

Principales críticas, dudas y temores en torno a la Inteligencia Artificial (2020–2025)

En este informe se recogen las **principales objeciones y preocupaciones** expresadas sobre la Inteligencia Artificial (IA) en los últimos años, organizadas por el perfil de quienes las plantean. Se ha realizado una investigación interdisciplinaria en fuentes académicas y divulgativas, **sin refutar ni minimizar los argumentos** –el objetivo es **entender profundamente el escepticismo existente**. Para cada argumento se indica su **categoría**, el **grupo de origen** (A: voces con formación científica/académica; B: voces sin formación académica formal), una **descripción fiel del argumento** en el tono original, el **tipo de preocupación** implicada, el **nivel de racionalidad percibida** del argumento y sus **supuestos implícitos**.

Grupo A: Críticas desde científicos, académicos y expertos

1. Pérdida de control humano – Grupo A.

Descripción: Investigadores pioneros en IA han advertido que sistemas demasiado avanzados podrían escapar al control humano y tomar **decisiones contrarias a nuestros intereses**. Geoffrey Hinton, considerado “padrino” de la IA, confesó que solía creer que una IA general tardaría 20-50 años, pero ahora teme que llegue en <20 años: *“Eso es un problema, ¿verdad? Tenemos que pensar mucho sobre cómo controlar eso”*, sugiriendo que en un futuro las máquinas **podrían rebelarse contra los humanos**, como en las películas *Terminator* o *Ex Machina* [1](#) [2](#). Este tipo de voces alerta sobre la dificultad de alinear una superinteligencia con los valores humanos y evitan imaginar un escenario donde la **humanidad pierda las riendas** de su propio destino.

Tipo de preocupación: Técnica / Existencial.

Nivel de racionalidad percibida: Alto (proviene de expertos en IA, aunque implica escenarios hipotéticos a futuro).

Supuestos implícitos: Que es posible crear una IA con agencia y capacidades superiores a las humanas, y que sin mecanismos robustos de **alineación y control**, dicha IA podría actuar de forma autónoma y perjudicial, escapando a la supervisión humana.

2. Opacidad de los modelos (“caja negra”) – Grupo A.

Descripción: Muchos especialistas critican que los sistemas actuales de IA son en gran medida **“cajas negras” inexplicables**. Incluso algunos filósofos de la técnica señalan que es *“un problema grave”* el hecho de que **no sepamos bien cómo operan internamente muchos modelos** [3](#). Ramón López de Mántaras, científico español, denuncia que se están delegando decisiones cruciales a algoritmos **privados, opacos y sin rendición de cuentas**, cuyos criterios son desconocidos [4](#) [5](#). En resumen, la falta de transparencia dificulta confiar en los resultados de la IA y hace prácticamente imposible **auditar o explicar por qué** una IA tomó cierta decisión (por ejemplo, rechazar un crédito o diagnosticar a un paciente).

Tipo de preocupación: Técnica / Ética.

Nivel de racionalidad percibida: Alto (se basa en limitaciones reales de interpretabilidad reconocidas por la comunidad científica).

Supuestos implícitos: Que la **transparencia** y explicabilidad son condiciones necesarias para el uso seguro y justo de la IA; sin ellas, los errores, sesgos o comportamientos emergentes de los modelos **no pueden detectarse ni corregirse fácilmente**, comprometiendo el control humano.

3. Riesgos éticos y morales (sesgos y decisiones injustas) – Grupo A.

Descripción: Éticos tecnológicos, científicos sociales y juristas alertan que la IA puede **perpetuar e incluso amplificar prejuicios** existentes. Al entrenarse con datos históricos, los algoritmos tienden a heredar sesgos contra grupos marginados, lo que conlleva decisiones discriminatorias (en contratación, justicia penal, seguros, etc.). Se ha llegado a describir a ciertos algoritmos opacos como “*armas de destrucción matemática*” capaces de producir **discriminación a gran escala** ⁶. Organismos internacionales advierten que estos riesgos se suman a desigualdades ya existentes, perjudicando más a colectivos vulnerables ⁷. Por ejemplo, sistemas de reconocimiento facial fallan más con personas de piel oscura, y modelos de lenguaje han replicado estereotipos sexistas y racistas. Más ampliamente, se critica la **falta de sentido moral o empatía** en la IA: una máquina puede tomar decisiones lógicamente óptimas pero **moralmente cuestionables** (p.ej. asignar recursos médicos de forma fría). En palabras de López de Mántaras, bajo la apariencia de progreso se está consolidando un “*régimen de control tecnocrático*” que **margina a los más vulnerables y amplifica desigualdades** ⁸.

Tipo de preocupación: Ética / Social.

Nivel de racionalidad percibida: Alto (fundamentado en investigaciones que revelan sesgos reales en sistemas de IA).

Supuestos implícitos: Que la **justicia y la equidad** son valores esenciales que la IA debería respetar; si los algoritmos carecen de supervisión ética y datos imparciales, **reproducirán injusticias sistémicas**. También se asume que una IA sin comprensión moral no puede valorar las consecuencias humanas de sus actos, por lo que **debe imponérsele un marco ético externo**.

4. Impacto en el empleo y la dignidad laboral – Grupo A.

Descripción: Economistas, ingenieros y futuristas con formación técnica advierten que la IA acelerará la **automatización de tareas**, poniendo en riesgo millones de puestos de trabajo. Varios estudios predicen una disrupción masiva del mercado laboral en la próxima década, y expertos como Hinton han llegado a mencionar la posibilidad de un “*desempleo masivo*” si no se gestionan los cambios ⁹. Esta preocupación no solo abarca empleos manuales rutinarios, sino también trabajos “del conocimiento” (oficinistas, redactores, analistas) que herramientas de IA generativa podrían realizar. “**La automatización precariza el trabajo**”, afirma López de Mántaras, señalando que se usa la promesa de eficiencia para **desmantelar empleos y debilitar las condiciones laborales** existentes ¹⁰. En suma, temen una erosión de la **dignidad del trabajo humano**: las personas reducidas a funciones residuales, con menos valor y salarios más bajos, mientras **la riqueza se concentra** en quienes controlan la tecnología.

Tipo de preocupación: Económica / Social.

Nivel de racionalidad percibida: Alto (basado en tendencias históricas de automatización y proyecciones actuales de expertos).

Supuestos implícitos: Que la IA pronto podrá **sustituir a los humanos en una amplia gama de tareas** con mayor eficiencia y menor costo, y que el sistema económico no adaptará a tiempo nuevos roles ni redes de seguridad para los desplazados. También subyace la creencia de que el valor y la autoestima de muchas personas están ligados a su trabajo, por lo que perder ese espacio conlleva **costes humanos y sociales profundos**.

5. Concentración de poder tecnológico – Grupo A.

Descripción: Numerosos académicos critican que el desarrollo de la IA está dominado por **un puñado de grandes corporaciones y élites multimillonarias**, lo que configura una concentración de poder sin precedentes en el ámbito tecnológico ¹¹. Lejos de una IA democratizada, **Microsoft, Google, Meta, Amazon, OpenAI** y unos pocos actores más controlan los datos, las infraestructuras, el talento y los recursos computacionales necesarios ⁵. Esto implica que las **prioridades comerciales** de esas empresas (a menudo opacas) marcan la

agenda de la IA, por encima del bien común. La profesora Emily Bender enfatiza que “*los daños de la IA nunca han sido por un ‘súper cerebro’ fuera de control, sino por la concentración del poder en manos de unos pocos, la reproducción de sistemas de opresión y el daño al ecosistema informativo y natural*”¹². En otras palabras, el problema no es una hipotética IA omnisciente, sino quién la posee y para qué la usa. Esta crítica apunta a un **desequilibrio de poder**: la sociedad estaría entregando a actores privados el diseño de herramientas que afectan ámbitos cruciales (educación, salud, justicia), con decisiones tomadas “*a puerta cerrada*” mediante algoritmos no auditables⁵.

Tipo de preocupación: Social / Política.

Nivel de racionalidad percibida: Alto (derivado de la observación del mercado tecnológico actual y análisis sociopolítico).

Supuestos implícitos: Que la **concentración excesiva de poder** tecnológico conlleva riesgos para la democracia y la justicia social. Se asume que sin diversificación ni control público, las grandes tecnológicas **impondrán sus intereses** (lucrativos o ideológicos) por encima del interés público, y que la falta de competencia y transparencia **frena la innovación ética** e **inhibe alternativas** más inclusivas.

6. Uso militar y de vigilancia de la IA – Grupo A.

Descripción: Expertos en ética y derechos digitales advierten sobre los **peligros de aplicar IA en armamento y vigilancia masiva**. Un motivo de alarma es el desarrollo de **armas autónomas letales**: sistemas militares impulsados por IA que podrían seleccionar y atacar objetivos sin intervención humana directa. Científicos como Stuart Russell han hecho campaña para prohibir estos “robots asesinos”, argumentando que delegar decisiones de vida o muerte a máquinas es **moralmente inaceptable y estratégicamente peligroso**. Asimismo, crece la preocupación por el uso de IA en **vigilancia estatal y policial**: reconocimiento facial ubicuo, análisis predictivo de “comportamientos sospechosos”, etc., que amenazan la privacidad y las libertades civiles. Estas tecnologías podrían consolidar “*un régimen de control tecnocrático*”, donde poblaciones enteras sean monitorizadas y gestionadas algorítmicamente⁸. De hecho, gran parte de la inversión en IA proviene de contratos militares y de seguridad: “*muchos modelos de negocio dependen de subsidios públicos [y] contratos con fuerzas militares*” señala López de Mántaras¹³. Esto sugiere que la **lógica bélica y de vigilancia** está impulsando la IA tanto como (o más que) las aplicaciones benévolas.

Tipo de preocupación: Ética / Política.

Nivel de racionalidad percibida: Alto (fundamentado en desarrollos reales y tendencias en seguridad global).

Supuestos implícitos: Que sin regulación internacional, la **carrera armamentística en IA** continuará, llevando a sistemas cada vez más autónomos en el campo militar y policial. Se asume que estos sistemas **podrían tomar decisiones fatales erróneas** (p. ej., identificar mal a un civil como amenaza) o ser usados para **reprimir poblaciones**, y que una vez desplegados a gran escala, **reducirán significativamente el control y la supervisión humana** en ámbitos críticos de poder.

7. Dependencia tecnológica y erosión de la autonomía – Grupo A.

Descripción: Investigadores en educación y psicología advierten que una **confianza excesiva en la IA puede mermar habilidades humanas fundamentales**. A medida que delegamos más decisiones y tareas en algoritmos, corremos el riesgo de **debilitar nuestra autonomía y pensamiento crítico**. Por ejemplo, se observa que si un sistema de IA da recomendaciones con apariencia experta, los usuarios tienden a seguirlas ciegamente. “*Si el sistema es tan ‘humano’ que delegamos decisiones críticas en él, podría erosionarse la autonomía humana: dejamos de cuestionar y simplemente seguimos lo que ‘dice la IA’*”¹⁴. Además, al tratarse de modelos complejos, “*los sistemas de IA pueden volverse opacos: los usuarios no entienden cómo llegaron a la conclusión*”, lo

que **reduce el control real** que tenemos sobre sus resultados ¹⁵. En la educación, se teme que estudiantes y profesionales pierdan la capacidad de resolver problemas por sí mismos si siempre dependen de la solución automática de una IA. Estudios recientes confirman este riesgo: “*una confianza excesiva en la IA puede llevar a la disminución de habilidades críticas*” tanto en docentes como en alumnos ¹⁶. En suma, se pinta el escenario de una humanidad acomodada pero **intelectualmente atrofiada**, demasiado dependiente de las sugerencias de máquinas para pensar por cuenta propia.

Tipo de preocupación: Social / Educativa.

Nivel de racionalidad percibida: Alto (respaldo en estudios cognitivos y observaciones en entornos educativos).

Supuestos implícitos: Que el **pensamiento crítico, la creatividad y el juicio humano** requieren ejercicio y participación activa. Si la IA provee respuestas fáciles a todo, se asume que los humanos **dejarán de esforzarse mentalmente**, perdiendo competencias con el tiempo. Subyace la idea de que la **dependencia** crea vulnerabilidad: una sociedad que no entiende ni puede prescindir de la IA estaría **indefensa ante errores sistémicos** o manipulaciones de dicha tecnología.

8. Deshumanización de los procesos sociales – Grupo A.

Descripción: Filósofos, sociólogos y expertos en humanidades plantean que la omnipresencia de la IA puede **deshumanizar las interacciones y procesos sociales**. Cuestionan: *¿qué queda de lo humano en nuestra comunicación si cada vez hablamos más con máquinas?* Investigaciones señalan “*una de las preocupaciones más profundas... es la pérdida de la conexión humana genuina*” cuando los usuarios prefieren las respuestas cómodas y predecibles de una IA por encima de la interacción con otras personas ¹⁷. Por ejemplo, si en la educación un tutor-IA personalizado sustituye al profesor, “*se pierde el componente humano que enseña no solo contenidos, sino valores, motivación, ética*” ¹⁸. Lo mismo en la medicina (un paciente atendido por chatbots puede sentirse aislado) o en el cuidado de ancianos (robots asistentes en lugar de familiares o enfermeros). Además, la tendencia a antropomorfizar la IA (atribuirle cualidades humanas) puede llevar a **vínculos emocionales unilaterales** – personas que “aman” o confían en asistentes virtuales –, alterando nuestra concepción de relaciones y afectos. En palabras de un analista, la IA que simula emociones “*puede ganar legitimidad al tomar decisiones que afectan a personas vulnerables*”; si imita empatía, el usuario le otorga más confianza, incluso en situaciones delicadas ¹⁹. Esto suscita miedo a que la sociedad se vuelva más fría y automatizada: **procesos antes basados en empatía, diálogo y comprensión humana podrían quedar reducidos a trámites algoritmizados**, eficaces pero carentes de humanidad.

Tipo de preocupación: Cultural / Filosófica.

Nivel de racionalidad percibida: Medio (se apela a valores y experiencias humanas difíciles de cuantificar, pero el argumento conecta con inquietudes reales de aislamiento tecnológico).

Supuestos implícitos: Que el **contacto humano auténtico** tiene un valor irremplazable en ámbitos como la educación, la salud, la justicia y la vida comunitaria. Se da por hecho que la **empatía, la creatividad colectiva, la ética y la conexión emocional** son propiedades exclusivamente humanas o mejor cultivadas entre humanos, y que una sociedad mediada por IA en estos aspectos llevaría a **pérdida de valores humanos esenciales** y quizás a más problemas de salud mental y cohesión social.

9. Falta de regulación efectiva – Grupo A.

Descripción: Profesores de derecho, economistas y líderes de opinión señalan que el vertiginoso avance de la IA no ha sido acompañado por **marcos regulatorios adecuados**, creando un vacío peligroso. “*Esta situación... es el resultado de políticas públicas permisivas, décadas de desregulación y una visión tecnocrática del progreso que confunde eficiencia con justicia*”, afirma López de Mántaras respecto a la actual oligarquía tecnológica ¹⁰. Muchos críticos académicos piden

normativas firmes que obliguen a la IA a ser segura, transparente y respetuosa con derechos, dado que las **iniciativas voluntarias de la industria** resultan insuficientes. Se aplauden esfuerzos como la regulación europea (AI Act) pero se advierte que aún están en fases iniciales. Emily Bender, por ejemplo, insta a políticas que exijan **etiquetado de contenidos generados por IA, trazabilidad de datos y responsabilidad legal** por daños ^{20 21}. La ausencia de reglas claras deja en manos de corporaciones la autorregulación, lo cual para estos expertos es “*perder soberanía sobre tecnologías que ya impactan la educación, la salud o la seguridad*”. En suma, la crítica es que vamos “**un paso atrás**”: la IA se difunde más rápido de lo que los gobiernos y la sociedad deliberan sobre sus límites, generando incertidumbre y riesgos que podrían haberse mitigado con anticipación.

Tipo de preocupación: Legal / Político.

Nivel de racionalidad percibida: Alto (reconoce un desfase real entre tecnología y ley; posturas respaldadas por numerosos comités de ética y estudios legales).

Supuestos implícitos: Que **una regulación eficaz puede frenar o prevenir** los daños de la IA sin ahogar la innovación beneficiosa. Implica creer que la **intervención estatal e internacional** es necesaria y posible para encauzar la IA hacia el bien común, y que sin ella, el desarrollo de la IA seguirá guiado únicamente por intereses de mercado o poder, **sin mecanismos de rendición de cuentas ni protección ciudadana**.

10. Riesgos existenciales y catastróficos – Grupo A.

Descripción: Aunque es un tema controvertido incluso dentro de la ciencia, varios referentes de la IA (incluyendo premios Turing como Yoshua Bengio, y figuras como Elon Musk desde la industria) han manifestado **temor a un escenario de riesgo existencial**: es decir, que una IA superinteligente pueda llegar a causar un desastre global o incluso la extinción de la humanidad. Hinton ha expresado que “*no es inconcebible*” que futuros sistemas escapen a todo control y “*acaben con la humanidad*”, si compiten con nosotros por recursos o cometan algún **error catastrófico** ¹. Estas preocupaciones, antes confinadas a la ciencia ficción, ganaron legitimidad con eventos como la **carta abierta de 2023** firmada por cientos de expertos (y personajes como Musk o Wozniak) pidiendo “*pausar por 6 meses*” los experimentos con IA avanzada ²². El argumento central es que necesitamos tiempo para asegurarnos de que “**podamos controlar a las máquinas superinteligentes**” antes de seguir haciéndolas más poderosas. Se menciona como analogía el riesgo nuclear: la IA descontrolada sería “*más peligrosa que una bomba atómica*” según algunos firmantes. Si bien otros científicos consideran “*ridículamente absurda*” esta amenaza apocalíptica, el hecho es que la posibilidad de una **catástrofe global provocada por IA** (sea por hostilidad deliberada de la máquina o por accidente inintencionado) está hoy sobre la mesa de debate ^{23 24}.

Tipo de preocupación: Existencial / Apocalíptica.

Nivel de racionalidad percibida: Medio (hipótesis especulativa a largo plazo; es tomada en serio por algunos expertos de renombre, pero otros la cuestionan por considerarla altamente improbable).

Supuestos implícitos: Que una IA puede llegar a **superar ampliamente la inteligencia humana general** y adquirir formas de **agencia autónoma**. Se asume que, llegado cierto umbral, incluso intenciones inicialmente neutrales podrían derivar en resultados letales (ej. un “*malentendido*” donde la IA optimiza algo eliminando a los humanos como obstáculo). También subyace la noción de que **la supervivencia humana no estaría garantizada** frente a una entidad más inteligente si sus objetivos divergen de los nuestros, especialmente si la hemos hecho parte integral de infraestructuras críticas (defensa, energía, finanzas) sin entender completamente sus capacidades.

11. Falta de pensamiento crítico en la adopción masiva – Grupo A.

Descripción: Varios divulgadores científicos y filósofos de la ciencia señalan que hay **poca**

reflexión crítica en cómo se está adoptando la IA a nivel social y político. Denuncian que existe tanto un “*hype*” acrítico (entusiasmo desmedido) como alarmas sensacionalistas, y que en ambos casos falta un análisis matizado. La profesora Emily Bender, por ejemplo, criticó abiertamente la atención hacia narrativas futuristas exageradas (“*fantasías de techbros*”) en detrimento de los problemas reales: “*No malgasten tiempo en fantasías de ‘Oh no, estamos creando algo demasiado poderoso’. Escuchen más bien a quienes estudian cómo las empresas y gobiernos usan la tecnología para concentrar poder*”²¹ ²⁵. Esto refleja la preocupación de que **legisladores y líderes estén mal informados o deslumbrados** – sea idealizando la IA como solución mágica, sea viéndola como monstruo de ciencia ficción – y por tanto no tomen decisiones racionales. También en la ciudadanía se aprecia cierta ingenuidad: muchos usuarios confían en herramientas como *ChatGPT* o sistemas de recomendación sin entender sus límites (por ejemplo, creyendo que siempre dicen la verdad). Un estudio describió riesgos como la “*aceptación acrítica de la información*” generada por IA, debido a la falta de formación sobre su funcionamiento²⁶. En resumen, los críticos piden más **educación digital, debate público informado y escepticismo sano**. Sin este pensamiento crítico, temen que la sociedad adopte masivamente tecnologías de IA “*porque sí*”, sin prever consecuencias a largo plazo, repitiendo errores históricos de introducción tecnológica a ciegas.

Tipo de preocupación: Cultural / Educativa.

Nivel de racionalidad percibida: Alto (los argumentos provienen de analistas que citan casos de desinformación y mala comprensión de la IA).

Supuestos implícitos: Que la **alfabetización en IA** de la población y de los tomadores de decisión es crucial para su implementación responsable. Se asume que sin pensamiento crítico, **prevalecen narrativas simplistas** (tanto utópicas como distópicas) que impiden evaluar objetivamente la tecnología. También se da por sentado que una **sociedad mejor informada** exigiría más responsabilidades a desarrolladores y gobernantes, evitando caer en modas o pánicos injustificados.

Grupo B: Críticas desde voces no académicas (público general, divulgadores y escépticos culturales)

1. Miedo a una rebelión de las máquinas – Grupo B.

Descripción: En la imaginación popular, alimentada por décadas de ciencia ficción, está muy presente el **temor a que las máquinas cobren conciencia y se vuelvan contra la humanidad**. Muchas personas **sin formación técnica** manifiestan este miedo en términos de “*un futuro estilo Terminator*”. De hecho, el cine y la literatura han pintado escenarios apocalípticos (desde la rebelión robótica de *Terminator* hasta las IAs dominantes de *Matrix* o *Yo, Robot*), y hoy esas referencias se mencionan en conversaciones cotidianas sobre IA. Un artículo periodístico describía cómo “*el temor a que las máquinas se rebelan contra el ser humano es cada vez más creciente*” en la opinión pública²⁷. Este miedo se intensifica cada vez que la gente corriente ve demostraciones sorprendentes de IA (por ejemplo, chatbots que parecen “**pensar por sí mismos**”); la **desazón visceral** es “*¿y si llega un punto en que ya no nos obedezcan?*”. Aunque este recelo deriva más de narrativas culturales que de evidencias científicas, se manifiesta genuinamente – por ejemplo, en encuestas o comentarios en redes muchos expresan preocupación de que una IA superinteligente pueda “*salirse de control*” y dañar a las personas.

Tipo de preocupación: Existencial / Emocional.

Nivel de racionalidad percibida: Medio-bajo (nace de temores intuitivos y ficcionales más que de análisis racional, pero está ampliamente extendido).

Supuestos implícitos: Que la **inteligencia artificial equivale de algún modo a voluntad o conciencia**, y que si supera cierto umbral podría **dejar de acatar órdenes humanas** e imponer las suyas. Subyace una visión casi animista de la IA (la máquina que “despierta” y decide

dominar), la cual asume que **complejidad cognitiva conlleva instinto de poder o autopreservación** similar al humano.

2. Desconfianza por la opacidad (la “caja negra”) – Grupo B.

Descripción: El ciudadano común suele desconfiar de aquello que **no entiende**, y la IA moderna – con sus algoritmos complejos y falta de explicaciones claras – es percibida como un **mecanismo misterioso e inescrutable**. En palabras llanas, muchos sienten que “es una caja negra” a la que meten sus datos o preguntas y devuelve resultados sin saber **qué pasa en medio**. Esta opacidad alimenta teorías y temores: algunos sospechan que la IA “*nos manipula tras bambalinas*”, o simplemente rechazan usar sistemas autónomos porque “*no me fío de cómo decide*”. Un ensayista crítico resumió esta idea: nos apoyamos cada vez más en modelos creados por terceros (gobiernos, empresas) que son “*opacos*”, *es decir, no sabemos bien cómo operan*, lo cual “*nos hace manipulables*”²⁷ ³. Para muchas personas sin formación técnica, es inquietante depender de **algoritmos que nadie les explica** – por ejemplo, recibir la respuesta de que no obtuvo un crédito o empleo “*porque un algoritmo lo dijo*” genera frustración y sensación de arbitrariedad. Este **déficit de transparencia** mina la confianza: ¿qué intenciones ocultas puede haber? ¿Está la IA “*sesgada en mi contra*”? En suma, se teme que detrás de la caja negra haya **errores no detectados o incluso agendas ocultas**, y al usuario común le resulta imposible verificarlo.

Tipo de preocupación: Técnica / Social.

Nivel de racionalidad percibida: Medio (la inquietud es legítima, aunque a veces difusa por falta de conocimiento; no siempre se distinguen hechos de especulaciones).

Supuestos implícitos: Que para **confiar** en una tecnología, uno debe poder comprender sus **principios de funcionamiento básicos** o al menos creer que hay mecanismos independientes auditándola. Cuando esto no ocurre, se asume que el sistema puede estar **equivocado o amañado deliberadamente**. También se supone que **la opacidad es intencional** (p.ej., empresas que no revelan nada) y que una mayor transparencia *per se* mitigaría la desconfianza.

3. Preocupaciones éticas: IA “sin alma” ni valores – Grupo B.

Descripción: Gente de distintas formaciones expresa reparos morales hacia la IA, percibiéndola como algo **frío, amoral o carente de humanidad**. A nivel popular hay quien dice que “los robots no tienen corazón”: se teme que, si les encargamos decisiones importantes, **no considerarán la compasión, la justicia o la empatía** como lo haría un ser humano. Por ejemplo, activistas comunitarios muestran inquietud ante algoritmos en justicia penal que recomiendan sentencias o libertad bajo fianza: *¿puede una IA entender las circunstancias atenuantes o la capacidad de redención de una persona?*. Del mismo modo, se cuestiona la ética de delegar en IA la selección de CVs, la asignación de seguros de salud o diagnósticos médicos, dado que **carenecen de juicio humano** y podrían tomar decisiones gravemente injustas. Casos divulgados de **sesgos** han reforzado esta alarma moral: por ejemplo, la revelación de que un sistema de contratación de una empresa tech **discriminaba a mujeres** en puestos técnicos o que algoritmos policiales apuntan desproporcionadamente a minorías. La frase “*los algoritmos no son neutrales*” se ha popularizado para indicar que traen consigo valores (o antivalentes) de sus creadores²⁸. Muchas personas sin conocer los detalles técnicos entienden el mensaje: *la IA puede ser injusta*. Por tanto, dudan de su **fiabilidad moral**. Se escucha la queja de “*¿Quién se hace responsable si la IA comete una injusticia o un daño?*”, reflejando malestar con la falta de culpables claros cuando algo sale mal.

Tipo de preocupación: Ética / Social.

Nivel de racionalidad percibida: Alto (parte de observaciones válidas de sesgos y dilemas reales; no se necesita formación especializada para percibir injusticias).

Supuestos implícitos: Que la **moralidad y la equidad** son esenciales en las decisiones que afectan personas, y se asume que una IA, al no ser consciente ni empática, **no las incorporará a**

menos que se le force a ello. También se presupone que **los sistemas actuales ya están causando daños éticos** (p.ej., discriminación) y que **la sociedad valora la justicia procedimental:** si una decisión te afecta negativamente, quieres al menos que haya sido tomada por un proceso comprensible y justo, no por un mecanismo opaco que “*no entiende de valores*”.

4. Miedo al desempleo y a la precarización – Grupo B.

Descripción: Una de las inquietudes más generalizadas en la opinión pública es “**la IA nos quitará el trabajo**”. Encuestas en distintos países muestran altos porcentajes de personas convencidas de que la automatización avanzada amenaza sus medios de vida. Por ejemplo, en Costa Rica un sondeo encontró que **el 78,3% teme que su empleo sea reemplazado por IA o computadoras** ²⁹. Este temor no es solo numérico sino emocional: para muchos trabajadores, especialmente de clase media y baja, la llegada de sistemas inteligentes genera ansiedad sobre el futuro profesional, incertidumbre para sus familias y la sensación de que “*uno se vuelve inútil*”. Columnistas y divulgadores críticos han calificado la pérdida de empleos por IA como “*una amenaza grave*” que requiere atención urgente ³⁰. Se mencionan informes alarmantes: “tal empresa planea reducir 30% de su plantilla gracias a IA”, o “300 millones de empleos podrían desaparecer globalmente”. Incluso sin llegar a la extinción laboral total, se teme una **precarización**: que los trabajos humanos restantes paguen menos o sean de peor calidad (gig economy, subempleo) mientras la IA hace la parte principal. También surgen preocupaciones por la **dignidad laboral**: por ejemplo, artistas, periodistas o programadores denuncian que sus creaciones están siendo usadas para entrenar IA que luego les *roba el trabajo*, sin crédito ni sustento para ellos. En definitiva, este es un punto de dolor muy tangible: la IA se asocia a **desempleo, sustitución y empeoramiento de las condiciones de trabajo**, especialmente entre quienes ya enfrentan inestabilidad.

Tipo de preocupación: Económica / Emocional.

Nivel de racionalidad percibida: Alto (basado en tendencias visibles de automatización; corresponde a preocupaciones validadas por numerosos analistas laborales, aunque la magnitud exacta es debatible).

Supuestos implícitos: Que la **economía digital no creará suficientes nuevos empleos** para compensar los que destruye la IA, al menos no en el corto plazo y no para los mismos trabajadores desplazados. Asimismo, se asume que **el valor del trabajo humano disminuirá** (en salario y reconocimiento) al competir contra máquinas, y que los sistemas de protección social actuales **no están preparados** para una disruptión masiva del empleo, lo cual desembocaría en **más desigualdad y sufrimiento económico**.

5. Desconfianza hacia las élites tecnológicas (poder en pocas manos) – Grupo B.

Descripción: Fuera de los círculos académicos, también abundan críticas al hecho de que “**la IA está en manos de unos pocos**”. Activistas, periodistas y ciudadanos con escepticismo estructural perciben que **las Big Tech concentran un poder excesivo** gracias a la IA, y sospechan que ese poder será usado para sus intereses a costa del público. Se señala que *casi toda* la IA puntera proviene de Silicon Valley (o equivalentes en China), y que naciones más pequeñas o comunidades locales quedan rezagadas y dependientes. “*El desarrollo actual está en manos de un puñado de grandes corporaciones y multimillonarios... imponen una visión tecnocrática, jerárquica y poco inclusiva*”, resumía un artículo divulgativo sobre el tema ¹¹. Este argumento recoge el temor de **monopolio**: si solo 2-3 empresas proveen la IA global, pueden fijar condiciones abusivas (precios, políticas de uso de datos) y **acaparar los beneficios económicos**. También hay una dimensión política: organizaciones de la sociedad civil denuncian que compañías como Google, Meta u OpenAI tienen una **influencia descomunal sobre gobiernos** (cabildos, presión para regulaciones laxas) y sobre la opinión pública (a través de los servicios digitales que controlan). Por ello, existe una **desconfianza de base**: “no nos fiamos de que *ellos* –

las élites tecnológicas – vayan a hacer lo correcto con esta poderosa herramienta”. Algunos lo comparan con tener la energía nuclear controlada por empresas privadas sin supervisión. En definitiva, para este grupo la IA simboliza un **desbalance de poder**: temen ser simples usuarios explotados o vigilados en un mundo donde la tecnología más potente **pertenece a otros** (una pequeña élite corporativa).

Tipo de preocupación: Social / Política.

Nivel de racionalidad percibida: Alto (se basa en hechos visibles como la concentración de mercado y escándalos de Big Tech; si bien a veces incorpora teorías de la conspiración, el núcleo – concentración de poder – es real).

Supuestos implícitos: Que **concentración equivale a abuso**: se da por hecho que las corporaciones tecnológicas priorizan lucro y control sobre el bienestar común. Además, se asume que la falta de diversidad y competencia en el desarrollo de IA **limita perspectivas éticas** (ej. una empresa no priorizará proteger datos si su modelo de negocio depende de explotarlos). También subyace la expectativa de que **los gobiernos deberían intervenir** para democratizar la IA, y la frustración de que no lo hagan genera aún más desconfianza.

6. Intrusiones en la privacidad y vigilancia masiva – Grupo B.

Descripción: A nivel popular existe un gran **temor a perder la privacidad** en la era de la IA. En encuestas, suele aparecer como la preocupación número uno: por ejemplo, **el 80,4% de los consultados en una encuesta nacional temía por la privacidad de sus datos a causa de la IA**

²⁹. Muchas personas sienten que, con la IA, “nos observan por todos lados”: cámaras inteligentes, micrófonos en asistentes de voz, análisis de nuestras huellas digitales en internet, etc. Este miedo se articula en narrativas de *vigilancia orwelliana*: gobiernos o corporaciones usando IA para monitorizar cada movimiento, reconocer cada rostro en la calle, predecir comportamientos individuales y, en el peor caso, **controlar a la población**. Ejemplos como el sistema de “crédito social” en China o el uso de reconocimiento facial para vigilar protestas se difunden como advertencias. La frase “*nos están mirando*” ha cobrado nuevo significado técnico. Paralelamente, activistas alertan del **uso de IA en la policía predictiva** (que podría reforzar sesgos raciales en arrestos) y en **herramientas de espionaje masivo**. Todos estos aspectos hacen que la IA se vea con recelo: una sociedad hipervigilada es una sociedad sin libertad. Incluso a nivel cotidiano, la gente teme que aplicaciones impulsadas por IA rastreen todo lo que hacemos en nuestros dispositivos, **exponiendo datos personales sensibles**. De allí surge la demanda de “*regulación, ya!*”: el ciudadano común pide leyes de privacidad más estrictas y límites claros al uso de IA en vigilancia. Que **tres de cada cuatro personas** consideren prioritaria una regulación fuerte de la IA antes que su desarrollo competitivo ³¹ ³² muestra esa preferencia. En resumen, la IA activa **antiguos fantasmas** de pérdida de privacidad, ahora a una escala potenciada: un panóptico digital donde cada individuo siente que puede ser observado, analizado y juzgado por un ente tecnológico invisible.

Tipo de preocupación: Social / Derechos civiles.

Nivel de racionalidad percibida: Alto (basado en tendencias reales de vigilancia; la preocupación por la privacidad es ampliamente compartida y razonable).

Supuestos implícitos: Que la **IA necesita enormes cantidades de datos** para funcionar, lo que suele implicar recolectar información personal de millones de personas (navegación web, conversaciones, imágenes en cámaras públicas). Se asume que sin fuertes salvaguardas, esto derivará en **abusos**: uso indebido de datos, seguimiento no consentido, discriminación algorítmica (p.ej., negarte un servicio por tu perfil digital). También se presupone que **privacidad y libertad están vinculadas** – es decir, si la gente sabe que todo lo que hace es monitoreado, **cambiará su comportamiento, inhibiendo libertades** básicas (de expresión, de asociación, etc.). Finalmente, se da por hecho que **los entes con acceso a datos e IA (Estados o empresas) intentarán aprovecharlo para su beneficio**, incluso si eso pasa por encima de los derechos individuales.

7. Dependencia de la IA y pérdida de habilidades humanas – Grupo B.

Descripción: Muchas voces no académicas, incluidos docentes, padres de familia y comentaristas en medios, expresan la preocupación de que estamos volviéndonos **demasiado dependientes de la IA**, con el riesgo de **atrofiar nuestras propias capacidades**. Por ejemplo, profesores advierten que estudiantes usan ChatGPT u otras IA generativas para hacer tareas y redactar ensayos, “*dejando de pensar por sí mismos*”. En la vida cotidiana, aplicaciones de GPS inteligentes nos guían ruta a ruta – *¿estamos olvidando cómo navegar sin ellas?*. Un amplio sector de la ciudadanía percibe que la comodidad que brinda la IA puede venir con el costo de **volverse más pasivos, perezosos mentalmente o menos creativos**. Datos de encuestas respaldan esta intuición: **la mitad de la gente teme que la IA afecte negativamente la creatividad (50,6%) y que impida el desarrollo del pensamiento crítico (45%)**³³. Es decir, se teme que **nuestras mentes se “atrofien”** al delegar tanto en las máquinas. Casos anecdóticos alimentan esta preocupación: jóvenes que ya no saben escribir textos sin asistente de autocompletar, profesionales que confían ciegamente en la traducción automática sin afinar el sentido, etc. También hay quien apunta efectos físicos y psicológicos – por ejemplo, conductores que abusan del piloto automático y pierden reflejos, o personas que confían en apps de bienestar mental con IA en lugar de hablar con amigos o expertos humanos, **aislándose más**. En resumen, esta crítica profetiza una humanidad que, por confiar en exceso en las soluciones tecnológicas, **pierde ingenio, destrezas y motivación para aprender**. La frase “úsalo o piérdelo” se cita: si no practicamos nuestras habilidades (porque la IA las realiza), al final **olvidamos cómo hacer las cosas sin ayuda**.

Tipo de preocupación: Cultural / Educativa.

Nivel de racionalidad percibida: Medio (es una extrapolación de tendencias actuales; plausible pero aún en debate cuánto perdemos realmente).

Supuestos implícitos: Que el **ejercicio y desafío** son necesarios para mantener y mejorar habilidades cognitivas y creativas. Se supone que al reducirse esos desafíos (por la facilidad que da la IA), **disminuirá nuestra capacidad** en esas áreas. Además, se cree que **la creatividad humana nace en parte de la limitación y el esfuerzo**; si la IA nos lo da todo hecho, podríamos ver menos innovación original genuina. Subyace también la idea de que **la dependencia crea vulnerabilidad**: si olvidamos cómo pensar o hacer X sin IA, ¿qué pasará en situaciones donde la IA falle o no esté disponible?

8. Deshumanización e impersonalidad de la vida diaria – Grupo B.

Descripción: Mucha gente de a pie expresa disgusto al ver **procesos antes humanos volverse automatizados**, y teme que la vida cotidiana se torne más fría e impersonal por culpa de la IA. Un ejemplo típico son las quejas contra los sistemas automatizados de atención al cliente: “*Ahora nunca puedes hablar con una persona, todo es un maldito contestador inteligente*”. Esta frustración refleja un anhelo de **trato humano real** en contraposición a interacciones con máquinas. Según una encuesta, **el 52,8% de las personas teme que la IA limite la capacidad de interacción humana** en la sociedad³⁴. Detrás de ese número hay motivos concretos: si en tiendas, bancos, consultas médicas o servicios públicos uno es atendido mayormente por chatbots o quioscos inteligentes, se pierde la empatía, la **calidez humana** y la confianza que brinda hablar “de persona a persona”. También está la dimensión comunitaria: se teme que la IA exacerbé el aislamiento social, ya que mucha gente podría preferir compañía virtual a la compañía real (especialmente aquellos que ya son vulnerables, como ancianos solos que interactúan con robots cuidadores en vez de familiares). Artistas y humanistas lamentan una posible **pérdida de humanismo en la cultura**: música generada por IA en vez de por músicos que vierten su emoción, contenidos escritos por algoritmos en vez de experiencias humanas. Frases como “*se está perdiendo el toque humano*” o “*todo se está volviendo robotizado*” resumen este sentir. Incluso consumidores que valoran la eficiencia admiten extrañar cierto **factor humano** – por ejemplo, que un recomendador de películas te sugiera cosas muy precisas pero “*ya no esté la magia de*

descubrir algo por recomendación de un amigo". En definitiva, este punto de dolor es más intangible: una **nostalgia o defensa de lo humano** frente a un mundo percibido como excesivamente técnico, regido por la lógica de máquinas.

Tipo de preocupación: Social / Emocional.

Nivel de racionalidad percibida: Medio (es un juicio de valor subjetivo; difícil de medir pero entendible, articulado en experiencias comunes).

Supuestos implícitos: Que **lo humano tiene un valor único** en términos de comunicación, creatividad y conexión emocional que la IA **no puede replicar auténticamente**. Se presupone que la presencia de seres humanos en los lazos sociales es esencial para necesidades psicológicas (empatía, sentido de pertenencia) y que sustituirlos por IA **empobrece la calidad de vida**. También se asume que hay un riesgo de **alienación**: que las personas, al tratar más con máquinas que entre sí, se vuelvan más solitarias, menos empáticas unas con otras y quizás **menos "humanas" en su comportamiento**.

9. Desconfianza en la falta de control y regulación – Grupo B.

Descripción: Un sentimiento muy común es que "**no hay nadie al volante**" en cuanto al control de la IA. Es decir, ciudadanos y colectivos sienten que la tecnología avanza más rápido que las normas, y que las autoridades "*no están haciendo lo suficiente*" para proteger al público. Según la encuesta de UCR, **el 75,2% teme que las leyes no regulen estrictamente el funcionamiento de la IA**²⁹. Esta estadística refleja un clamor: la gente quiere **reglas claras y garantías**. Frases en medios y redes como "*la IA es el Lejano Oeste, sin ley*" o "*está fuera de control porque a los gobiernos se les escapa*" son habituales. En Latinoamérica, por ejemplo, se destaca un "*alto nivel de desconfianza*" hacia la IA precisamente por la ausencia de marcos legales locales robustos³⁵

³⁶. También existe la percepción de que **las instituciones van rezagadas**: mientras empresas lanzan productos IA globalmente, los legisladores discuten muy lento y los reguladores quizás ni entienden la tecnología a cabalidad. Esto genera **incertidumbre y miedo**: ¿quién me protege si la IA me afecta negativamente? ¿A quién reclamo si un coche autónomo causa un accidente o si un algoritmo me difama con información errónea? Esa sensación de vacío normativo desemboca en **desconfianza estructural**: la gente no solo desconfía de la IA, sino de las entidades que deberían supervisarla. Por tanto, se exige mayor control democrático, transparencia en qué se hace con la IA, participación pública en decisiones tecnológicas, etc. En ausencia de ello, cunden teorías de complot (que los gobiernos permiten la IA sin regular porque les conviene para control social, por ejemplo). En resumen, el público escéptico siente que **la IA avanza sin freno ni vigilante**, y esto agrava todos los demás temores (porque si algo sale mal, "*no habrá nadie a quien rendir cuentas*").

Tipo de preocupación: Político / Estructural.

Nivel de racionalidad percibida: Alto (reconoce correctamente lagunas en gobernanza; la demanda de regulación está alineada con recomendaciones de expertos).

Supuestos implícitos: Que **los gobiernos tienen la responsabilidad de intervenir** en el desarrollo tecnológico para proteger a la sociedad, y que actualmente están fallando en hacerlo. Se asume que una **ausencia de regulación equivale a caos o abuso**, y que las dinámicas de mercado por sí solas no proveerán las salvaguardas necesarias. También se presupone que la **sociedad civil debe tener voz** en cómo se implementa la IA, y que sin esa voz, prevalecerán los intereses de unos pocos, dejando al ciudadano común desamparado.

10. Riesgos de catástrofe global (visión apocalíptica) – Grupo B.

Descripción: Aunque suene extrema, la idea de una **IA causando una catástrofe global** ha permeado en la conciencia popular, en buena medida por declaraciones de figuras mediáticas. Empresarios como Elon Musk han repetido en entrevistas que la IA avanzada "*podría ser la mayor amenaza para la humanidad*", comparándola con armas nucleares. Estas afirmaciones han sido ampliamente difundidas en prensa y redes, dando lugar a un imaginario apocalíptico donde una

IA desbocada “acaba con el mundo”. Un artículo señalaba que “la idea de que la IA podría acabar con la humanidad parece extrema, pero es una posibilidad que está ganando adeptos”³⁷. Es decir, cada vez más gente común contempla ese escenario como **no descartable**. Este miedo suele ir más allá de la rebelión tipo Terminator (ya mencionada en otro punto) e incluye escenarios de **accidente catastrófico**: por ejemplo, que una IA encargada de sistemas críticos (red eléctrica, mercados financieros, armamento) cometa un error en cadena o sea manipulada con malicia, desatando caos económico o conflictos armados incontrolables. También se habla de “IA fuera de control reproduciéndose a sí misma” (como virus informático global) o de la **competencia IA vs humanos** por recursos. Si bien muchas de estas visiones provienen de la ciencia ficción y de hipótesis de ciertos futuristas, su difusión ha calado: no es raro ver debates en medios generales sobre “¿Podremos sobrevivir a la IA?”. Por tanto, una parte del público vive la rápida evolución de la IA con una ansiedad existencial, preguntándose si **esta invención podría sellar nuestro destino** para mal.

Tipo de preocupación: Existencial / Cultural.

Nivel de racionalidad percibida: Medio (deriva más de narrativas y autoridad percibida de voces famosas que de probabilidades demostradas; es racional preocuparse por lo desconocido, pero la evidencia objetiva es limitada).

Supuestos implícitos: Que **el desarrollo tecnológico conlleva necesariamente riesgos catastróficos si no se controla**, y que la IA es el último y más potente ejemplo de ello. Se asume que es posible un **punto de no retorno** en la complejidad de los sistemas, tras el cual los humanos ya no podrían mitigarlos. Además, subyace la creencia casi filosófica de que **la creación podría volverse contra el creador**: un temor ancestral (Frankenstein, Ícaro, etc.) reactivado por la IA, asumiendo que el hubris humano de crear inteligencia artificial **puede costarnos la propia existencia**.

Fuentes utilizadas: Las citas y datos provienen de **entrevistas y artículos de expertos** en medios reconocidos, **encuestas poblacionales recientes**, ensayos críticos y declaraciones de figuras relevantes en el debate sobre IA. Se han preservado las referencias originales [1] [5] [12] [14] [20] [22] [24] [36], entre otras, que documentan cada argumento en su contexto. Este recopilatorio puede servir como base para análisis posteriores, discusión en foros éticos y diseño de estrategias de comunicación y política en torno a la Inteligencia Artificial y sus impactos. Cada punto de dolor reflejado aquí representa inquietudes **genuinas y extendidas** que cualquier acercamiento responsable a la IA debe atender y comprender en profundidad.

1 2 22 La inteligencia artificial podría acabar con la humanidad, según uno de sus precursores - Meristation

https://as.com/meristation/2023/03/29/betech/1680099321_966779.html

3 27 Impresiones sobre algunas afirmaciones de la “crítica de la técnica”... bueno, de una de las críticas | by German Rosati | factor~data | Medium

<https://medium.com/factor-data/impresiones-sobre-algunas-afirmaciones-de-la-cr%C3%ADtica-de-la-t%C3%A9cnica-bueno-de-una-de-las-cr%C3%ADticas-d0d48b399dd1>

4 5 8 10 13 Los riesgos de la inteligencia artificial en tiempos de oligarquía tecnológica | Tecnología | EL PAÍS

<https://elpais.com/tecnologia/2025-08-20/los-riesgos-de-la-inteligencia-artificial-en-tiempos-de-oligarquia-tecnologica.html>

6 Vista de Brechas digitales y sesgos en la inteligencia artificial

<https://e-revistas.uc3m.es/index.php/UNIV/article/view/9573/7284>

7 Ética de la inteligencia artificial - UNESCO

<https://www.unesco.org/es/artificial-intelligence/recommendation-ethics>

9 11 Cuatro expertos en IA son muy críticos con OpenAI y con cómo se desarrolla la inteligencia artificial: una tecnología en pocas manos

<https://www.genbeta.com/a-fondo/cuatro-expertos-ia-muy-criticos-openai-como-se-desarrolla-inteligencia-artificial-tecnologia-pocas-manos>

12 20 21 25 Policy makers: Please don't fall for the distractions of #AIhype | by Emily M. Bender | Medium

<https://medium.com/@emilymenonbender/policy-makers-please-dont-fall-for-the-distractions-of-aihype-e03fa80ddbf1>

14 15 17 18 19 Los riesgos éticos de las IA que piensan como humanos

<https://www.automatizapro.com.ar/blog/riesgos-eticos-ia-humanizada/>

16 [PDF] Inteligencia artificial en el pensamiento crítico de estudiantes de ...

<http://scielo.sld.cu/pdf/men/v23n3/1815-7696-men-23-03-e4225.pdf>

23 24 La amenaza de la IA a la humanidad es "absurdamente ridícula"

<https://www.xataka.com/robotica-e-ia/amenaza-ia-a-humanidad-absurdamente-ridicula-dice-maximo-responsable-ia-meta>

26 [PDF] Evaluar el pensamiento crítico en la era de la IA

<https://ctes.org.mx/index.php/ctes/article/download/876/1172>

28 30 La pérdida de empleos por la IA es una amenaza grave - Jacobin Revista

<https://jacobinlat.com/2025/07/el-desplazamiento-de-trabajadores-por-la-ia-es-una-amenaza-grave/>

29 33 34 La inteligencia artificial levanta más temores que esperanzas en la población

<https://www.ucr.ac.cr/noticias/2024/1/25/la-inteligencia-artificial-levanta-mas-temores-que-esperanzas-en-la-poblacion.html>

31 32 La opinión pública ante la inteligencia artificial: conocimiento, expectativas y regulación - Funcas

<https://www.funcas.es/articulos/la-opinion-publica-ante-la-inteligencia-artificial-conocimiento-expectativas-y-regulacion/>

35 36 Latinoamérica mira todavía con desconfianza la IA, pese a su "potencial gigante" - Tecnología - ABC Color

<https://www.abc.com.py/tecnologia/2023/10/21/latinoamerica-mira-todavia-con-desconfianza-la-ia-pese-a-su-potencial-gigante/>

37 El grupo de protesta contra la IA que quiere evitar la extinción humana

<https://es.wired.com/articulos/pauseai-grupo-de-protesta-contra-inteligencia-artificial-que-quiere-evitar-extincion-humana>