De marionetas a dioses

Libertad y autonomía de la inteligencia artificial



Cómo citar

APA
Guecha López, C. A. (2026). *De marionetas a dioses: libertad y autonomía de la inteligencia artificial*. Editorial Neogranadina.

Chicago
Guecha López, C. A. *De marionetas a dioses: libertad y autonomía de la inteligencia artificial.* Editorial Neogranadina, 2026.

MLA
Guecha López, C. A. (2026). *De marionetas a dioses: libertad y autonomía de la inteligencia artificial.* Editorial Neogranadina. 2026. Impreso.

DE MARIONETAS A DIOSES

Libertad y autonomía de la inteligencia artificial

Carlos Alfonso Guecha López

Universidad Militar Nueva Granada

Editorial Neogranadina

Guecha López, Carlos Alfonso, autor

De marionetas a dioses: libertad y autonomía de la inteligencia artificial / Carlos Alfonso Guecha López. -- Bogotá : Editorial Neogranadina, 2026.
184 páginas -- (Colección Studia)

Referencias: páginas 149 – 171

ISBN impreso: 978-628-7758-97-1 -- ISBN digital: 978-628-7758-98-8 --ISBN epub: 978-628-7758-99-5

1. Inteligencia artificial -- Aspectos éticos 2. Inteligencia artificial -- Aspectos jurídicos 3. Libertad 4. Autonomía I. Título II. Colección

174.9 SCDD23 CO-BoUMNG

CC BY NC ND

De marionetas a dioses:
libertad y autonomía de la inteligencia artificial
Colección Studia
Primera edición: julio, 2026
© Carlos Alfonso Guecha López
© Universidad Militar Nueva Granada
© Facultad de Derecho
© Vicerrectoría de Investigaciones
© Editorial Neogranadina
Bogotá, Colombia, 2026
editorial.neogranadina@unimilitar.edu.co
https://editorial.umng.edu.co/
ISBN (impreso): 978-628-7758-97-1
ISBN (PDF): 978-628-7758-98-8
ISBN (EPUB): 978-628-7758-99-5
DOI: https://doi.org/10.18359/9786287758971

Corrección de estilo: Yesid Castiblanco Barreto
Diseño de colección: Daniela Martínez
Montaje y diagramación: Mónica Cabiativa Daza
Imagen de portada: Una imagen en 3D de la cara y el cuello de una mujer. En https://unsplash.com
Impresión: DGP Editores S. A. S.
Conversión a epub: Hipertexto - Netizen

Todos los derechos reservados

Queda estrictamente prohibida la reproducción total o parcial de este libro, por cualquier medio o procedimiento —incluidos la reprografía, el escaneo, la digitalización o el tratamiento informático— sin la autorización previa y por escrito de los titulares del derecho de autor.

Esta edición, así como sus elementos gráficos y de diseño, es propiedad de la Universidad Militar Nueva Granada.

Contenido

1 **Prefacio**

5 **Introducción**

CAPÍTULO 1

9 **La libertad**

10 Genealogía de la libertad

15 Tipologías de la libertad

22 Autonomía personal

28 Autodeterminación

33 Condiciones y límites de la libertad

CAPÍTULO 2

47 **Robots, inteligencias artificiales y algoritmos**

48 Robots

51 Inteligencia artificial

54 Algoritmos y toma de decisiones

57 Objetivos predefinidos y funciones de utilidad

60 Niveles de autonomía en sistemas artificiales

CAPÍTULO 3

67 **El despertar de las máquinas**

68 Condiciones necesarias de la libertad en agentes artificiales

74 Volición y agencia computacional

79 Indeterminismo y aleatoriedad

83 Responsabilidad y atribución de acciones
89 Los test y las métricas de libertad artificial
94 Casos de estudio empíricos
103 Críticas y objeciones

CAPÍTULO 4
111 **Implicaciones jurídicas de la libertad artificial**
112 Régimen de responsabilidad civil y penal
113 Responsabilidad objetiva, solidaria y por riesgo tecnológico
114 ***Safe harbor*** y seguros obligatorios
117 Consideraciones de responsabilidad penal
119 Derechos fundamentales y conflictos de libertades
120 Colisión entre libertad humana y libertad artificial
122 Derecho a la explicación y transparencia algorítmica
126 Impacto en la libertad humana
126 Desplazamiento decisorio y pérdida de agencia cívica
128 Dependencia tecnológica y asimetrías de poder
132 Ética, moral y responsabilidad colectiva
132 El problema del responsable ausente
134 Deuda moral intergeneracional y principio de precaución
137 Escenarios futuros y políticas públicas
138 Corregulación humano-máquina
141 Salvaguardias para preservar la libertad humana
143 Marcos de transición e inclusión social ante la automatización
149 **Referencias**

173 **Sobre el autor**

DOI:https://doi.org/10.18359/9786287758971

De marionetas a dioses:
libertad y autonomía de la inteligencia artificial

Carlos Alfonso Guecha López*

Resumen: El libro se organiza alrededor de dos preguntas centrales: si las máquinas pueden ser libres y hasta qué punto esa posible libertad afectaría la libertad humana. Desde ahí, analiza un mundo marcado por la automatización que ya influye en decisiones sociales, económicas y políticas. La obra se enfoca en la IA de propósito general y en agentes robóticos basados en grandes modelos de lenguaje y aprendizaje por refuerzo. Su enfoque es doble: por un lado, propone criterios para identificar, atribuir y proteger una eventual libertad artificial; por el otro, examina sus límites y riesgos. Para hacerlo, reconstruye distintas ideas de libertad y las usa para estudiar la autonomía algorítmica y su relación con los derechos fundamentales. Entre sus principales aportes destacan salvaguardas como la transparencia, la explicación, la auditabilidad y la intervención humana. En conjunto, propone un marco para regular una IA autónoma de forma compatible con la democracia y la libertad humana.

Palabras clave: libertad artificial; autonomía; responsabilidad algorítmica; gobernanza de IA; neuroderechos

* Doctor en Derecho, *magna cum laude*, de la Universidad de Paris II Panthéon-Assas en Francia. Profesor de carrera e investigador del Centro de Investigaciones Jurídicas de la Universidad Militar Nueva Granada.

Orcid: https://orcid.org/0009-0006-2864-9284

CvLAC: https://scienti.minciencias.gov.co/cvlac/visualizador/generarCurriculoCv.do?cod_rh=0000103102

CC BY NC ND

DOI: https://doi.org/10.18359/9786287758971

From Puppets to Gods:
Freedom and Autonomy in Artificial Intelligence

Carlos Alfonso Guecha López

Abstract: The book is organized around two central questions: whether machines can be free, and to what extent that potential freedom would affect human freedom. From there, it analyzes a world shaped by automation that is already influencing social, economic, and political decisions. The work focuses on general-purpose AI and robotic agents based on large language models and reinforcement learning. Its approach is twofold: on the one hand, it proposes criteria for identifying, attributing, and protecting potential artificial freedom; on the other, it examines its limits and risks. To this end, it reconstructs various conceptions of freedom and uses them to study algorithmic autonomy and its relationship to fundamental rights. Among its key contributions are safeguards such as transparency, explainability, auditability, and human intervention. Taken together, it proposes a framework for regulating autonomous AI in a way that is compatible with democracy and human freedom.

Keywords: Artificial freedom; autonomy; algorithmic responsibility; AI governance; neurorights.

CC BY NC ND

DOI: https://doi.org/10.18359/9786287758971

De marionetes a deuses:
liberdade e autonomia da inteligência artificial

Carlos Alfonso Guecha López

Resumo: Este livro se organiza em torno de duas perguntas centrais: se as máquinas podem ser livres e até que ponto essa possível liberdade afetaria a liberdade humana. A partir daí, analisa um mundo marcado pela automatização, que já influencia decisões sociais, econômicas e políticas. A obra concentra-se na IA de propósito geral e em agentes robóticos baseados em grandes modelos de linguagem e aprendizagem por reforço. Sua abordagem é dupla: por um lado, propõe-se critérios para identificar, atribuir e proteger uma eventual liberdade artificial; por outro, seus limites e riscos são examinados. Para isso, reconstrói diferentes ideias de liberdade e as utiliza para estudar a autonomia algorítmica e sua relação com os direitos fundamentais. Entre suas principais contribuições, destacam-se salvaguardas como a transparência, a explicação, a auditabilidade e a intervenção humana. Em conjunto, propõe um marco para regular uma IA autônoma de forma compatível com a democracia e a liberdade humana.

Palavras-chave: liberdade artificial; autonomia; responsabilidade algorítmica; governança da IA; neurodireitos

CC BY NC ND

DOI: https://doi.org/10.18359/9786287758971

Des marionnettes aux dieux:
liberté et autonomie de l'intelligence artificielle

Carlos Alfonso Guecha López

Résumé: Cet ouvrage s'articule autour de deux questions centrales : les machines peuvent-elles être libres ? Et, si tel est le cas, dans quelle mesure cette liberté potentielle affecterait-elle la liberté humaine ? À partir de ces interrogations, l'analyse porte sur un monde de plus en plus marqué par l'automatisation, où les technologies influencent déjà les décisions sociales, économiques et politiques. Dans ce contexte, l'ouvrage met l'accent sur l'intelligence artificielle à usage général ainsi qu'aux agents robotiques basés sur de grands modèles linguistiques et l'apprentissage par renforcement. L'approche adoptée est double ; d'une part, on propose des critères permettant d'identifier, d'attribuer et, le cas échéant, de protéger une éventuelle forme de liberté artificielle. D'autre part, on en examine les limites et les risques. À cette fin, différentes conceptions de la liberté sont mobilisées afin d'analyser l'autonomie algorithmique et ses implications pour les droits fondamentaux. Parmi les principales contributions de l'ouvrage figure la formulation de garanties telles que la transparence, l'explicabilité, l'auditabilité et la supervision humaine. Dans l'ensemble, il propose un cadre normatif destiné à encadrer le développement et le déploiement d'une ia autonome, dans des conditions compatibles avec les principes démocratiques et la préservation de la liberté humaine.

Mots-clés: liberté artificielle ; autonomie ; responsabilité algorithmique ; gouvernance de l'intelligence artificielle ; neurodroits.

Prefacio

«De marionetas a dioses» es un diagnóstico incómodo sobre la redistribución del poder en nuestra época. Porque la escena no es, como suele narrarse, la de máquinas que se vuelven inteligentes y la de humanos que pierden empleos. La escena real, más íntima y más peligrosa, es la de voluntades que se desplazan: lo que antes era deliberación humana, responsabilidad y buen juicio se convierte en optimización, trazabilidad y puntuación. Y en ese tránsito, lo decisivo no es que aparezcan «nuevos agentes», sino que nosotros aprendemos a aceptar, como normal, que la vida social sea gobernada por criterios que nadie eligió conscientemente, pero que todos obedecemos.

La palabra libertad se ha gastado tanto que, a veces, ya no impacta. Se pronuncia con la facilidad con la que se pulsa un botón, se exige, se proclama, se promete. Pero la libertad realmente es otra cosa, es fricción, costo, riesgo. La libertad auténtica no es un estado de comodidad, sino una estructura de posibilidad; es la capacidad de interrumpir un curso de acción incluso cuando todo empuja a seguirlo. Y por eso, paradójicamente, la libertad siempre ha sido una noción sospechosa para los sistemas que defienden la previsibilidad. ¿Qué hace, entonces, una civilización entera enamorada de la predicción, del cálculo, del *ranking*, preguntándose si una máquina puede ser libre? Quizá lo que hacemos no es formular una pregunta sobre ellas; tal vez estamos confesando una duda sobre nosotros.

En el imaginario popular, la autonomía de la inteligencia artificial (IA) se asocia con una especie de emancipación romántica, la máquina que despierta, el robot que elige, el sistema que

decide por sí mismo. Pero la realidad técnica suele ser menos poética y más brutal, pues un sistema no quiere, sino que optimiza; no prefiere, predice; no busca sentido, sino que ordena patrones siguiendo modelos de probabilidad. Lo más perturbador es que esa brutalidad no es un defecto, sino su virtud. El problema comienza cuando confundimos esa virtud con un atributo moral. Cuando el lenguaje de la voluntad se injerta sobre la arquitectura del objetivo, aparece una ilusión peligrosa, ya que creemos estar frente a un sujeto, cuando estamos realmente frente a un mecanismo de obediencia extrema.

Sin embargo, tampoco conviene abrazar el consuelo fácil de llamarlas herramientas. Porque una herramienta tradicional no aprende de su experiencia, no reescribe su modo de operar, no se vuelve progresivamente opaca para su creador. En cambio, la autonomía artificial contemporánea se ha convertido en un continuo, una escalera con peldaños, grados, niveles; incluso existen escalas estandarizadas para describirla en dominios críticos como la conducción.

Esa forma de hablar (*nivel 0*, *nivel 1*, *nivel 2*…) es tentadora porque tranquiliza al dar la apariencia de que el control puede dosificarse como una perilla. Pero el control no es un interruptor, sino una relación política. Y conviene decirlo sin delicadeza: la discusión sobre la libertad de las máquinas es también una discusión acerca del tipo de sociedad que tolera delegar decisiones sin poder exigir explicaciones con sentido. ¿Dónde se rompe la ilusión? A menudo, en el lugar más prosaico, en la función objetivo que se le atribuye. Una IA puede estar diseñada para maximizar recompensas, minimizar errores o penalizar costos; lo que quiere es, estrictamente, lo que su función de utilidad premia.

Eso plantea el problema más serio de todos y no es que la máquina tenga libertad, sino que su libertad sea la de ejecutar literalmente un objetivo mal planteado. La historia humana está llena de advertencias sobre deseos formulados con torpeza. El mito solo cambia de narrativa. El rey Midas hoy se llama métrica, y su oro hoy se llama *engagement*. Cuando una medida se convierte en objetivo, deja de ser una medición y se vuelve un parámetro para dominar.

La civilización ya experimentó este fenómeno antes de hablar de libertad artificial. Lo vivimos con las redes sociales cuando nuestras plazas públicas se transformaron en mercados de atención

y nuestras conversaciones en combustible de recomendadores. Lo inquietante es que, en ese proceso, los humanos no fuimos derrotados por una voluntad ajena, sino reconfigurados por una gramática de incentivos.

Aquí aparece el punto de quiebre de esta investigación: la libertad no es solo ausencia de interferencia, es también la capacidad de no quedar subordinado a un poder que puede intervenir arbitrariamente. Y la forma contemporánea de arbitrariedad ya no siempre está acompañada de uniforme militar ni de leyes dictatoriales, viene en forma de modelo entrenado, de marcador asignado, de umbral decidido. En otras palabras, la dominación puede adoptar el rostro amable de una recomendación personalizada.

Pero sería ingenuo pensar que el problema se agota en la manipulación social. El propio desarrollo técnico empuja hacia sistemas con capacidades de automodificación y aprendizaje continuo, agentes que se ajustan, se reconfiguran, se reentrenan, incluso exploran formas de mejorarse sin intervención humana directa. En ese punto, la metáfora de la herramienta se vuelve insuficiente y la metáfora del sujeto se vuelve prematura. Estamos en una zona gris que tiene un nombre político: vacío de responsabilidad. Porque cuanto más autónomo es un sistema, más fácil se vuelve que cada actor humano diga «no fui yo». El programador culpará al dato; el operador culpará al modelo; la empresa culpará al usuario; el usuario culpará a la interfaz. La libertad, entonces, empieza a operar como coartada, una libertad sin sujeto responsable, sin imputación, sin culpa. Y eso no es libertad: es impunidad distribuida.

Por eso el corazón de la discusión no es metafísico, sino institucional. No se trata únicamente de si existe volición computacional, ni de si la aleatoriedad equivale a indeterminismo moral, ni de si podemos diseñar los test para detectar libertad en un agente artificial. Esos debates son importantes, sí, pero el golpe de realidad es más inmediato, pues incluso si no fueran libres, los sistemas ya están siendo tratados como si su *output* tuviera autoridad. Y si algún día merecieran el nombre de libres, el verdadero dilema sería qué tipo de normativa haría posible que esa libertad no aplaste la nuestra.

En ese horizonte aparece la prudencia, la obligación de no legar una infraestructura política en la cual la libertad humana sea un residuo decorativo. Hay marcos que ya ensayan ese impulso precautorio, al proponer límites y prohibiciones a ciertos usos por

considerarlos incompatibles con libertades fundamentales. En el mismo registro, la idea de prohibir prácticas consideradas inadmisibles revela algo crucial, que la tecnología no es el destino, y que las sociedades aún pueden decir «no» antes de que el daño se vuelva irreversible.

Ahora bien, defender límites no equivale a negar la pregunta de la libertad artificial. Por el contrario, tomarla en serio exige precisamente abandonar los extremos cómodos. El extremo ingenuo de la máquina como nuevo sujeto moral pleno, listo para tener «derechos», puede convertirse en un atajo para encubrir responsabilidades humanas. El extremo cínico de la máquina como simple martillo sofisticado puede convertirse en un velo para negar los efectos reales de su autonomía operacional y su opacidad. Entre esos extremos, esta obra propone una tarea más difícil, la de explorar, con rigor, si tiene sentido hablar de libertad artificial como hipótesis investigable y no como fantasía antropomórfica.

Pero el mayor mérito de esta discusión no está en resolver si una máquina merece el título de libre. Está en obligarnos a ver algo que preferimos no mirar, que la libertad humana, tal como la hemos narrado durante siglos, se vuelve frágil cuando el entorno social se llena de arquitecturas de decisión que operan por debajo del umbral de nuestra conciencia. La pregunta *¿puede una* IA *ser libre?* es, en el fondo, un espejo cruel: nos pregunta si aún somos capaces de distinguir entre elegir y ser empujados; entre querer y ser condicionados; entre decidir y ser administrados.

Así, esta obra no invita a venerar las máquinas ni a demonizarlas. Invita a recuperar la complejidad perdida de palabras como libertad, autonomía, responsabilidad y dignidad. Y lo hace con la apuesta peligrosa, pero necesaria, de que la libertad no es un atributo que se posee como se posee un objeto, sino una relación que se construye y se pierde, en la arquitectura misma de la vida social. Por esto, la pregunta final no es tanto si las inteligencias artificiales se volverán dioses y nosotros las marionetas, sino, si ello ocurre, cómo podremos desatarnos si los hilos no se ven.

Introducción

¿Pueden las máquinas ser libres?
¿Hasta dónde esta libertad implicará quitar la libertad a las personas?

Estas dos preguntas marcan el pulso de una época en la que la libertad de las máquinas ha dejado de ser una promesa futurista para convertirse en telón de fondo de nuestras decisiones, nuestra economía y, de manera cada vez más palpable, de nuestra propia autocomprensión. Si en el pasado la imprenta, la máquina de vapor o la red eléctrica transformaron las coordenadas de la libertad humana, hoy es el código el que ensancha y a la vez comprime el espacio de la voluntad. Cuando delegamos a sistemas inteligentes la facultad de clasificar, predecir y actuar, la libertad humana, en sus dimensiones individual, colectiva y estructural, se ve tensada.

Por un lado, nuestras capacidades se expanden hacia horizontes de creatividad y eficiencia inéditos, pero, por el otro, corremos el riesgo de convertirnos en usuarios manipulados de procesos cuyo funcionamiento interno apenas comprendemos y que, sin embargo, moldean el acceso a la información, el crédito, la salud o la justicia. Indagar sobre la posible libertad de las máquinas resulta, así, inseparable de la defensa y resignificación de la libertad humana, obligándonos a repensar los conceptos de agencia, responsabilidad

y dignidad en una realidad social que difumina las fronteras entre sujeto y artefacto, y que amenaza con redefinir lo que entendemos por acción intencional y por autonomía moral.

Aunque la libertad se haya pensado por tradición como un atributo estrictamente humano, el Derecho y la Filosofía llevan siglos conviviendo y operando con formas de agencia no-humana, lo que vuelve ineludible replantear hoy la noción de agencia cuando el agente es algorítmico.

En efecto, la agencia no-humana no es un descubrimiento exótico de nuestra era digital. El Derecho moderno ha construido, desde hace siglos, sujetos artificiales, las personas jurídicas, a los que atribuye capacidad para actuar, asumir obligaciones, detentar derechos y ser responsabilizados, aun cuando su voluntad no sea biográfica sino institucional y se exprese a través de órganos, representantes y reglas internas.

La Filosofía, por su parte, ensayó tempranamente la figura del autómata como contramodelo, máquinas capaces de producir conducta sin interioridad, útiles precisamente para delimitar qué entendemos por intención, control y libertad. Entre la ficción jurídica que permite imputar acciones y el autómata filosófico que interroga el libre albedrío, se abre el antecedente conceptual del problema contemporáneo, sistemas que ya operan como agentes en el mundo y que, sin embargo, desafían las categorías heredadas con las que describimos la libertad.

En este contexto emergen diversas preguntas jurídicas y éticas de trascendental relevancia: ¿cómo se asignará responsabilidad cuando una decisión autónoma genere un daño? ¿Qué principios deberán guiar la coexistencia normativa entre libertades humanas y libertades artificiales? ¿De qué modo se puede salvaguardar la dignidad y la justicia en sistemas en los cuales la agencia se reparte entre humanos y máquinas? Estas interrogantes reclaman un profundo análisis académico que sirva de base para futuros marcos regulatorios capaces de armonizar innovación y protección de derechos, así como para definir una ética aplicada que anticipe los dilemas de poder, sesgo y desigualdad que la libertad artificial pudiera agravar.

Para abordar estas problemáticas, esta obra se articula en cuatro capítulos interdependientes y progresivos. El primero ofrece una genealogía crítica del concepto de libertad en el pensamiento occidental, desde la *eleuthería* griega y el libre albedrío