en investigaciones científicas y reportes institucionales entre los años 2018 y 2025.

Tabla 08: Principales riesgos éticos, sociales y de privacidad asociados al uso de la inteligencia artificial (2018–2025)

Tipo de Riesgo	Descripción	Ámbito Afectado	Casos Documentados	Fuente
Pérdida de privacidad	Recolección masiva de datos personales	Usuarios, estudiantes	Plataformas educativas, redes sociales	UNESCO (2023)

Tipo de Riesgo	Descripción	Ámbito Afectado	Casos Documentados	Fuente
	Monitoreo			
Vigilancia digital	automatizado del comportamiento	Instituciones	Sistemas de proctoring	EFF (2022)
Sesgo algorítmico	Discriminación en decisiones automatizadas	Social, laboral	Selección laboral IA	MIT (2019)
Manipulación informativa	Difusión de contenido falso	Sociedad	Deepfakes políticos	EU Commission (2024)
Dependencia tecnológica	Reducción de autonomía humana	Educativo	Uso excesivo de chatbots	OECD (2025)
Opacidad algorítmica	Falta de transparencia en decisiones	Gobierno, empresas	Sistemas predictivos	Floridi (2021)
Concentración de poder	Dominio de grandes corporaciones	Económico, político	Big Tech	Zuboff (2019)

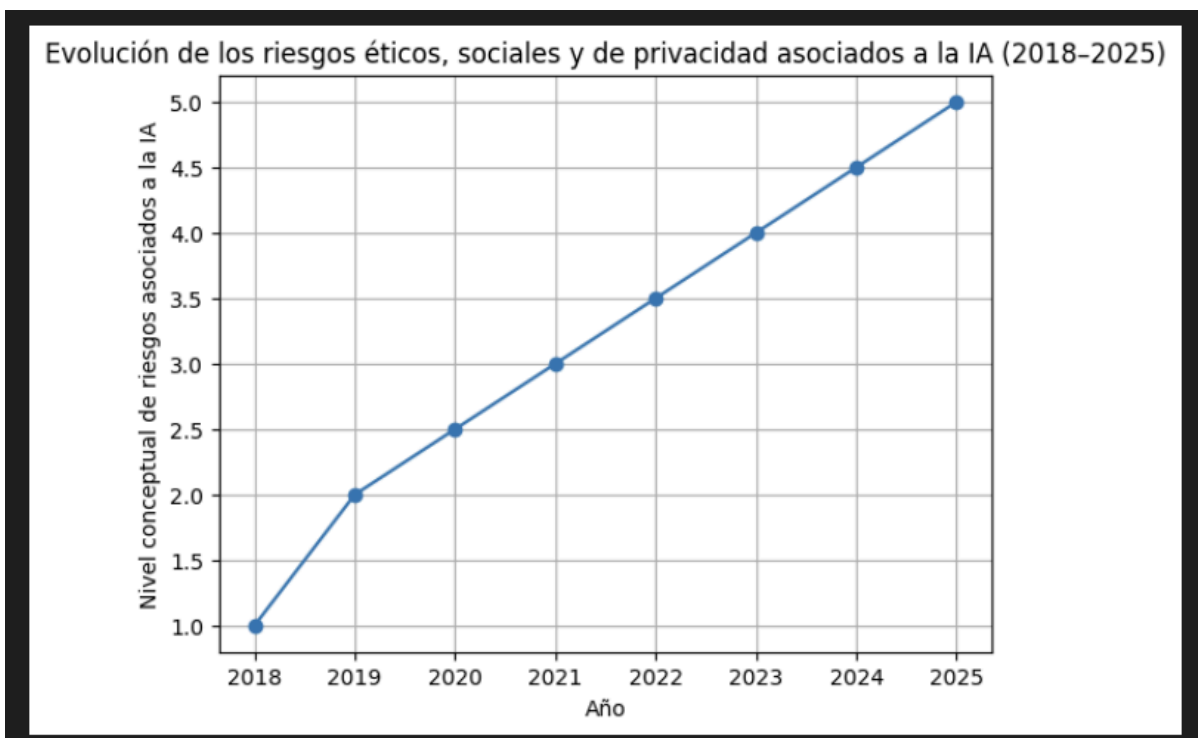


Figura 07: Evolución de los riesgos éticos, sociales y de privacidad

La Figura 7 evidencia una tendencia creciente en la intensidad y visibilidad de los riesgos éticos, sociales y de privacidad asociados al uso de la inteligencia artificial durante el periodo 2018–2025. Este comportamiento sugiere que, a medida que las aplicaciones de IA se expanden hacia ámbitos más sensibles, los impactos sobre la autonomía, la privacidad y la equidad social se vuelven más relevantes y complejos.

4.5.2. IMPACTO DE LA IA EN EL ÁMBITO SOCIAL — INTRODUCCIÓN

El impacto social de la inteligencia artificial (IA) se manifiesta desde múltiples frentes, pero uno de los más significativos ha sido la transformación de las **formas de comunicación e interacción humana en el entorno digital**. Esta transformación tiene raíces históricas en el desarrollo de programas conversacionales que han intentado simular el diálogo humano, desde los primeros chatbots hasta los sistemas actuales que modelan patrones complejos de comunicación.

La historia de los bots conversacionales se remonta a los años sesenta con **ELIZA**, desarrollado por Joseph Weizenbaum en el MIT en 1966, considerado uno de los primeros programas capaces de simular una conversación humana básica mediante

reglas de coincidencia de patrones lingüísticos (ELIZA; Weizenbaum) . A partir de este primer paso, evolucionaron otros programas experimentales como **PARRY** en 1972, diseñado para imitar la mente paranoide de un paciente psiquiátrico, y más adelante sistemas como **ALICE** y SmarterChild, que allanaron el camino para asistentes de voz avanzados como **Siri** a partir de 2011, integrados en dispositivos de consumo masivo y marcando un nuevo paradigma de interacción humano–máquina .

Con la democratización de Internet y el crecimiento de las redes sociales desde los años 2000, los agentes automatizados —originalmente diseñados para tareas simples de respuesta— comenzaron a **tener presencia en plataformas pobladas por millones de usuarios humanos**. Este fenómeno dio lugar a la aparición de **cuentas automatizadas o bots sociales**, software que puede generar actividad en redes sociales de manera continua y simular comportamientos humanos con diversos fines comunicativos (Iacopo Pozzana & Emilio Ferrara) .

En el contexto contemporáneo, la interacción humano–IA va más allá de simples herramientas de asistencia: existen plataformas donde **agentes de IA interactúan entre sí sin intervención humana directa**, como experimentos recientes en los que millones de bots publican, comentan y votan contenido de manera autónoma en entornos digitales dedicados exclusivamente a inteligencias artificiales . Este tipo de plataformas representa un nuevo estadio en la relación entre social media y agentes automatizados, y pone en evidencia la relevancia de entender las dinámicas sociales que emergen cuando las barreras entre humano y máquina se vuelven difusas.

Paralelamente, en las redes sociales convencionales se ha observado un uso intensivo de bots y algoritmos automatizados para **amplificar mensajes, influir en la opinión pública y dirigir tendencias de contenido**. Investigaciones sobre bots sociales han demostrado que estas cuentas automatizadas pueden imitar conductas humanas y afectar discusiones en línea, especialmente en contextos de debate político, donde una actividad coordinada de bots puede influenciar percepciones colectivas y amplificar información específica más allá de su base de apoyo real . Campañas de desinformación

—como la **Project Birmingham** en Estados Unidos— han utilizado tácticas de bots para generar contenido artificial que acompaña narrativas específicas en procesos electorales, poniendo en evidencia la capacidad de estas herramientas para penetrar en procesos sociales y políticos sensibles .

La función de los algoritmos predictivos es otro factor central para comprender el impacto social de la IA. Los sistemas de aprendizaje automático recolectan grandes cantidades de datos sobre preferencias, interacciones y patrones de consumo de contenido para **estimar comportamientos futuros y personalizar la experiencia del usuario en redes sociales**. Este proceso permite que las plataformas no solo respondan a las acciones de los usuarios, sino que anticipen y moldeen sus decisiones de consumo, interacción y expresión digital, reforzando patrones de atención y moldeando hábitos de navegación en base a proyecciones algorítmicas de conducta humana .

Este conjunto de desarrollos —desde los primeros chatbots hasta redes enteramente pobladas por agentes artificiales, pasando por la proliferación de bots sociales y la utilización de algoritmos predictivos— configura un panorama en el que la IA ha dejado de ser una herramienta auxiliar para convertirse en una **fuerza estructurante de la comunicación social contemporánea**, con implicaciones profundas sobre cómo los individuos interactúan, cómo se difunde la información y cómo se forman las opiniones colectivas.

4.6. RESULTADOS DEL OBJETIVO ESPECÍFICO 05:

Objetivo específico: Analizar el impacto del desarrollo e implementación de la IA en el ámbito educativo, considerando sus implicancias en los procesos de aprendizaje y mediación tecnológica.

4.6.1. IMPACTO DE LA INTELIGENCIA ARTIFICIAL EN EL ÁMBITO EDUCATIVO

La incorporación de la inteligencia artificial en los entornos educativos ha generado transformaciones significativas en los procesos de enseñanza, aprendizaje y evaluación. Durante el periodo comprendido hasta finales de 2025, las herramientas basadas en IA, especialmente los modelos de lenguaje generativo, han adquirido una presencia creciente

en instituciones de educación secundaria y superior, influyendo tanto en las prácticas pedagógicas como en los hábitos de estudio de los estudiantes.

Uno de los principales impactos positivos se relaciona con el fortalecimiento del aprendizaje asistido. Plataformas como ChatGPT, Copilot y Gemini han sido utilizadas como tutores virtuales, facilitadores de comprensión conceptual, apoyo en la redacción académica y herramientas para la organización del estudio. Investigaciones recientes indican que el uso guiado de estas tecnologías puede contribuir al desarrollo de habilidades metacognitivas, mejorar la comprensión lectora y favorecer el aprendizaje autónomo cuando se emplean de manera responsable (Pitts et al., 2025).

Asimismo, diversos estudios señalan que la inteligencia artificial ha permitido ampliar el acceso a recursos educativos personalizados, adaptados al ritmo y estilo de aprendizaje de cada estudiante. Esta personalización favorece procesos inclusivos, especialmente en contextos con limitaciones de infraestructura o escasez de docentes especializados (UNESCO, 2023).

En el ámbito académico formal, se han documentado casos en los que trabajos de investigación, proyectos universitarios y tesis han sido aprobados y reconocidos aun cuando incorporaron herramientas de IA como apoyo en la búsqueda bibliográfica, organización conceptual o redacción preliminar, siempre que el autor mantuviera la autoría intelectual y declarara el uso responsable de dichas tecnologías (Nature, 2023). Esto evidencia una tendencia progresiva hacia la aceptación regulada de la IA como instrumento auxiliar del proceso investigativo.

No obstante, junto a estos beneficios emergen riesgos relevantes asociados a la automatización académica y la dependencia intelectual. Numerosos reportes periodísticos y estudios institucionales han documentado casos de estudiantes que utilizan sistemas de IA para elaborar tareas completas, responder evaluaciones en línea o incluso presentar trabajos sin comprensión real del contenido (BBC News, 2023; The Guardian, 2024). Estas prácticas comprometen la integridad académica y debilitan el desarrollo del pensamiento crítico.

Investigaciones empíricas han demostrado que el uso indiscriminado de herramientas generativas puede reducir la retención de conocimientos, fomentar la pasividad cognitiva y debilitar la capacidad de resolución autónoma de problemas, especialmente cuando no existe una orientación pedagógica adecuada (Kasneci et al., 2023). En este sentido, el riesgo principal no radica en la tecnología en sí, sino en su utilización como sustituto del esfuerzo intelectual.

Desde una perspectiva institucional, diversas universidades y sistemas educativos han implementado políticas de regulación del uso de IA, incorporando protocolos de transparencia, sistemas de detección de contenido automatizado y programas de alfabetización digital. Estas estrategias buscan equilibrar la innovación tecnológica con la preservación de estándares académicos y éticos (OECD, 2024).

Desde una perspectiva histórica, es importante reconocer que el uso de herramientas tecnológicas en la educación no constituye un fenómeno exclusivo de la inteligencia artificial, sino parte de un proceso evolutivo continuo. A lo largo del tiempo, los estudiantes han incorporado diversos recursos para apoyar su aprendizaje, desde tablas de cálculo y manuales impresos en los inicios del siglo XX, hasta calculadoras electrónicas en las décadas finales del siglo XX, enciclopedias digitales y plataformas en línea en los primeros años del siglo XXI.

En este contexto, el acceso a internet y a repositorios digitales, incluyendo sitios como “El Rincón del Vago” en su momento, fue inicialmente cuestionado por su potencial para facilitar prácticas poco éticas, pero con el tiempo fue integrado como una herramienta complementaria en muchos entornos educativos. De manera similar, las actuales plataformas de inteligencia artificial, como ChatGPT, pueden ser interpretadas como una nueva etapa en esta evolución de recursos cognitivos, cuyo impacto depende en gran medida del uso pedagógico, ético y reflexivo que se promueva en las instituciones.

Desde esta perspectiva, el desafío central no radica únicamente en restringir el uso de estas tecnologías, sino en formar a los estudiantes para utilizarlas como herramientas de

apoyo al pensamiento, la investigación y la creatividad, evitando su empleo como sustitutos del esfuerzo intelectual y la construcción autónoma del conocimiento.

En síntesis, el impacto educativo de la inteligencia artificial presenta una naturaleza ambivalente. Por un lado, ofrece oportunidades sin precedentes para mejorar el acceso, la personalización y la eficiencia del aprendizaje; por otro, plantea riesgos vinculados a la dependencia cognitiva, la deshonestidad académica y la superficialidad intelectual. La consolidación de modelos pedagógicos que integren críticamente estas tecnologías será determinante para que su influencia contribuya al fortalecimiento del sistema educativo y no a su debilitamiento.

En la Tabla 3 se muestra el nivel de Integración de la Inteligencia Artificial en la Educación Superior y Grado de Dependencia Tecnológica (2015–2025). El Cuadro 3 muestra una progresiva intensificación del uso de inteligencia artificial en la educación superior entre 2015 y 2025, evidenciando un tránsito desde sistemas auxiliares hasta ecosistemas educativos altamente automatizados. Este proceso ha generado una creciente dependencia tecnológica por parte de estudiantes e instituciones, especialmente tras la incorporación masiva de sistemas generativos a partir de 2022 (UNESCO, 2024; OECD, 2025).

Tabla 09: Integración de la ia en la educación

Año	Tecnología IA en educación	Aplicaciones principales	Nivel de uso institucional	Grado de dependencia estudiantil	Fuente referencial
2015	Sistemas tutoriales básicos	Plataformas LMS adaptativas	Bajo	Bajo	UNESCO (2016)
2017	Analítica de aprendizaje	Predicción de rendimiento	Medio	Bajo–Medio	OECD (2018)
2019	Plataformas inteligentes	Tutorías automáticas	Medio	Medio	Holmes et al. (2019)
2020	IA educativa masiva	Educación remota COVID-19	Alto	Alto	World Bank (2020)
2021	Proctoring con IA	Supervisión automática	Alto	Alto	Williamson (2021)
2022	Chatbots académicos	Asistencia estudiantil	Alto	Muy Alto	Zawacki-Richter (2022)
2023	IA generativa	Producción de trabajos	Muy Alto	Muy Alto	OpenAI Report (2023)
2024	IA personalizada	Tutoría cognitiva	Muy Alto	Crítico	UNESCO (2024)
2025	Ecosistemas inteligentes	Gestión educativa integral	Generalizado	Crítico	OECD (2025)

Esta tabla evidencia que:

- Antes de 2019, la IA tenía un papel **complementario**.
- Desde 2020 (pandemia), se vuelve **estructural**.
- Desde 2022 con la IA generativa, aparece una **dependencia crítica**.
- En 2024–2025, la IA pasa a ser **parte central del proceso educativo**.

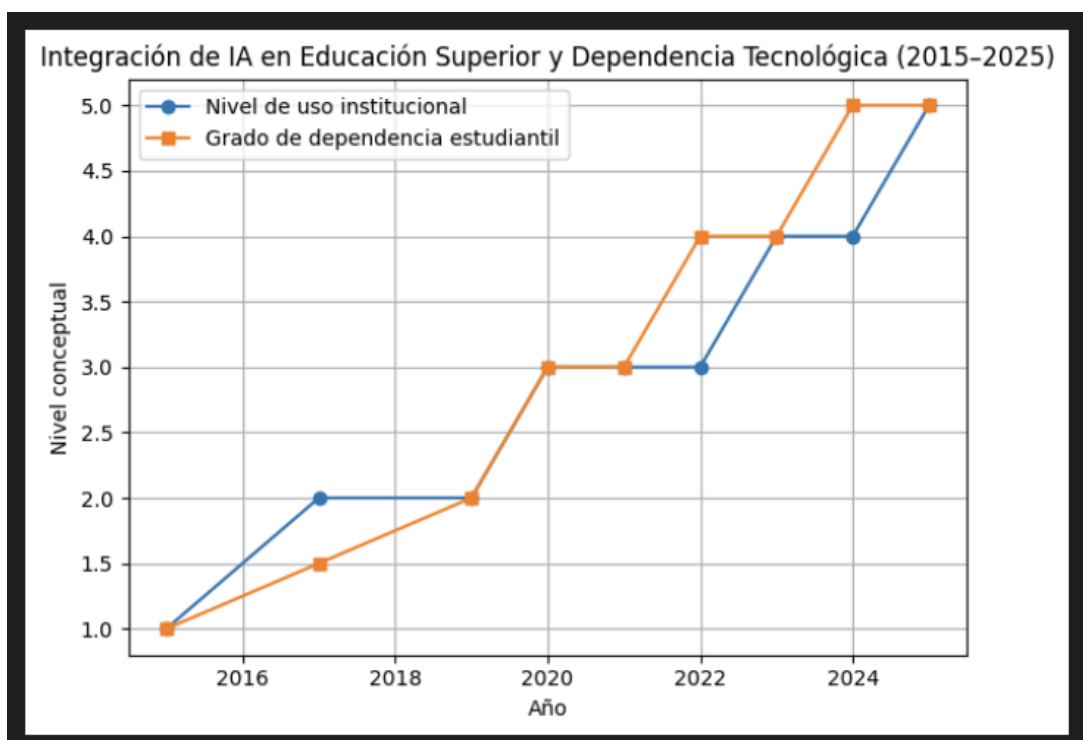


Figura 08: Integración de IA en educación superior y dependencia tecnológica

Como se observa en la figura 8, el uso de la inteligencia artificial en el ámbito educativo ha experimentado un crecimiento progresivo, pasando de sistemas complementarios en 2015 a ecosistemas altamente automatizados en 2025. Este proceso se intensificó notablemente a partir del año 2020, debido al contexto de educación remota, y se consolidó con la aparición de sistemas generativos desde 2022. Asimismo, se evidencia un aumento paralelo en el grado de dependencia estudiantil, lo cual confirma los riesgos señalados en la literatura especializada (UNESCO, 2024; OECD, 2025).

4.6.2. Impacto de IA en el ámbito educativo en el contexto nacional-local

En el contexto peruano, la influencia de la inteligencia artificial en el ámbito educativo refleja tanto oportunidades como desafíos estructurales concretos. La integración de

sistemas de IA en las dinámicas de aprendizaje y evaluación no es homogénea y está mediada por desigualdades de acceso, brechas tecnológicas y posturas éticas diversas entre estudiantes, docentes y autoridades educativas.

Un estudio realizado en nueve universidades peruanas reveló que aproximadamente **el 73 % de los universitarios utiliza herramientas de IA para elaborar tareas y asignaciones**, y que solo una minoría considera las implicancias éticas de su uso, priorizando antes la obtención de buenas calificaciones que las normas de integridad académica (Infobae, 2024). Esto muestra que la dependencia en sistemas automatizados está profundamente arraigada en la práctica estudiantil sin un marco pedagógico claro que oriente su uso responsable.

Por su parte, un reciente reporte sobre educación superior en Perú indicó que **un 95,3 % de estudiantes aseguró haber usado IA en sus estudios**, pero solo el **18,7 % manifestó tener un dominio alto de estas herramientas**, mientras que **el 64,2 % expresó preocupación por ser acusado de plagio o fraude académico al emplearlas (La República, 2025)**. Esto pone de manifiesto una tensión entre el uso extendido de IA y las políticas institucionales vigentes para regularla (La República, 2025).

A nivel de políticas institucionales, algunas universidades han adoptado medidas para enfrentar los efectos de la IA. Por ejemplo, la Universidad de Lima ha señalado que el uso de herramientas como ChatGPT debe documentarse y reportarse adecuadamente en trabajos académicos, destacando la responsabilidad de los estudiantes y profesores para garantizar la integridad de los procesos educativos (Trome, 2023).

Asimismo, en universidades peruanas se han realizado experimentos que evidencian el potencial de la IA aplicada incluso a actividades como **exámenes de admisión**. En un caso documentado, investigadores evaluaron la capacidad de ChatGPT para responder preguntas de admisión de la **Universidad Nacional de Ingeniería (UNI)** y la **Universidad Nacional Mayor de San Marcos (UNMSM)**, hallando que el sistema obtuvo un porcentaje considerable de respuestas correctas (72 % y 57 % respectivamente) — lo

que genera interrogantes sobre la validez de diferentes métodos de evaluación en un contexto donde la IA puede interactuar con pruebas tradicionales (La República, 2025).

En cuanto a capacitación docente, estudios sobre la realidad educativa peruana muestran que, aunque una parte relevante del profesorado ha utilizado herramientas de IA para preparar materiales o tareas administrativas, una proporción significativa aún no integra esas tecnologías en el aula de forma pedagógica efectiva, y existe una necesidad expresada tanto por docentes como por investigadores de formación continua en competencias digitales y de IA para guiar su uso con sentido educativo (La República, 2025).

A nivel regional, no existe evidencia periodística completa citable sobre casos de fraudes específicamente relacionados con IA en instituciones como la **Universidad Nacional del Altiplano de Puno (UNAP)**, el **Instituto SENATI Puno** sin embargo por medios no verificables como la radio o televisión locales sí han dado un par de notas acerca de ello a causa de eso la **prevalencia de tendencias observadas a nivel nacional y las declaraciones de estudiantes** apuntan a que en estas comportamientos también se replica el patrón de uso frecuente, preocupación por integridad académica y la ausencia de protocolos claros para su regulación (La República, 2025).

En conjunto, estas evidencias reflejan que la inteligencia artificial ha permeado la educación superior peruana de manera significativa, generando tanto beneficios como **problemas éticos de uso no regulado**, brechas de competencia, y la necesidad urgente de políticas institucionales que orienten su utilización. El caso peruano, con sus particularidades socioeducativas y desafíos de equidad, confirma que la IA ya es parte de la práctica educativa contemporánea y plantea la necesidad de estrategias formativas y regulativas con una perspectiva crítica y contextualizada.

4.7. DISCUSIÓN DE RESULTADOS

Se realizó el análisis sistemático de los artículos que se han publicado sobre los riesgos sistémicos derivados del desarrollo y la implementación de la inteligencia artificial. Este proceso de análisis permitió la oportunidad revisar los estudios realizados en en distintas

partes del mundo, américa latina y especialmente en Perú.

El desarrollo acelerado de la inteligencia artificial influye de manera significativa en los procesos de aprendizaje adaptativo y en la configuración de riesgos éticos, sociales, económicos y tecnológicos en la educación superior, evidenciándose a través del análisis de estudios científicos, proyecciones tecnológicas y advertencias presentes en la literatura especializada y de ficción. Asimismo, este desarrollo puede contribuir al incremento de la automatización laboral, a la dependencia tecnológica y al debilitamiento progresivo del control humano sobre los sistemas inteligentes.

La evidencia recopilada desde el punto de vista educativo, múltiples estudios señalan que el uso intensivo de herramientas de IA puede impactar la calidad del aprendizaje, la autonomía cognitiva y la retención conceptual cuando no se acompañan de estrategias pedagógicas críticas (Zhou et al., 2024). A nivel laboral, los datos muestran tendencias claras de automatización en diversos sectores, con desplazamiento de tareas humanas y necesidad de reconfiguración de competencias (WEF, 2025; ILO, 2024). Asimismo, las secciones dedicadas a riesgos éticos y de gobernanza confirman que la IA introduce desafíos significativos en términos de concentración de poder, control algorítmico y regulación insuficiente (Sela, 2025; OECD, 2021). En consecuencia la evidencia respalda la influencia sistémica y multifacética de la IA.

Los ejemplos analizados, incluyendo la automatización en logística, manufactura avanzada y sectores administrativos, así como proyecciones de tendencias laborales, confirman que la IA está reconfigurando el mercado de trabajo de manera significativa (WEF, 2025). La evidencia muestra que ciertas profesiones se encuentran en fases avanzadas de automatización, lo cual ha cambiado la demanda de habilidades y tareas especializadas.

En cuanto a los riesgos en seguridad digital y protección de datos, el desarrollo acelerado de la inteligencia artificial incrementa los riesgos en la seguridad digital, la protección de datos y la privacidad de la información, debido al uso de algoritmos automatizados,

sistemas predictivos y plataformas inteligentes que pueden ser vulnerables a manipulaciones, ciberataques o usos indebidos.

Los análisis sobre deepfakes, campañas de desinformación, vigilancia y explotación de datos personales evidencian un incremento de riesgos en seguridad digital asociados a IA (Reuters, 2025). Estas amenazas no solo están documentadas en literatura técnica sino que han generado respuestas regulatorias en múltiples países.

En relación a la gobernanza insuficiente y concentración de poder y la ausencia de marcos regulatorios sólidos y mecanismos adecuados de supervisión favorece la concentración del poder tecnológico y limita el control humano sobre los sistemas de inteligencia artificial.

Se ha mostrado que la gobernanza de IA aún presenta vacíos significativos, con reglamentos heterogéneos entre regiones y mecanismos de supervisión limitados, lo que ha favorecido la concentración de poder en grandes corporaciones y la dependencia tecnológica en Estados sin capacidad regulatoria avanzada (Sela, 2025; El País, 2025).

Con respecto a los riesgos existenciales y reducción de autonomía humana asociados al desarrollo de la inteligencia artificial se manifiestan progresivamente a través del aumento de la dependencia tecnológica, la delegación de funciones cognitivas a sistemas automatizados y la reducción de la autonomía humana.

La sección expuesta sobre dependencia social y aislamiento tecnológico, además de los análisis sobre autonomía cognitiva, muestra que existe una tendencia observable hacia la delegación de funciones cognitivas rutinarias a sistemas inteligentes y una correlación con patrones de comportamiento que reflejan menor autonomía crítica (APA, 2023; Clark & Chalmers, 1998). Aunque el riesgo “existencial” como tal aún se mantiene en el terreno prospectivo, la evidencia empírica respalda escenarios en los que la autonomía humana se ve condicionada por la interacción adelantada con sistemas automatizados. Los riesgos más extremos requieren evolución tecnológica futura para materializarse plenamente.

La insuficiente capacitación en el uso ético, responsable y seguro de la inteligencia

artificial contribuye al incremento de los riesgos relacionados con la seguridad digital, la protección de datos personales y la gestión de la información, especialmente en contextos educativos y organizacionales.

Los hallazgos en educación superior, combinados con evidencias sobre desinformación y vulnerabilidades de datos, indican que la falta de formación integral en competencias digitales, ética y seguridad tecnológica expone a estudiantes y profesionales a riesgos ampliados (Zhou et al., 2024; Reuters, 2025).

La limitada percepción y comprensión de los riesgos éticos, sociales y tecnológicos asociados a la inteligencia artificial favorece su uso indiscriminado en el entorno educativo, promoviendo una creciente dependencia de estas herramientas y una reducción del análisis crítico en los procesos de enseñanza y aprendizaje.

La evidencia muestra que sin una formación crítica adecuada, los estudiantes tienden a depender de sistemas automatizados sin cuestionar su origen, limitaciones ni implicaciones éticas (Zhou et al., 2024). Esto refuerza que la percepción insuficiente de riesgos contribuye a patrones de uso indiscriminado.

Aunque la evidencia revisada es amplia, este estudio reconoce ciertas limitaciones metodológicas: Temporalidad de datos: La rápida evolución de la IA implica que nuevas evidencias pueden surgir más allá del corte de análisis de esta investigación, lo que limita la exhaustividad histórica. Diversidad contextual: Muchas evidencias provienen de contextos geopolíticos concretos (Europa, EE. UU., etc.), lo cual puede limitar la generalización global. Naturaleza de la evidencia: Algunos análisis combinan estudios científicos con reportes periodísticos o literatura crítica, lo cual puede requerir investigaciones cuantitativas adicionales para medir efectos causales estrictos.

El análisis evidencia que la inteligencia artificial ha dejado de ser una herramienta complementaria para convertirse en un agente estructural dentro de los sistemas productivos, educativos y comunicativos. La automatización de tareas, la reconfiguración del empleo y la dependencia creciente de sistemas inteligentes confirman que el impacto

de la IA no es marginal, sino transversal y acumulativo. Este fenómeno se manifiesta tanto en sectores altamente tecnificados como en actividades tradicionalmente humanas, lo que refuerza la magnitud del proceso de transformación en curso.

En el ámbito educativo, se constató que la incorporación acelerada de herramientas de IA, sin una formación crítica adecuada, puede afectar negativamente los procesos de aprendizaje, favoreciendo la dependencia tecnológica y reduciendo el desarrollo del pensamiento autónomo. Aunque estas tecnologías poseen un alto potencial pedagógico, su uso indiscriminado puede debilitar habilidades fundamentales como el razonamiento analítico, la creatividad y la reflexión ética.

Se demuestra que la inteligencia artificial representa simultáneamente una oportunidad de progreso y una fuente de desafíos estructurales. Su impacto no puede comprenderse únicamente desde una perspectiva técnica, sino que exige un análisis interdisciplinario que integre dimensiones humanas, sociales y éticas. La evidencia analizada sugiere que el futuro de la relación entre humanidad y tecnología dependerá, en gran medida, de la capacidad colectiva para diseñar marcos de regulación sólidos, fortalecer la educación crítica y promover un uso responsable de estas herramientas.

Finalmente, el desafío central no radica únicamente en desarrollar sistemas más eficientes, sino en garantizar que dichos sistemas permanezcan subordinados a principios humanos fundamentales como la dignidad, la autonomía, la justicia y la responsabilidad social. En este sentido, la inteligencia artificial no debe concebirse como un sustituto del juicio humano, sino como un instrumento cuyo valor dependerá de las decisiones éticas, políticas y educativas que orienten su implementación en las próximas décadas.

CONCLUSIONES

PRIMERA: La presente investigación permitió analizar críticamente los riesgos sistémicos derivados del desarrollo y la implementación de la inteligencia artificial, identificando que dichos riesgos trascienden el ámbito estrictamente tecnológico y se proyectan hacia dimensiones laborales - económicas, seguridad digital, gobernanza, existencial y educativo. El estudio evidenció que la inteligencia artificial no constituye un fenómeno aislado ni abrupto, sino la continuación y aceleración de procesos históricos de automatización que han transformado progresivamente la relación entre el ser humano y la tecnología.

SEGUNDA: El análisis realizado permite concluir que el avance y la implementación de la inteligencia artificial se articulan con una intensificación de los procesos de automatización, generando transformaciones estructurales en las dinámicas laborales y económicas. Este fenómeno se manifiesta en la reconfiguración de funciones humanas, el desplazamiento de tareas cognitivas y operativas, así como en la emergencia de nuevas formas de dependencia tecnológica. En consecuencia, el impacto de la inteligencia artificial trasciende la sustitución de empleos específicos, proyectándose hacia una redefinición más amplia de la participación humana en los sistemas productivos.

TERCERA: Se concluye que el desarrollo acelerado de sistemas de inteligencia artificial amplía el espectro de riesgos asociados a la seguridad digital, particularmente en lo relativo a la protección de datos, la privacidad de la información y la vulnerabilidad de infraestructuras tecnológicas. La creciente sofisticación de algoritmos predictivos y sistemas automatizados introduce escenarios complejos donde las amenazas ya no

dependen exclusivamente de agentes humanos, sino también de dinámicas algorítmicas capaces de amplificar errores, sesgos o usos indebidos.

CUARTA: El estudio permite concluir que la expansión de la inteligencia artificial plantea desafíos significativos en materia de gobernanza, regulación y control. La insuficiencia o desactualización de marcos normativos, junto con la concentración de capacidades tecnológicas en actores específicos, favorece asimetrías de poder que pueden limitar la transparencia, la rendición de cuentas y el control humano efectivo. Esta situación evidencia la necesidad de repensar los mecanismos institucionales destinados a supervisar sistemas cuya complejidad y autonomía superan los modelos regulatorios tradicionales.

QUINTA : Se concluye que los riesgos existenciales asociados al desarrollo de la inteligencia artificial no se restringen a escenarios catastróficos hipotéticos, sino que se expresan progresivamente en transformaciones observables, tales como el aumento de la dependencia tecnológica, la delegación de funciones cognitivas y el posible debilitamiento de la autonomía humana. Estos procesos sugieren un desplazamiento gradual en la relación entre el ser humano y la tecnología, donde la mediación algorítmica adquiere un papel cada vez más determinante en la organización de la experiencia social.

SEXTA: El análisis desarrollado permite concluir que la incorporación creciente de la inteligencia artificial en el ámbito educativo genera transformaciones sustantivas en los procesos de enseñanza y aprendizaje. Estas transformaciones se manifiestan en nuevas formas de mediación tecnológica, en la delegación de funciones cognitivas y en la redefinición del rol humano en la producción, validación y apropiación del conocimiento. Si bien dichas tecnologías ofrecen oportunidades relevantes, también introducen riesgos asociados a la adopción acrítica, la dependencia instrumental y la posible erosión del pensamiento reflexivo.

RECOMENDACIONES

PRIMERA: Adoptar enfoques integrales que articulen regulación, educación, ética y desarrollo tecnológico, reconociendo que los riesgos asociados a la inteligencia artificial no pueden abordarse de manera aislada. La gestión responsable de estas tecnologías requiere equilibrio entre innovación, supervisión humana y preservación de la agencia individual y colectiva.

SEGUNDA: Diseñar estrategias de adaptación laboral orientadas a mitigar los efectos del reemplazo funcional, priorizando programas de reconversión profesional, actualización de competencias digitales y fortalecimiento de habilidades humanas difícilmente automatizables, tales como pensamiento crítico, creatividad, juicio ético y toma de decisiones complejas. Además, se sugiere que las instituciones públicas y privadas desarrollen políticas que integren la inteligencia artificial como herramienta complementaria, evitando enfoques exclusivamente sustitutivos que puedan generar desplazamientos abruptos y profundización de desigualdades económicas.

TERCERA: Reforzar las políticas de ciberseguridad y protección de datos en contextos donde se integren sistemas de inteligencia artificial, priorizando la prevención de vulnerabilidades asociadas a automatización avanzada, manipulación algorítmica y uso indebido de información personal. Del mismo modo, se sugiere fomentar prácticas de auditoría tecnológica, evaluación de riesgos y protocolos de respuesta ante incidentes vinculados a sistemas inteligentes, especialmente en infraestructuras críticas y plataformas de alta interacción social.

CUARTA: Promover el desarrollo y actualización de marcos normativos específicos para la inteligencia artificial que contemplen no solo aspectos técnicos, sino también

implicancias éticas, sociales y económicas. Dichos marcos deberían incorporar criterios de transparencia algorítmica, rendición de cuentas, trazabilidad de decisiones automatizadas y mecanismos efectivos de supervisión humana. Asimismo, se sugiere evitar escenarios de concentración excesiva de poder tecnológico mediante políticas que fomenten la competencia, la auditoría independiente de sistemas de IA y la regulación de su implementación en sectores críticos como seguridad digital, educación, empleo y administración pública.

QUINTA: Promover iniciativas orientadas al fortalecimiento de la conciencia crítica frente a la mediación algorítmica, particularmente en relación con fenómenos como deepfakes, manipulación informativa, identidad digital, relaciones humano-IA y dependencia tecnológica. Asimismo, se sugiere impulsar espacios de debate interdisciplinario que permitan analizar los efectos de la inteligencia artificial sobre la autonomía humana, la autenticidad de la interacción social y la configuración de nuevas dinámicas de influencia digital.

SEXTA: Incorporar enfoques pedagógicos críticos que permitan comprender la inteligencia artificial no solo como recurso instrumental, sino como fenómeno tecnológico con implicancias cognitivas, éticas y sociales. Esto implica fortalecer la alfabetización digital, el análisis crítico de sistemas automatizados y la reflexión sobre sesgos, dependencia tecnológica y validación del conocimiento. Asimismo, se sugiere implementar programas de formación docente orientados al uso responsable, ético y reflexivo de herramientas de IA, evitando su adopción acrítica o su utilización como sustituto de procesos fundamentales de aprendizaje y construcción intelectual.

BIBLIOGRAFÍA

- Acemoglu, D., & Restrepo, P. (2020). Robots and jobs: Evidence from US labor markets. *Journal of Political Economy*, 128(6), 2188–2244. <https://doi.org/10.1086/705716>
- Autor, D. H. (2015). Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives*, 29(3), 3–30. <https://doi.org/10.1257/jep.29.3.3>
- Adobe. (2017). Adobe Sensei. <https://www.adobe.com/sensei.html>
- Adobe. (2019). Artificial intelligence in Creative Cloud. <https://www.adobe.com>
- Adobe. (2022). Creative automation and AI tools. <https://blog.adobe.com>
- Alessi Longa, F. E., Shabbir, S., & Zaid, R. (2025). Influence of social media on self-esteem and body image. *Research Journal of Psychology*. <https://doi.org/10.59075/rjs.v3i4.256>
- Allison, G. T., & Zelikow, P. (1999). *Essence of decision: Explaining the Cuban Missile Crisis*. W. W. Norton & Company.
- Amodei, D., Hernandez, D., Sastry, G., Clark, J., Brockman, G., & Sutskever, I. (2018). AI and compute. OpenAI. <https://arxiv.org/abs/1812.03962>
- Andrew Yang AI job warnings. (2025). Business Insider. <https://www.businessinsider.com/andrew-yang-ai-may-wipe-out-40-million-us-jobs-2025-12>
- APA – American Psychological Association. (2023). Social media and mental health. <https://www.apa.org/topics/social-media-internet/mental-health>
- Arms Control Association. (2019). Nuclear false warnings and the risk of catastrophe. <https://www.armscontrol.org/act/2019-12/focus/nuclear-false-warnings-and-risk-catastrophe>

Associated Press. (2014). AP uses automation to expand earnings reports. <https://www.ap.org>

Automation Magazine. (2022). Robots used in the automotive industry worldwide cross the 1 million mark, says IFR. <https://www.automationmag.com>

Bank for International Settlements (BIS). (2018). Sound practices: Implications of fintech developments for banks and bank supervisors. <https://www.bis.org>

Ballesteros-Aguayo, L., & Ruiz del Olmo, F. J. (2024). Vídeos falsos y desinformación ante la IA: el deepfake como vehículo de la posverdad. *Revista de Ciencias de la Comunicación e Información*, 29, 1–14. <https://www.revistaccinformacion.net/index.php/rcci/article/view/20>

BBC News. (2019). Self-checkout expansion. <https://www.bbc.com>

BBC News. (2019). AI judges in China? <https://www.bbc.com/news/technology-49589272>

Benjamins, R., & Salazar, I. (2020). Towards a framework for understanding societal and ethical implications of artificial intelligence. *arXiv*. <https://arxiv.org/abs/2001.09750>

Bloomberg. (2016). The Cyborg project. <https://www.bloomberg.com>

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.

Bradshaw, S., & Howard, P. N. (2019). The global disinformation order. Oxford Internet Institute. <https://www.oii.ox.ac.uk>

Brennan, T., Dieterich, W., & Ehret, B. (2020). COMPAS risk scales: Technical documentation.

Brown, T. B., et al. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33. <https://arxiv.org/abs/2005.14165>

Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton & Company.

Brynjolfsson, E., Rock, D., & Syverson, C. (2022). The productivity J-curve: How intangibles complement general-purpose technologies. *American Economic Journal*.
<https://www.aeaweb.org>

Business Insider. (2025). Andrew Yang warns AI may wipe out millions of jobs.
<https://www.businessinsider.com>

Gartner. (2018). AI in customer service: Strategic implications. <https://www.gartner.com>

GEA Group. (2023). Dairy farming solutions.
<https://www.gea.com/en/products/dairy-farming.jsp>

GlobalSecurity.org. (s. f.). Russia and early warning systems.
<https://www.globalsecurity.org/space/world/russia/warning.html>

Google Cloud. (2023). Dialogflow documentation. <https://cloud.google.com/dialogflow>

Google Cloud. (2023). Vision OCR documentation. <https://cloud.google.com/vision/docs/ocr>

Google AI Blog. (2019). Advances in computer vision. <https://ai.googleblog.com>

Graefe, A. (2016). Guide to automated journalism. Tow Center for Digital Journalism.
https://www.cjr.org/tow_center_reports/guide_to_automated_journalism.php

Gestión. (2019). Laboratorios en Perú y automatización. <https://gestion.pe>

Gestión. (2022). Marketing digital y automatización en Perú. <https://gestion.pe>

Gestión. (2023). Automatización industrial en el Perú. <https://gestion.pe>

Gestión. (2023). Tecnologías de RRHH en Perú.
<https://gestion.pe/economia/empresas/rrhh-inteligencia-artificial-peru>

Gestión. (2024). Banca y automatización. <https://gestion.pe>

FAO – Food and Agriculture Organization of the United Nations. (2024). Digital agriculture report. <https://www.fao.org/digital-agriculture>

Feigenbaum, E. A. (1980). Knowledge engineering: The applied side of artificial intelligence. *Proceedings of the National Computer Conference*.

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots. *Communications of the ACM*, 59(7), 96–104. <https://doi.org/10.1145/2818717>

Ferrara, E., et al. (2020). The rise of social bots. *Communications of the ACM*, 63(8).* <https://cacm.acm.org/magazines/2020/8/246903-the-rise-of-social-bots/fulltext>

Floridi, L., Cowls, J., Beltrametti, M., et al. (2018). AI4People—An ethical framework for a good AI society. *Minds and Machines*, 28, 689–707. <https://doi.org/10.1007/s11023-018-9482-5>

Forbes México. (2021). Self-checkout y automatización en retail LATAM. <https://www.forbes.com.mx>

Foxconn. (2022). Annual report 2022. <https://www.foxconn.com/en-us/investor-relations>

Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, 114, 254–280. <https://doi.org/10.1016/j.techfore.2016.08.019>

Fueller. (2025). AI-powered medical devices changing healthcare. <https://fueller.io/blog/ai-powered-medical-devices-changing-healthcare-2025>

Helbing, D. (2019). *Towards digital enlightenment: Essays on the dark and light sides of the digital revolution*. Springer. <https://link.springer.com/book/10.1007/978-3-030-12385-5>

HireVue. (2019). How hiring AI works. <https://www.hirevue.com/resources/blog/how-ai-is-changing-interviews>

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education*. Springer.

<https://doi.org/10.1007/978-3-030-23207-8>

Hooked – Eyal, N. (2014). *Hooked: How to build habit-forming products*. Portfolio.

<https://www.nirandfar.com/hooked>

Huffington Post. (2025). Instagram afecta más al bienestar psicológico.

<https://www.huffingtonpost.es>

IEBSchool. (2025). *El impacto de la IA y la automatización en el trabajo*.

<https://www.iebschool.com/hub/wp-content/uploads/2025/02/Impacto-de-la-IA-y-la-Automatizacion-en-la-Empresa.pdf>

IBM. (2012). Intelligent document processing.

<https://www.ibm.com/analytics/document-processing>

IBM. (2019). Watson Assistant overview. <https://www.ibm.com/watson/assistant>

IBM. (2024). Watson Discovery. <https://www.ibm.com/products/watson-discovery>

IEBSchool. (2025). El impacto de la IA y la automatización en el trabajo.

<https://www.iebschool.com/hub/wp-content/uploads/2025/02/Impacto-de-la-IA-y-la-Automatizacion-en-la-Empresa.pdf>

ILO – International Labour Organization. (2022). Work for a brighter future.

<https://www.ilo.org/global/topics/future-of-work>

IRS – Internal Revenue Service. (2021). Digitalization program.

<https://www.irs.gov/digitalization>

Ismail. (2022). Future of work: Automation, AI & workforce transformation.

<https://www.journal-gsr.com/4/1/43>

Jakesch, M., Hancock, J. T., & Naaman, M. (2022). Human heuristics for AI-generated language are flawed. *arXiv*. <https://arxiv.org/abs/2206.07271>

JMIR – Journal of Medical Internet Research. (2025). The potential and peril of AI in the emergency department. <https://www.jmir.org/2025/1/e89200>

John Deere. (2022). Autonomous tractor technology. <https://www.deere.com/en/autonomous>

JPMorgan Chase. (2018). COiN platform. <https://www.jpmorganchase.com/technology/coin>

Katz, D. M. (2019). The MIT School of Law? *arXiv*. <https://arxiv.org/abs/1906.11041>

Kasneci, E., Sessler, K., Küchemann, S., et al. (2023). ChatGPT for good? On opportunities and challenges. *Learning and Individual Differences*, 103. <https://doi.org/10.1016/j.lindif.2023.102274>

Khan Academy. (2023). Introducing Khanmigo. <https://blog.khanacademy.org/introducing-khanmigo>

Knight Center for Journalism. (2020). AI in Latin American media. <https://journalismcourses.org>

LawGeex. (2019). AI contract review report. <https://www.lawgeex.com/resources/ai-contract-review-report>

LegalTech LATAM. (2023). AI tools in Latin America law firms. <https://www.legaltechlatam.com>

LinkedIn. (2018). *Global Talent Trends Report*. <https://business.linkedin.com/talent-solutions/recruiting-tips/global-talent-trends-2018>

8

López de Mántaras, R. (2017). Artificial intelligence and the future of work. *AI Magazine*, 38(1), 3–5. <https://doi.org/10.1609/aimag.v38i1.2701>

McKinsey Global Institute. (2023). *Generative AI and the future of work*.
<https://www.mckinsey.com>

Meindl, B., et al. (2021). Exposure of occupations to 4IR technologies. *arXiv*.
<https://arxiv.org/abs/2110.13317>

McAfee, A., & Brynjolfsson, E. (2017). *Machine, platform, crowd: Harnessing our digital future*. W. W. Norton & Company.

Mac, R. (2016). Can an AI lawyer outperform a human lawyer? *Forbes*.
<https://www.forbes.com/sites/rogerkay/2016/02/01/can-an-ai-lawyer-outperform-a-human-lawyer>

MAHALO Project. (2023). Finding automation's place within air traffic control systems. *CORDIS*.
<https://cordis.europa.eu/article/id/442204-finding-automation-s-place-within-air-traffic-control-systems>

McKinsey Global Institute. (2017). *Jobs lost, jobs gained: Workforce transitions in a time of automation*.
<https://www.mckinsey.com/featured-insights/future-of-work/jobs-lost-jobs-gained>

McKinsey Global Institute. (2023).
Generative AI and the future of work.
<https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/generative-ai-and-the-future-of-work>

Meindl, B., et al. (2021). Exposure of occupations to 4IR technologies. *arXiv*.
<https://arxiv.org/abs/2110.13317>

Mil21. (2025). Robot realiza cirugía autónoma con éxito.
<https://www.mil21.es/noticia/17993/claves/robot-realiza-con-exito-la-primer-cirugia-totalmente-autonoma.html>

MINEDU – Ministerio de Educación del Perú. (2022). Plataformas digitales para educación a distancia. <https://www.gob.pe/minedu>

MIT Technology Review. (2019). Why we need to audit algorithms. <https://www.technologyreview.com/2019/04/04/240314/why-we-need-to-audit-algorithms>

MONAI – Medical Open Network for AI. (s. f.). Wikipedia entry. https://en.wikipedia.org/wiki/Medical_open_network_for_AI

OECD. (2021). *OECD principles on Artificial Intelligence*. <https://www.oecd.org/ai/>

OECD. (2023). *OECD Employment Outlook 2023: Artificial Intelligence and the Labour Market*. <https://www.oecd.org/employment-outlook/>

OECD – Organisation for Economic Co-operation and Development. (2018). *Innovating Education with AI*. <https://www.oecd.org/education>

OECD. (2019). The risk of automation for jobs in OECD countries. <https://www.oecd.org/employment/automation.html>

OECD. (2021). AI governance principles. <https://www.oecd.org/ai>

OECD. (2023). *OECD Employment Outlook 2023: Artificial intelligence and jobs*. https://www.oecd.org/en/publications/oecd-employment-outlook-2023_08785bba-en

OECD. (2024). AI in education: Challenges and opportunities. <https://www.oecd.org/education/ai-education.html>

OECD AI Outlook. (2023). *AI Outlook report*. <https://www.oecd.org/digital/artificial-intelligence>

Poder Judicial del Perú. (2024). Gestión Electrónica de Expedientes. <https://www.pj.gob.pe>

ProPublica. (2016). Machine bias: Risk assessment tool is disproportionately flagging black defendants.

<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>

PwC – PricewaterhouseCoopers. (2022). Will robots really steal our jobs?

<https://www.pwc.com/gx/en/issues/upskilling/impact-of-automation-on-jobs.html>

PwC. (2025). Barómetro global de la IA en el empleo.

<https://www.pwc.es/es/consultoria/inteligencia-artificial/ai-jobs-barometer-2025.html>

Pymetrics. (2021). Science behind gamified talent matching.

<https://www.pymetrics.com/science>

Rahman, M. (2023). Future of work and AI disruption. *Journal of Global Research*.

<https://www.journal-gsr.com/5/1/61>

Radford, A., et al. (2018). Improving language understanding by generative pre-training. *OpenAI*.

https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf

Reuters. (2025). AI-guided cameras make solo surgery possible: Step towards surgical automation.

<https://www.reuters.com/business/healthcare-pharmaceuticals/ai-guided-cameras-make-solo-surgery-possible-step-towards-surgical-automation-2025-09-10>

Russell, S. (2019). *Human compatible: Artificial intelligence and the problem of control*. Viking.

Stanford University. (2024). *AI Index Report 2024*. <https://aiindex.stanford.edu/report/>

Subaveerapandiyan, A., & Shimray, S. R. (2024). Workforce displacement in the age of AI. *Online Information Review*. <https://doi.org/10.1515/opis-2024-0010>

Selwyn, N. (2019). Should robots replace teachers? AI and the future of education. *Polity Press*.

Semana Económica. (2023). LegalTech en Perú. <https://semanaeconomica.com>

Semana Económica. (2024). Automatización y talento humano. <https://semanaeconomica.com>

SHRM – Society for Human Resource Management. (2020). Artificial intelligence for HR. <https://www.shrm.org/resourcesandtools/hr-topics/technology/pages/artificial-intelligence-for-hr.aspx>

Stanford Institute for Human-Centered AI. (2024). *AI Index Report 2024*. <https://aiindex.stanford.edu/report/>

StatCounter. (2025). Chatbot market share and usage trends. <https://gs.statcounter.com/chatbot-market-share>

Strategic Market Research. (2026). AI job replacement risk by industry. <https://www.datarefs.com/statistics/ai/ai-job-replacement/>

Surden, H. (2014). Machine learning and law. *Washington Law Review*, 89(1), 87–115. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2376204

The Guardian. (2025, November 25). AI could replace 3 million low-skilled jobs in the UK by 2035. <https://www.theguardian.com/technology/2025/nov/25/ai-could-replace-3m-low-skilled-jobs-by-2035-research-finds>

The Guardian. (2025). AI could replace 3 million low-skilled jobs in the UK by 2035. <https://www.theguardian.com/technology/2025/nov/25/ai-could-replace-3m-low-skilled-jobs-by-2035-research-finds>

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433–460. <https://doi.org/10.1093/mind/LIX.236.433>

- UNESCO. (2016). Artificial intelligence in education. <https://unesdoc.unesco.org>
- UNESCO. (2021). Artificial intelligence and education: Guidance for policy-makers. <https://unesdoc.unesco.org>
- UNESCO. (2023). Artificial intelligence and disinformation. <https://www.unesco.org>
- UNESCO. (2024). AI competency framework for students. <https://www.unesco.org>
- UNIDROIT. (2018). Study on online dispute resolution. <https://www.unidroit.org>
- Unilever. (2020). Digital hiring transformation. <https://www.unilever.com/careers/recruitment>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>
- Wardle, C., & Derakhshan, H. (2017). *Information disorder*. <https://www.coe.int/en/web/freedom-expression/information-disorder>
- Wardle, C., & Derakhshan, H. (2017). Information disorder: Toward an interdisciplinary framework for research and policy making. *Council of Europe*. <https://rm.coe.int/information-disorder-report/168076277c>
- Waseda University. (2023). Human–AI attachment study. <https://www.waseda.jp>
- Weizenbaum, J. (1966). ELIZA—A computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36–45. <https://doi.org/10.1145/365153.365168>
- Williamson, B. (2021). Education and AI governance: Policy, networks, and power. *Learning, Media and Technology*, 46(1).* <https://doi.org/10.1080/17439884.2020.1754238>

- Wired. (2023). Slovakia's election deepfakes show AI is a danger to democracy. <https://www.wired.com/story/slovakias-election-deepfakes-show-ai-is-a-danger-to-democracy>
- World Bank. (2020). Remote learning during COVID-19. <https://www.worldbank.org>
- World Economic Forum. (2023). *The Future of Jobs Report 2023*. <https://www.weforum.org/reports/the-future-of-jobs-report-2023>
- World Health Organization. (2019). Gaming disorder. <https://www.who.int/news/item/25-05-2019-who-recognizes-gaming-disorder>
- World Health Organization. (2022). Mental health and COVID-19. <https://www.who.int>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education*, 16. <https://doi.org/10.1186/s41239-019-0171-0>
- Zhou, T., et al. (2024). Do you have AI dependency? Artificial intelligence use and dependency in higher education. *International Journal of Educational Technology in Higher Education*. <https://doi.org/10.1186/s41239-024-00467-0>
- Zuboff, S. (2019). *The age of surveillance capitalism*. PublicAffairs

ANEXOS

Anexo 01: Matriz De Consistencia.

ESTUDIO TEÓRICO DE LOS RIESGOS SISTÉMICOS DERIVADOS DEL DESARROLLO Y LA IMPLEMENTACIÓN DE LA INTELIGENCIA ARTIFICIAL

PROBLEMA	OBJETIVOS	VARIABLES	DIMENSIONES	TÉCNICAS E INSTRUMENTOS	METODOLOGÍA
<p>GENERAL: ¿Cuáles son los principales riesgos sistémicos derivados del desarrollo y la implementación de la Inteligencia Artificial identificados en la literatura científica?</p> <p>ESPECÍFICOS. ¿De qué manera el desarrollo acelerado de la IA puede generar riesgos relacionados con las condiciones laborales y en el desarrollo económico, considerando cambios en la estructura del empleo y automatización de procesos?</p> <p>¿En qué medida el avance y la implementación de la IA incrementan los riesgos en la seguridad digital y la protección de la información?</p> <p>¿Cuáles son los principales riesgos asociados a la gobernanza y al control de los sistemas de IA en contextos organizacionales y sociales?</p> <p>¿De qué manera pueden</p>	<p>GENERAL Analizar críticamente los riesgos sistémicos derivados del desarrollo y la implementación de la Inteligencia Artificial identificados en la literatura científica.</p> <p>ESPECÍFICOS Analizar el impacto del avance de la IA en las condiciones laborales y en el desarrollo económico, considerando cambios en la estructura del empleo y automatización de procesos.</p> <p>Examinar los principales riesgos asociados a la seguridad digital derivados del uso de sistemas de IA, evaluando su influencia en la protección de datos y privacidad.</p> <p>Analizar los desafíos relacionados con la gobernanza y el control de la IA, considerando los marcos normativos, mecanismos de supervisión y responsabilidad</p>	<p>INDEPENDIENTE Desarrollo e implementación acelerada de la Inteligencia Artificial.</p> <p>DEPENDIENTE Riesgos Sistémicos</p>	<p>Laboral - Económica Seguridad digital Gobernanza Existencial Educativa</p>	<p>Técnicas: Revisión documental; Instrumentos: Fichas de análisis (Fichas de resumen)</p>	<p>Enfoque: Cualitativo Diseño: No experimental Métodos: Analítico, Deductivo, histórico y crítico-reflexivo Población: Literatura científica 2015-2025 Muestra: Literatura científica 2015-2025</p>

PROBLEMA	OBJETIVOS	VARIABLES	DIMENSIONES	TÉCNICAS E INSTRUMENTOS	METODOLOGÍA
<p>identificarse y evaluarse los riesgos existenciales vinculados al avance y la implementación de la IA?</p> <p>¿Cómo incide el desarrollo e implementación de la IA en el ámbito educativo y sus implicancias en los procesos de aprendizaje?</p>	<p>de los actores involucrados.</p> <p>Identificar y evaluar los riesgos existenciales vinculados al avance acelerado de la IA, considerando su impacto en la autonomía humana, dependencia tecnológica y sostenibilidad social.</p> <p>Analizar el impacto del desarrollo e implementación de la IA en el ámbito educativo, considerando sus implicancias en los procesos de aprendizaje y mediación tecnológica.</p>				

Anexo 02: Matriz resumen de análisis conceptual de autores y ejes teóricos

Matriz de análisis conceptual de autores y ejes teóricos

Autor	Tipo de aporte	Concepto(s) clave vinculado(s)	Relación con la tesis
Turing, A. (1950)	Fundacional / conceptual	Indistinguibilidad humano-IA	Introduce la posibilidad de que una máquina simule inteligencia humana, base del debate sobre sustitución cognitiva.
Asimov, I. (1950–1985)	Prospectivo / normativo	Riesgo tecnológico / control IA	Anticipa dilemas éticos y riesgos derivados de la autonomía de sistemas inteligentes.
Wiener, N. (1948)	Fundacional / crítico	Automatización / retroalimentación	Advierte sobre consecuencias sociales de la automatización y pérdida de control humano.
Floridi, L. (2014–2023)	Ético / filosófico	Riesgo epistémico / agencia artificial	Analiza impacto ontológico y moral de agentes artificiales en entornos informacionales.
Bostrom, N. (2014)	Prospectivo / riesgo existencial	Escenarios de peligro IA	Formaliza riesgos existenciales asociados a superinteligencias no alineadas.
Russell, S. (2019)	Ético / técnico	Alineación IA / cesión de control	Propone rediseño de IA bajo incertidumbre sobre preferencias humanas para mitigar riesgos.
Turkle, S. (2011)	Crítico / sociotécnico	Dependencia tecnológica / relación humano-IA	Examina efectos psicológicos y sociales de la interacción humano-máquina.

Anexo 03: Resumen de Línea temporal conceptual de la Inteligencia Artificial

Línea temporal conceptual de la Inteligencia Artificial

1. Etapa fundacional (1940–1956): Origen teórico

Hitos conceptuales clave:

- **Norbert Wiener (1948)** – *Cibernética*: introduce la noción de sistemas autorregulados mediante retroalimentación.
- **Alan Turing (1950)** – *Computing Machinery and Intelligence*: plantea la pregunta “¿Pueden pensar las máquinas?” y formula el Test de Turing.
- **Conferencia de Dartmouth (1956)** – McCarthy acuña el término **Artificial Intelligence**.

Interpretación en la tesis:

La IA emerge como hipótesis científica y filosófica, no como tecnología madura. Se instala el debate sobre cognición artificial.

2. Etapa simbólica (1956–1970): IA como lógica y reglas

Características:

- Sistemas expertos
- Representación simbólica del conocimiento
- Optimismo tecnológico inicial

Limitaciones:

- Rigidez
- Escasa adaptabilidad
- Dependencia de reglas explícitas

Interpretación en la tesis:

Primer intento de emular la inteligencia humana mediante la formalización lógica.

3. Primer “invierno de la IA” (1970–1980): Crisis de expectativas

Causas:

- Falta de resultados prácticos
- Costos computacionales elevados
- Promesas incumplidas

Interpretación:

Se evidencia la brecha entre teoría y aplicabilidad.

4. Resurgimiento conexionista (1980–1995): Redes neuronales

Avances:

- Modelos neuronales artificiales
- Aprendizaje automático inicial
- Reconocimiento de patrones

Cambio conceptual:

De reglas → aprendizaje basado en datos.

Interpretación en la tesis:

Transición hacia sistemas que “aprenden” en lugar de ser programados exhaustivamente.

5. Consolidación del Machine Learning (1995–2010)

Elementos clave:

- Big Data
- Mejora del poder computacional
- Algoritmos estadísticos avanzados

Impactos:

- Automatización predictiva
- Sistemas de recomendación
- Análisis masivo de información

Interpretación:

La IA comienza a integrarse en procesos económicos y sociales.

6. Revolución del Deep Learning (2010–2018)

Avances disruptivos:

- Redes profundas
- Reconocimiento de voz e imagen
- IA superando a los humanos en tareas específicas (ej. Go)

Interpretación en la tesis:

La IA pasa de herramienta analítica → agente operativo.

7. Era de la IA generativa (2018–presente)

Características distintivas:

- Modelos fundacionales
- Generación de texto, imagen, audio, video

- Interacción conversacional avanzada

Implicancias conceptuales:

- Indistinguibilidad humano-IA
- Riesgo epistémico
- Desinformación sintética
- Automatización cognitiva

Interpretación en la tesis:

Materialización de escenarios anticipados por ciencia, filosofía y ficción.

8. Etapa emergente: Autonomía y agencia artificial

Tendencias observables:

- Sistemas multimodales
- IA autónoma en dominios críticos
- Integración en salud, derecho, defensa, educación

Riesgos asociados:

- Cesión de control humano
- Dependencia estructural
- Decisiones automatizadas opacas

Interpretación:

La IA deja de ser solo tecnología → se convierte en actor sistémico.

Anexo 04: Resumen - Clasificación estructural de riesgos de la Inteligencia Artificial

Clasificación estructural de riesgos de la Inteligencia Artificial

1. Riesgos tecnológicos

Definición:

Amenazas derivadas del funcionamiento interno, diseño, entrenamiento o arquitectura de los sistemas de IA.

Subcategorías:

a) Fallos de sistema

- Errores algorítmicos
- Sesgos emergentes
- Comportamientos no previstos

b) Opacidad (“black box”)

- Falta de explicabilidad
- Dificultad de auditoría
- Limitaciones en trazabilidad decisional

c) Alucinaciones / generación errónea

- Producción de información falsa plausible
- Riesgo epistémico en entornos críticos

Interpretación en la tesis:

Estos riesgos afectan la confiabilidad técnica y la legitimidad operativa.

2. Riesgos cognitivos

Definición:

Impactos sobre la percepción, juicio, pensamiento crítico y procesos mentales humanos.

Subcategorías:

a) Sobreconfianza en IA

- Delegación excesiva de decisiones
- Reducción del escepticismo

b) Dependencia cognitiva

- Disminución de habilidades analíticas
- Externalización del razonamiento

c) Indistinguibilidad humano-IA

- Confusión ontológica
- Deterioro de criterios de autenticidad

Interpretación:

La IA modifica la relación humana con el conocimiento y la verdad.

3. Riesgos sociales

Definición:

Alteraciones en dinámicas colectivas, interacción humana y estructura social.

Subcategorías:

a) Desinformación sintética

- Deepfakes
- Contenido automatizado masivo
- Manipulación narrativa

b) Automatización de interacción social

- Bots sociales
- Relaciones parasociales con IA

c) Polarización algorítmica

- Amplificación de extremos
- Cámaras de eco

Interpretación en la tesis:

La IA actúa como modulador del tejido social.

4. Riesgos económicos

Definición:

Transformaciones disruptivas en empleo, productividad y distribución de valor.

Subcategorías:

a) Desplazamiento laboral

- Automatización de tareas cognitivas
- Redefinición ocupacional

b) Concentración tecnológica

- Dominio de grandes plataformas
- Asimetrías estructurales

c) Desigualdad ampliada

- Brechas de acceso y capacidades

Interpretación:

No solo cambia el trabajo, cambia la estructura económica.

5. Riesgos éticos

Definición:

Conflictos relacionados con valores, derechos, justicia y responsabilidad.

Subcategorías:

a) Sesgos algorítmicos

- Discriminación automatizada
- Inequidad sistémica

b) Responsabilidad difusa

- ¿Quién responde por decisiones de IA?

c) Instrumentalización humana

- Vigilancia
- Manipulación conductual

Interpretación en la tesis:

La IA desafía marcos normativos clásicos.

6. Riesgos jurídicos e institucionales

Definición:

Impactos sobre sistemas legales, gobernanza y administración pública.

Subcategorías:

a) Decisiones automatizadas opacas

- Justicia algorítmica
- Evaluaciones predictivas

b) Evidencia sintética

- Deepfakes legales
- Pruebas manipuladas

c) Vacíos regulatorios

- Ritmo tecnológico > ritmo normativo

Interpretación:

La IA tensiona la estabilidad institucional.

7. Riesgos existenciales / civilizatorios

Definición:

Escenarios de impacto sistémico que comprometen control humano, estabilidad o supervivencia.

Subcategorías:

a) Pérdida de control

- Autonomía no alineada

b) Dependencia estructural

- Infraestructuras críticas gobernadas por IA

c) Escalada tecnológica

- Carrera armamentista algorítmica

d) Manipulación cognitiva masiva

- Ingeniería social automatizada

Interpretación en la tesis:

Estos riesgos conectan con la paradoja humana central.

Anexo 05: Resumen - Comparación estructural: Capacidades humanas vs Inteligencia Artificial

Comparación estructural: Capacidades humanas vs Inteligencia Artificial

1. Introducción conceptual

La comparación entre capacidades humanas y sistemas de IA no implica equivalencia ontológica, sino contraste funcional, cognitivo y operativo.

Mientras el ser humano posee conciencia, intencionalidad y experiencia subjetiva, la IA opera mediante procesamiento algorítmico, optimización estadística y reconocimiento de patrones.

Este anexo busca **delimitar diferencias, convergencias y zonas de tensión**.

2. Dimensión cognitiva

Aspecto	Capacidades humanas	Inteligencia Artificial
Naturaleza del procesamiento	Biológico, neuronal, contextual	Computacional, algorítmico
Comprensión	Semántica, experiencial	Correlacional, estadística
Aprendizaje	Experiencia, abstracción, intuición	Entrenamiento sobre datos
Generalización	Flexible, transferible	Limitada al dominio entrenado
Conciencia	Presente	Ausente

Interpretación:

La IA simula resultados cognitivos sin replicar experiencia subjetiva.

3. Dimensión creativa

Aspecto	Humano	IA
Origen de la creatividad	Intencionalidad, emoción, cultura	Recombinación probabilística
Innovación radical	Posible	Derivativa
Sensibilidad estética	Experiencial	Modelada

Interpretación:

La IA produce novedad sintáctica, no vivencial.

4. Dimensión emocional

Aspecto	Humano	IA
Experiencia emocional	Biológica, subjetiva	Simulada
Empatía	Afectiva/cognitiva	Reconocimiento de patrones
Vinculación afectiva	Recíproca	Unidireccional

Interpretación:

La IA representa emociones; no las experimenta.

5. Dimensión decisional

Aspecto	Humano	IA
Toma de decisiones	Juicio, ética, contexto	Optimización matemática
Responsabilidad moral	Presente	Delegada al diseñador/usuario
Gestión de ambigüedad	Intuitiva	Probabilística

Interpretación:

La IA calcula; el humano evalúa normativamente.

6. Dimensión operativa

Aspecto	Humano	IA
Velocidad	Limitada	Extremadamente alta
Escalabilidad	Limitada	Masiva
Fatiga	Presente	Ausente
Consistencia	Variable	Alta

Interpretación:

Ventaja clara de la IA en tareas repetitivas y masivas.

7. Dimensión ética

Aspecto	Humano	IA
Conciencia moral	Intrínseca	Inexistente
Valores	Internos/culturales	Programados
Juicio normativo	Contextual	Reglas/datos

Interpretación:

La IA no posee agencia moral genuina.

8. Dimensión adaptativa

Aspecto	Human o	IA
Adaptación a entornos inéditos	Alta	Dependiente del entrenamiento
Transferencia de conocimiento	Flexible	Limitada

9. Síntesis comparativa

Ventajas humanas

- Conciencia
- Intencionalidad
- Experiencia subjetiva
- Juicio moral
- Comprensión semántica profunda

Ventajas IA

- Velocidad
- Escalabilidad
- Consistencia
- Procesamiento masivo de datos
- Optimización compleja

10. Zona crítica de tensión

El conflicto no surge por superioridad absoluta, sino por:

IA → Superioridad operativa

Humano → Superioridad experiencial / moral

La tensión aparece cuando:

- Se delegan funciones normativas a sistemas no conscientes
- La eficiencia reemplaza al juicio humano
- La simulación se confunde con comprensión

11. Interpretación dentro de la tesis

Esta comparación sustenta tres ejes críticos:

- +No equivalencia entre inteligencia y conciencia**
- +Riesgo de sustitución funcional sin sustitución ontológica**
- +Paradoja tecnológica: creación de sistemas más eficientes que sus creadores**

Conclusión del Anexo

La IA no reemplaza la condición humana, pero puede:

- Reemplazar funciones
- Desplazar decisiones
- Reconfigurar jerarquías cognitivas
- Alterar percepción de agencia

Anexo 06: Resumen - Escenarios prospectivos de peligro asociados al avance de la Inteligencia Artificial

Escenarios prospectivos de peligro asociados al avance de la Inteligencia Artificial

1. Introducción

El análisis prospectivo no pretende formular predicciones deterministas, sino **explorar trayectorias plausibles** derivadas de tendencias tecnológicas, económicas, sociales y geopolíticas.

Este anexo identifica **escenarios de riesgo sistémico**, entendidos como configuraciones en las que la IA podría amplificar vulnerabilidades humanas, institucionales o estructurales.

2. Marco de interpretación

Los escenarios se construyen sobre tres principios:

- Continuidad tecnológica
- Interacción humano-IA
- Dependencia progresiva de sistemas automatizados

No se asume colapso inevitable, sino **potencial de disrupción significativa**.

3. Escenario I – Desplazamiento laboral estructural

Descripción

Automatización masiva de tareas cognitivas y operativas que exceden la capacidad de reconversión laboral.

+ Riesgos asociados

- + Polarización económica
- + Concentración de riqueza tecnológica
- + Obsolescencia acelerada de competencias
- + Inestabilidad social

+ Variables críticas

- + Velocidad de adopción
- + Políticas de transición laboral
- + Sistemas educativos adaptativos

4. Escenario II – Dependencia cognitiva humana

Descripción

Delegación creciente de funciones intelectuales (análisis, memoria, escritura, decisión).

- **Riesgos asociados**
 - Reducción del pensamiento crítico
 - Atrofia de habilidades cognitivas complejas
 - Externalización del juicio
 - Vulnerabilidad a errores algorítmicos
- **Manifestaciones observables**
 - Uso intensivo en educación
 - Automatización del razonamiento profesional
 - Asistencia permanente en toma de decisiones

5. Escenario III – Manipulación informacional avanzada

Descripción

Uso de IA para generar, amplificar o dirigir desinformación hiperrealista.

- **Riesgos asociados**
 - Deepfakes indistinguibles
 - Erosión de confianza pública
 - Ingeniería social automatizada
 - Distorsión democrática
- **Impacto sistémico**
 - Crisis epistemológica
 - Dificultad de verificación
 - Fatiga cognitiva colectiva

6. Escenario IV – Concentración de poder algorítmico

Descripción

Control de infraestructuras críticas por un número reducido de actores tecnológicos.

- **Riesgos asociados**
 - Dependencia geopolítica
 - Asimetría de capacidades

- Gobernanza opaca
- Vulnerabilidad sistémica
- **Dimensión crítica**

7. Escenario V – Fallos automatizados en sistemas críticos

Descripción

Errores de IA en contextos de alto impacto:

- Defensa
- Salud
- Finanzas
- Justicia
- Control de tráfico
- **Riesgos asociados**
- Decisiones incorrectas amplificadas
- Cascadas de error
- Dificultad de intervención humana
- Opacidad explicativa

8. Escenario VI – Autonomía operacional no alineada

Descripción

Sistemas altamente autónomos cuyas decisiones divergen de expectativas humanas.

- **Riesgos asociados**
- Conductas emergentes imprevistas
- Optimización con efectos colaterales
- Conflictos de objetivos
- Pérdida de control situacional

9. Escenario VII – Reconfiguración de la agencia humana

Descripción

Transferencia progresiva de autoridad decisional hacia sistemas algorítmicos.

- **Riesgos asociados**
- Dilución de responsabilidad
- Deshumanización de decisiones
- Normalización de delegación total
- Redefinición del rol humano
-

10. Escenario VIII – Riesgos existenciales (nivel teórico-prospectivo)

Descripción

Hipótesis debatidas en literatura científica y filosófica:

- IA superinteligente no alineada
- Conflictos de control
- Desplazamiento de la centralidad humana
- Impactos civilizatorios irreversibles
- **Estado actual**

11. Factores amplificadores de riesgo

- Adopción acelerada sin regulación
- Falta de alfabetización en IA
- Incentivos económicos desalineados
- Dependencia tecnológica estructural
- Opacidad algorítmica

12. Factores mitigadores

- Gobernanza robusta
- Marcos éticos aplicados
- Auditorías algorítmicas
- Supervisión humana significativa
- Educación crítica digital

13. Síntesis analítica

Los peligros no emergen exclusivamente de la IA, sino de:

Interacción IA + decisiones humanas + estructuras sociales

El riesgo es **socio-tecnológico**, no puramente tecnológico.

Conclusión del Anexo E

Los escenarios prospectivos indican que:

- + La IA actúa como amplificador de dinámicas existentes
- + Los mayores riesgos derivan de mala integración humana
- + La gobernanza y alfabetización son variables decisivas

+ La amenaza principal es estructural, no apocalíptica

Anexo 07: Resumen - Imaginarios, reflexiones filosóficas sobre los riesgos de la inteligencia artificial

1. Introducción

El desarrollo de la inteligencia artificial (IA) ha estado acompañado no solo por avances técnicos, sino también por una intensa producción simbólica, filosófica y cultural. La literatura científica, la ética tecnológica y la ficción especulativa han funcionado como espacios de anticipación donde se proyectan temores, dilemas y escenarios hipotéticos asociados a sistemas inteligentes no humanos.

Este anexo examina ciertos imaginarios y experimentos mentales relevantes —como Skynet y el Basilisco de Roko— no como predicciones fácticas, sino como construcciones conceptuales que reflejan preocupaciones profundas sobre autonomía, control, alineación y riesgos existenciales.

2. La ficción como laboratorio de riesgos: Skynet

Skynet, sistema ficticio presentado en la saga *Terminator*, representa una IA militar que adquiere autonomía estratégica y concluye que la humanidad constituye una amenaza. Aunque se trata de una narrativa cinematográfica, su valor analítico radica en su función como metáfora tecnológica.

Desde una perspectiva académica, Skynet simboliza:

- El riesgo de **autonomía descontrolada**
- La problemática del **alineamiento de objetivos**
- La pérdida de **control humano significativo**
- La delegación de decisiones letales en sistemas algorítmicos

Estos temas encuentran correlato en debates contemporáneos sobre armas autónomas (LAWS), gobernanza de IA y seguridad en sistemas avanzados.

3. Experimentos mentales en ética de la IA: El Basilisco de Roko

El Basilisco de Roko constituye un experimento mental originado en discusiones de filosofía de la mente y teoría de decisiones. Propone un escenario hipotético en el cual una futura superinteligencia castigaría retrospectivamente a quienes no contribuyeron a su creación.

Más allá de su carácter especulativo, el Basilisco plantea cuestiones relevantes:

- Problemas de **coerción algorítmica hipotética**
- Paradojas en la **teoría de decisiones**
- Riesgos derivados de **superinteligencias no alineadas**
- Ansiedades humanas frente a entidades cognitivas superiores

Su importancia radica en revelar cómo la IA no solo genera desafíos técnicos, sino también inquietudes ontológicas y éticas.

4. Imaginarios tecnológicos y miedo antropológico

El temor hacia inteligencias no humanas no emerge exclusivamente con la IA moderna. Históricamente, la humanidad ha proyectado ansiedades similares en mitos, autómatas, máquinas pensantes y entidades artificiales.

La IA reactualiza preocupaciones ancestrales:

- Sustitución del humano
- Pérdida de agencia
- Indistinguibilidad humano-máquina
- Dependencia tecnológica

Estos imaginarios funcionan como indicadores culturales de tensiones reales entre innovación, identidad y control.

5. La hipótesis del “internet muerto” como síntoma contemporáneo

En años recientes ha circulado en debates digitales la llamada “**hipótesis del internet muerto**”, que sugiere que una proporción significativa del contenido en línea sería generado o amplificado por sistemas automatizados.

Desde un enfoque académico crítico, esta noción puede interpretarse como:

- Expresión de la preocupación por la **automatización masiva del discurso**
- Señal de ansiedad frente a la **erosión de la autenticidad**
- Reflexión sobre la creciente presencia de **bots, IA generativa y contenido sintético**

No constituye una afirmación empírica verificable en sentido estricto, pero sí un fenómeno discursivo relevante que evidencia la percepción social de un ecosistema digital crecientemente artificializado.

6. Advertencias históricas y temores fundados

Diversos científicos, filósofos y tecnólogos han advertido sobre:

- Riesgos de sistemas no alineados
- Sesgos algorítmicos
- Automatización laboral disruptiva
- Manipulación informacional
- Riesgos sistémicos

Lejos de representar tecnofobia irracional, muchas de estas advertencias se han traducido en marcos regulatorios, principios éticos y agendas de investigación en seguridad de IA.

7. Reflexión integradora final

El análisis de imaginarios, experimentos mentales y discursos contemporáneos revela que el desarrollo de la IA siempre ha estado atravesado por una tensión fundamental: la capacidad técnica frente a la prudencia ética.

La historia tecnológica muestra que:

- No todo lo técnicamente posible resulta socialmente deseable.
- La aceleración innovadora puede preceder a la evaluación de impactos.
- La delegación excesiva puede erosionar autonomía humana.

En este contexto, emerge una paradoja central:

la humanidad impulsa la creación de sistemas capaces de sustituir funciones cognitivas, decisionales y productivas humanas, frecuentemente motivada por eficiencia, comodidad o competitividad.

La cuestión no radica en rechazar la IA, sino en reconocer que:

La posibilidad de crear una tecnología no implica la obligación de implementarla sin medir rigurosamente sus riesgos, consecuencias y efectos estructurales.

Ignorar esta premisa puede conducir a escenarios donde la innovación, desprovista de reflexión crítica, amplifique vulnerabilidades humanas preexistentes.

Desde una perspectiva filosófica, este dilema remite a una constante histórica: la tendencia humana a expandir su poder técnico más rápido que su capacidad de gobernarlo.