Conscientious objection in the design of artificial intelligence: when the law questions the algorithm

Angel Jeancarlo Coaquira-Flores Investigador Independiente. Lima, Perú.

The disruptive expansion of Artificial Intelligence (AI) has had a cross-cutting impact on all areas of human activity, offering significant advantages but also substantial risks to human rights. Thus, among the problems arising from its misuse are mass surveillance, social manipulation, and the application of intelligent systems for military purposes. Thus, in response to this problem, a debate has arisen about the relevance of recognizing the right to conscientious objection in the design and development of technologies, particularly in the field of AI. Consequently, this article examines the legal nature of conscientious objection from a doctrinal and normative perspective, thus making it possible to analyze its possible configuration as an ethical defense mechanism against potentially harmful technological projects. Finally, we propose essential elements that could characterize a possible conscientious objection in the technological field, with an emphasis on AI, advocating for specific regulation that reconciles the moral integrity of professionals with the public interest.

Palabras clave: Objeción de conciencia; Inteligencia artificial; Tecnología; Derechos Civiles.

Abstract

The disruptive expansion of Artificial Intelligence (AI) has had a cross-cutting impact on all areas of human activity, offering significant advantages but also substantial risks to human rights. Thus, among the problems arising from its misuse are mass surveillance, social manipulation, and the application of intelligent systems for military purposes. Thus, in response to this problem, a debate has arisen about the relevance of recognizing the right to conscientious objection in the design and development of technologies, particularly in the field of AI. Consequently, this article examines the legal nature of conscientious objection from a doctrinal and normative perspective, thus making it possible to analyze its possible configuration as an ethical defense mechanism against potentially harmful technological projects. Finally, we propose essential elements that could characterize a possible conscientious objection in the technological field, with an emphasis on AI, advocating for specific regulation that reconciles the moral integrity of professionals with the public interest.

Keywords: Conscientious objection; Artificial intelligence; Technology; Civil rights.

Introducción

La tecnología per se carece de una naturaleza intrínsecamente buena o mala; su valoración depende, en última instancia, del uso que las personas decidan darle. A lo largo de la historia, se ha podido apreciar que diversas innovaciones han transformado profundamente la vida humana, dejando huellas imborrables en el desarrollo social, económico y cultural. Sin embargo, en el presente, consideramos que la tecnología con mayor potencial para marcar un antes y un después en la historia de la humanidad es la IA. Gracias a las amplias y prometedoras ventajas que puede ofrecer, los sistemas inteligentes han logrado incidir en prácticamente todos los ámbitos donde el ser humano despliega sus actividades, consolidándose como eje central para la configuración del

futuro y planteando, al mismo tiempo, interrogantes sin precedentes sobre sus implicancias éticas, jurídicas y sociales.

Entre los problemas más preocupantes derivados del uso indebido de la IA, se encuentran la vigilancia masiva, la manipulación social, la aplicación de sistemas inteligentes con fines militares, entre otros. Si bien existen diversos instrumentos orientados a mitigar los efectos nocivos de estas tecnologías, la realidad demuestra que el abuso y las consecuencias perjudiciales continúan materializándose, evidenciando un persistente y preocupante divorcio entre el desarrollo tecnológico y la capacidad del derecho para regularlo eficazmente. En este contexto, cobra especial relevancia una situación singular: el caso de un desarrollador o diseñador de IA que, al advertir que la creación de determinada tecnología puede poner en riesgo derechos humanos y generar dilemas éticos de gran envergadura, se enfrente a la necesidad de materializar una posible objeción de conciencia como mecanismo de defensa ética frente a tales escenarios.

Consecuentemente, resulta imperioso analizar el probable impacto negativo de la IA, a fin de que, a la luz del examen de la figura de la objeción de conciencia, sea posible esbozar la naturaleza y características esenciales que debería revestir una eventual objeción de conciencia en el ámbito tecnológico, con especial énfasis en el desarrollo de la IA.

Aproximación a la naturaleza de la objeción de conciencia

Hervada (1984) indica que la libertad de pensamiento, religión y conciencia representan el núcleo fundamental de la libertad cívica, ya que se encuentran vinculadas al desarrollo espiritual inherente al ser humano. En esa misma línea, el autor, al analizarlas individualmente, expone que: 1) la libertad de pensamiento se refiere al cúmulo de creencias y valoraciones que el individuo formula sobre las diversas realidades del mundo y de la existencia, que orientan su conducta conforme a sus fines naturales; 2) la libertad religiosa tutela principalmente el acto positivo de adhesión a Dios en una relación diagonal; y 3) la libertad de conciencia entendida como la valoración que realiza la razón práctica de una persona sobre si una acción, ya sea próxima a ejecutar, en curso o ya realizada, es moralmente correcta o incorrecta, ampara tanto a la conciencia cierta como a la invenciblemente errónea, protegiendo así la autonomía moral del individuo frente a presiones externas o imposiciones legales.

En esa misma línea, Aparisi y López (2006) mencionan que la objeción de conciencia se configura como una concreción *ad extra* de la libertad de conciencia, lo cual comprende la garantía, por parte del Estado y de los ciudadanos, de que el juicio individual y las acciones que de él se desprendan se desarrollen libremente, sin obstáculos ni restricciones de ningún tipo. Añaden que la libertad de conciencia solo puede ser ejercida por las personas de manera individual y que no se realiza en abstracto.

Conforme a lo señalado, acerca del término “objeción de conciencia” se ha desarrollado un abanico de definiciones, entre las cuales encontramos: Have y Patrão (2021) la conciben como aquel privilegio moral o derecho de una persona a rehusarse a realizar un deber profesional debido a convicciones de índole moral, ética o religiosa; en esa misma línea, Magwentshu et al. (2023) conceptualizan a la objeción de conciencia como aquella facultad que posee una persona para rechazar su participación en actos que perciba como contrarios a sus convicciones morales, religiosas o filosóficas. Asimismo, Kassner y Lefkowitz (2012) señalan que la objeción de conciencia consiste en que un agente se niega a acatar una norma o regla autoritativa que le corresponde, debido a que su cumplimiento vulneraría alguno de sus principios o convicciones más fundamentales. Por su parte, Sánchez (2024), como resultado del análisis de la naturaleza y evolución de la objeción de conciencia, la define como aquella conducta individual de carácter no violento por la que se omite obedecer una norma jurídica específica, motivada por razones de conciencia, las cuales, en ciertos casos, pueden tener un fundamento jurídico.

Así pues, con base a lo expuesto, podría sostenerse que la objeción de conciencia es pasible de

conceptualizarse como aquel derecho subjetivo, de ejercicio personal, que faculta a su titular a abstenerse de realizar un acto impuesto por una norma jurídica o por obligaciones profesionales, cuando su ejecución suponga una grave contradicción con sus convicciones morales, éticas, religiosas o filosóficas esenciales, siempre que dicha negativa se materialice de forma pacífica y, en los casos previstos, cuente con un amparo legal.

Siguiendo esa lógica, al abordar la naturaleza jurídica de la objeción de conciencia, Sánchez (2024), luego de analizar las dos tesis fundamentales -en primer lugar, aquella postura que rechaza reconocer a la objeción de conciencia como un derecho fundamental y, en segundo lugar, la doctrina que defiende el reconocimiento de la objeción como derecho fundamental-, considera como más adecuado al enfoque que sostiene que la objeción forma parte del contenido del derecho fundamental a la libertad de conciencia, tal como lo reconoce el ordenamiento jurídico español en sus dimensiones ideológicas o religiosas.

Bajo ese enfoque, la autora sostiene que si se reconoce a la libertad de conciencia -entendida como el derecho de toda persona a formar libremente sus convicciones y a actuar conforme a ellas- como un derecho fundamental, y de ella se deriva el reconocimiento de la objeción de conciencia, entonces resulta viable establecer un derecho general a la objeción de conciencia. Consecuentemente, la autora señala que de esta afirmación se desprenden las siguientes consecuencias: 1) Cualquier ciudadano puede gozar de este derecho, sin necesidad de la *interpositiolegislatoris*; 2) al formar parte del conjunto de derechos fundamentales, se advierte que la objeción de conciencia se encuentra sujeta a una protección jurídica especial, como el recurso de amparo; y 3) se presupone la capacidad de sujetar al legislador al respeto del contenido esencial del derecho.

Asimismo, Sánchez (2024) desarrolla las notas características de la objeción de conciencia de la siguiente manera: 1) Preexistencia de un deber jurídico, el cual debe ser personal, concreto y legalmente impuesto; 2) Coexistencia de un deber de conciencia, dado que el objetor se enfrenta a un conflicto inevitable y directo entre el cumplimiento del deber establecido por la norma y las convicciones dictadas por la conciencia, de modo que optar por una de estas opciones implica necesariamente rechazar la otra; ante esta disyuntiva, el objetor elige desobedecer la norma para mantenerse fiel a los principios de su conciencia; 3) Comportamiento omisivo, en el sentido de que respecto de los dos presupuestos anteriores se requiere su exteriorización mediante la materialización de una conducta, pero en el caso de la objeción de conciencia, esta se manifiesta a través de un acto de omisión por parte del objetor quien, por razones de conciencia, rehúsa realizar una determinada acción que, en condiciones normales, le sería jurídicamente obligatoria; 4) Necesidad de límitesen su ejercicio. Al igual que los demás derechos fundamentales, el derecho a la objeción de conciencia no sería absoluto, teniendo como límites el orden público y los derechos de terceros; 5) Reconocimiento jurídico, dado que la objeción de conciencia puede ser legal o ilegal, según el sistema jurídico reconozca su ejercicio o no. Su reconocimiento puede darse de forma condicional (cuando la ley o Constitución reconocen los efectos jurídicos a la mera declaración de objeción, sin importar los motivos que la sustentan, basándose únicamente en la convicción personal manifestada externamente) o de forma condicionada (cuando se requiere el desarrollo de un procedimiento legal para verificar la legitimidad y autenticidad de los motivos invocados por el objetor); 6) Ausencia de contenido político, pues a través de la objeción de conciencia no se busca reformar ninguna norma; en ese sentido, la presente figura no constituye una táctica política ni una forma de influir o ejercer presión sobre la opinión pública.

En relación con los principales ámbitos donde se ejerce la objeción de conciencia, se identifican los siguientes:

1. Ámbito sanitario: la objeción de conciencia se originó por la necesidad de resguardar a aquellos profesionales de la salud que, tras los fallos Roe v. Wade y Doe v. Bolton en 1973, se negaban a practicar abortos. Si bien, en un principio la objeción de conciencia abarcaba únicamente el aborto, con el discurrir del tiempo, su ámbito de aplicación se ha ampliado hacia la fertilización *in vitro*, la

intervención destinada a suprimir la capacidad de procreación, las decisiones sobre el final de la vida y la experimentación orientada al análisis de células madre (Jones-Nosacek, 2021). Según señala Card (2020), la objeción de conciencia puede plantearse en casi todas las ramas de la práctica médica. El autor enumera posibles escenarios de su aplicación: un padre que se opone a vacunar a su hijo, un farmacéutico que rechaza suministrar anticonceptivos de emergencia debido a sus convicciones religiosas, una estudiante de medicina que rehúsa intervenir en una práctica de disección anatómica en el laboratorio, ya que, como integrante del pueblo Navajo, su cosmovisión ancestral le impide tener contacto con cuerpos sin vida; entre otros.

A su vez, Symons (2023) menciona que, aunque la objeción de conciencia médica ha recibido críticas, existe una justificación inicial para admitirla en el ámbito sanitario, dado que la conciencia desempeña un papel fundamental en la construcción del sentido y propósito vital de cada individuo, además de funcionar como un mecanismo que permite reaccionar frente a las aberraciones morales en el ejercicio de la medicina.

2. Ámbito militar: en el año 295 d.C., durante una campaña de reclutamiento militar en Numidia (territorio que actualmente corresponde a Argelia), un joven llamado Maximiliano se negó a incorporarse al ejército romano, alegando que su fe cristiana le impedía participar en actos de violencia, debido a que tales acciones contradecían las enseñanzas de Cristo; por esta razón fue condenado a muerte, convirtiéndose en el primer objetor de conciencia registrado (Heval, 2013).

En la actualidad, la oposición a participar en actividades de carácter militar puede manifestarse principalmente de dos formas:

a) La objeción de conciencia absoluta, que implica un rechazo generalizado a cualquier forma de participación en conflictos armados; en estos casos la persona acepta cumplir un servicio civil auténtico en reemplazo del servicio militar, pero se niega a desempeñar cualquier función vinculada con las fuerzas armadas del Estado. Yiannaros (2018) advierte que bajo ningún motivo se tiene que confundir la objeción de conciencia absoluta con la objeción de conciencia total, que se emplea para referirse a casos en los que los objetores rechazan toda forma de servicio obligatorio, tanto el de carácter militar como sus posibles sustitutos civiles.

b) Y, la objeción de conciencia relativa, que comprende la oposición de forma selectiva a ser parte de determinados conflictos armados, pudiendo interpretarse como un alejamiento del principio central de pacifismo, que rechaza toda manifestación de violencia y el uso de la fuerza letal (Yiannaros, 2018).

Por otro lado, Toomey (2020) expone que la criminalización de la objeción de conciencia al servicio militar cuenta con una vasta trayectoria y se ha configurado como una práctica habitual, especialmente en países que se encuentran en conflicto o que prevén la posibilidad de un conflicto futuro. El autor destaca, entre los casos más resaltantes, el de la República de Corea, donde se ha criminalizado y encarcelado a los objetores, justificando este actuar en la necesidad de imponer el servicio militar obligatorio para garantizar la defensa del Estado, basándose en el conflicto armado que tuvo lugar en la península coreana en 1950. Consecuentemente, desde el fin de dicha guerra en 1953, la República de Corea ha impuesto penas de prisión a objetores de conciencia al servicio militar en una proporción alarmante, superando los 19,300 jóvenes -en su mayoría testigos de Jehová- reclutados a comienzos de sus veinte años y condenados a un total acumulado que excede los 36,700 años de reclusión (Toomey, 2020).

Asimismo, conviene mencionar que, con el acelerado desarrollo de los sistemas inteligentes, han surgido contextos en los que ciertos profesionales han manifestado reservas éticas y jurídicas frente a la potencial peligrosidad de estas tecnologías. Estos supuestos podrían, eventualmente, configurar el ejercicio del derecho de objeción de conciencia en el ámbito tecnológico. En este sentido, a continuación, se examinará el proceso de diseño y generación de las inteligencias artificiales, así como los riesgos que actualmente plantean.

IA: Concepto, ciclo de vida (ACV) y ámbito de aplicación

Sheikh et al. (2023) resaltan la dificultad inherente a la tarea de conceptualizar la IA y subrayan que, hasta la fecha, no se cuenta con una definición consensuada del término. Más bien, a partir de la revisión de conceptos clásicos de la IA, tales como:

1. La definición amplia, según la cual IA se asocia con el uso de algoritmos. Esto podría generar complicaciones debido a que extendería el concepto a otras funciones simples como el uso de una calculadora personal o la lectura de instrucciones en un libro de cocina.

2. La definición estricta, que entiende la IA como el intento de las máquinas de reproducir la inteligencia humana; sin embargo, algunos especialistas argumentan que las aplicaciones actuales son todavía relativamente básicas y, en consecuencia, no representarían una manifestación auténtica de IA. La presente postura no es adecuada ya que adoptarla implicaría negar la existencia actual de la IA y subestimar su impacto real.

3. La definición común, que establece la IA como un determinado tipo de tecnología que faculta a las máquinas para reproducir diversas capacidades humanas de alta complejidad. No obstante, esta conceptualización se restringe a reformular el término de IA utilizando diferentes expresiones.

Tomando en cuenta, además, dos consideraciones importantes: primero, evitar restringir el significado de la IA a un área específica de la tecnología; y segundo, comprender que el concepto cambiará con el tiempo, los autores sostienen que la IA debe entenderse como aquellos sistemas capaces de exhibir un comportamiento inteligente, mediante el análisis de su entorno y la toma de decisiones con cierto grado de autonomía, orientado al cumplimiento de objetivos determinados.

En consonancia con ello, International Business Machines Corporation (IBM, 2023) sistematiza los diversos tipos de IA de la siguiente forma:

1. En función de sus capacidades encontramos: a) IA estrecha: también conocida como IA débil, que se configura como un sistema de IA limitado a funciones específicas y carente de sensibilidad (Kavitha & Akila, 2022); b) IA general: la cual, aunque en la actualidad sea únicamente un concepto teórico (IBM, 2023), abarca aquellos sistemas de IA que cuentan con una inteligencia comparable a la humana, con la capacidad de comprender, aprender y aplicar información en múltiples tareas y actividades (Sufyan et al., 2023); c) Super IA: la cual es también conocida como superinteligencia artificial (IBM, 2023) y se refiere a un sistema hipotético de IA diseñado a través de software, con un alcance intelectual que supera la inteligencia humana, cabe añadir que, en su nivel más fundamental, este tipo de IA presenta funciones cognitivas avanzadas y un nivel de razonamiento que supera al de cualquier individuo humano (Mucci & Stryker, 2023).

2. Según sus funcionalidades hallamos: a) IA de máquina reactiva: en el ámbito de la IA, las máquinas reactivas generan una salida idéntica cada vez que reciben una misma entrada (Rashid & Kausik, 2024); conviene aclarar que este tipo de IA no posee memoria y se encuentra diseñada para la realización de una tarea concretamente determinada (IBM, 2023); b) IA de memoria limitada: la cual es una de las más empleadas y, además, se caracteriza porque este tipo recurre a información histórica para prever determinados patrones de variación en los datos (Saghiri et al., 2022); c) IA con teoría de la mente: llegado a este punto, autores como Cuzzolin et al. (2020) sostienen que para que las máquinas inteligentes puedan integrarse de manera fluida con entornos desarrollados por y para humanos, es imperioso incorporar la teoría de la mente en su diseño; siendo así, esta clase podría enmarcarse en la IA general y, aunque aún no se encuentre desarrollada, una IA con capacidad de Teoría de la Mente sería capaz de interpretar los pensamientos y emociones de otras entidades (IBM, 2023); d) IA autoconsciente: la cual es estrictamente teórica y en caso de concretarse, tendría la facultad de percibir sus condiciones internas y atributos, así como entender las emociones y pensamientos de los seres humanos (IBM, 2023), por su parte, Barnes y Hutson (2024) advierten que, desde una perspectiva política, la generación de este tipo de sistemas

demandaría marcos normativos firmes que mitiguen posibles riesgos y aseguren impactos positivos; además, deberían salvaguardar la transparencia, el respeto a los derechos humanos, el desarrollo inclusivo, entre otros.

En concordancia con lo anterior, a pesar de que no existen directrices específicas para el ACV de la IA (Plociennik et al., 2025), De Silva y Alahakoon (2022) descomponen el ciclo en tres fases (diseño, desarrollo e implementación) y diecinueve etapas; en ese sentido, se expondrá dicha propuesta a continuación:

1. En cuanto a las fases, el científico de IA/datos, responsable de la *fase de diseño*, tiene a su cargo la formulación del problema y la conceptualización de la solución. Asimismo, debe identificar los datos representativos y disponibles, culminando esta etapa con una propuesta prescriptiva del problema, una descripción de la solución y los conjuntos de datos que entregará al especialista encargado del desarrollo.

Posteriormente, a lo largo de la *fase de desarrollo*, el científico de IA tiene como objetivo transformar la formulación de un problema en un modelo prototípico de IA. Finalmente, durante la *fase de implementación*, el ingeniero de IA convierte este modelo prototípico de IA en una solución operativa y estandarizada, accesible para todas las partes interesadas y los usuarios finales.

2. En esa misma línea, respecto a las etapas, el ciclo abarcaría:

a) Identificación y formulación del problema: el ciclo de vida se inicia con la identificación, análisis y definición del problema, considerando su naturaleza, el contexto en el que se presenta, los objetivos perseguidos, los actores involucrados, los sistemas implicados, los procesos asociados y los datos disponibles.

b) Revisión de data y ética de la IA: durante esta fase, la formulación del problema, el enfoque seleccionado, la posible solución y los conjuntos de datos requeridos deben ser analizados y validados para identificar eventuales riesgos para la seguridad, en coordinación con la evaluación preliminar de riesgos, así como para asegurar su cumplimiento con los estándares éticos y legales. Cabe indicar que se recomienda que esta revisión ética sea llevada a cabo por especialistas en ética y no por miembros del equipo de IA con competencias técnicas avanzadas.

c) Revisión de la literatura técnica sobre algoritmos de IA, aplicaciones y modelos pre entrenados: la definición del problema establece el marco contextual que hace viable el análisis y examen de investigaciones previas, sistemas existentes, soluciones propuestas y bibliotecas empleadas en escenarios comparables.

d) Preparación de datos: usualmente las fuentes de datos determinadas durante la etapa de formulación del problema se han acumulado orgánicamente, carecen de un formato unificado y se encuentran dispersas entre sí. En ese sentido, los autores recomiendan que, en lugar de usar datos aislados para la generación de modelos de IA, se adopte un enfoque que planifique y construya una infraestructura unificada en datos, como un almacén, lago o centro de datos, que integre de forma centralizada aspectos como el acceso, la propiedad, la administración, la ética de datos, los metadatos, la gobernanza y la regulación, antes de proceder con el desarrollo de modelos de IA.

e) Exploración de datos: la presente fase se enfoca en los datos reales, en contraposición con la etapa anterior, cuyo foco principal es la estructuración de los datos. Cabe mencionar que esta etapa comienza con una comparación con los puntos de referencia de la industria y las líneas base algorítmicas, documentadas en la literatura, donde se ha tratado cuestiones similares. Finalmente, es necesario destacar que la estructura de datos definida en la etapa de preparación se rellena con datos reales mediante técnicas como análisis de correlación, visualización de datos, verificación de relaciones entre puntos de datos y atributos, entre otros.

f) Adquisición de datos externos: en vista de que tanto la etapa de preparación como de exploración de data pueden evidenciar las deficiencias de los datos existentes -lo que imposibilita la gestión de modelos de IA-, es necesario analizar oportunidades para la adquisición externa de datos. Así, la contratación de intermediarios y proveedores de data se configura como una estrategia a corto plazo ampliamente utilizada.

g) Procesamiento de datos: etapa en la cual se garantiza que toda la data obtenida para la generación del modelo o la aplicación de IA se pueda incorporar de manera precisa en su algoritmo, manteniendo al mínimo la afectación de su exactitud, valor informativo y calidad.

h) Construcción del modelode IA inicial: la presente etapa da inicio al desarrollo de un modelo de IA a través de la determinación de un algoritmo idóneo que refleje las capacidades de IA requeridas para la aplicación en cuestión. Asimismo, en la actualidad se puede verificar que todas las aplicaciones prácticas pueden encuadrarse en una o más de las siguientes capacidades: predicción, clasificación, asociación y optimización.

i) Aumento de datos: tomando como referencia los resultados y la evaluación del modelo original, la mejora de datos comprende superar las deficiencias del conjunto de datos que afectan negativamente el desempeño del modelo.

j) Desarrollo de un referente de evaluación:durante el desarrollo del primer modelo de IA, es necesario considerar un referente adecuado para su evaluación. Este criterio de referencia usualmente se fundamenta en una heurística basada en el sentido común, como la experiencia humana en la resolución del mismo problema (por ejemplo, detección de anomalías) o el nivel de acierto de un experto al identificar dichas anomalías con los mismos datos de entrada.

k) Construcción de múltiples modelos de IA: como se desprende de la etapa anterior, mediante la gestión del referente de evaluación se determina lo que falta en el primer modelo y, como resultado, los modelos subsiguientes incrementan progresivamente su nivel de complejidad.

l) Evaluación de métricas primarias: las métricas primarias se establecen en función del parámetro de rendimiento previamente definido. Es esencial comprender tanto el significado de cada métrica como su método de cálculo. Esta información puede hallarse en diversas fuentes, desde la documentación de interfaces de programación hasta estudios de caso. Cabe señalar que una métrica de evaluación eficaz debe ser precisa, robusta, escalable e interpretable.

m) Explicabilidad del modelo de IA: la explicabilidad de modelos usualmente se clasifica en métodos intrínsecos, que son mecanismos inherentes al algoritmo que actúan conjuntamente con la funcionalidad del sistema de IA, y en métodos extrínsecos, que generalmente consisten en técnicas de posprocesamiento aplicadas a los resultados generados por el modelo.

n) Evaluación de métricas secundarias: durante esta etapa, el científico de IA hace entrega de un modelo funcional y en etapa de prototipo al ingeniero de IA/ML. Es así que, más allá de demostrar de forma efectiva alguna de las cuatro habilidades características de la IA al ejecutar determinada tarea, el modelo debe mostrar eficiencia computacional adecuada para su implementación y funcionamiento a mayor escala. Cabe mencionar que en esta fase no se calcula únicamente un conjunto de métricas secundarias como el rendimiento computacional, complejidad temporal o métricas de convergencia, sino que también se desarrollan y evalúan métricas complementarias enfocadas en evaluación de riesgos de privacidad, ciberseguridad, confianza, robustez, explicabilidad, interpretabilidad, usabilidad e implicancias sociales relacionadas.

o) Implementación de modelos de IA y evaluación de riesgos:esta fase, también llamada servicio, puntuación o producción de modelos, corresponde al momento en que el modelo evaluado se pone en funcionamiento para su aplicación práctica. Su alcance es más limitado que el proceso de “operacionalización”, que se detalla más adelante, y su implementación involucra a un número

menor de especialistas y usuarios en lugar de toda la organización.

p) Revisión posterior a la implementación: considerando tanto el sector industrial como el alcance del proyecto, un comité directivo desarrollará una evaluación técnico-ética integral centrada en los conjuntos de datos y el modelo de IA, para analizar las métricas utilizadas y su efectividad. En esta etapa, también se gestionan aspectos relacionados con cumplimiento normativo, estandarización, documentación posterior a la puesta en marcha, contratos y acuerdos de nivel de servicio.

q) Operacionalización mediante pipelines de IA: esta fase consiste en adaptar capacidades de automatización del software DevOps, reconocidas por su gran eficacia, para gestionar la implementación del modelo de IA.

r) Hiperautomatización de procesos y sistemas:actualmente, la hiperautomatización ha cobrado mayor notoriedad en entornos industriales y de trabajo del conocimiento, buscando combinar la eficiencia de los procesos automatizados con el aumento de productividad que ofrece la IA. Aunque la IA tiene como objetivo respaldar, facilitar y fortalecer acciones y decisiones humanas, la hiperautomatización aspira a automatizar también procesos decisionales y operativos, siempre dentro del marco ético y normativo establecido.

s) Monitoreo y evaluación del desempeño: una vez implementado el modelo, esta fase final del ciclo de vida se centra en su monitoreo y evaluación. Entre los principales criterios de evaluación se incluyen la representación de la propia tecnología, la variedad de personas y contextos en que se emplea, así como el valor generado por la misma.

En continuidad con lo planteado, puede observarse que, en la actualidad, debido a los aparentes beneficios y ventajas que se desprenden del despliegue de la IA, cada vez son más los ámbitos que adoptan la posibilidad de automatizar procesos mediante sistemas inteligentes. Entre los casos más representativos se encuentran: 1) el ámbito educativo, donde la IA ha llegado a automatizar diversas actividades administrativas propias de la labor docente y ha demostrado ser capaz de satisfacer los requerimientos específicos de los estudiantes, ofreciendo currículos personalizados y evaluaciones adaptativas (Ifenthaler et al., 2024); 2) el ámbito financiero, en el cual se ha implementado pronóstico financiero basado en IA y minería de texto, gestión de riesgos, calificación crediticia mediante aprendizaje automático y la detección de fraudes financieros mediante IA y minería de datos, entre otros (Vuković et al., 2025); 3) el ámbito sanitario, donde los sistemas inteligentes han generado un impacto positivo en la asistencia sanitaria, destacando su contribución al perfeccionamiento de diagnósticos, personalización del tratamientos, análisis predictivo, automatización de procesos y desarrollo de la robótica, entre otros (Olawade et al., 2024).

En suma, el vertiginoso desarrollo de la IA ha generado escenarios que, hasta hace poco, solo podían concebirse dentro de la ciencia ficción. No obstante, es importante destacar que los efectos, ya sean beneficiosos o perjudiciales, dependen en gran medida tanto de quienes utilizan los sistemas como, en muchos casos, de quienes los desarrollan. En ese sentido, resulta imprescindible analizar los posibles abusos que podrían derivarse del uso de la IA.

Peligros derivados del diseño, uso y abuso de la IA

La acelerada injerencia de la IA en las esferas en las que el ser humano se desarrolla ha traído consigo no solo oportunidades de innovación sino, también, riesgos de alto impacto jurídico, ético y político. Estos peligros que no son meramente hipotéticos, han llegado a ser reconocidos y regulados a través de un bagaje de marcos regulatorios emergentes. A continuación, se detallan algunos de los principales:

1. Uso militar de la IA: mientras que la tecnología evoluciona con rapidez, su impacto en la seguridad nacional y la esfera bélica ha generado tanto expectativas como serias preocupaciones.

Así, el Pentágono emplea sistemas inteligentes como una estrategia compensatoria frente al aumento de la competencia geopolítica (Hoffman & D'Amelio, 2024). Aunque la implementación de IA en contextos bélicos todavía se encuentra en sus etapas iniciales, su acelerado desarrollo ofrece oportunidades significativas para los ejércitos a nivel global. Sin embargo, las crecientes capacidades de los sistemas militares con IA han generado inquietud en determinados activistas, quienes han solicitado limitaciones o incluso la prohibición de ciertas armas autónomas (Osimen et al., 2024). Cabe mencionar que, durante el 2023, los desarrolladores ucranianos declararon que sus drones ya ejecutaban ataques de forma autónoma contra fuerzas rusas, sin intervención humana directa (Hambling, 2023).

Por su parte, Basuchoudhary (2025) indica que, como producto de la IA en el ámbito militar, es posible desarrollar un software de asesoramiento de combate con IA que agilice la toma de decisiones del combatiente mientras refuerza la seguridad de las tropas; predecir el desarrollo de la cadena de suministro para las fuerzas de tarea multidominio; y optimizar la preparación operativa al acelerar la rotación y asociación del personal con los equipos necesarios, entre otros.

En contraparte, Erskine y Miller (2024) sostienen que, a pesar de que estos sistemas basados en IA ofrecen un abanico de ventajas, como la capacidad de procesar enormes volúmenes de información, identificar patrones imperceptibles para los responsables humanos, emitir recomendaciones y predicciones, además de reaccionar ante amenazas con una rapidez y eficacia extraordinarias, aún quedan desafíos por atender. Estos autores relatan que, como resultado de un taller enfocado en el análisis del impacto de sistemas inteligentes en el futuro de la guerra, se identificaron cuatro complicaciones: a) el desplazamiento del juicio humano por decisiones automatizadas mediante IA sobre el recurso a la fuerza, lo cual plantea posibles consecuencias para la teoría de la disuasión y el riesgo de una escalada involuntaria del conflicto; b) las eventuales consecuencias derivadas del sesgo inducido por la automatización; c) la opacidad algorítmica -entendida como la imposibilidad de que un sistema de IA justifique, de manera clara y accesible, la razón detrás de una decisión concreta ante personas con escasos conocimientos técnicos (Glikson & Woolley como se citó en Guo et al., 2025)- y sus consecuencias para la legitimidad democrática e internacional de un Estado, dado que esta requiere una justificación clara y comprensible al momento de decidir recurrir a una guerra; la ausencia de transparencia genera serias inquietudes cuando son las máquinas las que informan, calculan o incluso ejecutan de manera autónoma tales decisiones; y d) la posibilidad de que los sistemas inteligentes agudicen los problemas en los procesos decisionales dentro de las organizaciones.

2. Vigilancia masiva y menoscabo de libertades civiles: uno de los usos más críticos de la IA es el fortalecimiento de los sistemas de vigilancia y monitoreo. Así, de acuerdo con el índice de vigilancia global (GSI), al menos 75 de los 176 países del mundo se encuentran invirtiendo activamente e IA y utilizándola con fines de control y supervisión, especialmente en ámbitos como las ciudades inteligentes, el reconocimiento facial y la vigilancia policial automatizada (Saheb, 2022).

Ahora bien, autores como Fontes et al. (2022) aseguran que tecnologías como el reconocimiento facial en tiempo real en espacios públicos o las tecnologías de vigilancia de sanitaria -en particular, las herramientas digitales destinadas al rastreo de contactos-facilitan la labor de las autoridades públicas en la aplicación de la ley, al permitir el rastreo de desplazamientos individuales y la obtención de data para monitorear y anticipar patrones de conducta social. Por consiguiente, son vistas como instrumentos con potencial para dar respuesta a crisis sociales, como las derivadas de la delincuencia o de emergencias sanitarias, incluyendo las pandemias.

Sin embargo, estudiosas del tema como Barkane y Buka (2025) señalan que, a través de numerosas pesquisas, se han destacado los impactos nocivos de este tipo de sistemas respecto a los derechos fundamentales, en especial la dignidad de la persona, el respeto a la vida privada y a los datos personales, la igualdad de trato, así como las libertades de expresión y de reunión pacífica. Es así que las autoras advierten que los sistemas de identificación biométrica remota representan una grave vulneración del derecho a la privacidad, al comprometer la autodeterminación individual, el

conocimiento de la propia identidad y la integridad psicológica. Asimismo, destacan que este tipo de tecnologías puede conducir a actos discriminatorios y afectar principios como la igualdad y la justicia, especialmente cuando se basa en datos sesgados o propensos a errores. Finalmente, consideran la posibilidad de que estas herramientas tecnológicas sean utilizadas por las autoridades públicas como instrumentos para implementar estrategias de vigilancia masiva, aun vulnerando valores democráticos.

3. Manipulación social y la automatización de la desinformación: la cada vez mayor difusión de sistemas sociotécnicos en línea enfocados en la influencia racional de la conducta mediante IA -tales como la publicidad de microsegmentación, las redes sociales y los algoritmos de búsqueda personalizados- ha propiciado modos innovadores de relacionarse con los usuarios, recopilar sus datos y, en ciertos casos, condicionar su conducta. No obstante, estos métodos y herramientas también han suscitado preocupaciones respecto a su capacidad para manipular, dado que permiten influir y dirigirse a las personas de forma masiva, con un nivel de sutileza, automatización y alcance sin precedentes. (Ienca, 2023)

En sintonía con lo expuesto, Williamson y Prybutok (2024) resaltan que las posibles consecuencias de la manipulación a través de IA son variadas y sumamente alarmantes. Así, dentro de los riesgos más significativos encontrarían el menoscabo de la autonomía personal, la alteración de conductas y normas sociales, así como el aumento de las desigualdades sociales. Asimismo, los autores mencionan que el perjuicio no siempre es evidente, puesto que los “empujoncitos” –sutiles intervenciones de diseño que incentivan a los usuarios a optar por alternativas ventajosas– diseñados para promover decisiones beneficiosas, podrían considerarse formas de manipulación, aun cuando persigan fines positivos, tales como hábitos de vida más saludables o el equilibrio económico.

Para finalizar, tal y como se ha expuesto, si bien en torno a la IA se reconocen un sinnúmero de ventajas, resulta cada vez más evidente el campo minado sobre el que camina la humanidad al incorporar estos sistemas en prácticamente todas las esferas de la actividad humana. No obstante, se observan indicios preliminares de que algunos desarrolladores de IA han planteado objeciones éticas o morales frente a determinados avances tecnológicos.

Objeción de conciencia e IA: ¿U n freno al desarrollo tecnológico?

Aunque a primera vista podría pensarse que la oposición de los desarrolladores de IA frente al avance de un determinado proyecto, ante la posible vulneración de bienes jurídicos, solo puede contemplarse en escenarios hipotéticos, la realidad demuestra todo lo contrario, como se observa en casos como los siguientes:

1. El proyecto Maven: se trata de una iniciativa en la que Google autorizó al Pentágono a utilizar determinadas tecnologías con fines militares. El proyecto tenía como objetivo inicial apoyar el procesamiento eficiente de las grabaciones de video recopiladas por los drones empleados por la fuerza militar estadounidense, mediante la detección de vehículos y otros objetos, el seguimiento de sus desplazamientos y la entrega del análisis resultante al Departamento de Defensa (BBC Mundo, 2018a).

En respuesta, los trabajadores, al considerar que la empresa, con su actuar, no solo podría dañar su marca de forma irreparable, sino que también comprometía la confianza de los usuarios y desatendía su deber moral y ético, firmaron una carta abierta dirigida al director ejecutivo, Sundar Pichai, solicitando el retiro de la empresa del denominado proyecto Maven (BBC Mundo, 2018a). Cabe indicar que, si bien solo 3,100 empleados de un total de 70,000 trabajadores firmaron la carta, este acto refleja la inquietud que tienen respecto a las decisiones adoptadas por la empresa, recordando su histórico lema *“Don’t be evil”*, es decir, no seas malvado (Fernández, 2018).

2. El caso de Timnit Gebru: en 2018, Gebru, autora de un artículo publicado en el que demostró una

mayor tasa de error en los sistemas de reconocimiento facial aplicados a mujeres de piel más oscura, recibió un mensaje de un alto directivo de Google mediante el cual se le solicitaba retractarse o retirar su nombre del artículo. La investigación sostenía que las compañías tecnológicas podrían adoptar medidas más eficaces para garantizar que los sistemas de IA orientados a replicar el lenguaje escrito y oral humano no reproduzcan ni profundicen sesgos de género históricos ni fomenten expresiones ofensivas. (Wong, 2020)

A pesar de que Gebru trató de negociar con Google, ofreciendo retirar su nombre a cambio de una explicación detallada de las objeciones de la empresa y un diálogo para mejorar el manejo de estos temas en el futuro -y acordando retirarse posteriormente, en caso de que la empresa se negara-; la respuesta del gigante tecnológico fue no solo el rechazo, sino también el envío de un correo a su equipo informando que se había aceptado su renuncia. Como consecuencia de la abrupta salida de Gebru, más de 1,200 trabajadores de Google y más de 1,500 personas vinculadas al ámbito académico y a la sociedad civil suscribieron una carta en señal de protesta. (Wong, 2020)

3. El casode Amazon Rekognition: en el pasado, Amazon ha demostrado su disposición a colaborar con el Estado para materializar determinados fines. Así, se pudo verificar que Amazon Web Services fue empleada por Palantir, empresa de datos vinculada al ICE, en tareas de detención y reportación. Posteriormente, con el advenimiento y gestión de Amazon Rekognition -sistema que proporciona capacidades avanzadas de análisis y reconocimiento facial en imágenes y videos proporcionados por los usuarios-, capaz de identificar, examinar y comparar rostros, siendo útil en diversas áreas como la verificación de identidad, la contabilización de personas y la seguridad pública, surgieron críticas provenientes de diversos organismos que alertaban sobre el potencial peligro de esta tecnología. Ante esta situación, los empleados de Amazon, enviaron una carta a Jeff Bezos, solicitando la interrupción de las relaciones colaborativas entre la empresa y autoridades gubernamentales o policiales, en vista de que el sistema podría emplearse con fines contrarios a los principios de la compañía (BBC Mundo, 2018b).

4. El proyecto Dragonfly: Jack Poulson, quien formó parte del equipo de investigación e IA de Google, centro esfuerzos en optimizar la precisión de los sistemas de búsqueda (Gallagher, 2018). A inicios de agosto de 2018, Poulson manifestó su preocupación a los directivos de Google tras la publicación de un informe de The Intercept, que revelaba que la compañía se encontraba desarrollando de manera confidencial una aplicación de búsqueda para dispositivos Android destinada al mercado chino. Este sistema, conocido internamente como Dragonfly, fue diseñado para censurar contenidos que el régimen chino considera sensibles, como información sobre disidentes políticos, libertad de expresión, democracia, derechos humanos y manifestaciones pacíficas (Gallagher, 2018). Como consecuencia, tras una conversación con sus superiores y ante la imposibilidad de seguir trabajando en Google, presentó su renuncia (Gallagher, 2018).

Durante su entrevista de salida, se le comunicó que la empresa podía perdonar sus posturas políticas y enfocarse en sus contribuciones técnicas, siempre y cuando no cometiera una falta grave, como hablar con la prensa (Poulson, 2019). Bajo esta perspectiva, quedan claros dos aspectos: primero, que las grandes empresas tecnológicas continúan creando tecnologías de monitoreo empleadas para vulnerar derechos humanos; y segundo, que los empleados que se niegan a cumplir determinadas órdenes relacionados con el desarrollo de este tipo de tecnologías se encuentran en una situación de indefensión, por lo que deberían contar con mecanismos de protección legal.

Es importante reconocer que, más allá de los contextos en los que la ilegalidad es evidente -como la vigilancia masiva o el uso bélico de la IA-, existe un terreno ambiguo en el que la objeción de conciencia se enfrenta a la innovación, a la presión empresarial y a la competitividad global. En tales casos, la coexistencia de objeción e innovación plantea un desafío más complejo, pues la cuestión no consiste únicamente en el cierre evidente de derechos, sino en la gestión de la tensión entre la conciencia individual y los objetivos empresariales vinculados al avance económico y tecnológico.

Cuando esto ocurre, la situación se torna aún más compleja, ya que en el ámbito de los negocios las empresas velan por la eficiencia (Handoyo et al., 2023), la rapidez (Kolev et al., 2024) y la innovación constante (Garrido-Moreno et al., 2024) por encima de otras consideraciones. Entre estas características, la innovación es la que con mayor frecuencia despierta tensiones de orden ético (McCausland, 2023). En consecuencia, en contextos de alta competitividad global, donde las empresas quieren estar a la vanguardia tecnológica, el objetor será visto como un obstáculo al avance.

Por ello, resulta imperante considerar posibles estructuras normativas y organizativas que permitan que el derecho a la objeción de conciencia se ejerza legítimamente sin convertirse en un pretexto para el marginamiento ocupacional y sin constituir una barrera absoluta para el avance tecnológico. De este modo, más que una incompatibilidad insalvable, la relación entre objeción de conciencia e innovación debe entenderse como una compatibilidad condicionada, viable en la medida en que existan marcos claros que aseguren tanto el respeto a las convicciones individuales como el desarrollo empresarial.

Como respuesta a los potenciales peligros que pueden generarse como producto de un desarrollo inconsciente de nuevas tecnologías, se ha gestado un cúmulo de instrumentos que podemos agrupar de la siguiente manera:

1. Marcos corporativos, los cuales se caracterizan por no poseer fuerza legal y por ser elaborados por las propias empresas. Como ejemplos encontramos: a) los principios de la IA de Google, entre los cuales se halla el principio de desarrollo y despliegue responsables, mediante el cual se estipula que, dado que la IA, en tanto tecnología transformadora aún en desarrollo, implica complejidades y riesgos en constante cambio, se procura gestionarla de manera responsable a lo largo de su ciclo de vida, incorporando aprendizajes a medida que la IA avanza y sus aplicaciones evolucionan (Google, s.f.); o b) los principios de Microsoft para una IA responsable, que comprenden los principios de justicia, fiabilidad y seguridad, privacidad y seguridad, inclusión, transparencia y responsabilidad (Microsoft, s.f.).

2. Normativavinculante, que posee obligatoriedad jurídica y carácter coercitivo, cuyo principal exponente es el Reglamento Europeo de IA.

3. Marcos internacionales, conjunto de directrices de alcance transnacional, flexibles y surgidos del consenso entre organismos internacionales. Entre los principales ejemplos se encuentran:

a) Los principios de la IA de la Organisation for Economic Co-operation and Development (OCDE), que constituyen el primer referente intergubernamental en la materia que abarcan cinco principios basados en valores: crecimiento inclusivo, desarrollo sostenible y bienestar; derechos humanos y valores democráticos, incluida la equidad y la privacidad; transparencia y explicabilidad; robustez, seguridad y protección; y responsabilidad (OECD, s.f.).

b) La Recomendación sobre la ética de la inteligencia artificial de la (UNESCO), que aborda la ética de la IA como una reflexión normativa y sistemática, sustentada en un marco integral, global, multicultural y dinámico de valores, principios y acciones independientes; siendo así, este marco busca orientar a las sociedades en la gestión responsable de los impactos, tanto previsibles como imprevistos, que las tecnologías de IA pueden generar sobre las personas, las comunidades, el medio ambiente y los ecosistemas, proporcionando además una base para decidir su aceptación o rechazo (UNESCO, 2022).

c) En tercer lugar, el Llamamiento de Roma sobre la Ética de la IA, documento a través del cual se busca fomentar una responsabilidad colectiva entre organismos internacionales, estados, instituciones y empresas privadas, con miras a construir un futuro en el que la innovación digital y el avance tecnológico mantengan a la humanidad en el centro; cabe añadir que la iniciativa se estructura en torno a tres áreas de impacto (ética, educación y derechos) y seis principios rectores

(transparencia, inclusión, responsabilidad, imparcialidad, confiabilidad y seguridad y privacidad) (RenAIssance Foundation, 2024).

A pesar de los esfuerzos existentes, el posible derecho emergente de todo desarrollador de IA a abstenerse de ejecutar un deber profesional fundamentado en convicciones morales, éticas o religiosas se encontraría en un estado crítico de indefensión. Esta situación puede verificarse en uno de los casos más famosos de objeción de conciencia en el ámbito tecnológico: el caso de Jack Poulson. En ese sentido, a continuación, se analizará la posible viabilidad de reconocer este derecho en el campo tecnológico, especialmente en la IA, a la luz de la normativa peruana.

Para ello, tomando como base el estudio sobre objeción de conciencia al final de la vida, elaborado por la doctora Sánchez (2020), se puede indicar que, a nivel jurídico, la objeción de conciencia se tutela a través de:

1. El artículo 6 de la Convención Americana sobre Derechos Humanos, establece de manera expresa que, en el ámbito del servicio militar obligatorio, es procedente invocar la excepción por motivos de conciencia. Asimismo, su artículo 12 denominado “Libertad de Conciencia y de Religión”, señala que ninguna persona puede ser sometida a medidas restrictivas que limiten su libertad de mantener o cambiar su religión o sus creencias. Del mismo modo, estipula que la libertad de expresar la propia religión o creencias solo podrá ser restringida por las limitaciones establecidas por la ley que resulten necesarias para salvaguardar la seguridad, el orden, la salud o la moral pública, así como los derechos y libertades de las demás personas.

2. En el Perú, a nivel constitucional, la objeción de conciencia se tutela a través del artículo 2 inciso 3, el cual precisa que toda persona tiene derecho al ejercicio de la libertad de conciencia y de religión, ya sea de manera individual o en asociación con otros. De igual manera, el artículo 14 se garantiza que la enseñanza religiosa se imparta con pleno respeto a la libertad de conciencia.

3. De la misma forma, el artículo 4 de la ley N° 29635, Ley de libertad religiosa, conceptualiza la libertad de conciencia como la negativa de una persona a cumplir un determinado deber legal debido a sus convicciones morales y religiosas. Además, señala que la objeción de conciencia se ejerce cuando una persona se encuentra impedida de cumplir una obligación legal por un imperativo moral o religioso, grave e ineludible, que cuenta con el reconocimiento de la entidad religiosa a la que pertenece. En este punto, concordamos con la autora, quien advierte una imprecisión en la redacción, dado que la objeción de conciencia puede sustentarse no solo en motivos morales y religiosos, sino también en fundamentos de carácter filosófico, axiológico o relacionados con la justicia. Por otro lado, destaca que, pese a tratarse de una norma sobre libertad religiosa, esta no solo se limita a convicciones de naturaleza estrictamente religiosa, sino que también abarca consideraciones morales. En ese sentido, la autora sostiene que la intención del legislador es otorgar una protección amplia a la libertad de conciencia, sin restringirla exclusivamente al ámbito religioso.

Asimismo, Sánchez (2020) sostiene que, con la promulgación del Reglamento de la Ley de Libertad Religiosa, se subsanaron algunas de estas imprecisiones, puesto que, en lo relativo a la objeción de conciencia regulada por dicha ley, esta se circunscribe a convicciones de carácter religioso, fundamentadas en la doctrina de la fe profesada y debidamente reconocidas por la autoridad de la entidad religiosa correspondiente, siempre que no vulneren los derechos fundamentales, la moral ni las buenas costumbres. Por otra parte, la autora advierte que, si bien resulta pertinente incorporar las convicciones morales dentro de una norma que regula un derecho específico, es necesario precisarlas con mayor rigor. Ello se debe a que, a diferencia de la objeción de conciencia religiosa -que exige el reconocimiento del imperativo por parte de la entidad de fe correspondiente-, en el caso de la convicción moral existe incertidumbre respecto de qué organismo podría verificar dicho imperativo no religioso. En consecuencia, se requiere un análisis más detallado sobre su forma de acreditación, puesto que la Ley de Libertad Religiosa no lo prevé, lo que podría desnaturalizar el sentido del artículo.

Del análisis normativo, se concluye que en el Perú el derecho a la objeción de conciencia se encuentra en una fase incipiente, estrechamente vinculada al factor religioso y cuya existencia no se concibe al margen de este; sin embargo, ello no supone su desconocimiento.

Ahora bien, atendiendo al desarrollo doctrinal y normativo del derecho a la objeción de conciencia, así como a las particularidades del ecosistema tecnológico contemporáneo, es posible esbozar determinados elementos esenciales que deberían caracterizar el ejercicio de una probable objeción de conciencia en el ámbito del diseño y desarrollo de tecnologías, especialmente de IA:

1. Fundamento ético-jurídico claro: la objeción de conciencia debe sustentarse en convicciones éticas, morales o religiosas firmes y debidamente sustentadas, que generen un conflicto grave e insuperable con la actividad requerida. Su reconocimiento encontraría respaldo en el marco del derecho internacional de los derechos humanos, particularmente en la libertad de pensamiento, de conciencia y de religión, así como en la jurisprudencia comparada que ha delimitado sus alcances en diversos contextos profesionales.

2. Delimitación del ámbito de aplicación: el ejercicio de la objeción de conciencia debería restringirse a aquellas actividades del ciclo de vida de la IA, ya sea en su diseño, entrenamiento, implementación o mantenimiento, que impliquen una intervención directa e inevitable en procesos que el objetor perciba como contrarios a sus convicciones.

3. Proceso estructurado y verificable: la objeción de conciencia debería presentarse por escrito, indicando motivos concretos y su vinculación con los principios de conciencia. Su evaluación correspondería a un órgano imparcial, como un comité ético- tecnológico, que garantice un análisis justo y evite decisiones arbitrarias. Es menester advertir que el proceso debe resguardar la confidencialidad y promover un diálogo transparente entre la entidad responsable del proyecto y el objetor de conciencia.

Cabe agregar que, si del análisis del proyecto se verifica la existencia de un impacto negativo en los derechos humanos, el comité tendría la obligación de denunciar dicho acto.

4. Respeto al interés público y a los derechos de terceros: en caso de que el ejercicio de la objeción de conciencia pudiera afectar los derechos de terceros, la negativa a participar deba complementarse con mecanismos de reemplazo o redistribución de funciones que aseguren la continuidad del proyecto cuando este responda al interés público o atienda necesidades esenciales como la seguridad o la salud.

5. Protección contra represalias: se debe garantizar que el ejercicio legítimo de la objeción de conciencia no derive en despidos, perjuicios en la trayectoria profesional ni situaciones de discriminación.

6. Carácter excepcional y no abusivo: la objeción de conciencia debe considerarse una medida excepcional, reservada para casos en las que existan conflictos de conciencia de significativa gravedad, evitando que se emplee como una excusa para eludir obligaciones comunes o entorpecer el normal desarrollo y progreso de la labor tecnológica.

En suma, la eventual incorporación de la objeción de conciencia en el ámbito tecnológico exige un equilibrio entre la protección de la integridad moral del profesional y la preservación del interés general, lo que exige una regulación particular que articule principios éticos, jurídicos y técnicos.

Conclusiones

El rápido avance de los sistemas inteligentes, junto con los beneficios derivados de su despliegue, ha conferido a la inteligencia artificial una pseudo-omnipresencia, ya que no existe ámbito alguno

en el que no tenga un cierto nivel de injerencia. No obstante, aunque en un principio se consideró que este tipo de tecnologías podrían emplearse únicamente con fines benignos, la realidad demuestra lo contrario. Las tecnologías cognitivas han demostrado tener un gran potencial para ser utilizadas con fines nocivos para la humanidad, incluyendo la manipulación social, la vigilancia masiva mal empleada, los sistemas de armas letales autónomos, entre otros.

De este modo, se ha podido apreciar que a lo largo de todo el ciclo de vida de la IA se manifiestan determinados filtros de ética que buscan un desarrollo tecnológico conforme a directrices propias del *softlaw* y marcos jurídicos; con el paso del tiempo, se ha evidenciado una tensión creciente entre la innovación tecnológica y las convicciones de quienes participan en su diseño. Esta tensión da lugar a un escenario en el que la objeción de conciencia podría configurarse como un mecanismo necesario para salvaguardar la integridad profesional y la coherencia ética de los desarrolladores.

En este sentido, los programadores y diseñadores de IA enfrentan problemas muy concretos: la presión empresarial orientada a la eficiencia y la competitividad global, así como la posibilidad de que sus creaciones sean empleadas en contextos que comprometen derechos fundamentales. Frente a tales dilemas, la objeción de conciencia surge como una respuesta legítima, aunque su efectividad dependerá de la existencia de un marco jurídico claro que la regule y proteja frente a posibles represalias.

En consecuencia, consideramos que la eventual implementación de la objeción de conciencia en el campo del diseño y la creación de tecnologías, particularmente en lo relativo a la IA, podría comprender elementos tales como: un fundamento ético-jurídico claro, una delimitación del ámbito de aplicación clara, un proceso estructurado y verificable, con carácter excepcional y no abusivo, que además proteja al objetor frente a represalias.

Cabe destacar que, de materializarse la figura de la objeción de conciencia en el diseño y desarrollo de la inteligencia artificial, su alcance no debería limitarse únicamente a este ámbito, sino extenderse a cualquier área tecnológica. Tal proyección se fundamenta en el crecimiento acelerado, y en ocasiones desmedido, de las tecnologías, así como en la continua aparición de innovaciones disruptivas, como la computación cuántica, entre otras, que requieren respuestas éticas y jurídicas con la misma urgencia y profundidad.

Referencias

Aparisi, A. & López, J. (2006). El derecho a la objeción de conciencia en el supuesto del aborto: de la fundamentación filosófica-jurídica a su reconocimiento legal. *Persona y Bioética, 10*(1), 35-51.

Barkane, I. & Buka, L. (2025). Prohibited AI surveillance practices in the Artificial Intelligence Act: promises and pitfalls in protecting fundamental rights. En V. Galis, H. Gundhus & A. Vradis. (Eds.). *Critical Perspectives on Predictive Policing: Anticipating Proof?* (pp. 110–129). Edward Elgar Publishing. https://doi.org/10.4337/9781035323036.00011

Barnes, E. & Hutson, J. (2024). AI and the Cognitive Sense of Self. *Journal of Intelligent Communication, 3*(1), 35-51. https://doi.org/10.54963/jic.v3i1.320

Basuchoudhary, A. (2025). AI and Warfare: A Rational Choice Approach. *Eastern EconomicJournal, 51*, 74–86. https://doi.org/10.1057/s41302-024-00280-7

BBC Mundo (05 de abril de 2018a). Qué es el polémico Proyecto Maven que Google desarrolla con el Pentágono y por el que sus empleados están en pie de Guerra. *British Broadcasting Corporation*.https://www.bbc.com/mundo/noticias-43662384

BBC Mundo (25 de junio de 2018b). La dura carta de los empleados de Amazon en la que piden a Jeff Bezos que deje de vender tecnología al gobierno de Donald Trump. *British Broadcasting Corporation*. https://www.bbc.com/mundo/noticias-44601388

Card, R. (2020). *A New Theory of Conscientious Objection in Medicine: Justification and Reasonability*. Routledge.

Cuzzolin, F., Morelli, A., Cîrstea, B. & Sahakian, B. (2020). Knowing me, knowing you: theory of mind in AI. *Psychological Medicine, 50*(7). https://doi.org/10.1017/S0033291720000835

De Silva, D. & Alahakoon, D. (2022). An artificial intelligence life cycle: From conception to production. *Patterns, 3*(6). https://doi.org/10.1016/j.patter.2022.100489

Erskine, T. & Miller, S. (2024). AI and the decision to go to war: future risks and opportunities. Australian *Journal of International Affairs, 78*(2), 135-147. https://doi.org/10.1080/10357718.2024.2349598

Fernández, Y. (09 de abril de 2018). *Qué es Project Maven y por qué 3.100 empleados de Google le piden a la empresa que lo abandone*. Xataca. https://www.xataka.com/empresas-y-economia/que-es-project-maven-y-por-que-3-100-empleados-de-google-le-piden-a-la-empresa-que-lo-abandone

Fontes, C., Hohma, E., Corrigan, C. & Lütge, C. (2022). AI-powered public surveillance systems: why we (might) need them and how we want them. *Technology in Society, 71*. https://doi.org/10.1016/j.techsoc.2022.102137

Gallagher, R. (13 de septiembre de 2018). *Senior Google Scientist Resigns Over "Forfeiture of Our Values" in China*. The Intercept. https://theintercept.com/2018/09/13/google-china-search-engine-employee-resigns/

Garrido-Moreno, A., Martín-Rojas, R. & García-Morales, V. (2024). The key role of innovation and organizational resilience in improving business performance: A mixed-methods approach. *International Journal of Information Management, 77*. https://doi.org/10.1016/j.ijinfomgt.2024.102777

Google (s.f.). *AI Principles: Our AI Principles*. https://ai.google/principles/

Guo, C., Liu, H., Song, F. & Guo, J. (2025). The double-edged sword effects of algorithmic opacity: The self-determination theory perspective. *Acta Psychologica, 260*. https://doi.org/10.1016/j.actpsy.2025.105600

Hambling, D. (17 de octubre de 2023). *Ukraine's AI Drones Seek And Attack Russian Forces Without Human Oversight*. Forbes. https://www.forbes.com/sites/davidhambling/2023/10/17/ukraines-ai-drones-seek-and-attack-russian-forces-without-human-oversight/?sh=69da7b6b66da

Handoyo, S., Suharman, H., Ghani, E. & Soedarsono, S. (2023). A business strategy, operational efficiency, ownership structure, and manufacturing performance: The moderating role of market uncertainty and competition intensity and its implication on open innovation. *Journal of Open Innovation: Technology, Market, and Complexity, 9*(2). https://doi.org/10.1016/j.joitmc.2023.100039

Have, H. & Patrão, M. (2021). Conscientious Objection. En H. Have & M. Patrão. (Eds.). *Dictionary of Global Bioethics* (pp. 347–348). Springer. https://doi.org/10.1007/978-3-030-54161-3_169

Hervada, J. (1984). Libertad de conciencia y error sobre la moralidad de una terapéutica. *Persona y Derecho, 11*, 13-53. https://doi.org/10.15581/011.32653

Heval, Ö. (2013). *Conscientious Objection to Military Service in International Human Rights Law*. Palgrave Macmillan.

Hoffman, F. & D'Amelio, A. (2024). AI's impact on war's enduring nature. *Orbis, 68*(1), 72-91. https://doi.org/10.1016/j.orbis.2023.11.006

Ienca, M. (2023). On Artificial Intelligence and Manipulation. *Topoi, 42*, 833–842. https://doi.org/10.1007/s11245-023-09940-3

Ifenthaler, D., Majumdar, R., Gorissen, P., Judge, M., Mishra, S., Raffaghelli, J. & Shimada, A. (2024). Artificial Intelligence in Education: Implications for Policymakers, Researchers, and Practitioners. *Technology, Knowledge and Learning, 29*, 1693–1710. https://doi.org/10.1007/s10758-024-09747-0

International Business Machines Corporation (2023). *Understanding the different types of artificial intelligence*. https://www.ibm.com/think/topics/artificial-intelligence-types

Jones-Nosacek, C. (2021). Conscientious Objection, Not Refusal: The Power of a Word. *The Linacre Quarterly, 88*(3), 242–246. https://doi.org/10.1177/00243639211008271

Kassner, J. & Lefkowitz, D. (2012). Conscientious Objection. En R. Chadwick. (Ed.). *Encyclopedia of Applied Ethics* (2da ed.). Academic Press. https://doi.org/10.1016/B978-0-12-373932-2.00373-2

Kavitha, V. & Akila, C. (2022). A deep learning approach toward prediction of mental health of Indians. En S. Jain, P. Jain, K. Pandey & K. Phooi. (Eds.). *Artificial Intelligence, Machine Learning, and Mental Health in Pandemics: A Computational Approach* (pp. 141-165). Academic Press. https://doi.org/10.1016/B978-0-323-91196-2.00009-0

Kolev, K., Dykes, B., Kang, S., Hughes-Morgan, M. & Ferrier, W. (2024). Fast or slow? A meta-analysis on the performance implications of decision, implementation, and response speed. *Journal of Management Scientific Reports, 3*(1). https://doi.org/10.1177/27550311241291976

Magwentshu, M., Chingwende, R., Jim, A., Rooyen, J., Hajiyiannis, H., Naidoo, N., Orr, N., Menzel, J. & Pearson, E. (2023). Definitions, perspectives, and reasons for conscientious objection among healthcare workers, facility managers, and staff in South Africa: a qualitative study. *Sexual and Reproductive Health Matters, 31*(1). https://doi.org/10.1080/26410397.2023.2184291

McCausland, T. (2023). Ethics in Innovation. *Research-Technology Management, 66*(6). https://doi.org/10.1080/08956308.2023.2257107

Microsoft (s.f.). *Microsoft Responsible AI: Principles and approach*. https://www.microsoft.com/en-us/ai/principles-and-approach

Mucci, T. & Stryker, C. (18 de diciembre de 2023). *What is artificial superintelligence? IBM*. https://www.ibm.com/think/topics/artificial-superintelligence

Olawade, D., David-Olawade, A., Wada, O., Asaolu, A., Adereni, T. & Ling, J. (2024). Artificial intelligence in healthcare delivery: Prospects and pitfalls. *Journal of Medicine, Surgery, and Public Health, 3*. https://doi.org/10.1016/j.glmedi.2024.100108

Organisation for Economic Co-operation and Development (s.f.). *AI principles*. https://www.oecd.org/en/topics/ai-principles.html

(2022). *Recomendación sobre la ética de la inteligencia artificial*. UNESCO.

Osimen, G., Newo, O. & Fulani, O. (2024). Artificial intelligence and arms control in modern warfare. *Cogent Social Sciences, 10*(1). https://doi.org/10.1080/23311886.2024.2407514

Plociennik, C., Watjanatepin, P., Van Acker, K. & Ruskowski, M. (2025). Life Cycle Assessment of Artificial Intelligence Applications: Research Gaps and Opportunities. *Procedia CIRP, 135*. https://doi.org/10.1016/j.procir.2025.01.079

Poulson, J. (23 de abril de 2019). *I used to work for Google. I am a Conscientious Objector*. The New York Times. https://www.nytimes.com/2019/04/23/opinion/google-privacy-china.html

Rashid, A. & Kausik, A. (2024). AI revolutionizing industries worldwide: A comprehensive overview of its diverse applications. *Hybrid Advances, 7*. https://doi.org/10.1016/j.hybadv.2024.100277

RenAIssance Foundation (2024). *The Call*. https://www.romecall.org/the-call/

Saghiri, A., Vahidipour, M., Jabbarpour, M., Sookhak, M. & Forestiero, A. (2022). A Survey of Artificial Intelligence Challenges: Analyzing the Definitions, Relationships, and Evolutions. *Applied Sciences, 12*(8). https://doi.org/10.3390/app12084054

Saheb, T. (2022). Ethically contentious aspects of artificial intelligence surveillance: a social science perspective. *AI Ethics, 3*(2). https://doi.org/10.1007/s43681-022-00196-y

Sánchez, R. (2020). La objeción de conciencia frente a la eutanasia: un análisis biojurídico. *Apuntes de Bioética, 3*(1), 89-97. https://doi.org/10.35383/apuntes.v3i1.398

Sánchez, R. (2024). *Objeción de conciencia, dignidad humana y vida humana: Presupuestos y análisis a la luz de las sentencias del TEDH y de la CIDH*. Editorial Tirant Lo Blanch.

Sheikh, H., Prins, C. & Schrijvers, E. (2023). Artificial Intelligence: Definition and Background. En H. Sheikh, C. Prins & E. Schrijvers. (Eds.). *Mission AI: The New System Technology* (pp. 15-41). Springer. https://doi.org/10.1007/978-3-031-21448-6_2

Sufyan, M., Shokat, Z. & Ashfaq, U. (2023). Artificial intelligence in cancer diagnosis and therapy: Current status and future perspective. *Computers in Biology and Medicine, 165*. https://doi.org/10.1016/j.compbiomed.2023.107356

Symons, X. (2023). Why Conscience Matters: A Theory of Conscience and Its Relevance to Conscientious Objection in Medicine. *Res Publica, 29*, 1-21. https://doi.org/10.1007/s11158-022-09555-2

Toomey, L. (2020). The Right to Conscientious Objection to Military Service: Recent Jurisprudence of the United Nations Working Group on Arbitrary Detention. *Human Rights Law Review, 19*(4), 787–810. https://doi.org/10.1093/hrlr/ngz031

Vuković, D., Dekpo-Adza, S. & Matović, S. (2025). AI integration in financial services: a systematic review of trends and regulatory challenges. *Humanities and social sciences communications, 12*(562). https://doi.org/10.1057/s41599-025-04850-8

Williamson, S. & Prybutok, V. (2024). The Era of Artificial Intelligence Deception: Unraveling the Complexities of False Realities and Emerging Threats of Misinformation. *Information, 15*(6). https://doi.org/10.3390/info15060299

Wong, J. (04 de diciembre de 2020). *More than 1,200 Google workers condemn firing of AI scientist TimnitGebru*. The Guardian. https://www.theguardian.com/technology/2020/dec/04/timnit-gebru-google-ai-fired-diversity-ethics

Yiannaros, A. (2018). Refusing to Kill: Selective Conscientious Objection and Professional Military Duties. *Journal of Military Ethics, 17*(2), 108-121. https://doi.org/10.1080/15027570.2018.1548995

References

1. Aparisi, A. & López, J. (2006). El derecho a la objeción de conciencia en el supuesto del aborto: de la fundamentación filosófica-jurídica a su reconocimiento legal. Persona y Bioética, 10(1), 35-51.
2. Barkane, I. & Buka, L. (2025). Prohibited AI surveillance practices in the Artificial Intelligence Act: promises and pitfalls in protecting fundamental rights. En V. Galis, H. Gundhus & A. Vradis. (Eds.). Critical Perspectives on Predictive Policing: Anticipating Proof? (pp. 110–129). Edward Elgar Publishing. https://doi.org/10.4337/9781035323036.00011
3. Barnes, E. & Hutson, J. (2024). AI and the Cognitive Sense of Self. Journal of Intelligent Communication, 3(1), 35-51. https://doi.org/10.54963/jic.v3i1.320
4. Basuchoudhary, A. (2025). AI and Warfare: A Rational Choice Approach. Eastern Economic Journal, 51, 74–86. https://doi.org/10.1057/s41302-024-00280-7
5. BBC Mundo (05 de abril de 2018a). Qué es el polémico Proyecto Maven que Google desarrolla con el Pentágono y por el que sus empleados están en pie de Guerra. British Broadcasting Corporation. https://www.bbc.com/mundo/noticias-43662384
6. BBC Mundo (25 de junio de 2018b). La dura carta de los empleados de Amazon en la que piden a Jeff Bezos que deje de vender tecnología al gobierno de Donald Trump. British Broadcasting Corporation. https://www.bbc.com/mundo/noticias-44601388
7. Card, R. (2020). A New Theory of Conscientious Objection in Medicine: Justification and Reasonability. Routledge.
8. Cuzzolin, F., Morelli, A., Cîrstea, B. & Sahakian, B. (2020). Knowing me, knowing you: theory of mind in AI. Psychological Medicine, 50(7). https://doi.org/10.1017/S0033291720000835
9. De Silva, D. & Alahakoon, D. (2022). An artificial intelligence life cycle: From conception to production. Patterns, 3(6). https://doi.org/10.1016/j.patter.2022.100489
10. Erskine, T. & Miller, S. (2024). AI and the decision to go to war: future risks and opportunities. Australian Journal of International Affairs, 78(2), 135-147. https://doi.org/10.1080/10357718.2024.2349598
11. Fernández, Y. (09 de abril de 2018). Qué es Project Maven y por qué 3.100 empleados de Google le piden a la empresa que lo abandone. Xataca. https://www.xataka.com/empresas-y-economia/que-es-project-maven-y-por-que-3-100-empleados-de-google-le-piden-a-la-empresa-que-lo-abandone
12. Fontes, C., Hohma, E., Corrigan, C. & Lütge, C. (2022). AI-powered public surveillance systems: why we (might) need them and how we want them. Technology in Society, 71. https://doi.org/10.1016/j.techsoc.2022.102137
13. Gallagher, R. (13 de septiembre de 2018). Senior Google Scientist Resigns Over “Forfeiture of Our Values” in China. The Intercept. https://theintercept.com/2018/09/13/google-china-search-engine-employee-resigns/
14. Garrido-Moreno, A., Martín-Rojas, R. & García-Morales, V. (2024). The key role of innovation and organizational resilience in improving business performance: A mixed-methods approach. International Journal of Information Management, 77. https://doi.org/10.1016/j.ijinfomgt.2024.102777
15. Google (s.f.). AI Principles: Our AI Principles. https://ai.google/principles/
16. Guo, C., Liu, H., Song, F. & Guo, J. (2025). The double-edged sword effects of algorithmic opacity: The self-determination theory perspective. Acta Psychologica, 260. https://doi.org/10.1016/j.actpsy.2025.105600
17. Hambling, D. (17 de octubre de 2023). Ukraine’s AI Drones Seek And Attack Russian Forces Without Human Oversight. Forbes. https://www.forbes.com/sites/davidhambling/2023/10/17/ukraines-ai-drones-seek-and-attack-russian-forces-without-human-

oversight/?sh=69da7b6b66da
18. Handoyo, S., Suharman, H., Ghani, E. & Soedarsono, S. (2023). A business strategy, operational efficiency, ownership structure, and manufacturing performance: The moderating role of market uncertainty and competition intensity and its implication on open innovation. Journal of Open Innovation: Technology, Market, and Complexity, 9(2). https://doi.org/10.1016/j.joitmc.2023.100039
19. Have, H. & Patrão, M. (2021). Conscientious Objection. En H. Have & M. Patrão. (Eds.). Dictionary of Global Bioethics (pp. 347–348). Springer. https://doi.org/10.1007/978-3-030-54161-3_169
20. Hervada, J. (1984). Libertad de conciencia y error sobre la moralidad de una terapéutica. Persona y Derecho, 11, 13-53. https://doi.org/10.15581/011.32653
21. Heval, Ö. (2013). Conscientious Objection to Military Service in International Human Rights Law. Palgrave Macmillan.
22. Hoffman, F. & D'Amelio, A. (2024). AI's impact on war's enduring nature. Orbis, 68(1), 72-91. https://doi.org/10.1016/j.orbis.2023.11.006
23. Ienca, M. (2023). On Artificial Intelligence and Manipulation. Topoi, 42, 833–842. https://doi.org/10.1007/s11245-023-09940-3
24. Ifenthaler, D., Majumdar, R., Gorissen, P., Judge, M., Mishra, S., Raffaghelli, J. & Shimada, A. (2024). Artificial Intelligence in Education: Implications for Policymakers, Researchers, and Practitioners. Technology, Knowledge and Learning, 29, 1693–1710. https://doi.org/10.1007/s10758-024-09747-0
25. International Business Machines Corporation (2023). Understanding the different types of artificial intelligence. https://www.ibm.com/think/topics/artificial-intelligence-types
26. Jones-Nosacek, C. (2021). Conscientious Objection, Not Refusal: The Power of a Word. The Linacre Quarterly, 88(3), 242–246. https://doi.org/10.1177/00243639211008271
27. Kassner, J. & Lefkowitz, D. (2012). Conscientious Objection. En R. Chadwick. (Ed.). Encyclopedia of Applied Ethics (2da ed.). Academic Press. https://doi.org/10.1016/B978-0-12-373932-2.00373-2
28. Kavitha, V. & Akila, C. (2022). A deep learning approach toward prediction of mental health of Indians. En S. Jain, P. Jain, K. Pandey & K. Phooi. (Eds.). Artificial Intelligence, Machine Learning, and Mental Health in Pandemics: A Computational Approach (pp. 141-165). Academic Press. https://doi.org/10.1016/B978-0-323-91196-2.00009-0
29. Kolev, K., Dykes, B., Kang, S., Hughes-Morgan, M. & Ferrier, W. (2024). Fast or slow? A meta-analysis on the performance implications of decision, implementation, and response speed. Journal of Management Scientific Reports, 3(1). https://doi.org/10.1177/27550311241291976
30. Magwentshu, M., Chingwende, R., Jim, A., Rooyen, J., Hajiyiannis, H., Naidoo, N., Orr, N., Menzel, J. & Pearson, E. (2023). Definitions, perspectives, and reasons for conscientious objection among healthcare workers, facility managers, and staff in South Africa: a qualitative study. Sexual and Reproductive Health Matters, 31(1). https://doi.org/10.1080/26410397.2023.2184291
31. McCausland, T. (2023). Ethics in Innovation. Research-Technology Management, 66(6). https://doi.org/10.1080/08956308.2023.2257107
32. Microsoft (s.f.). Microsoft Responsible AI: Principles and approach. https://www.microsoft.com/en-us/ai/principles-and-approach
33. Mucci, T. & Stryker, C. (18 de diciembre de 2023). What is artificial superintelligence? IBM. https://www.ibm.com/think/topics/artificial-superintelligence
34. Olawade, D., David-Olawade, A., Wada, O., Asaolu, A., Adereni, T. & Ling, J. (2024). Artificial intelligence in healthcare delivery: Prospects and pitfalls. Journal of Medicine, Surgery, and Public Health, 3. https://doi.org/10.1016/j.glmedi.2024.100108
35. Organisation for Economic Co-operation and Development (s.f.). AI principles. https://www.oecd.org/en/topics/ai-principles.html
36. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (2022). Recomendación sobre la ética de la inteligencia artificial. UNESCO.
37. Osimen, G., Newo, O. & Fulani, O. (2024). Artificial intelligence and arms control in modern

warfare. Cogent Social Sciences, 10(1). https://doi.org/10.1080/23311886.2024.2407514
38. Plociennik, C., Watjanatepin, P., Van Acker, K. & Ruskowski, M. (2025). Life Cycle Assessment of Artificial Intelligence Applications: Research Gaps and Opportunities. Procedia CIRP, 135. https://doi.org/10.1016/j.procir.2025.01.079
39. Poulson, J. (23 de abril de 2019). I used to work for Google. I am a Conscientious Objector. The New York Times. https://www.nytimes.com/2019/04/23/opinion/google-privacy-china.html
40. Rashid, A. & Kausik, A. (2024). AI revolutionizing industries worldwide: A comprehensive overview of its diverse applications. Hybrid Advances, 7. https://doi.org/10.1016/j.hybadv.2024.100277
41. RenAIssance Foundation (2024). The Call. https://www.romecall.org/the-call/
42. Saghiri, A., Vahidipour, M., Jabbarpour, M., Sookhak, M. & Forestiero, A. (2022). A Survey of Artificial Intelligence Challenges: Analyzing the Definitions, Relationships, and Evolutions. Applied Sciences, 12(8). https://doi.org/10.3390/app12084054
43. Saheb, T. (2022). Ethically contentious aspects of artificial intelligence surveillance: a social science perspective. AI Ethics, 3(2). https://doi.org/10.1007/s43681-022-00196-y
44. Sánchez, R. (2020). La objeción de conciencia frente a la eutanasia: un análisis biojurídico. Apuntes de Bioética, 3(1), 89-97. https://doi.org/10.35383/apuntes.v3i1.398
45. Sánchez, R. (2024). Objeción de conciencia, dignidad humana y vida humana: Presupuestos y análisis a la luz de las sentencias del TEDH y de la CIDH. Editorial Tirant Lo Blanch.
46. Sheikh, H., Prins, C. & Schrijvers, E. (2023). Artificial Intelligence: Definition and Background. En H. Sheikh, C. Prins & E. Schrijvers. (Eds.). Mission AI: The New System Technology (pp. 15-41). Springer. https://doi.org/10.1007/978-3-031-21448-6_2
47. Sufyan, M., Shokat, Z. & Ashfaq, U. (2023). Artificial intelligence in cancer diagnosis and therapy: Current status and future perspective. Computers in Biology and Medicine, 165. https://doi.org/10.1016/j.compbiomed.2023.107356
48. Symons, X. (2023). Why Conscience Matters: A Theory of Conscience and Its Relevance to Conscientious Objection in Medicine. Res Publica, 29, 1-21. https://doi.org/10.1007/s11158-022-09555-2
49. Toomey, L. (2020). The Right to Conscientious Objection to Military Service: Recent Jurisprudence of the United Nations Working Group on Arbitrary Detention. Human Rights Law Review, 19(4), 787–810. https://doi.org/10.1093/hrlr/ngz031
50. Vuković, D., Dekpo-Adza, S. & Matović, S. (2025). AI integration in financial services: a systematic review of trends and regulatory challenges. Humanities and social sciences communications, 12(562). https://doi.org/10.1057/s41599-025-04850-8
51. Williamson, S. & Prybutok, V. (2024). The Era of Artificial Intelligence Deception: Unraveling the Complexities of False Realities and Emerging Threats of Misinformation. Information, 15(6). https://doi.org/10.3390/info15060299
52. Wong, J. (04 de diciembre de 2020). More than 1,200 Google workers condemn firing of AI scientist Timnit Gebru. The Guardian. https://www.theguardian.com/technology/2020/dec/04/timnit-gebru-google-ai-fired-diversity-ethics
53. Yiannaros, A. (2018). Refusing to Kill: Selective Conscientious Objection and Professional Military Duties. Journal of Military Ethics, 17(2), 108-121. https://doi.org/10.1080/15027570.2018.1548995