

Sospecho que el ordenador se equivocó. Comentario a Mark Coeckelbergh, Filosofía política de la inteligencia artificial*

(2023) Cátedra
Madrid, 224 pp.

J. Ignacio Criado Grande
Universidad Autónoma de Madrid
ORCID ID 0000-0002-9184-9696
ignacio.criado@uam.es

Cita recomendada:

Criado Grande, J. I. (2024). Sospecho que el ordenador se equivocó. Comentario a Mark Coeckelbergh, Filosofía política de la inteligencia artificial. *Eunomía. Revista en Cultura de la Legalidad*, 26, pp. 433-440
DOI: <https://doi.org/10.20318/eunomia.2024.8521>

Recibido / received: 25/02/2024
Aceptado / accepted: 09/03/2024

Algunos temas han centrado la agenda política, social, cultural, filosófica, etc., en este inicio de milenio, pero quizá ninguno de ellos haya captado la atención como lo ha hecho la Inteligencia Artificial (IA), al menos, en los últimos años. Si bien no es una cuestión nueva, no es menos cierto que el reciente interés por todas las aristas de la IA es revelador de las muchas incógnitas, preocupaciones y oportunidades, y por qué no decirlo también, fantasías, que genera un concepto que ha alcanzado el nivel de fenómeno social. Todo ello, en un contexto mucho más digitalizado respecto del momento en que las ideas originales sobre la IA acertaron sólo a vislumbrar las primeras utopías/distopías en los años sesenta del siglo pasado. Dicho lo anterior, pocos libros recientes como el de Mark Coeckelbergh, *La Filosofía Política de la Inteligencia Artificial. Una Introducción*, desentrañan con tanto alcance algunos de los debates y conceptos filosóficos, pero también políticos o jurídicos, más destacados que están presentes en torno a la IA, a través del diálogo con una nutrida variedad de autores de la historia filosófica y política.

Comienza el libro con una introducción en la que se rememora *El Proceso* de Kafka a partir de una historia que termina con la frase de un miembro de un cuerpo

* Este documento se ha realizado con el apoyo del Proyecto PID2022-136283OB-I00, MCIN/AEI/10.13039/501100011033 y el FSE+.



policial (y que inspira el título de este mismo documento): «sospecho que el ordenador se equivocó». En dicha situación real, una persona negra es detenida gracias al uso de un sistema de reconocimiento facial basado en algoritmos e IA, sin que nadie del cuerpo policial de Detroit supiera realmente el motivo, provocando que el presunto «malhechor» (un ciudadano cualquiera sin la menor sombra de sospecha) fuera excarcelado sólo mucho tiempo después (eso sí, afortunadamente, sin acabar como el protagonista accidental de la novela kafkiana un siglo atrás). Este caso reciente condensa algunos de los miedos más íntimos en torno a la IA: un conjunto de herramientas tecnológicas basadas en la utilización masiva de datos, acerca de las personas y sus contextos de actividad, que integrados con sistemas algorítmicos y grandes capacidades de computación, permitirían aumentar exponencialmente las capacidades humanas (razonamiento, predicción, etc.), en diferentes áreas, como la vigilancia (Zuboff, 2019) y, eventualmente, abrir el camino a una inteligencia superior (momento de singularidad). Sin embargo, no todo es tan evidente como parece, y la multiplicidad de ramificaciones derivadas de la aplicación de la IA en la robótica o la automatización requieren, a juicio del autor, una atención desde una perspectiva filosófica-política, que apenas se ha desarrollado más que embrionariamente.

Resulta fundamental identificar el punto de partida del libro, que nada tiene que ver con muchas de las aproximaciones a la IA al uso, ancladas en un notorio, cuando no espurio, determinismo tecnológico (tanto utópico como pesimista). El autor belga, uno de los más originales de su generación en la escena filosófica presente a la hora de adentrarse en los debates sociales y políticos más complejos, se posiciona ante la necesidad de un enfoque «no instrumental de la tecnología: la tecnología no es solamente un medio para alcanzar un fin, sino que también moldea ese fin». (Y este posicionamiento sobre la IA, que el autor sintetiza en la frase «La IA es política de cabo a rabo», denota su alcance más allá de lo técnico y su ausencia de neutralidad en términos de política y poder. De hecho, esta posición entronca con científicos sociales que han construido el corpus teórico sociotécnico a la hora de entender de forma no determinista la relación entre tecnologías y sociedad, economía o política (Pinch y Bijker, 1984; Castells, 1997). En todo caso, esta aproximación suena mucho más original en el plano filosófico, donde no abundan este tipo de propuestas.

Lo anterior también conforma el contexto a partir del cual se atisba el propósito del libro, a saber: crear un diálogo entre la reflexión sobre política y la reflexión sobre tecnología. Y esto, con el objetivo de servir de cauce para abordar algunos de los problemas clave de nuestro tiempo (tales como el cambio climático, las desigualdades, el envejecimiento, las migraciones, la guerra, el autoritarismo o las pandemias), debatiendo conceptos filosófico-políticos concretos, incluyendo libertad, igualdad y justicia, democracia, poder o cambio climático y poshumanismo/transhumanismo. Como se puede observar, la lógica interna de funcionamiento del libro se sostiene en esa recreación de la IA como una realidad que desafía algunas de las construcciones conceptuales más significativas del campo filosófico-político. Y esas nociones filosófico-políticas se desarrollan en cada uno de los capítulos que componen el libro, lo que nos permite sugerir algunos de sus hallazgos clave en los siguientes párrafos.

«Libertad: manipulación por parte de la IA y esclavitud robótica» es el sugerente título del segundo capítulo del libro. Dentro del mismo el autor hace referencia a varias preguntas clave: «¿Qué significa la libertad cuando la IA ofrece nuevas formas de tomar, manipular e influir en nuestras decisiones? ¿Cuán de libres somos cuando realizamos trabajos digitales para grandes y poderosas empresas? Y, ¿conlleva la sustitución de trabajadores por robots la continuación del pensamiento esclavista?» Cada una de esas preguntas y otras se van destilando en el capítulo a

partir de diferentes concepciones de libertad, que llevan aparejadas cuestiones sobre la IA, algo que sucede cuando se habla de libertad negativa o libertad positiva. A modo de ejemplo, en el caso de la libertad negativa se aborda el impacto de sistemas de IA para el reconocimiento facial o biométrico en ámbitos básicos de privacidad o la propia privación de libertad (como en el supuesto señalado previamente). También se ofrece el caso de la *policía predictiva* para comprender potenciales sesgos en las decisiones o incluso momentos *kafkianos* con decisiones no justificadas o no comprensibles. De hecho, se lleva el argumento al extremo, cada vez más plausible, apuntando los peligros de la proliferación de regímenes autoritarios o totalitarios, lo que algunos ya han llamado *datacracias*, basados en la extirpación de libertades fundamentales como resultado de un control exhaustivo de los datos sobre cada vez más espacios de la vida de las personas.

No es menos interesante la aproximación a conceptos novedosos como el *nudging* (empujoncitos) o los derechos morfológicos y su relación con la IA. En el caso de la libertad positiva, el autor apunta cómo la pretensión de alterar el comportamiento de las personas a través de la explotación de las oportunidades que brinda la combinación de los conocimientos de la psicología cognitiva con la IA resulta todo un desafío. Los sistemas de recomendación de las grandes plataformas tecnológicas del momento nos dan empujoncitos hacia determinadas elecciones (sobre música, libros, películas, comida, relaciones sociales, etc.), ¿por qué no también hacia otro tipo de comportamientos sociales como una dirección más amistosa respecto del medioambiente o una visión más positiva sobre los servicios públicos? Imaginemos las implicaciones éticas que se abren ante tales elecciones. También se apunta el debate sobre los límites de la libertad de expresión en plataformas sociales digitales ante la proliferación de noticias falsas, *deepkafes* o desinformación como arma política. Finalmente, el debate sobre la libertad a modificar el propio cuerpo y nuestras limitaciones biológicas controlando nuestra morfología (transhumanismo) o nuestra mente (neurotecnologías) es objeto de atención. Tal libertad se entiende que estaría «muy cercana a la noción liberal clásica de libertad como autonomía (libertad positiva) y, especialmente de la no interferencia (libertad negativa)» (Coeckelbergh, 2023, p. 50).

Continúa el autor, en el tercer capítulo, «Igualdad y justicia: sesgo y discriminación por la IA», con otras preguntas clave. En concreto, aquí trata de dar respuesta a ¿cuáles son los efectos políticos de la IA y de la robótica (normalmente no intencionados) en términos de igualdad y justicia? ¿Incrementa la automatización y la digitalización ligadas a la robótica las desigualdades en la sociedad? ¿La toma de decisiones automatizada por parte de la IA conduce a una discriminación injusta, sexismo y racismo, y si es así, por qué? ¿Cuál es el significado de justicia y equidad utilizado en estas discusiones? Esta parte del libro se concentra en los sesgos discriminatorios (género, raza, orientación sexual, estatus socioeconómico, etc.) ligados a los algoritmos y cómo tienen un impacto directo e indirecto en la justicia e igualdad. Los sesgos se pueden encontrar en diferentes momentos del ciclo de vida de los algoritmos, desde los datos de entrenamiento, el diseño del algoritmo por el equipo de desarrollo, los datos en los que se aplica el algoritmo o los resultados derivados de la interacción con los seres humanos. Desde un plano teórico, se aborda el debate de la IA y las desigualdades desde la perspectiva liberal clásica (igualdad de oportunidades), la justicia distributiva, la justicia como equidad o redistributiva, así como los enfoques posmarxistas o posmodernos de la identidad.

Aquí se sostiene que hay quienes consideran la posible perpetuación de los sesgos y desigualdades sociales una vez implementada la IA, no sólo por los aspectos señalados, sino también por quienes están detrás de la propiedad de estas tecnologías y sus características intrínsecas. Este proceso puede ser desencadenado

tanto por gobiernos y administraciones públicas, por ejemplo, en la determinación de ayudas sociales o el acceso a determinados diagnósticos o tratamientos médicos, como por otros actores económicos como los bancos (créditos) o aseguradoras (seguros de vida, coche, etc.), empresas repartidoras de comida a domicilio, taxis o buscadores de contenidos en la web. La implementación de la sociedad calificada (*scored society*) y los sistemas de clasificación de personas en función de su comportamiento, estatus socioeconómico o características personales (que recientemente han sido prohibidos por la regulación europea de IA), activa todas las alarmas en relación con los sesgos desigualitarios, sobre todo, en países sin regulaciones o una visión regulatoria demasiado laxa. También lo hace un sistema de propiedad de las grandes corporaciones ligadas a la IA, que se sitúan en sociedades occidentales acomodadas, así como el hecho de que «teniendo presente el trasfondo del colonialismo histórico, se puede mirar críticamente a la IA actual y a otras prácticas técnicas refiriéndose al peligro (¿o realidad?) del neocolonialismo» (Coeckelbergh, 2023, p. 73). Un neocolonialismo de los datos y de la explotación de materias primas necesarias para su desarrollo que, junto al proceso de neofeudalismo digital entre bloques regionales, son dos de las macrotendencias más inquietantes desde el plano de las desigualdades a escala internacional.

Por su parte, en el capítulo cuatro del libro se abordan los nuevos desafíos de la IA para la «Democracia: cámaras de eco y totalitarismo de las máquinas». En particular, las preguntas que se presentan en esta sección son algunas como ¿Destruye la democracia la vigilancia mediante IA? ¿Nos encontramos de camino hacia un tipo de fascismo y colonialismo de datos? Lo anterior parte de la premisa según la cual la IA puede ser utilizada, no sólo para la vigilancia total, sino también para la manipulación masiva de votantes y procesos electorales, tal y como se habría demostrado empíricamente con la empresa *Cambridge Analytica* durante las elecciones presidenciales de EE. UU. o el referéndum del Brexit, ambos en 2016. Todo ello sirve para discutir las diferentes concepciones sobre la democracia y sus tensiones respecto al uso de la IA, desde las concepciones platónico-tecnocráticas, los ideales de democracia participativa y deliberativa o las más recientes aproximaciones agonistas. Sea cual sea la aproximación, el autor propone, no sin razones, que «se teme que las tecnologías de IA conduzcan a un nuevo tipo de ágora político» (Coeckelbergh, 2023, p. 81) que, siguiendo uno de los libros más recientes e influyentes sobre el tema (Zuboff, 2019), hace que nos encontremos dentro de un «capitalismo de vigilancia», que representa una amenaza individual y para la propia democracia, rememorando la jaula de hierro weberiana. Se trata de una evolución del capitalismo que expropia el conocimiento humano a través de la captura de acciones, comportamientos, e incluso, deseos y emociones, e impone un control basado en la concentración de conocimiento en grandes corporaciones (EE. UU.) o gobiernos (China). Esta última cuestión seguramente traspasará esta primera parte del siglo, si no es que no lo hace en su conjunto.

Lo anterior abre la puerta a un debate de calado sobre las diferentes aproximaciones a través de las cuales se puede abordar la tensa relación entre IA y democracia. Por un lado, se encuentra el debate sobre las cámaras de eco, el populismo y la polarización (que no se menciona, pero es parte y resultado de todo lo anterior). De hecho, los procesos democráticos se encuentran cada vez más enfrentados por la transmisión de bulos y desinformación, así como discursos de odio o, directamente, acciones en redes sociales digitales que se orientan a construir realidades inventadas, que son consumidas por una parte de la población y tienden a visibilizar discursos cada vez más extremos. Más grave todavía resulta enfrentarse a las lecciones sobre totalitarismo de Arendt (2017) y sus derivadas en relación con la IA (banalidad del mal). Aquí se sugiere que «la IA facilita que dispongamos de una versión digital del Gran Hermano en la que cada ciudadano se encuentra bajo

vigilancia constante mediante telepantallas» (Coeckelbergh, 2023, p. 106). El sistema de crédito social en China representaría la materialización de un escenario potencialmente totalitario, que incluso podría llegar más lejos, dadas las capacidades de control y manipulación masiva a través de las burbujas epistémicas. En definitiva, y este tema puede que sea el más trascendente del libro, en el mundo actual se estarían generando las condiciones psicosociales y socioepistémicas subyacentes para el surgimiento del totalitarismo, tal y como las mencionara Arendt respecto al nazismo. Véliz (2020, p. 114) asume que la «recopilación de datos puede matar» y compara la captura de datos personales actual, que anula nuestra privacidad y transparenta hasta lo más profundo de nuestro ser, tal y como ya asumiera Han (2013), con el uso de registros por los nazis para perpetrar el genocidio judío. Quizá no estemos lejos de experimentar una versión distópica totalitaria que contenga los anticuerpos que requieren las democracias para resistir ante unas amenazas crecientes y más serias.

A partir de aquí, el autor se adentra en una dimensión, todavía más si cabe, política en su capítulo «*El poder: la vigilancia y la (auto)disciplina por medio de los datos*». Las preguntas que guían este capítulo son algunas como las siguientes: ¿Cómo se puede usar la IA para sancionar la observancia de las leyes y autodisciplinarse? ¿Cómo afecta al conocimiento, cambia y moldea las relaciones de poder existentes entre humanos y máquinas, pero también entre los propios humanos e, incluso, dentro de los humanos? ¿Quién se beneficia de ello? Este capítulo contiene una referencia explícita a la obra de Foucault, también Butler y otros, si bien entronca con clásicos de la reflexión filosófica sobre la tecnología como Winner (1980) o el más reciente Sattarov (2019). Aquí se insiste en cómo la IA ya estaría influyendo en la construcción de la subjetividad humana, es decir, qué tipo de sujetos somos o a qué comunidades pertenecemos. Ahí aparece el carácter performativo de las tecnologías basadas en IA, ya que «como seres vivientes, móviles y situados, realizamos (*perform*) el yo y el poder, y la IA desempeña un papel en estas realizaciones (*performances*), por ejemplo, codiriéndolas» (Coeckelbergh, 2023, p. 126). En realidad, todo ello es un prelude de la afirmación de que el poder se realiza (*is performed*) mediante las tecnologías.

Efectivamente, el autor desarrolla su propuesta y plantea algunos hallazgos teóricos y conceptuales recientes de interés sobre IA y poder. Por ejemplo, cuando se compara la IA con el poder nuclear, que justificaría la nueva carrera internacional por el poder, mencionando a Bartolletti (2020). O también cuando se cita a Dyer-Witthof Kløsen y Steinhoff (2019), quienes observan al capitalismo actual, bajo el prisma marxista, «poseído» por la cuestión de la IA, que es «un instrumento de poder del capital y de la explotación». Asimismo, también se vuelve a mencionar el «colonialismo de datos» (en palabras de Couldry y Mejías, 2019) como expresión actual de la explotación y el poder que, igual que el colonialismo histórico se apropió de territorio y recursos para obtener beneficios, ahora explota seres humanos a través de la apropiación de datos. También asocia la sociedad disciplinaria de Foucault (1977), que funciona como el panóptico de Bentham, pero dentro de todos los aspectos de la vida humana, a aplicaciones para el cuidado de la salud, la meditación, la búsqueda de relaciones sociales, etc., de manera que se podría afirmar que «la IA y la ciencia de datos producen un yo cuantificado, a través de una práctica de autorastreo» (Coeckelbergh, 2023, p. 140). En definitiva, nos encontramos ante «tecnologías del yo», que implican no solo un ejercicio de dominación y castigo de otros, sino también de poder (autovigilancia, autorastreo, autocuidado y autocastigo) sobre uno mismo. Pero el autor, haciendo referencia a otras obras propias (Coeckelbergh, 2019a y 2019b), abre la puerta a otras posibles realizaciones (*performatividades*), que permiten elaborar una visión crítica de la IA orientadora de estrategias políticas emancipadoras del ser humano. Sin embargo, no queda del todo

clara la posible materialización de dicha perspectiva en el proceso político y social actual.

El capítulo sexto se pregunta ¿Qué sucede con los no humanos? Política medioambiental y poshumanismo. En este caso, las preguntas a desarrollar son algunas como: ¿cuentan políticamente solo los humanos? ¿Es la IA una amenaza o una oportunidad para lidiar con el cambio climático? ¿Pueden tener los sistemas de IA y los propios robots un estatus político de, por ejemplo, ciudadanía? ¿Cuáles son las implicaciones políticas si un ser superinteligente se pone al mando? ¿Es el fin de la libertad humana, la justicia y la democracia? En esta parte final del libro se abre la reflexión a debates más recientes de la teoría filosófica y política, con un despliegue de referencias a autores que excede las posibilidades de este texto. Sin embargo, algunas ideas son especialmente sugerentes, como la necesidad de cuestionarse la visión tradicional antropocéntrica de la política o las implicaciones de incluir en el *demos* a no humanos. Así, conviene conocer las diferencias entre poshumanismo y transhumanismo (tratando este último de las «mejoras de las capacidades humanas y que (al menos en una variante) considera la IA como reemplazo de los humanos o que toma el control del poder de los humanos» (Coeckelbergh, 2023, p. 159). Mientras el poshumanismo supone una suerte de «deconstrucción de lo humano» y las perspectivas dualistas y jerárquicas.

Lo anterior abre el debate acerca de la política y su relación con los animales y la naturaleza no humana, así como la posibilidad de un estatus político para la propia IA. Por un lado, se mantiene la creencia en que la IA puede facilitar el combate contra el cambio climático o ayudar en la mejor gestión y conservación de los hábitats naturales de los animales. Pero al mismo tiempo, el impacto de la IA sobre el medioambiente puede ser devastador, como lo confirman los ingentes recursos que son necesarios para el procesamiento de datos o la energía requerida para el entrenamiento de algoritmos, que generan una huella de carbono insoportable incluso a corto plazo. Con todo, se afirma que «un giro no antropocéntrico hacia la política de la IA no significaría contemplar a los animales, al medioambiente y al clima como una consideración adicional en un marco ya existente, sino que constituiría un cambio fundamental sobre la idea misma de lo político, que se ha expandido hasta incluir a los no humanos» (Coeckelbergh, 2023, p. 168). Una posición igualmente original se refiere al hecho de asumir que algunos sistemas de IA obtengan estatus político, dejando abierta la pregunta sobre si incluir potencialmente a los no humanos tecnológicos. Dónde situamos entonces los derechos de los robots y qué implicaciones tiene este asunto para lo político son aspectos que completan este bloque del libro, apuntando algunas ideas acerca, por ejemplo, del debate en torno al papel de la IA como agente político y de gobierno. Con todo, el mensaje final queda para el poder real de la IA, que consiste en lo que hacemos a nivel individual en nuestra vida cotidiana con los dispositivos a través de los que los algoritmos median nuestra relación con este tipo de sistemas. Quizá esta visión centrada en la necesidad de educar y formar a las personas sobre las oportunidades y desafíos de la IA queda inexplorada, y podría ser una de las líneas de acción más directas, y provechosas, para embridar algunos de los efectos no deseados de la IA.

El capítulo final, «Conclusión: tecnologías políticas», completa el libro con un resumen de sus ideas y hallazgos principales, permitiendo aquí debatir con el autor algunos de los aspectos clave contenidos en el mismo, así como quizá algunas de sus oportunidades de desarrollo en el futuro. La contribución del libro de Coeckelbergh es multifacética. Primero, el libro es una aproximación exhaustiva a diferentes ideas y conceptos clave de la filosofía y la política que se revisitan no sólo para conocer su encaje en la IA, sino también para identificar cómo la IA puede alterar su propia naturaleza. Así, la apelación a los términos libertad, igualdad y justicia, democracia o

poder, no es un mero ejercicio retórico, sino que trata de profundizar en sus desafíos actuales bajo la nueva luz que ofrece la IA. Evidentemente, no se ofrecen soluciones cerradas sobre cómo enfrentar las versiones actuales de los grandes retos sociales (desigualdades, hambre, envejecimiento, migraciones, guerra, autoritarismo o pandemias), pero sí se reflexiona de manera directa sobre algunas de sus derivadas contemporáneas.

En segundo lugar, la conexión entre la reflexión filosófica política y análisis político aplicado se desarrolla con éxito en muchos de los temas que se abordan. En realidad, las cuestiones que se analizan en el texto pueden ayudar no sólo a personas que se desempeñan en círculos académicos, sino que también pueden colaborar para que responsables y decisores públicos puedan construir soluciones informadas acerca de cuestiones prácticas, por ejemplo, cómo orientar y priorizar valores enfrentados en una nueva regulación o estrategia sobre IA, desde qué puntos de vista es posible acometer el debate sobre las nuevas maneras de control de los trabajadores mediante IA o cómo entender las contradicciones entre las políticas ambientales sostenibles y el desarrollo económico que puede generar el despliegue de la IA en determinados sectores de actividad y ámbitos geográficos.

En tercer lugar, este libro ofrece una guía para quienes desean analizar las implicaciones de las tecnologías, en general, y la IA, en particular, desde aproximaciones no deterministas. Si bien lo anterior se encuentra más o menos aceptado en otras disciplinas, quizá queda mucho más camino en determinados debates ámbitos filosóficos, pero también aplicados, en torno a la IA en las humanidades, así como en las ciencias sociales, económicas o jurídicas. Ya se mencionó que ese punto de partida que insiste en el carácter no instrumental de la tecnología es esencial para entender qué sucede, y qué sucederá en el futuro, ante nuevos impulsos tecnológicos por venir. A modo de ejemplo, la computación cuántica seguramente supondrá un salto cualitativo para el que necesitamos contar desde ya con el suficientemente aparatado teórico y conceptual con el que entender sus implicaciones.

El futuro estudio de las tecnologías en nuestras sociedades seguirá requiriendo de *La Filosofía Política de la Inteligencia Artificial*. Pronto será necesario acercarse también a la dimensión empírica de la implementación de este tipo de tecnologías en diferentes ámbitos, poniendo especial énfasis en el papel de los poderes públicos para acometer una regulación y puesta en práctica de nuevas aplicaciones basadas en IA, lo que requeriría situar siempre a los derechos humanos (y quizá también de los no humanos) en el centro. Aunque es evidente que se están desarrollando diferentes aproximaciones sobre la IA en diferentes partes del mundo, no es menos cierto que la supervivencia de la especie humana es posible que dependa de que se establezcan unas reglas mínimas que sean comunes en todos los contextos geográficos del planeta. Por consiguiente, en qué medida la IA también pueda ser un facilitador de instituciones políticas transnacionales, que lleguen incluso más allá de la Tierra, sería otro de los ámbitos a explorar en los próximos años. En suma, en un mundo donde se habla de una política posdemocrática y muchos agoreros adelantan su superación, quizá vale la pena imaginar nuevas formas de política democrática que puedan nutrirse de sistemas de IA, siempre que también tengan integrados esos valores comunes y cívicos desde su ideación y diseño.

Bibliografía

- Arendt, H. (2017 [1951]). *Los orígenes del totalitarismo*. Alianza.
Castells, M. (1997). *La era de la información: economía, sociedad y cultura*. Alianza.



- Coeckelbergh, M. (2023). *La Filosofía Política de la Inteligencia Artificial. Una Introducción*. Cátedra.
- Couldry, N. y Mejías, U. (2019). *The costs of connection. How data is colonizing human life and appropriating it for capitalism*. Stanford University Press.
- Han, B. C. (2013). *La sociedad de la transparencia*. Herder.
- Pinch, T. y Bijker, W. (1984). The Social Construction of Facts and Artefacts: or How the Sociology of Science and the Sociology of Technology might Benefit Each Other. *Social Studies of Science*, 14 (3), 399-441.
- Sattarov, F. (2019). *Power and Technology: A Philosophical and Ethical Analysis*. Rowman & Littlefield.
- Véliz, C. (2020). *Privacy is Power. Why and How You Should Take Back Control of Your Data*. Bantam Press.
- Winner, L. (1980). Do Artifacts Have Politics? *Daedalus*, 109 (1), 121-136.
- Zuboff, S. (2019). *The Age of Surveillance Capitalism. The Fight for a Human Future at the New Frontier of Power: The Fight for a Human Future at the New Frontier of Power*. Profile Books.