



Resonancias
Revista de Filosofía

N° 17 / Julio 2024



Resonancias

Revista de Filosofía

RESONANCIAS. REVISTA DE FILOSOFÍA N° 17 • AÑO 2024 • ISSN: 0719-790X

DIRECTOR: Felipe Álvarez Osorio, Universidad de Chile, Chile.

EDITOR EJECUTIVO: Javier Castillo Vallez, Universidad de Chile, Chile.

ASISTENTE EDITORIAL: Nicolás Rojas Cortés, Universidad de Chile, Chile.

SECRETARIO DE REDACCIÓN: Hugo Alarcón Acuña, Universidad de Chile, Chile.

EDICIÓN: Patricia Ayala

DISEÑO: Rosana Espino

Sumario

ARTÍCULOS

Megan Sara Zeinal-Werba	Responsabilidad y lenguaje en la obra de Agustín de Hipona	7-30
Martin Becker-Lorca	En torno a diferentes usos de la nada: místico y filosófico-político	31-55
Felipe Zegers-Quiroga	El principio supremo de la moral kantiana como conocimiento práctico	57-78

DOSSIER: INTELIGENCIA ARTIFICIAL

Tamara Chibey-Rivas & Roberto Rubio	Editorial: Inteligencia artificial: cuatro planteos filosóficos	81-84
Diana Inés Pérez	La IA entre nosotros: ¿qué debemos hacer?	85-100
Gabriela Arriagada-Bruneau	Una mirada crítica a la ética de la IA: de preocupaciones emergentes y principios orientadores a un desvelar ético	101-120
Martín Gendler	Datificación social e inteligencia artificial: ¿hacia un nuevo “salto de escala”?	121-141
Javier Blanco	Cognición computacional	143-155

RESEÑAS

Nicolás Antonio Rojas-Cortés

David Lapoujade. *La alteración de los mundos, Versiones de Philip K. Dick.* Ciudad Autónoma de Buenos Aires, Argentina: Cactus, 2021.

159-165

Hugo Alarcón-Acuña

Gorgias. Trad. Benjamín Ugalde. *Elogio de Helena.* Santiago: Democracia y Libertad, 2023.

167-170

The logo consists of a square frame containing the letters 'V' and 'A'. The 'V' is black and the 'A' is red.

ARTÍCULOS

Responsabilidad y lenguaje en la obra de Agustín de Hipona

Responsability Language in Augustine of Hippo

MEGAN SARA ZEINAL-WERBA¹

Universidad Católica del Uruguay/Universidad de Montevideo, Uruguay
megan.zeinal@ucu.edu.uy

Fecha de Recepción: 08/05/2023

Fecha de Aceptación: 04/01/2024

Resumen

En el siguiente artículo se recoge una singular mirada del lenguaje, desde una dimensión pragmática que sintetiza aportes relativos a la eficacia del lenguaje en la obra agustiniana. Desde un enfoque centrado en la responsabilidad y la autoapropiación de la acción lingüística, Agustín de Hipona se vierte a la detección de mecanismos volitivos y cognitivos que intervienen en los modos de uso e intercambio verbal para elaborar una singular y compleja reflexión sobre responsabilidad lingüística. Entre las relaciones y consecuencias más visibles de cada acto de habla en específico, expone cómo en la enunciación se conjugan múltiples variables de raíces profundas, que no solo están conectadas con la dimensión del *ethos*, sino que contribuyen a su constitución y determinación cotidiana para una reflexión comunitaria. Por lo que el siguiente análisis

CÓMO CITAR ESTE ARTÍCULO:

En APA: Zeinal-Werba, M. S. (2024). Responsabilidad y lenguaje en la obra de Agustín de Hipona. *Resonancias*, (17), 7-30. DOI: 10.5354/0719-790X.2024.70597

En MLA: Zeinal-Werba, M.S. "Responsabilidad y lenguaje en la obra de Agustín de Hipona." *Resonancias*, n.º 17, julio de 2024, pp. 7-30. DOI: 10.5354/0719-790X.2024.70597

Palabras clave: responsabilidad, lenguaje, signo, acción, performatividad

Keywords: Responsibility, language, sign, action, performativity

¹ Docente e Investigadora. Universidad Católica del Uruguay /Universidad de Montevideo. ANII Uruguay (Agencia Nacional de Investigaciones). Doctora en Filosofía Universidad Católica de Chile. ORCID: <https://orcid.org/0009-0003-9616-7154>

ético-discursivo a los escritos del filósofo de Hipona conjugará estos aspectos y habilitará la apreciación de una noción de responsabilidad verbal basada en fundamentos de performatividad del lenguaje y sentido de transmisión en sus obras.

Abstract

A unique vision at language is collected in Saint Augustine, seen from a pragmatic dimension that synthesizes the main contributions related to a sign efficacy. From an approach focused on responsibility and self-appropriation of action, the philosopher turns to the detection of volitional and cognitive mechanisms that intervene in the modes of use and verbal exchange. Among the most visible relationships and consequences of each specific speech act, it exposes how multiple variables with deep roots are combined in the enunciation, which are not only not disconnected from the dimension of ethos, but also contribute to its constitution and daily determination for a moral reflection. Therefore, the following ethical-discursive analysis of the philosopher's writings will combine these aspects and enable the appreciation of a first notion of verbal responsibility based on foundations of performative of language and meaning in transmission.

I. Introducción

La mixtura de una reflexión sobre la expresión verbal como elemento transmisor y la conciencia de la valía del signo lingüístico como forma dominante y hegemónica para el intercambio, toman un peso esencial en la obra agustiniana. A lo largo de su vida, Agustín de Hipona fue madurando progresivamente una perspectiva empírica y estrecha con la lengua que le permitió meditar acerca de la importancia del cuidado de la palabra, por las consecuencias y efectos que tienen los actos lingüísticos para la determinación de cualidades esenciales a la existencia humana. Su visión aumentó hasta llegar a preguntarse: “¿quién, efectivamente, podrá descubrir por qué se dice de tal manera cualquier cosa que se diga?”² En el presente artículo se recoge la singular mirada del lenguaje, vista desde una dimensión pragmática que sintetiza los principales aportes relativos a una reflexión moral. Desde un enfoque centrado en la responsabilidad y la autoapropiación de la acción, el filósofo se vierte a la detección de mecanismos volitivos y cognitivos que intervienen en los modos de uso e intercambio verbal. Entre las relaciones y consecuencias más visibles de cada acto del habla en específico, expone cómo en la enunciación se conjugan múltiples variables de raíces profundas, que no solo no están desconectadas de la dimensión del *ethos*, sino que contribuyen a su constitución y determinación cotidiana. Por lo que el

² Ret., I, 2.

siguiente análisis discursivo a los escritos del filósofo conjugará estos aspectos y habilitará la apreciación de una primera noción de responsabilidad verbal, basada en fundamentos de convivencia y sentido de la transmisión.

Como parte de este aporte, el pensamiento agustiniano congenia una comprometida visión acerca de la representación y del obrar con relación al exceso y faltas de la lengua (*peccata oris*), también con relación a ciertas virtudes y bondades que pueden ser expresadas en la afabilidad lingüística. Desde su perspectiva, esta preocupación no se centra en un recaudo de información que deviene de una motivación únicamente intelectual, sino que emprende una apreciación propia de la experiencia intensamente vivida y registrada. Asimismo, la investigación de las expresiones y la búsqueda de indicios que realiza sobre la tradición oral de su tiempo resultan un recurso metodológico fundamental para analizar el curso de una mirada crítica acerca de las formas y contenidos que valoró y rechazó; porque al percatarse de ciertos efectos del habla se vio comprometido a escribir y a revisar los presupuestos e intenciones implícitas a la realización de las palabras que, contiguamente, le resultan tan significativas como ordenadoras de la vida. Pues, Agustín, prontamente se da cuenta de que tanto la modulación de la volición, como la elección de elementos cualitativos de la dicción penden, en última instancia, de espinosas complejidades relacionales radicalmente humanas.

La propuesta de presentar algunas concepciones y problemáticas de la eficacia del lenguaje con relación al concepto de responsabilidad inaugura un nexo que no solo sirve para observar la implicancia determinante de la palabra en el alma, sino que también delinea una reflexión complementaria entre las determinaciones de la lengua y la acción moral en su dimensión colectiva. Algunos escritos, como *Confesiones*, *Sobre la Doctrina Cristiana* o *La ciudad de Dios* contienen dispersos esfuerzos por la divulgación de una comprensión adecuada de las consecuencias de cada práctica discursiva en el plano intersubjetivo, y abogan por un adecuado reconocimiento y uso de las modulaciones alegóricas, espirituales y literales de la lengua. La interacción humana considerada como un sistema de significaciones será definida en virtud de dos criterios que determinan un eje de referencia performativo: por un lado, el de la naturaleza y las características de la transmisión; y por otro, la naturaleza y particularidad de las formas que permiten la diferenciación de cada locución³. En lo que respecta al terreno de lo visible de los efectos lingüísticos, Agustín explora y se arriesga a intuiciones que asumen la conformación de una especie de ética discursiva plena de referencias cercanas que contribuyen a la vigilia de la locución y al continuo ejercicio de comprensión de un sentido constituido a partir de determinadas prácticas verbales. La vigilia del lenguaje como *praxis* en continua transformación y

³ Brown, P., *Biografía de San Agustín de Hipona*. (Trad. de Tovar, S & Tovar, M.R). Madrid: Ediciones Castilla, 1970, p. 345.

las apreciaciones al respecto que comparte asumen significados dinámicos, por medio de los cuales demuestra que la potencia del signo contiene mensajes poderosos y susceptibles de rearticularse en cualquier momento y lugar, según el nivel de conciencia. Dicho de otro modo, el filósofo da cuenta de un mundo de prácticas sociales sedimentadas en los mecanismos interiores como en convenciones intersubjetivas, en donde los orígenes arbitrarios de las asociaciones entre significados y significantes deberían ser olvidados. Lo que intenta es proveer de argumentos volitivos y de interrogantes profundas cada acción verbal. Busca demostrar los alcances y secuencias de la causalidad de cada efecto lingüístico, no solo por la limitación de las facultades cognitivas, sino también sus mismos factores de valoración se diversifican continuamente e irreversiblemente. Puesto que en su obra se aprecia la firme convicción de que el lenguaje como actividad, lejos de ser una realidad autónoma, está inserto en una trama de valoraciones ético-antropológicas que determinan la calidad de vínculos y relaciones psicosociales.

A fin de generar un análisis más completo, resulta preciso conectar la progresión de cuatro aspectos centrales que podría plantearse en torno al análisis de la palabra, en la visión agustiniana: el origen, el peso, la declinación social y la disposición anímica. De modo que Agustín observará estas cuatro cuestiones, sumando la potencia relacional para evaluar la cualidad de la expresión. Según sus palabras, toda expresión se compromete en la estructuración de una configuración que conllevaría un gesto indicativo de consecuencias visibles e invisibles en juego: “[...] porque cada hombre, a la manera de una letra en el discurso, forma como el elemento de la ciudad y del Estado”⁴. Las palabras, desde su ambigua naturaleza, para representar alumbran la insuperable opacidad de los procesos internos que se constituyen. En el estudio de la complejidad y de la extrema diversidad de la expresión lingüística se hace evidente que todo signo verbal opera, en cierta medida, enigmáticamente, a modo de rompecabezas que desafía la inteligencia humana y la impulsa a penetrar los misterios de la donación, iniciando un retorno al origen que podría resultar psíquicamente reparador para el alma. Porque las reglas que dirigen la acción exceden misteriosamente los actos en los cuales nos las representamos, y de aquellos en los que nos decidimos a cumplirlas, así como de los que consisten, justamente, en sus efectivos cumplimientos. Por lo cual, a pesar de que, en su análisis ya note que el número de factores a tener en cuenta al momento de esbozar una consideración de responsabilidad lingüística, su perspectiva incluya y esté ligada a la comprensión de una tensión existente entre las infinitas modulaciones que aparecen entre los mecanismos de transmisión y recepción, así también observará la variabilidad del contexto y la disposición interior por medio de la cual la palabra se ofrece. Su hipótesis de la eficacia del lenguaje aparece subyacentemente relacionada con

⁴ *Civ. Dei* IV, III.

una serie de encrucijadas relativas a los conceptos de *creatio*, *conversio*, *formatio*, que aparecen respectivas al devenir de las realidades espirituales y materiales de la acción humana. En *De Genesi ad litteram*, Agustín compone esta triada que le proporciona elementos para describir un esquema que explica la formación de la acción desde lo metafísico hasta las realizaciones y conformaciones físicas de la palabra, como un planteo que no solo estratifica, sino que toma las modulaciones y los procesos que son parte de la unicidad de la acción desde una visión compleja y procesual. Porque “florece la palabra cuando agrada la idea”⁵. De un modo especular, todo interlocutor transita entre un interior y un exterior, cuando la palabra entra e instala a través de la eficacia signica el mensaje de una tensión de que va haciéndose material y vitalmente manifiesta.

El lenguaje como acción significativa

Prontamente y por medio de diversas experiencias, Agustín da cuenta de que el lenguaje acciona en el alma, pero ¿cuál es la realización que ejecuta?, y aún más, ¿cuál es la propiedad que determina la actividad de su ejercicio? Es la significación. Como ya se mencionó en el capítulo I, tras el análisis y la definición del signo, el filósofo abre el paradigma conceptual de la significación al pronunciar que no solo las palabras son signos, sino que hay un mundo de acciones y gestos a los que nombra y asume como “significables”. Así, dice: “[...] pues llamamos universalmente signos a todas las cosas que significan algo”⁶. De igual modo lo organiza en *De Magistro* VI, 8: “¿Te parece bien que llamemos significables aquellas cosas que pueden significarse con signos y no son signos?”. De allí que los actos lingüísticos también puedan ser catalogados como significables o acciones significantes. Pero, ¿cómo interpretar su significado?, “¿qué diferencia hay entre estos signos?”⁷. El de Hipona deja la cuestión abierta para enfatizar que el juego dialéctico desplegado con Adeodato en *De Magistro* es más serio de lo que aparenta, porque el mismo recorrido laberíntico tiene un sentido que se performa en el curso del diálogo y que asoma ser revelado. Así cuando sugiere: “[...]en efecto (verás que), el que habla muestra exteriormente el signo de su voluntad por la articulación del sonido”⁸. Al afirmar que el conocimiento de las cosas significadas es mejor que los signos mismos, el filósofo empieza a ponderar y a traslucir el peso de la realidad a la que estos signos remiten: la realidad anímica. De un modo especular a la procesión de la formación, todo interlocutor transita entre

⁵ *Trin.* IX, VIII.

⁶ *Mag.* VI, 9.

⁷ *Mag.*, VI, 8.

⁸ *Mag.* I, 2.

un interior y un exterior, cuando la palabra entra e instala el mensaje de una tensión de significados que va haciéndose material y vitalmente manifiesta.

Los signos de la intención definen el discurso en el sujeto que lo pronuncia; y también los actos de recepción, interpretación y audición se hallan atravesados por quien lo obra; es decir, al concretizarse racionalmente, el lenguaje se relaciona con el sentido del alma de manera inmediata y no deja de inscribirse como una acción bajo los efectos del proceso de significación atravesados por la volición. De modo que, tras el reconocimiento de que toda oración resulta significativa, el filósofo se aboca a describir y descubrir los polifacéticos aspectos volitivos que cada acción verbal convoca en lo práctico. Y de esta manera, no solo amplió los límites de definición del signo y los asoció a la expresión verbal, sino que además proclamó un modo de visión sobre la performatividad lingüística ligada a una construcción significativa del deseo en la acción. A través del reconocimiento de que cada acción es un signo (y por tanto lenguaje), le resultó fundamental la indagación de su fuerza motriz subyacente, o del deseo por el que es movida⁹. Porque la estructura poliédrica de la significación no es arbitrariamente indicativa, sino que, en su mayor hondura, releva los aspectos de contenidos anímicos. Así afirma: "... como bien sabes, ya ves cuánto menos se han de estimar las palabras que aquello para lo que sirven, puesto que el uso de las palabras debe ser antepuesto a las palabras mismas"¹⁰. Declaraciones que muestran una concepción de racionalidad práctica que refleja una original posición que constantemente liga la reflexividad filosófico-lingüística de la vida a la palabra. Su dramática perspectiva de una relación tan ensamblada entre lenguaje y *praxis* expone el sustrato de una reflexividad, que no solo vincula cualquier signo o indicio de conocimiento a su aplicabilidad concreta, sino que, además, exige una rigurosidad consecuente al servicio de que esta canalice una significancia adecuada con la realidad anímica. Este hábito justifica que su intuición sobre la lengua sea, ante todo, entendida en vistas de un exigente conocimiento práctico, ni técnico, ni científico, ni contemplativo, sino una guía indispensable para experimentar la realidad y hacerla dialogar con un sentido interior al alma.

De este modo, de las palabras, puestas en varias frases y en sus lugares y oídas repetidas veces, iba coligiendo yo poco a poco los objetos que significaban y, vencida la dificultad de mi lengua, comencé a dar a entender mis querer por medio de ellas. Así fue como empecé a usar los signos comunicativos de mis deseos con aquellos entre quienes vivía y entré en el fondo del proceloso mar de la sociedad.¹¹

⁹ Bochet, I. *Saint Augustin et le désir de Dieu*. Paris: Institut d'études augustiniennes, 1982, pág. 44.

¹⁰ *Mag.* IX, 26.

¹¹ *Conf.* I, VIII, 13.

Ahora bien, dado que el lenguaje sirve como instrumento para traslucir ciertos mecanismos y contenidos de la voluntad y ciertas formas de querer que el hombre configura, el autor asume que cuando la intención es trasparente, las “palabras se oyen como signos de las cosas que tenemos en el ánimo” y son descriptas como los “signos comunicativos de deseos” presentes.¹² Por otro lado, y de manera inicialmente contradictoria, aun habiendo confirmado estas declaraciones y habiendo vueltos sobre ellas, afirma que la palabra de ningún modo llega a manifestar enteramente lo que tenemos en el espíritu, porque su tecnología no es capaz de colmar la representación absoluta de esta realidad, pues ningún signo tiene el potencial de manifestar ni transparentar por completo el pensamiento, ni de abarcar toda la complejidad interior de las facultades humanas. De modo que una de las primeras dificultades en las que el filósofo repara con respecto a la revelación y la eficacia del signo, redundaba en una ambivalencia contenida por cierta equivocidad o incompletitud respecto a la posibilidad de representación del lenguaje y su promesa misma de reflejo.¹³

La palabra como todo signo, produce una doble y conjunta correspondencia que refiere a la acción respectiva a hacer presente lo ausente. De allí se deduce que Agustín llama signos “a todo lo que se emplea para dar a conocer alguna cosa”¹⁴. La primera, en representar por medio del signo-palabra al referente; la segunda, en hacer presente a la memoria un recuerdo más remoto que el que el signo en su primera expresión revela. La fuerza del signo es la de hacer presente lo ausente a la mente y, como signos, los actos del habla incitan o no a la memoria del alma. Ahora bien, la potencia del signo dialoga con la posibilidad o imposibilidad de rememoración del alma, según la disposición y la dirección hacia la que se ejecute dicho signo. El pensamiento ha llamado a las cosas y las llama a través del lenguaje (con los signos), incitando a una operación contigua por medio de la que entiende su fuerza performativa. Puede entenderse que hay un acto del habla que acompaña en simultáneo y es contemporáneo al acto de significar que hace que la teoría del lenguaje agustiniano pueda presentar intuiciones de realización por medio de la incitación de la memoria. Su eficacia se efectúa mutuamente, pero en dos movimientos similares que operan en facultades de naturaleza distinta. Por lo tanto, tenemos dos partes: el signo que hace presente lo que se quiere decir, y también una fuerza que hace presente la *admonitio* orientadora que tiene el poder de determinación que habilita un recuerdo. Sería un error considerar que la fuerza del signo no es análoga al poder del modo de operar de la acción verbal. Son dos fuerzas que operan desde una analogía, pero

¹² *Mag*, XIII, 42.

¹³ *Mag*, XIII, 42.

¹⁴ *De doc. Chris.* Prólogo, II, 2.

con dos movimientos diferenciados en su campo de acción.¹⁵ Lo común es que en ambas, en esta relación entre signo y acción verbal, se configuran o proyectan tensiones nucleares que pendulan entre lo fáctico y la representación, entre lo necesario y lo contingente, lo psicológico y lo trascendental, como polaridades o fuerzas subyacentes por medio de las cuales se hace manifiesta exterior e interiormente la palabra.

El acontecimiento fundamental del lenguaje, como el final del *De Magistro* lo asume, es justamente la presuposición de un mensaje o sentido que, si bien no es ajeno a la significación, tampoco es la significación misma, porque su mensaje tiene un contenido trascendental que lo excede. Los actos del habla y los modos de expresión transparentan y determinan la forma en la que alma se relaciona con lo que está presente en una profundidad anímica de la que el lenguaje es potencialmente signo: el *verbum cordis*. El verbo interior o verbo del corazón aparece en la obra agustiniana como una realidad potencialmente dialogante. El acto de responder anímicamente a la invocación presente en la locución es tanto el dinamismo regenerador del sentido como el de la posibilidad de aprehenderlo; por ello, su captación es descrita como un acto primordial o fundamento de la vida. Toda *admonitio* signa un eje de coordenadas desde donde se trazan las relaciones de sentido escondidas bajo la potencia del verbo. Cada una de ellas lleva consigo un tono existencial y volitivo desde donde surgen las relaciones de significación y de relación. La palabra nos convoca a reunirnos en el ser; el vocativo es el propio dominio de la esencia lingüística que aparece anterior a la noción de signo, puesto que es la condición de posibilidad de su constitución, de una llamada que repite y responde vocativamente. Chrétien, al captar esta particularidad se pregunta: “¿Cabe pensar esa llamada de otro modo que como un verbo que venga a sobrecogernos y a requerirnos? La admonición recoge al hombre de la dispersión; donde surge la llamada, arde la palabra” (25). Se asume en la llamada agustiniana del lenguaje un *logos* que se pronuncia silenciosamente en cada cosa sensible (Merleau-Ponty, 1988). La voz llama inquietando la regeneración del verbo; es decir, engendra los discursos, generando un orden de concordancia que suscita la aparición del origen, lo reúne y pone en evidencia lo primordial de la relación que lo constituye. De modo que pensar el lenguaje desde la mirada de una interpelación no constituye un rasgo contingente, sino la esencia de su movimiento unitivo. Este acto de llamar y el acto de responder a la llamada es la dimensión profunda del sentido de la performatividad (por ser una acción concomitante a la expresión misma) en la que se designa la potencia lingüística de recordar esta acción. Como derivado de esta productividad del signo verbal, es que Agustín comienza a dimensionar la arquitectura de una responsabilidad lingüística para el despliegue de la conciencia. Hablar de responsabilidad en su

¹⁵ *De doc. Chris.* Prólogo, II, 2.

obra implica la comprensión de este registro eficaz del lenguaje por medio del que potencialmente se responde, de alguna manera, a lo que en lo visible o invisible nos lanza una interpelación.

Del mismo modo, la interpretación aloja esta arbitrariedad y variabilidad según quien la asume; la condición de la respuesta siempre debe estar libremente asignada a la dirección en que esté puesta la atención, pues eso hace y determina el sentido que tiene esa experiencia para el alma. Humanamente se requiere la conciencia de asumir la fuerza relacional implicada en el lenguaje y su implicancia para la dirección de la vida. Con esta visión, el de Hipona jamás supone que la experiencia lingüística pueda quedarse en una mera vivencia subjetiva de ese acontecimiento, sino que, desde una mirada auténticamente filosófica, intenta describirla como un acontecimiento que potencialmente ordena o dispone a oír esta convocatoria interna. El sentido de la responsabilidad en la obra de Agustín está dado por lo que hace el hombre frente a la posibilidad de contacto con la verdad interior y lo trascendental de la escucha; es decir, está muy asociado a la concepción de una veracidad que ordene la acción del lenguaje y la intención, en congruencia con el origen relacional del verbo. Por eso, el hiponense exhorta: “guarda el orden, y el orden te guardará a ti”¹⁶. Su reflexión sobre el lenguaje ético muestra que, aun siendo sus formas de enunciación muchas veces un imperativo moral, el modo más propio se fundamenta en juicios deontológicos y axiológicos que no solo no anulan la condición de libertad, sino que se fundamentan en la misma para la determinación de la vida. La libertad es la que rige la acción frente a la posibilidad del habla y, por lo tanto, rige también la responsabilidad como la respuesta que la deriva. Ese es el bien que se origina en una llamada, incluso en un grito: que abre la posibilidad al hombre de responder con su vida y a registrar los mecanismos y formas de ofrecerse. En el corresponder a la palabra dada es que el hombre escucha propiamente la interpelación del habla, y a través de ella se cruzan decisiones fundamentales relativas a la existencia misma del hablante y su identidad. Lo que se halla de modo latente en la expresión verbal es una llamada que nos quiere en acción recíproca y transforma la autocomprensión. Porque cuando se empieza a captar la admonición detrás del signo, ya de algún modo se ha respondido; se abre la mirada cuando captamos la llamada como tal, esta ya ha convertido el ojo. La palabra incita y por medio del deseo actúa como motor de la acción de respuesta de una disponibilidad que habla por sí misma, porque de algún modo: “escuchar la llamada es estar ya respondiendo a la llamada” (Chrétien 22), y de otro modo sería inaudible.

¹⁶ *In Epistolam Ioannis homiliae*, VII, 8.

La aporía presente en las fuentes agustinianas: es relativa a que la llamada detrás de todo signo de lenguaje presenta una imposibilidad congénita a verdaderamente corresponder. Porque, a pesar de ser el lenguaje un recurso para responder la palabra no llega a destino.¹⁷ Esto se debe a que Agustín distingue la diferencia radical entre la llamada originaria (divina) que resuena en toda criatura y la llamada humana. Con lo que respecta al quehacer humano, la llamada siempre convoca la fragilidad y limitación. De modo que la potencia de la voz está determinada por las dimensiones del ser. A su vez, la resonancia de la llamada convoca un requerimiento volitivo que se estructura dentro de la dinámica y de las lógicas de un orden dado. La respuesta a la *admonitio* se constituye según como responde o se corresponde. Así, cada palabra tiembla y resuena entre dos abismos: en el de distinguir el origen de la llamada que la hace posible y en el de advertir la posibilidad de albedrío que se juega en la respuesta.

Pero, si bien la admonición es una llamada que nos presupone como seres a una respuesta, y Agustín concibe la experiencia lingüística como una continua llamada, significa que convoca de antemano la voz del ser; que nosotros como criaturas seamos anticipados y prometidos en ese grito inicial no implica que pueda anticiparse nuestra propia escucha ni nuestra propia respuesta. La respuesta solo puede darse desde un terreno imprevisible y libre, de ahí se despliega su verdadera potencia para la conversión del ser.

En esa medida, ya hemos respondido a dicha convocatoria primera que fundamenta todas las otras, y cualquier respuesta o responsabilidad provienen de este llamado al que el hombre debe responder con libertad. Esta condición abre los objetos a lo que en el alma atrae y provoca deseo, en cuanto a que permite al hombre elegir a qué responder: “La vida cotidiana es algo más que coexistencia neutral; es un sentirse atraído, comprometido, reclamado por propuestas, estímulos o ideales que vienen a nuestro encuentro y nos invitan a reconocerlos y a responder a su llamada” (Goma 297). El ser humano tiene la posibilidad de elegir y de autodeterminarse, y eso conlleva el poder de responder con su propio lenguaje y hacer de su lenguaje un signo vital. La palabra figura y determina en contexto, y cada elección, por más sutil que parezca, la voluntad elige un camino determinado y, al mismo tiempo, elige querer o no asumir y experimentar tal estructura de la libertad. Las *Confesiones*, como instancia verbal de registro de la aparición de la conciencia para sí misma plasmada, pueden también ser interpretadas como una autoexaminación o un acto de responsabilidad del alma frente a sí misma que asume este conflicto de poder de elegir. De esta lectura se deriva la concepción agustiniana de libertad como autodeterminación de la criatura; es decir, de un ser que adhiere a su voluntad y asume su ser en una

¹⁷ De *doc. Chris.* VI, 6 ; De *Mag.* XI, 36.

libertad enmarcada. Para el filósofo, la elección pertenece a una zona íntima del alma que no está determinada por ninguna necesidad, sino estructurada por el poder alojado en sí misma, que le requiere ser autora de sus propios actos para ordenar su tendencia¹⁸.

El hombre es, de algún modo, responsable de la conversión y plasmación de su deseo. ¿Pero cómo regir esta conversión? ¿Hacia dónde orientarla? Cuando Agustín sugiere un orden amoroso que de alguna manera forja el núcleo de la identidad, habla de un estrato performativo y profundo de la persona que configura la vida y que reside particularmente en su forma y elección de amar y su significación moral (Fernández Bleites 266). Para desplegar su concepción de la expresión verbal acude al amor como acto espontáneo y espiritual, lúcido respecto al mundo interior como término intencional que configura el *ethos* de la persona. El amor está en la base de la vida moral, no solo desde la perspectiva de su configuración, sino también en la medida en que posibilita la articulación del autoconocimiento. El acto amoroso ordena, crea un orden y se ofrece como un dinamismo, un movimiento o una fuerza intencional. De ahí el aforismo y exposición de la Epístola de *San Juan II*, 14: “cada uno es tal cual es su amor”. Sin embargo, el nudo está en el carácter identificativo y constitutivo del *ordo amoris* como una voluntad que dirige la atención del espíritu. Hay una dependencia profunda entre conversión e *imitatio*, según la cual la voluntad reconoce el bien que anhela y toma la decisión de seguirlo, de repetir la «forma» de la vida que se presenta ante sí. La conversión (*conversio*) es, precisamente, el último momento de la *imitatio*, entendida como una *via similitudinis* o vía de la similitud. Porque no se puede escuchar verdaderamente si no es respondiendo a la promesa de nuestra identidad (Chrétien 41). La libertad agustiniana se diferencia del libre albedrío en ese punto, a saber: que no se define por el acto mismo de elección o por la potencia electiva, sino que su experiencia es la de una adhesión al bien por medio de la que cobra o adopta su forma y proporción en un orden dado. Ahora bien, queda ver como gestiona y comunica o representa cada hombre esta donación y su contigua debilidad.

Libertad como responsabilidad

La articulación entre iniciativa de donación divina y recepción-aceptación humana permite destacar una particular estructura dialógica de la moral figurada por Agustín, la cual puede ser comprendida como un proceso de integración

¹⁸ *Serm.*, 169, 13.

paulatino entre el despertar del alma al llamado y su posibilidad de respuesta, o como un proceso de articulación y resolución interna entre la verdad que se recibe y la conciencia moral. Para explicar más precisamente la relación entre la voluntad y las responsabilidades que Agustín desarrolla en este punto relativo al lenguaje, es preciso entender que la responsabilidad emana de la libertad o, en otras palabras, podría interpretarse que la responsabilidad es la carga de la libertad humana. La responsabilidad como la potencia y la habilidad de responder supone una exigencia moral que recorre inexplicitamente todo el pensamiento agustiniano. La libertad aparece tácitamente como autonomía, para darse a sí misma la dirección de sentido desde la cual cada uno elige las acciones y así, desde el cimiento más profundo, el ser humano puede decidir lo que es o quiere ser, esto es, porque puede elegir lo que ama. Toda alma se define por lo que ama y de ahí la figuración de que el “peso” es amor en cada acción que se realiza. En el centro de esta teoría toma sentido el célebre “Ama y haz lo que quieras”¹⁹, ya que la libertad no solo es parte y componente de la condición amorosa, sino que hace significativo su sentido como motivo existencial para responder a la vida.

El ser sí mismo se muestra como conciencia del propio poder y responsabilidad; como libre, pero con una libertad no absoluta sino derivada, que puede denominarse una libertad hacia una autodeterminación. El hombre debe responder a esta elección que lo determina, saberse autor de su elección como respuesta al orden dado. En resumen, la responsabilidad del discurso no debe limitarse a adoptar un punto de vista moral ajeno a formas de vida concretas, donde puedan realizarse efectivamente, sino que en su distensión cobra especial relevancia el potencial de reflexividad y disponibilidad hacia la vida interior. La responsabilidad no tiene que ver con una operación psíquica que tome en cuenta el cálculo de los efectos o consecuencias de lo que se dice o de cómo se desarrollará el significado hermenéutico de los interlocutores y oyentes, porque lo que gobierna los efectos mantiene siempre, como toda acción, una cuota insuperable de imprevisibilidad. Incluso la auténtica significación muere en esta imprevisibilidad hermenéutica de la acción y en la región de disimilitud interpretativa de la diferencia de una alteridad. En todo caso, la responsabilidad pasa por la conciencia reflexiva del peso y la convivencia continua con la imprevisibilidad misma, pero pende de la dirección con la que la acción fue inicialmente realizada. Por otro lado, resulta esencial poner de relieve las consecuencias y tensiones que se sintetizan en la teoría agustiniana del lenguaje, y que a su vez se representan como garantes de la responsabilidad moral. La responsabilidad es aquí enfocada y direccionada desde la perspectiva del proceso lingüístico, en tanto que el hombre debe aceptar activamente por medio de la palabra su autoría en la acción.

¹⁹ Cfr. *Ep. Ioan. VII, 8: Amat et quod vis fac.*

“El que te hizo sin ti, no te justifica sin ti”²⁰ se muestra como una frase que invita a la reflexión sobre una tensión entre gracia y libertad, en la que aparece mediando una acción lingüística. Con esa afirmación, el de Hipona diferencia el plano ontológico-creacional del plano ético-discursivo y, a la vez, los compara. Mientras que el acto creacional es puramente divino, toda acción posterior humana ya adhiere un componente ético que requiere la consideración de la acción libre y de la responsabilidad humana como tal. Su justificación lingüística deriva al hombre como agente de acción, a diferencia del acto creacional, que es de absoluta autoría divina.²¹ Sin embargo, el hombre debe asumirse y reconocerse en la acción que escoge.

La plenitud de la libertad en el hombre no solo consiste en elegir dejarse mover por *caritas* como amor corresponsal, sino en moverse por sí mismo, también hacia la dirección de esa forma del amor, la cual implica una superación del *cupiditas* por medio de una adhesión gozosa a la voluntad divina. La voluntad experimentada como certeza íntima e inmediata puede entenderse como fuente motriz que desde la propia experiencia interior lo describe como que: “No hay nada que sienta yo tan firme e íntimamente como que tengo voluntad propia y que por ella me muevo a procurar el goce de alguna cosa”²². De modo que la vida actúa regida por estas fuerzas íntimas, complejas, de la voluntad, a través de las cuales el sujeto justifica su poder de actuar y su ser como causa de sí. La autodeterminación aparece como propiedad esencial de la voluntad, como un estímulo directriz que permanecerá a lo largo de toda la obra de Agustín y que se deja resumir en la fórmula agustiniana: “la voluntad del hombre es la causa de sus actos”²³.

Confieso algo innegable: todos tenemos esta voluntad [...]. Porque nada hay, en efecto, que sienta yo tan firme e íntimamente [como] que tengo voluntad y que por ella me muevo a procurar el goce de alguna cosa. Y en realidad no encuentro qué cosa pueda llamar mía si no es mía la voluntad por la que quiero y no quiero.²⁴

Sin embargo, la articulación de la acción se complejiza cuando Agustín la describe como una facultad desdoblada de donde mana su vida consciente, sus palabras, sus acciones y que opera dialógicamente. En *Confesiones* VIII, capítulo IX, 21, *De trin.*, XIV, 6, 8 y en *De lib. arb.*, I, XII, 25, se encuentra atestiguada la

²⁰ Cfr. *Qui ergo fecit sine te, non te iustificat sine te.*

²¹ *Lib. arb.* I, 1.

²² *Lib. arb.* III, I, 3.

²³ *Civ. Dei.* V, IX, 3: *quoniam et humanae voluntates humanorum operum causae sunt; lib. arb.*, I, 12, 25: *Fateor, negari non potest habere nos voluntatem; III, 1, 3: non enim quidquam tam firme atque intime sentio, quam me habere voluntatem, eaque me moveri ad aliquid fruendum; quid autem meum dicam, prorsus non invenio, si voluntas qua volo et nolo non est mea.*

²⁴ *Lib. arb.* I, 12, 25; III, 1, 3.

paradoja que aloja la voluntad. La experiencia interna de ser uno movido por sí mismo es tan nítida que la voluntad tiende a identificarse con el propio ser sí mismo o su conciencia: “no encuentro qué cosa pueda llamar mía, si no es mía la voluntad por la que quiero y no quiero”²⁵.

Agustín sitúa ahora el conflicto de la voluntad en el interior del alma humana misma y formula una sola ley para los asuntos de la voluntad que intuye bajo la fórmula: “querer y ser capaz de querer”, no son la misma cosa. Basado principalmente en el Sermón 163, al igual que San Pablo, Agustín experimentó en su vida el conflicto interior entre dos voluntades, entre querer y poder (*velle-posse*), cuando teniendo ya claridad respecto del camino que había de seguir no logra ese deseo total o absoluto de la voluntad que consiste en una disyuntiva entre dos leyes que se prefiguran en el alma contrarias al individuo²⁶. Esta experiencia vital le lleva a comprender que la conversión completa de la voluntad pasa por el reconocimiento de la propia debilidad y yuxtaposición de las fuerzas del alma. Porque este no querer totalmente, el querer en parte y en parte no querer, es una cierta flaqueza que lo acompaña²⁷. El problema se plantea en *La Epístola a los Romanos*, y lo que en Pablo aparece como dos leyes antagónicas, Agustín lo describe como un carácter doble de la voluntad, expresado mejor en latín con los términos *velle* (quiero) y *nolle* (no quiero), siendo ambas voluntades, como se ha dicho ya, activas y afirmativas de un modo de querer. Hablará no de dos leyes, sino de dos voluntades que interpreta como un monstruoso desdoblamiento de la voluntad sobre sí misma.

¿De dónde este monstruo? ¿Y por qué así? Manda, digo, que quiera □ y no mandara si no quisiera□, y, no obstante, no hace lo que manda. Luego no quiere totalmente; luego tampoco manda toda ella; porque en tanto manda en cuanto quiere, y en tanto no hace lo que manda en cuanto no quiere, porque la voluntad manda a la voluntad que sea, y no otra sino ella misma. Luego no manda toda ella; y esta es la razón de que no haga lo que manda. Porque si fuese plena, no mandaría que fuese, porque ya lo sería. No hay, por tanto, monstruosidad en querer en parte y en parte no querer, sino cierta enfermedad del alma; porque elevada por la verdad, no se levanta toda ella, oprimida por el peso de la costumbre. Hay, pues, en ella dos voluntades, porque, no siendo una de ellas total, tiene la otra lo que falta a esta.²⁸

La facultad de elección, tan decisiva para el *liberum arbitrium*, se asocia aquí a la elección entre *velle* y *nolle*, entre querer y no querer. El *nolle* puede ser traducido

²⁵ *Lib. arb.* III, I, 3.

²⁶ *Sermón* 163, 10, 12.

²⁷ *Conf.*, VIII, IX, 21.

²⁸ *Conf.*, VIII, IX, 21.

como una ausencia de voluntad que actúa y tiene su propia fuerza, o como la enunciación de una resistencia que desdobra el deseo. De modo que donde haya una voluntad hay siempre dos voluntades, ninguna de las cuales está completa, y lo que tiene una de ellas le falta a la otra. Así, Agustín afirma que la actividad de la voluntad consiste en un querer-no querer simultáneo por el cual en todo acto de querer existe al mismo tiempo una resistencia, un no querer. En toda *voluntas* hay un yo quiero y un no quiero implicados²⁹. Así, “No se honraría nuestra libertad si se creyera, como presumiblemente, que de la intelección correcta resulta automáticamente la acción correcta” (Apel 114). Esto dado que, en sus múltiples aspectos, el amor agustiniano aparece como un regente central de la vida y de las relaciones; y hasta cierto punto es posible relacionar el *ethos* con el *ordo amoris*. Se presenta en ocasión a esta división, la posibilidad interpretativa de una interesante distinción entre el *ordo amoris* descriptivo y el *ordo amoris* normativo vinculado al desarrollo moral de la persona. Dado que la voluntad puede moverse en sentido opuesto a una verdad de razón que le haya sido revelada, la libertad es lo que resulta sustancial para analizar el terreno de la expresión verbal. Cualquiera que sea el poder que el intelecto pueda tener sobre el espíritu, se trata de un poder exigente y ambiguo; lo que el intelecto no podrá probar nunca al espíritu es que no debe simplemente adherir a su razón, sino también querer estar. Los razonamientos y entendimientos no movilizan por sí mismos a la acción, sino que es necesario el libre querer de la voluntad sostenido y reelegido.

Arendt hace un análisis crítico de esta cuestión agustiniana en su obra *La Vida del Espíritu*, y añade a la escisión de estas dos voluntades problemáticas y antagónicas una calificación sobresaliente y diferencial que establece Agustín respecto a la interpretación de la epístola paulina: ambas dimensiones del querer y ser capaz están estrechamente ensambladas. Toda acción implica la presencia de la voluntad, pero también la capacidad de ejecutarla en el terreno de la acción pública debe estar presente como instrumento de operación. La autora lee el lado de este conflicto de dos leyes antagonistas interiores bajo la condición de una fuerza de acción debilitada, entendiendo que lo que se desdobra en la condición humana es la unidad entre el deseo y el poder de actuar que ha perdido integridad. Agustín lucha a la vez contra la anulación del esfuerzo personal y contra la confianza desmedida en las fuerzas humanas, y presenta al ser humano, por una parte, como esclavo, dotado de una facultad que no puede ejercer plenamente en libertad y que a la vez se reconoce deseante de la misma.

Desde la perspectiva de Arendt, la aporía se presenta cuando se intenta buscar una causa exterior a la voluntad, porque esta nos llevaría a una regresión al infinito. Y, de hecho, en palabras de la autora, intentar superficialmente resolver

²⁹ *Romanos*, 7, 15-23.

el misterio de la voluntad escindida con una solución al estilo paulino, que recurre al don de la gracia intentando acabar con un problema, solo provocaría otros problemas más complejos. El conflicto de la voluntad como facultad del espíritu se encuentra en ella misma, y su riqueza filosófica permite la expansión de la conciencia y la maduración del pensamiento y de la acción en un desdoblamiento de la voluntad misma. La división «monstruosa» de la voluntad no se resuelve, sino que decanta en la elección mediante la importancia que atribuye al amor como peso del alma, como una ley de gravitación propia derivada de las doctrinas del amor y de la libertad humana para la vida relacional. Sin la necesidad de previos contenidos buenos o malos, en la misma voluntad se aloja un conflicto dual que Arendt reconoce como conflicto de voluntades plurales. El yo *volente* se divide por naturaleza. En todo caso, la resolución a este conflicto no dependerá —según la autora— de la gracia divina, como en San Pablo, sino, como veremos más adelante, de la ley agustiniana del amor. El amor viene a ser el factor decisivo, la fuerza que une y que actúa como una fuerza vinculante que integra y une la dualidad escindida en el alma. El acto amoroso es un acto de manifestación y afirmación para la libertad que permite decantar interiormente esta pluralidad.

En relación a este punto, también en adhesión, Chavez presenta un estimulante artículo relativo a la *admonitio* o llamado sónico relacionado con la idea de libre albedrío, la libertad y el proceso de liberación que tiene lugar en el transcurso histórico-biográfico. Según la autora, las palabras de Agustín en *De Magistro* sitúan al discípulo ante diversas acciones libres, señalando la atención, la experiencia, el juicio y el asentimiento. Se encuentran en nuestro espíritu: el ser, el saber y el querer. Estos tres elementos son inseparables de toda vida humana. “El yo soy, contiene implícito el sabiendo y queriendo, sé que soy y quiero, y quiero ser y saber”³⁰ La tríada aparece también en el *De Trinitate* como memoria, intelecto y voluntad. De modo que la analogía no debe interpretarse como la simple comparación, sino que el sentido de la semejanza se da a partir de la función unificadora de la voluntad con relación a las otras facultades del alma. La voluntad unifica estas facultades; es ella quien dice a la memoria lo que debe retener, y es ella también quien dice al intelecto lo que este debe seleccionar para conocer. Es ella misma la que ejerce la fuerza sobre el espíritu para llevar a cabo la acción que moviliza y le da unidad al resto de las dimensiones y facultades por medio de una acción que es necesariamente lingüística. En otros términos, la intelección en los fundamentos de la responsabilidad no conduce a la realización de acciones específicas ni puede ser garantía de estas, sino que es el deseo el que contribuye a direccionar la vida misma en función del llamado del lenguaje que reclama o incita a una respuesta.

³⁰ Chavez, P(2017). Palabras que liberan. El enseñar como *admonitio* en *De magistro* de Agustín de Hipona, *Topicos*, (53), México, pág. 304.

Responsabilidad lingüística

La reflexión de una responsabilidad lingüística está vinculada a todos los problemas expuestos que aparecen relativos a la libertad y a la potencia humana de responder con la acción a la admonición originaria que inicialmente funda todo signo y lenguaje. Hablar de responsabilidad lingüística sugiere que cada sujeto podría disponerse o no a actuar en la línea de la *admonitio* originaria, *in via* hacia la verdad como camino hacia la libertad plena por medio del registro, captación y escucha de esos mensajes. Algo que demuestra que tanto la concepción de la “vida buena” (*beata vita*), como las cuestiones puramente evaluativas que atañen a la esfera de la eticidad inmanente se hallan relacionadas a la vida del lenguaje. En este aspecto es que la sabiduría práctica no requiere solo de la reflexión teórica sobre principios morales y consecuencias, sino también una adhesión de la voluntad al llamado originario que se manifiesta en la palabra interior. Se trata de ver otra cara del discurso, como una *praxis* argumentativa, narrativa e interpretativa que revela a modo de especular la condición de la voluntad y su posibilidad de responder al llamado originario que para Agustín justifica el signo.

Por otra parte, cuando Chavez sugiere que un sentido de responsabilidad agustiniana podría ser considerado desde un carácter heterónomo y teónimo, en cuanto su criterio moral se encuentra no en la sola razón, sino en la relación entre la conciencia humana y la fisionomía de la creación, lo hace para sugerir la articulación dual de estos aspectos de la respuesta. El hombre debe responder a más de una realidad a la vez, y de ahí la complejidad real de la elección. Este mismo aspecto quizás haya contribuido para algunos agustinólogos a una polarización de la reflexión crítica del lenguaje. Esta doble dimensión que articula la admonición supone un proceso de elección y discernimiento que involucra, tanto la captación de las dimensiones de lo revelado como la gestualidad de la respuesta humana a dicha donación y revelación. Semejante a la dialéctica entre libertad y gracia, la respuesta humana a este respecto podría entenderse del siguiente modo: todo comienza con la gracia, que sale al encuentro del hombre, llamándolo, moviéndolo, esperándolo; la voluntad humana responde poniéndose o no en disposición a recibirla, escucharla. De ahí la figura de la responsabilidad agustiniana: justamente implica hacerse cargo de esa incitación al retorno, junto con el alcance de su libertad puesta en acción relacional. La voluntad indica el querer del alma y determina su deseo de orientarse o alejarse hacia aquello que descubre en el lenguaje con la memoria e inteligencia, por lo que los verdaderos alcances de esta libertad han de buscarse en un proceso de interiorización, en el

entrar el alma dentro de sí y reconocer el tesoro de su operatividad.³¹ La posibilidad o imposibilidad de corresponder a lo que se evoca significa disponerse o no a la bondad de la donación. En definitiva, la aproximación hacia una responsabilidad lingüística en la obra de Agustín, necesariamente, tomará en cuenta dos aspectos que aparecen conjugados: los esfuerzos por una comprensión radical y activa de la potencia del lenguaje, así como también de la propia libertad y sus alcances relacionales.

Hay también en la filosofía del lenguaje agustiniana un inherente sentido de responsabilidad/responsividad diferente, que más bien se hace respectivo a la alternante capacidad de advertir y actuar en función del compromiso implicado en el efecto de todo acto lingüístico. Incluso, la fundamentación de la filosofía es atravesada por una ética de la responsabilidad que se halla formulada por el sentido de un planteo pragmático-trascendental y dialógico al que el hombre debe advertir. Solo aquello que responde tiene verdaderamente sentido. La voz es, pues, la fuente de un sentido personalizado; detrás de ella hay un sujeto persona; pero no se trata de una “metafísica de la presencia”, de los sentidos preexistentes e inamovibles, ni de algo fantasmal, sino de un constante devenir del sentido permanentemente generado por el acto-respuesta, que se va modificando en el tiempo. Es decir, el sentido se hace manifiesto como una respuesta a lo dado. En su análisis ya es notorio que el número de factores a tener en cuenta al momento de esbozar una consideración de responsabilidad del lenguaje está ligado a la tensión existente entre transmisión y recepción, como también al contexto y a la intención con la que la palabra interiormente se ofrece. Su sentido se imbrica con la acción y adquiere el poder contenido en la misma. Como todo acto, el verbal posee el valor de un acto ético, sumado a la fuerza persuasiva y expansiva de todo llamado en potencia.

Desde otro ángulo, el desarrollo de la noción de responsabilidad en la obra de Agustín se inserta en el contexto de las reflexiones sobre el mal que atravesarán una parte significativa y recurrente de su producción filosófica sobre la libertad y el libre albedrío. Cuando en *De lib. arb.*, Agustín intenta resaltar una doctrina de la responsabilidad, su aproximación consistirá en un problema de dos enfoques: antropológico y ético. En ambos casos se evidencia la ruptura de Agustín con la antropología dualista y la lógica con una impronta determinista. La responsabilidad tiene que ver con un llegar a ser uno mismo, a responder al verbo del corazón y con desplegar verdaderamente su ser ante el poder que lo fundamenta. En su particularidad pueden observarse fuertes rasgos de autonomía moral que se dejan

³¹ Cfr. *Trin.*, X, XI, 17. “...más propio decir que los tiempos son tres: presente de las cosas pasadas, presente de las cosas presentes y presente de las futuras. Porque estas son tres cosas que existen de algún modo en el alma, y fuera de ella yo no veo que existan: presente de cosas pasadas (la memoria), presente de cosas presentes (visión) y presente de cosas futuras (expectación).”

ver en el propio proceso de desarrollo del acto lingüístico. La responsabilidad es entendida como “capacidad de” y “disposición” para responder y aportar razones con el fin de fundamentar acciones y decisiones. El hombre responsable es aquel capaz de deliberar y gestionar su acción atendiendo a su situación concreta con los alcances que despliega. La continua decisión sobre aquello que es conveniente para que la vida humana sea íntegra y acorde al orden de su naturaleza. Para Agustín, la vida consciente, revisada y responsable trae consigo una dulce comunicación interior, que no cesa de acrecentarse. Lo que verdaderamente da sentido es el amor *caritas*, que se une y renueva para vincularse. Justamente es la caridad (*caritas*) la que funda el sentido del lenguaje y da la respuesta fundante de la responsabilidad lingüística y vital. Cualquier sentido de responsabilidad verbal, por tanto, debe estar asociado y fundamentado a la idea de que los seres humanos, en cuanto seres racionales, no podemos renunciar a la competencia lingüístico-dialógica sin lesionar nuestro propio sentido de la existencia. Por ello, al momento de una pronunciación de entradas y argumentos que van en la línea de una ética de la responsabilidad de la acción, Agustín menciona los efectos de la palabra, como también los efectos individuales y colectivos de los actos verbales. Su análisis, se inicia en el descubrimiento de un entendimiento complejo y sentido más hondo del movimiento que genera el signo, para luego incluir el reconocimiento de las esferas de acción personal en la interacción. Ahora bien, como algo indispensable de observar, Agustín incita a reparar en los problemas, las consecuencias y las fricciones del uso colectivo de la palabra, abogando por un cuidado verbal que devendría en una especie de incitación hacia una corresponsabilidad lingüística y comunitaria creciente. Porque si bien la libertad se ajusta al carácter propio de la existencia humana con un fin y un orden en el marco del cual se ha de escribir la historia personal, la escritura de la propia vida no se realiza en soledad, sino en la interrelación permanente, tras la cual es posible visualizar la importancia de la determinación relacional de los sujetos que forman el entramado social. La unidad de responsabilidad no designa enteramente el yo aislado, sino que comprende todas las instancias relacionales con las que dialoga para conjugarlas y examinarlas. La responsabilidad está orgánicamente ligada al acto concreto (proviene de él), pero, al mismo tiempo, está ahí, siempre con un otro, resguardada en un ser con el otro. De modo que con el análisis de los efectos de la sombras y matices de la responsabilidad personal y colectiva en la palabra podría, también, formularse una teoría translingüística —o una lingüística que va más allá del análisis de los elementos formales de la lengua, hacia las relaciones que se establecen—. La realidad del lenguaje como acción, en la versión agustiniana, implica el reconocimiento y registro de la pluralidad de lenguajes sociales y de discursos heterogéneos que se constituyen como medios muy dinámicos y hospitalarios del ser. Todo miembro de la comunidad hablante se enfrenta a la palabra, y no en tanto libre de aspiraciones y valoraciones ajenas, despoblada de voces; sino que la recibe por medio de la voz y escucha del otro, saturada de juicios y memorias atravesadas por el signo. La palabra llega al

contexto del hablante a partir de otros múltiples contextos subyacentes, y colmada de sentidos ajenos; incluso su propio pensamiento la encuentra ya poblada.

Sin embargo, para el de Hipona, una postura y sentido de la responsabilidad adecuado proviene de la posibilidad de autoapropiación del signo implicado y del entendimiento de su dimensión activa. De modo que, cuando se trata de un sujeto colectivo o comunidad, el de Hipona tiende a manifestar que a través del tipo de expresión dominante en la comunidad hablante puede observarse decisivamente el carácter de un pueblo. Esto puede verse en la distinción que realiza en Ciudad de Dios respecto de las dos ciudades fundadas en dos amores distintos (Babel y Jerusalén), pero que a su vez se diferencian por el pecado lingüístico de la soberbia relatado. Así, sus exhortaciones interpelan a todo locutor hacia un sentido de pertenencia de la propia dicción, como también hacia una concientización de los efectos que se juegan en los mecanismos expresivos de una sociedad. Su observación y crítica, por oposición, apuntaría a una gradual comprensión del poder transgeneracional y epidémico de la transmisión verbal a nivel público. Al reconocer los efectos de la palabra en el alma, se vuelve visible también la posibilidad de utilizar su potencia para cooperar activamente en la determinación. En los primeros libros de *De Civ. Dei*, Agustín incita a figurar una progresiva idea de corresponsabilidad o responsabilidad solidaria de las faltas y las bondades de un sujeto moral inmerso en una realidad colectiva. Para ello, el hombre debe corresponder primero al peso de la libertad y a sus efectos, que atañen al intercambio verbal y, por lo tanto, hacerse responsable de desarrollar una lucidez basada en el conocimiento del amor en acción y de la libertad que comprende toda vida. Su perspectiva trata de un desarrollo moral vinculado al ordenamiento del amor y de la tendencia de la interioridad como mecanismo de revisión volitiva, para luego manifestarse exteriormente en la expresión. Así dice que:

“si alguien pronuncia una palabra airada, o maltrata a otro, no podría hacerlo si no se moviera la lengua o la mano al influjo de la voluntad, que en cierto modo manda; y esos miembros, aun sin haber ira alguna, son movidos por la misma voluntad”³².

Esto se debe a que la tendencia de la voluntad determina la formación de intenciones por las que se asumen diversos modos de querer y obrar. La expresión verbal cumple un papel fundamental en esta conversión anímica que reivindica un cierto orden del amor, y el enriquecimiento de dicho orden habilita reflectivamente el mejoramiento del *ethos* y de su expresión. A través de la diversidad de oficios relativos a la lengua, Agustín también emprendió algunas

³² *Civ. Dei.*, 14, 19: 969.

consideraciones acerca del papel que caben a la pedagogía y a la educación en la transformación de la realidad a partir de los condicionamientos presentes en los hábitos lingüísticos de su sociedad.³³ La enseñanza de la lengua y el uso que de ella le preocupan frecuentemente y, por ello, en *De Magistro* establece algunas menciones y competencias que debe tener el docente que imparte esta área de conocimiento, porque indudablemente existe una responsabilidad de creencia, de transmisión, que hiere, corta, desgarrar, engaña o manipula. En la lectura, la escritura y la oralidad, de poco o nada sirven los conocimientos gramaticales sin un espíritu de criterio descriptivo y funcional en todas las dimensiones de resonancia³⁴. Lo que Agustín buscará, ante todo, serán los mecanismos internos que cada alma posee para articular la relación con la propia palabra, que comprometa una cooperación conjunta hacia una transmisión revitalizadora. Se redirige la pauta de acción a través de la transformación del amor y del deseo, como una dimensión que, al transformarse, trastorna todo el conocer, querer y obrar de un sujeto. Mientras que la transformación es incitada a través del signo, el cambio profundo del alma humana conlleva una conversión moral de la persona y de la dirección de sus actos. De modo que la admonición o llamado que performa el signo no solo convierte, sino que también alerta, concientiza, golpea el corazón y lo abre a un estado de vigilia de la acción. Pero ¿cómo es posible tal giro?

La afirmación de la responsabilidad implica ya una conversión interior, un modo distinto de habitar el lenguaje. Se parte de una autoapropiación lingüística como el punto de partida más radical de toda reflexión sobre la acción y su eco. Porque el lenguaje está ontológicamente atravesado por actos y mecanismos por medio de los cuales se incita a asunción de la responsividad. Toda interrogación al respecto evoca exploraciones discursivas que incluyen la exigencia ética de cooperar en la transformación de las condiciones de comunidad que apunten a un *ideal* de transmisión como propósito directivo o creador de su mismo devenir. La voluntad es la causa de sus actos, dice expresamente Agustín en *De Civ. Dei*, V, IX, y a través de sus mecanismos internos es la articuladora y la encargada de elegir de qué modo expresarse, incluso de hacer significativa la captación del sentido. Lo cierto es que la llamada del signo aparece como una vía posible hacia la redirección anímica de un recuerdo profundo del presente humano: su naturaleza vincular y libre a la vez. En el ángulo de esta paradoja se verifica la tensión de una unidad del querer y conocer en la acción. El amor no solo aparece en el acto conativo de la tendencia, sino que también se hace manifiesto en el valor en las funciones intencionales del sentir, en los actos de preferir y postergar, de jerarquizar y psíquicamente ordenar, proferir que señaladamente en el amor conforman una tendencia.

³³ *De doc. Chris.*, I, II.2.

³⁴ *Conf.* IV, II.

Ser autor y locutor implica estar vivo y especulativamente relacionado con la potencia creatriz. Comprender la vida de las palabras en la obra de Agustín es, a la vez, discernir la verdad de una experiencia interna, de un relato, de una narración, de una biografía, tarea que emprende el propio autor en las *Confesiones*. Al escribir sus tratados registra las consecuencias vitales y la respuesta significativa que gestualmente acarrearán. Esto es algo que se relaciona con la advertencia cualitativa de los efectos y consecuencias de cada acción verbal.³⁵ La perspectiva agustiniana funda un sentido de la libertad como relación y límite que subyace a la idea del hacerse a sí mismo, de autodeterminarse a través de la acción. Hay una correlación entre gracia y libertad que constituye y determina el espíritu humano. De esta manera sintetiza Agustín tal imbricación: “La ley se dio para que la gracia pudiera ser buscada; la gracia se dio para que la ley pudiera ser cumplida”³⁶. En la narración bíblica del paraíso que Agustín analiza, se muestra que el comienzo de la libertad es, al mismo tiempo, el nacimiento del límite. Mediante el mandato, Dios muestra al hombre su libertad, a la vez que su límite y, por ello, el despertar de la libertad es, en palabras de Giannini: “despertar como conciencia del límite” (*Vida* 16). Paradójicamente, la prohibición, el límite y la obediencia constituyen al amante que se define a sí mismo por su amor y por elegir tender hacia un objeto, y esto también implica que se elija concomitantemente desatender otros deseos. Determinar una estructura de jerarquía y orden es una de las formas del amor que implica la conciencia de un límite y su enunciación en la vida. La expresión *ordo amoris* agustiniana denota, precisamente, la ordenación que deriva del amor como una llamada a crear desde un reconocimiento y libre asunción por la persona que le requiere comprometerse en un deseo. La realidad afectiva debe una fundamentación que intenta mostrar una dimensión de convergencia entre libertad y condicionamiento que incluye la conciencia de una limitación y atención.

De modo que el sentido de responsabilidad se relaciona a una dimensión tan gnoseológica como estructural de la acción que implica el conocimiento de los límites humanos, y les hace el lugar adecuado en el reconocimiento, porque el modo de su configuración transforma la autocomprensión del sujeto. Tras reconocerse en el poder y en la necesidad limitante de realizar elecciones discursivas y de establecer criterios modales, el locutor se convierte en un sujeto que realiza (por medio del lenguaje) su condición de agente libre, en tanto que dispone de un abanico de opciones diferentes e independientes al momento de expresarse (Meis, 1982). La acción verbal, cualquiera que esta sea, es un acto ético que determina y opera sobre el mundo y sobre la alteridad, y sus efectos son manifestaciones que habilitan a contraer una responsabilidad concreta para con el mundo

³⁵ Cfr. *Civ. Dei*, V, X.

³⁶ *Del espíritu y la letra*, 19,34.

y con el otro. Como un espejo agencial que, incluso desde su limitación, invoca a responder³⁷.



Referencias bibliográficas

- Agamben, Giorgio. *El sacramento del lenguaje*. Arqueología del juramento. Homo sacer II 3. Traducción de Mercedes Ruvituso. Buenos Aires : Adriana Hidalgo Editora, 2010.
- Aubin, Pierre, *Le probleme de la conversion Etude sur un terme commun á l'hellenisme et au christinisme des trois premiers siecles* (Beauchesne et ses fils, Paris 1963), pág.122.
- Arendt, Hannah. *El concepto de amor en San Agustín*. Encuentro, 2009.
- . "Lying in Politics", 1-49. En Arendt, Hannah. *Crises of the Republic*. Hartocut Brace & Company, 1972.
- . *La vida del espíritu*. Centro de Estudios Constitucionales, 1984.
- Apel, Karl Otto. "Primordiale Mitverantwortung. Zur transzendentalpragmatischen Begründung der Diskursethik als Verantwortungsethik". En: Apel/Burckhard, 2001, pp. 97-121.
- Bajtín, Mijaíl. *Hacia una filosofía del acto ético. De los borradores y otros escritos*, 7-81. Barcelona : Anthropos, 1997.
- Bochet, I. *Saint Augustin et le désir de Dieu*. Paris : Institut d'études augustiniennes, 1982.
- Blondel, M. *L'action. Essai d'un critique de la vie et d'une science de la pratique*. 1983, p. 445.
- Chavez, Pamela. "Palabras que liberan. El enseñar como *admonitio* en *De magistro* de Agustín de Hipona". *Tópicos*, 53, 2017.
- Brown, P. *Biografía de San Agustín de Hipona*. (Trad. de Tovar, S & Tovar, M.R). Madrid: Ediciones Castilla, pág. 345,1970.
- Chrétien, Jean Louis. *La llamada y la respuesta*. Caparrós, 2002.
- . *Saint Augustin e les actes de parole*. Epiméthée, 2002.
- . *Le regard de l'amour*. Paris : Desclée de Brouwer, 2000.
- Derrida, Jacques. *Historia de la mentira*. Prolegómenos, 1995.
- Dodaro, Robert. "Eloquent Lies, Just War, and the Politics of Persuasion: Reading Augustine's City of God in a Postmodern World". *Augustinian Studies* vol. 25, 1994.
- Escoto, Juan. *Commentaire sur l'Evangile de Jean*, I, 27. Ed. et trad. E. Jeuneau, Paris, 1972.
- Fernández Beites, Pilar. "Libertad y ordo amoris: una alternativa al existencialismo". En Garcia de Leaniz, Ignacio (ed.). *De nobis silemus: Homneaje a Juan Miguel Palacios*. Encuentro, 2010.
- Giannini, Humberto. *Vida inauténtica y curiosidad*. Departamento de Filosofía Universidad de Chile, 1971.
- Godefroy, Jean. "Mensonge". En Vacant, A ; Mangenot, E. & Amann, E. (ed.). *Dictionnaire de Théologie Catholique*. Letouzey et Ané, X, col. 555-569, Paris, 1928.
- Goma, Francisco. "Scheler y la ética de los valores". En Camps, Victoria (eds.). *Historia de la ética*, vol. 3, Crítica, 2000, pp. 296-326.
- Gramigna, Remo. "La mentira en San Agustín". *DeSigniS*, V.25, 44/56, 2016.
- Griffiths, Paul. *Lying: An Augustinian Theology of Duplicity*. Brazos Press, 2004.
- Horn, Christoph. *Agustín de Hipona. Una introducción*. IES, 2012.
- Magnavacca, Silvia. "El cuidado del amigo". En *La palabra y el silencio. El caso de Alipio*. Humanidades: Revista de la Universidad de Montevideo, n.º 1, junio 2017, pp. 17-28.
- Mann, William. E. "To Catch a Heretic: Augustine on Lying". *Faith and Philosophy* 20/4, 2003.

³⁷ *Civ. Dei*, I, I.

- Meis, Anneleise. "A. La libertad como gracia". En San Agustín. *De Spiritu et Littera*, 1982 : 77-95.
- Merleau-Ponty, Maurice. *Le visible et l'invisible*. Gallimard, 1988.
- Pegueroles, Juan. "Metafísica de San Agustín". En J. Oroz Reta y J. A Galindo Rodrigo (Dir.) *El pensamiento de San Agustín para el hombre de hoy* (I). Edicep, 1998, pp. 274-276.
- Pierantoni, Claudio. "El *verbum cordis* formabile agustiniano y la imagen trinitaria en el hombre". *Teología y Vida* vol. LII, 2011.
- Scheler, Martin. *El resentimiento en la moral*. Caparrós, 1998.
- Schuld, Joyce. *Foucault and Augustine: Reconsidering Power and Love*. University of Notre Dame Press, 2003.
- Svensson, Manfred. (2010) Scientia y sapientia en De Trinitate XII: San Agustín y las formas de la racionalidad. *Teología y vida* [artículo de revista] vol. 51, n.º 1-2 (2010), p. 79-103.
- Watson, Gerard. "St. Augustine's Theory of Language". *The Maynooth Review* vol. 6, 4/20, 1982.



En torno a diferentes usos de la nada: místico y filosófico-político

About Different Uses of Nothingness: Mystical and Philosophical-Political

MARTIN BECKER-LORCA¹

Pontificia Universidad Católica de Chile, Chile
mbeckerl@uc.cl

Fecha de recepción: 15/05/2023

Fecha de aceptación: 17/03/2024

Resumen

Este artículo describe tres usos del concepto de la nada. Primero, en el ámbito de la mística cristiana, analizamos el tratamiento que de esta categoría hace el monje dominico del siglo XIII Meister Eckhart; luego, dentro del ámbito filosófico-político, revisamos los usos que de esta noción hacen el filósofo de la liberación argentino Enrique Dussel y el poeta y teórico cultural norteamericano Fred Moten. Nuestra intención es investigar posibles influencias del empleo de la nada en la mística cristiana sobre el ámbito filosófico-político. Postulamos que en los tres casos se da una dinámica de la nada que concluye con una *inversión*, en la cual se pasa desde la experiencia de la nada como defecto o privación

CÓMO CITAR ESTE ARTÍCULO:

En APA: En torno a diferentes usos de la nada: místico y filosófico-político. *Resonancias*, (17), 31-55. DOI: 10.5354/0719-790X.2024.70672

En MLA: “En torno a diferentes usos de la nada: místico y filosófico-político.” *Resonancias*, n.º 17, julio de 2024, pp. 31-55. DOI: 10.5354/0719-790X.2024.70672

¹ Martin Becker Lorca es doctor en Estudios de Religión, University of California Santa Barbara, profesor adjunto de la Facultad de Teología de la Pontificia Universidad Católica de Chile y profesor miembro del Centro de Estudios de la Religión UC. Sus áreas de investigación son la Filosofía de la religión, la Fenomenología, la teoría de la Secularización y el pensamiento decolonial. <https://orcid.org/0000-0002-2823-3696>.

Palabras clave: ser, nada, Eckhart, Dussel, Moten

Keywords: Being, nothingness, Eckhart, Dussel, Moten

hacia la nada como exceso, es decir, de lo sub-ontológico a lo trans-ontológico. Consideramos que estos tres usos de la nada están en función del rédito que esta inversión posibilita.

Abstract

This article describes three uses of the concept of nothingness. First, in the field of Christian mysticism, I analyze the treatment of this category by the 13th century Dominican monk Meister Eckhart; then, within the philosophical-political sphere, I study the uses of this notion by the Argentine liberation philosopher Enrique Dussel and the North American poet and cultural theorist Fred Moten. My intention is to investigate possible influences of Christian mysticism's use of nothingness on the philosophical-political field. I argue that in all three cases there is a dynamic of nothingness that ends with an *inversion*, in which one passes from the experience of nothingness as a defect or deprivation to nothingness as excess, that is, from the sub-ontological to the trans-ontological. I demonstrate that these three uses of nothing are based on the revenue that this inversion makes possible.

Introducción

Este artículo describe y yuxtapone tres usos diversos del concepto de la nada, con el objetivo de investigar posibles influencias que el uso místico de la nada ha ejercido sobre un uso de tipo filosófico-político. Pero antes de presentar la tesis de este trabajo, y asumiendo la variedad de sentidos y ambigüedad inherente de la categoría de la nada, enmarcamos el siguiente artículo bajo las siguientes premisas generales. De entrada, advertimos que estas premisas forman un esquema simplificado que *nosotros* vamos a emplear para intentar explicar luego ciertos usos del concepto de la nada y que, por tanto, no asumimos que los autores que vamos a tratar en este artículo puedan reconocerse necesariamente en él. Además, si bien al construir este esquema vamos a mencionar ciertas figuras de la historia de la filosofía para ejemplificar alguna de sus partes, con esto no afirmamos que los filósofos que nombremos estarían de acuerdo con el esquema en su totalidad. Proponemos este esquema simplificado, entonces, como una herramienta que nos pueda ayudar a iluminar la paradoja inherente en ciertos discursos sobre la nada que encontramos en épocas y ámbitos diversos.

Frecuentemente el *ser* se ha entendido como fundamento sobre el cual se construye una realidad jerárquicamente ordenada que queda estructurada desde este fundamento. A esta jerarquía (lo fundado) se le ha concebido, a veces,

como una especie de cadena o escalera ontológica². Nos topamos aquí con una cierta ambigüedad: en ocasiones al ser como fundamento se lo entiende como estando más allá de lo fundado; otras veces, al ser se lo identifica con lo fundado —exigiendo, por tanto, un fundamento distinto a él—. Pero ya sea que al *ser* se lo conciba como único fundamento de la realidad (es decir, como fundándose a sí mismo) o como un fundamento que a su vez requiere ser fundado, en ambos casos un orden jerárquico de la realidad queda *ontológicamente* establecido y justificado.

Sobre todo, en el segundo caso, esto es, cuando al ser se le identifica con la jerarquía, al *no-ser* se lo entiende, entonces, como aquello que queda por sobre o por debajo de la jerarquía ontológica: por sobre, como si la fundara o sostuviera desde arriba, pero asumiendo que aquello que sostiene (desde arriba) a lo fundado está más allá de la jerarquía, o sea, más allá del ser. Este es el caso, por ejemplo, de Platón, para quien el Bien (τὸ ἀγαθόν) estaría más allá del ser o de la esencia (“ἐπέκεινα τῆς οὐσίας”) (Platón 6, 509b); o es el caso de Plotino (V, 4, 1), para quien el Uno trasciende al ser. Lo que funda desde arriba, en cuanto es más que el ser, ya no se le puede predicar la categoría de ser. Si bien, ni Platón ni Plotino emplean esta terminología, podemos decir nosotros que el Bien y el Uno corresponden a una cierta nada con respecto al ser, una cierta nada que *excede* al ser. En cambio, por debajo, está aquello que por *defecto* aún no *es* propiamente: una simple nada o algo que está muy próximo a ella. Este es el lugar—o, mejor dicho, el no-lugar—que para algunos de la tradición neoplatónica ocupa la materia (ὕλη); y podemos nombrar aquí también a San Agustín, para quien la condición de posibilidad de la perversión de la naturaleza humana reside en la *nada* desde la cual Dios creó al humano³.

Mencionemos brevemente tres características que definen esta cadena o escalera ontológica. Primero, en un sentido espacial y siguiendo la lógica neoplatónica (simplificada) de las emanaciones⁴, esta escalera ontológica genera una lógica de proximidad y distanciamiento: lo que está más próximo al “techo” (a la fuente de la emanación que se encuentra en lo alto), tendría más ser, sería más

² Una versión de esta jerarquía ontológica, de gran influencia en el medioevo europeo y que se remontaría a la tradición platónica y neoplatónica, es la que se conoce como la “gran cadena del ser” o la *Scala Natura*. Para una historia de esta idea, ver la obra clásica de Lovejoy, *The Great Chain of Being*.

³ “Pero una naturaleza no puede pervertirse por el vicio salvo si ha sido creada de la nada. Y por esto, el ser naturaleza le viene de haber sido creada por Dios, pero el apartarse de lo que es de haber sido hecha de la nada”; “[...] aunque Dios modeló al hombre del polvo de la tierra, la misma tierra y toda la materia terrena fue originada completamente de la nada, y cuando el hombre fue creado, dotó su cuerpo de un alma creada de la nada” (San Agustín 433, 428).

⁴ La noción de emanación traduce los términos griegos *προεῖναι* y *ἀπορρεῖν*, que significan, literalmente, un desbordamiento. Recordemos, sin embargo, que el esquema que estamos describiendo no busca ser una versión fiel de la tradición neoplatónica. Para una versión más compleja y atenta a la lógica apofática de la doctrina de las emanaciones, ver Sells (1994).

real⁵; lo que está, en cambio, más distanciado del “techo”, aquí abajo, tendría menos ser, menos realidad. Segundo, en un sentido temporal, aquello que más es, es lo que más perdura, lo más estable, lo que no cambia en el tiempo y permanece en un presente eterno más allá del tiempo y su devenir. Por el contrario, lo efímero en su constante devenir denotaría una cierta pobreza ontológica propia de aquello que está distante de la fuente. Por último, con respecto a la categoría de cantidad, si bien en la jerarquía ontológica (esto es, en lo fundado) se da una diversidad y multiplicidad de entes (cosas que son), el fundamento en cambio, en su absoluta simplicidad, indiferenciación e indeterminación está más allá de toda multiplicidad.

Ahora bien, en cuanto hay “algo” que escapa a la jerarquía ontológica tanto por arriba como por debajo—siguiendo con la imagen de la cadena o escalera ontológica—, esto implica, entonces, que lo indiferenciado puede ser, por un lado, el Uno, como una *nada* excesiva de arriba que derrama su realidad y da forma al intelecto, al alma y al mundo (para la tradición neoplatónica, por ejemplo), sin que ella (la nada en cuanto lo Uno) tenga forma; y, por el otro lado, lo indiferenciado puede ser la pura materia, como la nada defectuosa de abajo.

Esta ambigua similitud—o en palabras del filósofo italiano Sergio Givone, esta “admirable y terrible proximidad”⁶—entre lo más alto y fundamento de la jerarquía con lo más bajo de la jerarquía, es decir, una similitud entre la indiferencia de la nada excesiva-de-arriba y la indiferencia de la nada carente-de-abajo, es la que, consideramos, otorga una clave de interpretación para ciertos usos de la *nada*. Postulamos, por tanto, que este esquema de una jerarquía ontológica y su correlato, por un lado, una *nada* que en su exceso funda la jerarquía y, por el otro, una *nada* que en su defecto queda afuera de la jerarquía —y la ambigua similitud que estas dos “nadas” implican— puede explicar ciertos usos de la nada, tanto en textos de la mística (cristiana) como también en ciertos discursos que llamamos filosófico-políticos.

En las siguientes tres secciones de este artículo nos dedicamos a describir por separado el uso que tres autores hacen de la categoría de la nada. En la primera

⁵ Dicho en un lenguaje más técnico, lo que está más próximo a la fuente es más plenamente o *participa* en mayor grado del ser.

⁶ En la sección acerca de Plotino del libro *Historia de la nada*, Givone (2010) escribe: “hay que vérselas aquí con una doble meontología, una doble ontología de la nada. La nada es la potencia de lo Uno. Pero es también la potencia (antes bien, la impotencia) de la materia y del no ser. En la vía ‘hacia arriba’, es el principio de la negación a través del cual la forma se libera de esto y de aquello y se constituye como libertad. Por el contrario, en la vía ‘hacia abajo’, la nada es el principio negativo que priva a la forma de ley y medida, despotenciándola, desfigurándola, aniquilándola. Lo Uno y la materia están distanciados de modo abismal, en los polos opuestos de la subida y el descenso, del retorno y el extravío, del epístrofe y de la catástrofe. Pero asimismo están en una admirable y terrible proximidad” (88). Consideramos que esta “proximidad” es la que precisamente inaugura una cierta “lógica” que puede iluminar algunos discursos sobre la nada.

sección describimos los cuatro sentidos de la nada propuestos en un sermón del monje dominico del siglo XIII Meister Eckhart, los cuales forman una especie de escalera en la que se asciende hasta llegar a la nada que “es” Dios. En la segunda sección, vemos como para el filósofo de la liberación argentino Enrique Dussel la nada denota la exterioridad al sistema encarnada en la persona oprimida, en donde confluyen lo trans-ontológico, lo sub-ontológico y lo teológico. De este modo, el rostro de la persona oprimida, aquella que se la excluye como si fuese nada, es el lugar de la epifanía del Otro divino. En la tercera sección, seguimos al poeta y teórico cultural norteamericano Fred Moten en su reflexión acerca la nada absoluta. En el contexto de la dicotomía entre *ser* como anti-negro y la negritud como una *nada* absoluta, Moten invierte la relación clásica entre el ser y la nada, y entonces es el ser el que depende del no-ser y no a la inversa; generando así las condiciones para liberar la experiencia de la negritud (no-ser) del yugo de un mundo anti-negro (ser).

En la conclusión nos dedicamos a comparar estos tres usos de la nada, esbozando una *analogía apofática* entre ellos. Argumentamos que estos autores en sus usos de la nada —a pesar de sus diferencias— asumen el sentido del ser como fundamento de un sistema jerarquizado y, como contraparte del ser, consideran la nada en sus dos acepciones, esto es, como una nada en exceso, más allá del ser y una defectuosa, menos que el ser. Asumiendo la ambigüedad de estas dos “nadas” que este esquema contiene, es esencial para este trabajo el demostrar que en los tres casos se da una lógica o dinámica de la nada que concluye con una *inversión*, en la cual se pasa desde la experiencia de la nada de abajo hacia la nada de arriba, es decir, de lo sub-ontológico a lo trans-ontológico⁷. Gracias a esta inversión se puede postular un cierto beneficio o rédito de la experiencia casi intolerable de la nada. Por lo tanto, a pesar de la diferencia de épocas y lugares, y sobre todo a pesar de la variedad de ámbitos de estos usos de la nada (místico y filosófico-político), postulamos que estas tres descripciones de la experiencia de la nada están en función del rédito que esta inversión posibilita⁸.

⁷ Mencionamos al paso que la valoración de la materia por parte del autor neoplatónico Yámblico en el contexto de la teúrgia, puede ser interpretada como un antecedente antiguo de esta *inversión*, en donde se hace uso de la materia para precisamente acceder al Uno, o sea, se hace uso de la nada-de-abajo para acceder directamente a la nada-de-arriba. Sobre la teúrgia, ver de Yámblico (2003).

⁸ Para revisar ejemplos de una estrategia similar a la de este artículo, de emplear la tradición mística para iluminar discursos políticos contemporáneos, ver, por ejemplo, Laclau (2006). Sobre el uso de la mística con respecto a discursos filosóficos contemporáneos, ver, por ejemplo, Marion (1977), Carlson (1999, 2008), Haas (2009). Sobre una lectura creativa de la mística en relación con discursos filosófico-políticos contemporáneos, dentro de un contexto decolonial, ver, Yountae (2017).

Meister Eckhart: Dios es nada

En un sermón acerca del texto bíblico de Lucas —“Saulo se levantó del suelo y, con los ojos abiertos, nada veía” (Hch, 9,8)— Eckhart comenta:

Me parece que esta palabra tiene cuatro sentidos. Un sentido es éste: cuando se levantó del suelo, con los ojos abiertos, nada veía y esa nada era Dios; puesto que, cuando ve a Dios, lo llama una nada. El segundo [sentido es]: al levantarse, allí no veía nada sino a Dios. El tercero: en todas las cosas nada veía sino a Dios. El cuarto: al ver a Dios veía todas las cosas como una nada (Eckhart 2018 118).

Siguiendo la premisa de una jerarquía ontológica fundada desde más allá del ser, podemos conjeturar que estos distintos sentidos de no-ser —de nada— también pueden ordenarse ascendentemente. Seguimos en este análisis a Duane Williams (2011), quien parte desde el cuarto sentido mencionado por Eckhart y va ascendiendo por esta escalera o “jerarquía ontológica de nada” (*ontological hierarchy of ‘nothings’*—Williams 86⁹) desde la nada más distante de Dios hacia la nada que es Dios.

Pero antes de ascender por medio de esta escalera de cuatro peldaños, hay un sentido de la nada que precede al momento en que la luz divina ciega a Pablo, a saber, cuando él veía las *cosas como cosas*¹⁰. Para poder ver las cosas como cosas ya está operando un cierto *no*, una negación que distingue una cosa de otra. En el contexto de la pregunta acerca de qué es lo que se quema en el infierno, Eckhart propone el siguiente símil:

Toma un carbón encendido y ponlo en mi mano. Si dijera que el carbón me quemó la mano, sería una injusticia. Si dijera la verdad, lo que me quema es una negación, porque el carbón contiene algo que mi mano *no* tiene. Este *no* es lo que me quema [...] Del mismo modo, digo, el hecho de que Dios y todos los que están delante de Su rostro tienen a causa de su verdadera bienaventuranza algo que *no* tienen los que están separados de Dios, este mismo *no* atormenta a las almas en el infierno más que la propia voluntad o cualquier fuego. En verdad digo, en la medida en que él *no* se adhiere a ti, en esa medida eres imperfecto. Por lo tanto, si quieres ser perfecto, debes deshacerte del *no* (Eckhart 2009 109, citado en Williams 87).

⁹ Para todas las citas publicadas en idiomas distintos al español, las traducciones son nuestras.

¹⁰ Tenemos entonces cinco sentidos distintos de la nada. A modo de anticipo, y siguiendo a Williams, los podemos enunciar del siguiente modo: 1) la nada que permite ver cosas como cosas; 2) cosas siendo nada (*things being nothing*); 3) cosas siendo nada sino (sólo) Dios (*things being nothing but God*); 4) nada sino (sólo) Dios (*nothing but God*); 5) Dios como nada (*God as nothing*). Al respecto ver Williams (86).

Lo que daña la mano no es el carbón ardiendo, sino una negación: el carbón tiene algo que mi mano *no* tiene. Si mi mano se deshiciera de ese *no* y fuese como el carbón, ardería sin mayor problema. Pero el hecho de que yo tenga una mano distinta al carbón se debe precisamente a este *no*, a esta negación, que constituye la identidad de mi mano como diferente al carbón. Ambas, la mano y el carbón, contienen algo que la otra *no* contiene. Por tanto, la identidad de algo se constituye diferenciándose de otra cosa, y en esta diferencia se encuentra un “no ser otra cosa” que precisamente la define como algo particular. En este sentido, toda distinción es negación.

Este primer sentido del *no* (que podemos imaginar como el suelo donde se posa la escalera) no se opone al ser, sino que es parte del ser y es la condición de posibilidad de que experimentemos cosas, es decir, de que haya una diversidad de entes en el ser. Por lo tanto, lo no-carbón no hace referencia a una nada absoluta, sino que denota, por ejemplo, a esta mano. Remontando esta acepción o uso de la negación al *Sofista* de Platón, Williams comenta:

El otro, según el pensamiento de Platón, significa que las cosas son tanto no ser como ser, o que son tanto nada como que son algo. Este ser *no* y por tanto *otro*, o ser *otro* y por tanto *no*, da origen a las cosas como cosas. Es en este sentido que todas las nadas son cosas, es decir, cada *no* es *otro* (Williams 88).

Ahora bien, con respecto al primer paso en este camino ascendente de unificación con Dios, según Eckhart, hay que deshacerse de este *no* —de esta negación— que crea la multiplicidad de cosas. Como primer escalón, entonces, cuando Pablo vio la luz divina, él transitó de ver *cosas como cosas* a ver *cosas como nada*; o, según la expresión de Williams, Pablo pasó de ver “toda nada como cosa” (esto es, la nada que nos permite distinguir cosas como cosas) a ver “toda cosa como nada” (Williams 88). Para comprender este primer peldaño, consideramos útil adentrarnos en la lógica (neoplatónica) que conlleva comprender a Dios como Uno, y a la práctica humana que le corresponde.

El Uno en su simpleza es pura indistinción. En cuanto en las cosas particulares hay un *no* que las distingue del Uno, entonces el poder ser uno con el Uno implica deshacerse de ese *no* que nos distingue como cosas particulares y que nos “aleja” del Uno¹¹. Si el *no* (de la primera negación) nos hace ser algo, entonces el camino hacia Dios como Uno implica dejar de ser algo y convertirse en nada por medio de una segunda negación que niega la primera.

¹¹ En cuanto nunca dejamos de estar en el Uno, este “alejarse” no debe ser entendido de modo espacial (como estando afuera y lejos del Uno), sino que es un alejarse de la experiencia de ser uno con el Uno.

Nos introducimos aquí en la lógica o dinámica del “ser separado” (*abgescheidenheit*) de todo *esto* o *aquello*; lo que se busca es devenir nada, llegar a ser indistinto en el Uno, vaciarse—en cuanto estar vaciado de toda cosa implica estar lleno de Dios—. “Y debes saber”, comenta Eckhart, “estar vacío de todas las criaturas es estar lleno de Dios, y estar lleno de todas las criaturas es estar vacío de Dios” (Eckhart 2018 171). Pero más que nuestra renuncia voluntaria a desear cualquier cosa, esta lógica requiere entregarnos y abandonarnos a Dios, para que Dios vaya deshaciendo ese *no* que nos distingue y separa de la indistinción divina.

Ahora bien, en cuanto fuera de Dios no hay nada, esto implica que solo Dios es. Si solo Dios realmente es, entonces todas las cosas en sí mismas son nada¹². Por tanto, cuando Pablo vio a Dios desde la unidad de Dios, vio todas las cosas como nada. Según Eckhart, Pablo “veía a todas las criaturas como una nada, pues él [Dios] tiene en sí el ser de todas las criaturas. Es un ser que tiene en sí a todos los seres” (Eckhart 2018 123). Esto no significa que Pablo no vio cosas; sino que precisamente las vio como nada en relación a Dios; es decir, como no autosuficientes, como absolutamente dependientes de Dios.

En síntesis, hasta ahora hemos explorado dos sentidos relativos de la nada. Williams explica:

El primer *no* es relativo, en cuanto que es el fundamento relacional de las cosas con respecto a las cosas. El segundo *no* es también relativo, pero es el fundamento relacional de las cosas con respecto a Dios. En la primera instancia, las cosas *no* son en relación con otras cosas. En la segunda instancia, las cosas *no* son en relación con Dios. Este *no* impide que las cosas sean cosas. Por lo tanto, todas las cosas se ven como nada (Williams 90).

Ascendiendo al segundo escalón, cuando Pablo “en todas las cosas nada veía sino a Dios”, la nada se refiere ahora a Dios como el “negador” de la negación de la negación. Si en la formulación del primer escalón (“veía todas las cosas como una nada”) se enfatiza la negación de la negación diferenciadora que nos permite ver todo como relativo y dependiente al Uno divino, en este segundo paso, la nada se refiere a Dios como el que niega la negación de la negación. Williams lo formula de la siguiente manera: “en todas las negadas cosas negadas, sólo se veía

¹² El poeta, escritor, sacerdote y revolucionario nicaragüense Ernesto Cardenal sintetiza esta enseñanza: “Si llamamos existencia a la que tienen todas las cosas, Dios no existe. Y si llamamos existencia a la de Dios, entonces ninguna otra cosa existe. Sencillamente es tan diferente de todo cuanto existe que es como si no existiera. O bien, si Él existe, todo lo demás es nada ante Él. En cierto sentido, pues, Dios no existe, y en cierto sentido sólo Él existe” (Cardenal 53).

al negador” (*in all negating things negated, only the negater was seen*) o “en todas las nadas nadeadas, sólo se veía al nadador” (*in all nothings nothinged, only the nothings was seen—92*). Si el panteísmo ve lo absoluto en lo relativo y entiende que “Dios está en todo”, en este segundo escalón Eckhart estaría diciendo, según Williams, lo contrario: ver lo relativo en lo absoluto, o sea, “todo está en Dios” (Williams 91).

En el tercer escalón, cuando Pablo “al levantarse, allí no veía nada sino a Dios”, ya no se ven cosas (ni *en ellas* al negador de la negación de las cosas negadas), sino que tan solo se ve al negador—ya no lo relativo visto en el absoluto, sino que solo lo absoluto—. Según Williams, entre el segundo y el tercer paso se articulan dos sentidos de la nada que han dividido en bandos opuestos la historia de la interpretación de la nada: por un lado, la nada como un rechazar relativo (*relative rejection*, por ejemplo, el segundo sentido de la nada cuando se ven las cosas como nada sin [*without*] Dios) y, por el otro lado, la nada como un negar absoluto (*absolute denial*, por ejemplo, el tercer sentido de la nada cuando se ve nada además [*besides*] de Dios). El idioma griego reconoce esta distinción sutil, y marca estos dos significados de la nada empleando dos palabras distintas: “*me on*” denota un rechazo relativo, subjetivo, que niega una voluntad o un pensamiento; “*ouk on*”, en cambio, se refiere a una negación absoluta, objetiva, en donde lo que se niega es un hecho o una afirmación (Williams 97). Según Williams, los filósofos analíticos frecuentemente entienden la nada como negación absoluta; el pensamiento religioso y poético, en cambio, la entiende como un rechazo relativo. Anticipamos la importancia de esta distinción —entre nada relativa y absoluta— para la comparación de los distintos usos de la nada que llevaremos a cabo más adelante (habrá, sin embargo, que permitir que el sentido y valoración de estas dos nadas pueda cambiar). Por ahora es relevante constatar que para Eckhart, según Williams, la clave de interpretación reside en la nada relativa.

Por último, cuando “Pablo se levantó del suelo, con los ojos abiertos, nada veía y esa nada era Dios”, nos encontramos con la enseñanza de Eckhart, repartida en varios sermones, de *Dios como nada*. Por ejemplo, el sermón 54 dice: “Pero si Dios no es bondad, ni ser, ni verdad, ni uno, entonces, ¿qué es Él? Es pura nada: no es ni esto ni aquello. Si piensas en algo que Él podría ser, Él no es eso” (Eckhart 2009 287). Volviendo a la experiencia de Pablo, según el análisis de Williams, ahora el negador es a su vez negado. Nos encontramos con una triple negación: la negación del negador que niega las cosas negadas. Williams expresa toda la secuencia de la visión de Pablo en esta fórmula densa: “las negaciones son negadas por el negador negado, o las nadas son nadeadas por la nada nadeada” (*the negations are negated by the negater negated, or the nothings are nothinged by the nothings nothinged—93*).

Por su puesto que cuando Dios es nada, el sentido de la nada no es el de una negación absoluta de Dios, sino de una nada relativa; lo que se está negando es que “Dios sea algo”, un “esto” o un “aquello”. Williams aclara: “Eckhart no se refiere a ‘nada’ como la nulidad total de la aniquilación. Si bien Dios es ‘nada’, es decir, ‘no-cosa’ [*no-thning*], Dios, sin embargo, es—como Nada” (Williams 94). Entonces, Dios es nada, siempre cuando esa “nada” no se la interprete como *nihil negativum*. Dios no es ni algo ni absolutamente nada¹³.

De hecho, en este rechazo relativo de Dios hay que ver una exaltación. Esta “nada” final, en la cúspide de la jerarquía, es una afirmación de lo que está más allá de las palabras. En el sermón 62, Eckhart escribe: “Dios es nada: no en el sentido de no tener ser. No es *esto* ni *aquello* de lo que se puede hablar: es el ser por encima del ser. Él es un ser sin ser” (Eckhart 2009 317). Dios es nada de los que es, ya que es fundamento (*Grund*) del ser y, por tanto, en su exceso, está por encima del ser. La nada se entiende aquí como una nada relativa, que denota lo que tiene más ser que el ser.

En la cúspide de la escalera, el proceso de “nadeamiento” del alma se correlaciona con Dios como nada: “Cuando el alma llega a lo uno y allí entra en un rechazo puro de sí misma, encuentra a Dios como en una nada” (Eckhart 2018 123). La nada antropológica se condice con la nada divina y, en la nada, la unidad sin diferencia deviene posible.

En síntesis, en esta escalera ascendente encontramos los siguiente sentidos de la nada: como suelo donde se posa la escalera, la nada se refiere a un *no* o negación en toda cosa que la determina como algo particular distinto a otra cosa; como primer escalón, la nada se refiere a la negación de la negación diferenciadora, se refiere, por tanto, al acto de negación en la que solo Dios queda, o dicho de otra forma, solo Dios es, y toda cosa es, en sí misma, una nada en relación a Dios; segundo paso, la nada se refiere a que en la negación de las cosas negadas, solo se ve el negador; cuarto, solo se ve al negador sin las cosas; quinto, el negador es a su vez negado, Dios es nada de lo que es, ya que es fundamento del ser y, por tanto, en su exceso, está más allá del ser.

Nos hemos movido desde la nada que nos aleja de la simplicidad del Uno, hacia la nada que “es” el Uno más allá del ser. Siguiendo nuestra imagen de la escalera, hemos partido desde la nada de abajo, deficiente en su ser, hacia la nada de arriba que, si bien sostiene a todos los entes, en su exceso, está más allá del ser,

¹³ En este punto nos topamos con la queja de los analíticos: “Dios es nada o algo, Él no puede, lógicamente hablando, ser ambos” (Williams 95). Pero aquí nos sirve la doble negación de la teología apofática, en la que se niega tanto la afirmación como la negación: Dios no es algo ni tampoco nada. Con esta doble negación se supera la naturaleza proposicional del lenguaje y este deviene oración o himno. Es decir, más que querer decir algo *sobre* Dios, las palabras se emplean para hablar *a* Dios (ver Marion 1999).

esto es, “es” nada. Es a este movimiento desde la nada de abajo hacia la de arriba que deseamos llamar una *inversión*. Si bien esta palabra no la emplea Eckhart, y desde la experiencia del alma sería más correcto hablar de una confluencia, fusión o unidad entre el alma y Dios, consideramos que la palabra inversión sirve para enfatizar el movimiento (siguiendo la imagen propuesta de una escalera) de abajo hacia arriba¹⁴. Más que entrar en las discusiones técnicas de la validez de una unión mística sin diferencias, nos interesa acentuar el cambio de posición: el pasar de estar fuera de la jerarquía ontológica por defecto a estar fuera de ella por exceso y, con ello, constatar el cambio de valoración y las posibilidades “positivas” que la posición final permite en comparación a la posición inicial.

En las siguientes dos secciones, nos queda ahora investigar posibles influencias de este complejo y rico uso de la ambigüedad constitutiva del concepto de nada propia de esta corriente del misticismo cristiano medieval en el ámbito político-filosófico contemporáneo.

Enrique Dussel: la nada como liberación

Quisiéramos comentar, al menos, dos usos relevantes de la nada en la filosofía de la liberación de Enrique Dussel. Primero, en cuanto el sistema (o el mundo) se le identifica con el ser, surge entonces una *nada* que marca una exterioridad al sistema totalizante. Segundo, interpretando la idea cristiana de la *creatio ex nihilo*, para Dussel la nada (*nihil*) denota la contingencia del mundo creado, posibilitando el proyecto revolucionario de transformación de la realidad. En ambos casos, la nada cumple una función que podríamos llamar “positiva” ya que desde la nada como exterioridad o como contingencia se pueden cuestionar la legitimidad del sistema para cambiarlo. Dentro de los límites de este artículo, tan solo describiremos brevemente el segundo uso y nos abocaremos al primero.

Según Dussel, la nada de la *creatio ex nihilo* hace referencia a la libertad incondicionada del Absoluto (Dussel 2011 163), el cual libremente crea pero permaneciendo más allá de su creación; libertad que es como una “nada” desde el punto de vista de lo creado: “Detrás de la obra humana se encuentra, como la nada de mundo, la libertad pro-ductora del Otro; detrás del cosmos como totalidad,

¹⁴ Podría dar la impresión de que la noción de *ascensión* (o retorno) capta mejor este movimiento de abajo hacia arriba. Pero aquí no nos referimos a un ascenso gradual, en el cual uno va transitando por realidades cada vez más sublimes trepando por la jerarquía ontológica, sino que, por el contrario, es como si en el suelo o fondo de la jerarquía hubiese—por alguna extraña razón—una conexión directa con la cúspide de la jerarquía. Es en este sentido que, de pronto, la nada defectuosa de abajo se *invierte* y deviene la nada excesiva y sublime de arriba.

desde la nada originaria, se encuentra la Libertad creante del Otro absolutamente absoluto” (Dussel 1980 69-70). La libertad incondicionada del creador la experimentamos como la contingencia de la creación. De hecho, el postular una cierta necesidad del modo de darse de la creación sería limitar la libertad productora del creador. Contra la tendencia divinizadora del sistema, que en cuanto divino quiere eternizar el *statu quo* del sistema, la nada de la *creatio ex nihilo* nos invita a cuestionar la pretendida estabilidad del sistema, y entender la creación como contingente y por tanto transformable:

La teoría meta-física de la creación es la garantía teórica de la revolución liberadora; es la formulación más acabada de que ningún sistema es eterno, porque todo, aun el Sol y la Tierra, es contingente (puede no ser) y posible (en su tiempo no fue) (Dussel 2011 162).

Pasemos al segundo sentido. Como recién mencionamos, para Dussel todo sistema político tiende a totalizarse y a divinizar el orden político del poder. Esta tendencia del sistema de concebirse como una totalidad divina implica la negación o cancelación de la alteridad. Una vez que el sistema se ha divinizado, es posible oprimir al hermano o hermana en nombre precisamente del derecho divino; o, a la inversa, la persona que oprime a alguien se absolutiza a sí misma como parte de un sistema divinizado: “por una lógica infalible, el que niega a su hermano se afirma como único, como Señor, como dominador” (Dussel 1980 131).

Pero para poder visibilizar esta correlación entre concebirse como totalidad y negar la alteridad (o viceversa), se necesita acceder a un lugar de exterioridad al sistema desde donde, precisamente, la pretensión de totalidad y la negación de la alteridad se hagan patentes.

Gracias a sus lecturas del filósofo lituano-francés Emmanuel Levinas, Dussel articula la posibilidad de una exterioridad al sistema, encarnada esta en las personas oprimidas, pobres, huérfanas, discriminadas sexual y racialmente. Cuando somos confrontados cara-a-cara a estas personas, esta experiencia nos impone una responsabilidad irrenunciable de transformar la situación de opresión y exclusión que les aflige.

Para Dussel, siguiendo a Levinas, el ser y la ontología se identifican con el sistema que tiende a absolutizarse y que reproduce la lógica imperial de la dominación: “...el discurso ontológico es la justificación teórica de la praxis dominadora” (Dussel 1980 67). Ya desde Heráclito algunos filósofos han considerado al *ser* mismo como guerra. Al respecto, el filósofo puertorriqueño Nelson Maldonado-Torres comenta: “La ontología se convierte así en una racionalización del mundo de la guerra en el que reina la política del imperio” (2008 180).

Por lo tanto, la exterioridad al sistema de dominación que se revela en el rostro del oprimido apunta más allá del ser. En contraste a la diferencia ontológica, esto es, la diferencia entre el ente y el ser del ente, habría una diferencia trans-ontológica (nombre que le da Maldonado-Torres) que se refiere a la diferencia entre el ser y la exterioridad al ser—exterioridad que Levinas y Dussel llaman lo *metafísico*—.

Al entender la exterioridad metafísica como exclusión, según la interpretación de Maldonado-Torres, Dussel estaría confundiendo los niveles trans- y sub-ontológicos, y habría transferido categorías que en Levinas funcionan a un nivel de su fenomenología genética (que constituyen al sujeto) al nivel de la experiencia *del* sujeto:

Dussel (con)funde aquí el “más allá del Ser” con el no-ser. En nuestros propios términos, Dussel fusiona la diferencia trans-ontológica y la sub-ontológica. La diferencia sub-ontológica se refiere a la distinción entre el Ser y aquello a lo que se niega su propio ser. Más precisamente, destaca la apariencia del no-ser / no-Otro, o lo que yo llamaría el sub-alterno. El sub-alterno no es el Otro qua Otro, ni siquiera el alter ego —el Otro como yo mismo— sino ese “otro” que es no-Otro; no es el irremplazable y amado, o el otro reemplazable y respetado, sino el Otro eliminable [...] El sub-alterno es el ser humano concreto que es representado como menos que una “cosa” o como un animal (Maldonado-Torres 2008 182).

Pero más aún, en la exterioridad del oprimido, no solo confluyen lo trans- y lo sub-ontológico, sino que también lo *teológico*. El encuentro con la persona oprimida puede desenmascarar la idolatría del sistema, en cuanto el rostro del oprimido es el lugar de la epifanía del Otro divino. Es en el contraste con la divinidad revelada en el rostro sufriente del oprimido que la incongruencia e injusticia del sistema supuestamente divino se hace manifiesta. En este encuentro surge la religión en su sentido metafísico. Dussel escribe:

la re-ligión, en su sentido meta-físico, no es conocimiento ni saber (como para las ontologías), sino proximidad real: cara-a-cara, presentida, previvida, pregustada en el rostro del Otro. El infinito se avanza en sus huellas no son meros fenómenos ópticos sino *epifanía* meta-física: el rostro de la viuda (mujer pobre), del huérfano (niño suplicante), del pobre y extranjero (el hermano sin poder). En el rostro del pobre (no del igual *en* el sistema, sino del *exterior* al poder del sistema), en sus ojos que interpelan, es donde se encuentra la enigmática y misteriosa epifanía (1980 56).

El Dios que se muestra en los ojos que interpelan de la persona oprimida, no es legitimación del sistema, sino que denota aquella promesa escatológica que abre una exterioridad al sistema desde donde se le des-diviniza y, por tanto, posibilita su transformación.

El evaluar la legitimidad de esta interpretación de Dussel¹⁵ (o la crítica que Maldonado-Torres hace de su lectura a Levinas) trasciende los límites de este artículo. Nos interesa tan solo notar el sentido de esta nada. Paradójicamente, la “nada” significa algo: la exterioridad al sistema, que desde la perspectiva del sistema se considera una nada. Este espacio metafísico, en cuanto más allá del ser, es una *nada* en la que se confunde lo menos-que-ser (un otro indiferente y eliminable) con lo más-allá-que-ser (lo divino) o, mejor dicho, en el rostro de la nada menos-que-ser se revela la nada-más-que-ser; y en esta (con) fusión de planos de la nada, se esconde su potencial liberador.

Ya habiendo explicado algunos términos, concluimos esta sección con una cita densa de Dussel, donde él explicita el uso de la nada que hemos expuesto hasta ahora:

Negada la divinidad del fetiche, ¿qué es lo que hay *más allá* del horizonte del sistema? *Más allá* del ser de la ontología, del límite de la Totalidad establecida no hay *nada* del sistema. La *nada*, negativamente, no es algo del sistema: es lo que está *más allá* (meta) y *más arriba* (anó). La *nada*, afirmativamente, es lo metafísico: el Otro que el sistema, pobre *en* el sistema, ‘nada’ como otro con respecto al sistema mismo [...] Lo que está más allá que la Totalidad, *nada*, es lo *real* más allá del ser: lo misterioso, lo místico, el Otro. Por ello, *la condición práctica de la posibilidad de afirmar teóricamente al Otro es ser ateo del sistema, de la Totalidad vigente*, del dios-fetiche. De lo que se trata ahora es de saber internarse en la *nada*, en el “no-ente” (*oúk ón*), realmente, prácticamente. Es decir, debemos describir cuáles son las condiciones de posibilidad práctica del acceso a la realidad del Otro (Dussel 1980 51).

¿Es posible relacionar el sujeto místico—del tipo descrito por Meister Eckhart—que deviene nada para ser llenado por Dios con el sujeto sub-alterno oprimido que, en cuanto una nada desde la perspectiva del sistema totalizado, deviene el lugar de la epifanía del Dios absolutamente Otro? La referencia de Dussel a “lo místico” parece señalar una continuidad entre su uso de la nada y la tradición mística cristiana ejemplificada en la sección anterior por Eckhart.

En la versión de Dussel de la lógica de la nada, el momento que hemos llamado una *inversión* se puede interpretar—siguiendo a Maldonado-Torres—como una (con) fusión de lo menos-que-ser con lo más-allá-que-ser, la cual, al mismo tiempo, según Dussel, se identifica con lo divino. En el rostro de la persona

¹⁵ Interpretación que contempla, al menos, los siguientes puntos: primero, identificar el sistema con el ser; segundo, entender el ser como una versión ideologizada de la lógica de dominación imperial de la guerra; tercero, considerar a los oprimidos como exterioridad del sistema, más allá del ser y lugar de revelación del absolutamente Otro divino; cuarto, desde la exterioridad poder relativizar la pretensión absolutizante y develar como fetiche al sistema divinizado.

oprimida, es decir, en la nada-menos-que-ser, se revela la nada-más-que-ser; y en la incapacidad de decidir entre fusión o confusión de planos que ocurre en la experiencia de la nada, se esconde precisamente su potencial liberador. Liberación que implica convertirse, en nombre del absoluto divino revelado en el rostro sufriente de la persona oprimida, en un ateo al sistema divinizado. Este gesto de ir más allá de Dios en nombre de Dios nos recuerda el famoso rezo de Eckhart: “Por eso rogamos a Dios que nos vacíe de Dios” (2008 108)¹⁶.

Ahora bien, en contra de la interpretación de Eckhart propuesta por Williams, la cual considera que el sentido más elevado de la nada —esto es, la “nada que es Dios”— sería un caso de una nada relativa, la referencia de Dussel al *oûk ón* sugiere una valoración positiva de la nada absoluta. Es con miras a iluminar la lógica de la nada desde la diferencia entre la nada relativa y la absoluta, que pasamos ahora al uso de la nada propuesto por Fred Moten.

Fred Moten: misticismo de la nada absoluta en la carne

Dentro de los Estudios negros (*black studies*) ha habido un giro hacia la ontología, el cual busca acceder a la fuente última de la violencia anti-negra del mundo moderno. Sin embargo, en cuanto se considera al *ser* como anti-negro, más que una ontología, estos discursos desean articular una paraontología, esto es, un horizonte trascendental que logre escapar de la ontología clásica, considerada fundamento del mundo anti-negro. Dentro de los límites de este artículo queremos tan solo esbozar—superficialmente—algo del uso que el pensador y poeta Fred Moten hace de la categoría de la nada dentro de un debate al interior del pensamiento radical negro entre el afro-pesimismo y el optimismo negro —siendo Moten un representante de este último—.

Lo que los afro-pesimistas y optimistas negros comparten es una comprensión profunda de la relación entre la negritud (*blackness*) y la ontología (Warren 222). Ambos interpretan el abismo insalvable entre la negritud y la humanidad/blanca anti-negra en términos de la dicotomía mutuamente excluyente entre la nada y el ser.

¹⁶ “*Darum bitten wir Gott, dass wir ‘Gottes’ ledig werden*” (Eckhart 2020 555). Como lo nota Sells (1), es relevante constatar como habitualmente los editores (como en este caso) le agregan comillas a la segunda mención a “Dios”, queriendo con ello distinguirlo de la primera. Pero con esta interpolación al texto original se pierde precisamente el gesto apofático del discurso de Eckhart, en donde hasta el nombre de Dios puede ser un obstáculo para ser uno con Dios.

Parecido a Dussel, los afro-pesimistas desmitifican la ontología, despojándola de su supuesta “pureza” y neutralidad defendida por la tradición occidental, y exponen la ontología como el producto de procesos políticos¹⁷. La ontología no precede a lo político; y el creer que sí lo hace, es a menudo la base de una lógica emancipadora defectuosa (Warren 223).

Fred Moten comienza su reflexión “optimista” desde el no-lugar de la bodega del barco de esclavos:

Buscaré comenzar a explorar no solo la ausencia sino también el rechazo del punto de vista [*standpoint*], para explorar realmente y habitar y pensar lo que Bryan Wagner (2009: 1) llama “existencia sin posición” [*existence without standing*] desde ningún punto de vista porque esto es lo que significaría realmente permanecer en la bodega del barco [...] ¿Qué sería, más profundamente aún, qué es, pensar desde ningún punto de vista; pensar fuera del deseo de un punto de vista? (Moten 738).

Si bien la filosofía de la liberación y el giro decolonial han reivindicado el lugar de enunciación de todo discurso para cuestionar la supuesta universalidad—y la *hybris* del punto cero—de la filosofía europea¹⁸, para el pensamiento negro, en cambio, ni siquiera el privilegio del *lugar* del “lugar de enunciación” es posible. Pero Moten no solo padece la ausencia de punto de vista, sino que además lo rechaza para poder pensar la condición del negro desde ese no-lugar de la bodega del barco en algún lugar incierto entre África y América. Es en esta conexión entre negritud y esclavitud, en la imposibilidad de constituirse en sujeto humano¹⁹ en un mundo que considera la negritud como “mercancía en forma corpórea” (Warren 223), donde Moten se refiere a la nada:

¹⁷ Tanto Dussel como la teoría radical negra han sido influenciados por el pensamiento de Franz Fanon. Para Fanon, en un mundo colonial y blanco, “el negro no es un hombre. Hay una zona de no-ser, una región extraordinariamente estéril y árida, una rampa esencialmente despojada, desde la que puede nacer un auténtico surgimiento. En la mayoría de los casos, el negro no ha tenido la suerte de hacer esa bajada a los verdaderos infiernos” (42). ¿Pero cómo puede la zona del no-ser convertirse en un auténtico surgimiento? ¿Cómo se puede obtener algún beneficio de ser condenado al infierno? ¿Estará aquí operando la economía de la lógica de la nada donde al final se da una cierta inversión “productiva” y beneficiosa?

¹⁸ Con respecto a la “*hybris* del punto cero”, ver Castro-Gómez (2010). Dussel (2017) explica la *hybris* del punto cero de la siguiente forma: “En el Renacimiento los pintores trazaban primeramente la línea del horizonte y fugaban hacia un punto todos los objetos que pintarían, Ese ‘punto cero’ será, inadvertidamente, el ojo mismo del pintor, pero como su negativo. El pintor no aparece en el cuadro, pero el cuadro es orquestado desde ese punto de fuga, no visto como el ojo del pintor, pero que en realidad está omnipresente en la obra. Es la ausencia del *ego cogito* de todo *cogitatum*. El *cogitatum*, evidentemente, es el ser-colonial” (268, nota 22).

¹⁹ La imposibilidad de constituirse en sujeto humano la podemos entender en relación a dos nociones kantianas. Al negro se le excluye de ser parte del *sujeto transcendental* y, con ello, se le niega una posición en el campo político, relegándolo a ser una masa social indiferenciada, condenado a una cierta muerte social en vida. Y además la negritud “proporciona la violación permanente del imperativo categórico de Kant: la negritud es puro medio [*pure means*]” (Warren 223).

Si el esclavo es, al fin y al cabo, nada [*nothing*], lo que queda es la necesidad de una investigación de esa nada [*nothingness*]. ¿Qué es la nada, es decir, la negritud, del esclavo que no es reducible a lo que hizo, aunque lo que hizo es irreducible en él? (Moten 744).

Es en la experiencia de la esclavitud en donde confluyen estas tres preguntas: “¿Qué es la nada? ¿Qué es la cosidad [*thingliness*]? ¿Qué es la negritud? (Moten 742). Para pensar esta convergencia, Moten emplea la distinción del filósofo japonés Nishida Kitarō entre una nada relativa y otra absoluta. Moten emplea el concepto budista de *Mu* desarrollado por Kitarō, esto es, una “subjetividad vacía” (743), para comprender la nada absoluta y así poder pensar la negritud. Moten interpreta, yendo más allá que Kitarō, la práctica de *Mu* como un “misticismo de la carne” (*Mu is a practice of mysticism in the flesh—753*). Es la experiencia de cosidad radical la que lleva a este misticismo de la nada absoluta en la carne.

Ahora bien, la posibilidad de una paraontología depende de pensar la interacción entre negritud, cosidad y la *nada absoluta*. De hecho, la diferencia entre el afro-pesimismo y el optimismo negro, reside, para Moten, en que el afro-pesimismo se mueve al nivel de la nada relativa, en cambio él—como representante del optimismo negro— quiere acceder a la radicalidad de la nada absoluta. Es decir, su optimismo surge desde la nada absoluta, desde una radicalización del horizonte (relativo) del afro-pesimismo. Intentemos distinguir estas dos nadas.

Implicito en la nada relativa está la oposición entre la “nada y algo, donde la nada se entiende simplemente como un modo para velar (como si fuera una librea epidérmica) algún ser superior y, por lo tanto, es relativa en lugar de absoluta” (Moten 755). La nada relativa es *relativa* en cuanto permanece subordinada al ser. La nada (relativa) entendida, por ejemplo, como “ausencia” o “vacío” *de algo*, queda subordinada precisamente al horizonte de la presencia de ese algo ausente.

Por ejemplo, sobre la dependencia de la nada al ser, Paul Tillich escribe: “El no ser depende del ser que niega [...] No podría haber negación si no hubiera una afirmación precedente para ser negada” (2000, 40)²⁰. Si bien concedemos que la nada en cuanto negación queda sujeta a aquello que niega, preguntamos, ¿es la

²⁰ Tillich es consciente de la posibilidad de pensar una nada independiente al ser (por ejemplo, el *Urgrund* y el *Ungrund* de Schelling). De hecho, las dos palabras griegas que significan nada, ya permitirían pensar esta posibilidad: “El *ouk on* es la ‘nada’ que carece de toda relación con el ser; el *me on*, en cambio, es la ‘nada’ que tiene una relación dialéctica con el ser” (1972 244). Para Tillich, el seguir el camino del *ouk on* implicaría una renuncia a pensar la nada en su misterio: “La tentativa ontológica de rehuir el misterio del non-ser adopta la estrategia de intentar despojarlo de su carácter dialéctico. Si se sitúa el ser y la nada en un contraste absoluto, el non-ser queda excluido del ser bajo todos los aspectos; todo queda excluido (243). Contra Tillich, el proyecto de Moten ve en el pensar la nada absoluta una posibilidad de liberar la negritud del yugo de un ser-anti-negro.

nada siempre consecuencia de una negación? O podría ser al revés, ¿que la negación sea una consecuencia de la nada?²¹ Si es así, la nada no sería dependiente de aquello que niega, sería, por tanto, una nada no subordinada al ser.

Según Moten, la nada absoluta es “una nada sin reservas, independiente del deseo de presentarse en y para la óptica convencional en la que alguien es delineado e identificado” (Moten 761). Si bien la nada imaginada como lo oscuro está subordinada a la luz, Moten se pregunta, “en ausencia de lo que se toma por luz, en ausencia del pensamiento, del esquema, que se llama cuerpo, ¿cómo describimos la nada extraordinaria, o absoluta?” (772).

Por tanto, a diferencia de los afropesimistas que ven la negritud como un estatus ontológico que es introducido, o incluso producido, por la anti-negritud moderna, Moten argumenta que “la negritud es ontológicamente anterior al poder logístico y regulador que se supone que la trajo a la existencia” (739). De hecho, para Moten, la negritud no solo es anterior a la ontología de la esclavitud moderna, sino que es anterior a toda ontología. Si no se accede a esta negritud anterior a toda ontología, la negritud queda determinada por los prejuicios de un ser anti-negro; es decir, la negritud queda relegada a una nada relativa. No basta con describir la condición de no-ser en la que es condenada la negritud por la ontología anti-negra, ya que con esto se está re-instalando la mirada normativa anti-negra, subordinando el no-ser al ser.

Con Moten estamos complejizando la distinción propuesta por Williams (vista en la sección sobre Eckhart) entre un rechazar relativo y un negar absoluto. Si—según Williams—la posibilidad de entender a “Dios como nada” depende de interpretar la nada de modo relativo, Moten nos desafía a explorar la nada absoluta como un modo de pensar la negritud más allá de constituir ella una “falta”, “ausencia”, “privación”, “vacío”, pero también más allá de constituir ella un “exceso”—ya que tanto la privación como el exceso quedan dentro del horizonte del ser (anti-negro)—. Si para Williams el discurso místico y poético piensa la nada como relativa y la filosofía analítica tacha de absurdo cualquier discurso sobre la nada al pensarla como absoluta, Moten, en cambio, nos exige pensar desde una poética mística de la carne, precisamente, este absurdo de una nada absoluta.

La clave interpretativa está en comprender que la *nada* no solo puede decirse de un modo relativo o absoluto, sino que, además, cada uno de estos modos de decir la nada puede tener a su vez sentidos distintos: la nada relativa puede referirse a una “realidad dependiente del ser” o a una “realidad distinta al ser”; la nada absoluta, a su vez, puede significar una “realidad independiente del ser” o una “negación de toda realidad”. Por ejemplo, con respecto a la nada relativa,

²¹ Sobre una concepción de la nada que no es consecuencia de una negación, ver Heidegger 1976.

en el primer escalón en la interpretación del sermón de Eckhart, en cuanto solo Dios es, las cosas “son” vistas como nada, o sea, las cosas son nada en cuanto *dependientes del ser* (o dependientes de Dios como fuente del ser); en cambio, en el cuarto escalón, ahora Dios mismo es nada, en el sentido de que Dios es *distinto al ser* (“está” más allá del ser).

Asumiendo que el sentido de la nada relativa—a causa de su ambigüedad— puede quedar subordinado al ser, hace que Moten reivindique el uso de la nada absoluta; no en el sentido de una “negación de toda realidad” (*absolute denial*), sino que para pensar una realidad no solo distinta al ser (*relative rejection*) sino que además independiente del ser (nada absoluta)²².

De hecho, la reflexión en torno a la nada absoluta lleva a Moten a invertir la relación clásica entre el ser y la nada. Si—como hemos visto—el pensamiento de Moten gira entorno a la dicotomía entre *ser* como anti-negro y la negritud como *nada* absoluta, entonces, es el ser el que depende del no-ser y no a la inversa. El ser es relativo y la nada es absoluta. La negritud como nada absoluta es el anti-fundamento (el *Ungrund*) del ser anti-negritud. Según King-Ho Leung:

“La blanquitud [*whiteness*] no es más que una relación con la negritud” [...] El ser qua antinegritud sólo puede *ser* como negación de la negritud misma; sin negritud, el ser no sería en absoluto porque ya no existiría como “anti-negritud”, como negación y producto de la negritud. Lo que encontramos en la paraontología de Moten es, por lo tanto, una inversión del tradicional privilegio metafísico o incluso onto-teológico del ser sobre la nada, donde la nada se convierte en el centro o incluso en el “fundamento” [*ground*] de todo. El ser sólo “es” en virtud de estar en relación antitética con la nada absoluta; sólo existe como “anti-nada” (como anti-negritud) [...] En otras palabras, el ser como anti-negritud es siempre meramente relativo, mientras que la negritud como nada es absoluta (Leung 9).

Pero parece, entonces, que lo fundado oprime a su fundamento. Y, sin embargo, al tener el *ser* la *nada* como fundamento, reside aquí también la posibilidad del desfondamiento del ser. Lo “optimista” de esta exploración de la convergencia entre negritud y nada absoluta radica en las posibilidades que devela la nada. En esto podemos encontrar una semejanza con la nada como exterioridad del sistema que vimos en Dussel. En la nada absoluta se esconde la posibilidad de otro mundo, de un mundo más allá. Sobre la pobreza, escribe Moten:

²² Esta búsqueda de pensar en una realidad que no esté subordinada a la lógica del ser la vemos también en el uso de Dussel de la expresión *oùk ón* (Dussel 1980 51), citada al final de la segunda sección de este artículo.

La pobreza en este mundo se manifiesta en una especie de acceso poético a lo que es del otro mundo que permanece inaudito, inadvertido, no reconocido en este [...] queda por considerar precisamente qué es lo que tienen los que no tienen nada. ¿Qué es esa nada que tienen o a la que tienen acceso? ¿Qué viene de eso? ¿Y cómo opera tenerla en relación con la pobreza? (Moten 776)

La reflexión de Moten exige—como en el caso de Dussel—ir más allá de la ontología. Más que querer acceder a una ontología de la negritud anterior a la anti-negritud del mundo moderno, Moten quiere prescindir por completo de toda ontología. Y en esta esperanza de escapar de los confines de la ontología aparece el misticismo. Acerca de esta conexión entre misticismo y el pensamiento de Moten, Warren comenta:

Es precisamente este misticismo el que proporciona nuestro escape fugitivo de los confines de la ontología (o dominación metafísica si seguimos a Heidegger). Así, el misticismo no es enemigo de la libertad; es el único aspecto de la existencia que da esperanza [...] Su anhelo de escapar de la brutalidad física de la ontología da como resultado una forma de “misticismo en la carne” (Warren 222).

Si bien esta referencia al misticismo (de la carne) no apunta necesariamente a Eckhart, consideramos que la yuxtaposición de estos tres usos de la nada que hemos llevado a cabo en este trabajo ha mostrado una cierta lógica similar; y, por tanto, se hace necesario terminar este trabajo explicitando ciertas preguntas y tensiones que se originan al comparar el uso místico y el político-filosófico de la nada.

Conclusión

Al haber explorado la lógica de la nada en los discursos místicos de Eckhart, por un lado, y en los político-filosóficos de Dussel y Moten, por el otro, nos surgen ciertas preguntas. ¿Cómo llamar a estas conexiones e influencias?: proyecciones, transferencias, secularización, sacralización, afinidad electiva.

La cuestión que más nos implica e interesa, sin embargo, se experimenta como un cierto “desatino” que surge al yuxtaponer las dos experiencias humanas de convertirse en una nada: la del discurso místico de un vaciarse y devenir nada para ser llenado por Dios, por un lado, y la de ser relegado a una pobreza material y muerte social, a una condición de no-ser con respecto al sistema político-social

imperante, por el otro. ¿Hay algo del rédito (de la ganancia) de la lógica de la nada mística que se pueda transferir a la nada de la exterioridad colonial y de la negritud? Si bien Eckhart puede predicar el camino del “puro ser separado”, ¿por qué se siente desatinado el imaginar una prédica en la que se afirme y fomente la condición de no-ser de la negritud o del oprimido en un mundo anti-negro y colonial? Y, sin embargo, Moten parece hacerlo. Tanto Dussel como Moten ven en la *nada* el lugar—o mejor dicho el no-lugar—desde donde se puede destruir un mundo que está destruyendo al mundo.

Intentemos iluminar este “desatino” al explicitar algunas semejanzas y diferencias de estos diversos discursos de la nada. Comparando estas distintas lógicas de la nada que hemos descrito, postulamos una cierta semejanza *lógica* o *estructural* entre ellas. Como hemos visto, en todas ellas se da una dinámica en la experiencia de la nada que puede conducir a una cierta *inversión*, que nos transporta desde la experiencia de la nada como falta o defecto hacia la nada como exceso. Es decir, el vaciamiento o desapego de las cosas en el alma (Eckhart) o la muerte social y política en el contexto de un sistema colonial o anti-negro (Dussel, Moten) pueden conducir a ser uno llenado por Dios (Eckhart) o generar la condición para que surja una exterioridad desde donde se pueda cuestionar y así transformar el sistema opresor (Dussel, Moten).

Ahora bien, el análisis se torna más complicado cuando trasladamos la atención desde la lógica y la estructura hacia los *términos* que ellas ponen en relación. Si bien un polo en la relación está relativamente claro y que podemos denominar el polo humano —el alma, la persona oprimida, la persona negra—; es el segundo polo el que permanece indecible e incierto: Dios antes o más allá de la creación; lo absolutamente Otro que se revela en el rostro de la persona oprimida; la posibilidad de una negritud más allá de un mundo anti-negro. Tanto Eckhart como Dussel llaman *divino* a este segundo polo que el vaciamiento del polo humano manifiesta. En el caso de Moten, la ausencia de una referencia a lo divino debe complementarse con su uso explícito del concepto de *misticismo* (de la carne). En otras palabras, si bien en Moten la inversión propia de la lógica de la nada no se reviste de un lenguaje teísta, sí describe esta lógica aludiendo al misterio propio de la mística.

Al constatar en estos tres usos de la nada una similitud en la lógica y la estructura, y además una similitud en la imposibilidad de identificar el segundo término que esta lógica implica, consideramos que la comparación de estos usos constituye—o al menos se asemeja a—lo que Thomas Carlson (2000) ha llamado una *analogía apofática*. En este tipo de analogías, según Carlson, “los dos términos no pueden ser identificados ni distinguidos entre sí. En ambos casos el término ‘en sí mismo’ permanece estrictamente incognoscible e inefable, y, por lo tanto, desde la perspectiva de un sujeto cognoscente en el lenguaje, ni idéntico ni

distinto” (211). La descripción de los tres usos de la nada que hemos hecho nos permite postular una analogía apofática entre ellos, en cuanto en cada uno de estos usos hay un polo que es estrictamente incognoscible y misterioso, es decir, se muestra ocultándose; y en cuanto incognoscibles, estos tres polos son, para nosotros, ni identificables ni distinguibles entre ellos.

Intentemos ahora señalar ciertas diferencias entre estos usos de la nada. Si bien hasta ahora en nuestra comparación hemos asumido dos polos—por ejemplo, en Eckhart, el alma y Dios—; implícito hay un tercer término, a saber, el mundo. De hecho, si el alma está llena de las cosas del mundo, esto implica que no hay espacio para Dios. La relevancia de este tercer término se hace aún más patente en los textos de Dussel y Moten. Para ellos, es precisamente el mundo—como una totalidad de referencias y significados—lo que reduce a ciertas personas a la condición de nada. Consideramos que la mayor diferencia entre los usos de la nada de Eckhart, por un lado, y los de Dussel y Moten, por el otro, radica en el *lugar* que se le asigna a la “agencia” que pone en movimiento la dinámica de la nada: ya sea el sujeto humano o el mundo.

Si bien, según la lógica del misticismo descrita por Eckhart, la persona no puede ser ella misma el agente de su convertirse en *nada*, y debe, de modo pasivo, aprender a dejarse vaciar de las cosas por Dios; consideramos, sin embargo, que habría una cierta *decisión* inaugurando el camino del ser-separado. Si bien esta decisión del alma no genera el vaciarse mismo y que, por tanto, la decisión consiste más bien en *responder* a un llamado previo de Dios; sostenemos que esta respuesta implica una decisión: un querer dejar de querer, una voluntad que se esfuerza para dejar a Dios que haga en uno²³. En el caso de Dussel y Moten, en cambio, la condición de ser nada no es precedida por ninguna decisión de la persona; es el *mundo* el que ya ha tomado la “decisión” por ella: es el sistema-mundo el que se constituye como totalidad precisamente reduciendo a parte de la población—contra su voluntad—a la condición de nada. Y es esta ausencia de decisión la que explicaría, en parte, la diferencia de estos modelos y aclararía, también, la sensación de “desatino” que nos produce el confundir la lógica de la nada del discurso místico con la del contexto de la negritud y la decolonialidad.

Sin embargo, una vez que el sujeto humano se encuentra arrojado a la condición de nada—ya sea voluntaria o involuntariamente por ser oprimido por un sistema fetichizado o anti-negro—, entonces el *beneficio* de la lógica de la nada, descrito por Eckhart, se repite también en el ámbito político-social. Así, el estar arrojado a la condición de nada deviene la condición de posibilidad para

²³ Por su puesto, la excepción sería la experiencia de Pablo descrita en el sermón de Eckhart que revisamos en la primera sección. En este caso, Dios habría irrumpido sin consentimiento previo.

que pueda irrumpir la transformación del sistema opresor. Y esta irrupción—tal como en Eckhart—toma el lenguaje de lo divino (Dussel) y lo místico (Moten).

Volviendo a las diferencias, comparemos estos usos de la nada con respecto a la distinción entre la nada relativa y la absoluta. Según Williams, cuando la “nada es Dios” el sentido de esta nada es relativo. ¿Qué significa que el sentido sea relativo? Significa que la nada no se refiere a que no haya absolutamente nada, sino que denota “algo” indeterminable; es decir, lo que se niega es que Dios sea tanto algo determinado como una nada absoluta. Estas fórmulas apofáticas nos permiten entrar en contacto con “algo” más allá del ser, con una nada excesiva, la de arriba, que desde más allá del techo funda la escalera del ser.

Pero analicemos esta nada relativa ahora con respecto a su dependencia al ser. Esta nada más allá del ser (la de arriba, Dios —para Eckhart), ¿queda de algún modo determinada por el *ser* que ella misma supera? Es decir, si bien está claro que la nada defectuosa (la de abajo) depende del *ser* (en cuanto ella se entiende privación de ser), no queda claro si el “más allá” de la nada de arriba queda sujeta o no a aquello que trasciende. Por lo tanto, Dios como nada relativa (en el sentido propuesto por Williams) es, con respecto al ser, ¿relativo o absoluto (o sea, dependiente o independiente)?

Es esta ambigüedad la que tampoco nos permite distinguir enfáticamente entre los usos de la nada de Dussel y Moten. Como vimos, ambos desean pensar una nada absoluta (esto es, independiente del ser). Si bien la nada absoluta (*oûk ón*) propuesta por Dussel, claramente supera lo relativo de la nada defectuosa (la de abajo)—esto es, la persona oprimida deja de experimentarse como una falta del sistema y, por tanto, definida negativamente desde el sistema—, surge la sospecha de si tras la inversión (o la fusión de lo sub-ontológico, trans-ontológico y lo teológico), esta nada excesiva y exterior al sistema (la de arriba) podría quedar paradójicamente ligada al sistema. Dicho de otra forma, según la perspectiva que nos ofrece Moten, también la nada que es más que el *ser*, sigue, de algún modo, determinada por el *ser* que ella misma supera. Si bien la *inversión* permite a Dussel pasar de lo sub-ontológico a lo trans-ontológico y teológico, no nos queda claro si acaso lo “trans” o el “teos” implicaría—paradójicamente—quedar todavía determinado por lo ontológico.

Ahora bien, también Moten puede ser objeto de esta misma crítica. Cuando él invierte el orden clásico, y hace depender el ser de la nada, ¿queda el fondo (aunque sea entendido como abismo) completamente desligado de aquello que funda? Si bien concedemos que, al afirmar que el ser (anti-negro) depende de la nada absoluta (negritud) —esto es, que el fondo del ser es un abismo—, esto posibilita desfondar, liberarse y transformar las estructuras del ser (anti-negro); nos surge la pregunta: ¿qué hay en la negritud (como nada absoluta y fondo del ser) que permitió fundar una estructura del ser como anti-negro. ¿No queda el fondo

(aunque sea un abismo) de algún modo emparentado con aquello que funda? Si es así, ¿es la nada entonces realmente absoluta?

Nos topamos quizás con los límites de toda “inversión” de términos. Si bien el invertir el orden clásico y pensar al *ser* dependiente de la nada puede producir cambios reales (por ejemplo, transformaciones en el orden político social opresor), nos parece dudoso que tal inversión de los términos permita superar la lógica y la estructura. En este artículo hemos explorado discursos que invierten la jerarquía ontológica, y desde el no-ser de abajo convocan—casi de modo teúrgico—a la nada de arriba. Pero, en cuanto la nada de arriba tiende a fundar precisamente al sistema jerárquico, no queda claro que ella sea una superación de las jerarquías. Dicho de otro modo, con la inversión, quizás el contenido de las jerarquías cambie, pero la lógica que distribuye el poder en jerarquías queda intacta²⁴.

Como dijimos en la introducción, todo lo que hemos expuesto ha estado bajo una primera premisa, a saber, que el *ser* es fundamento de una realidad jerárquicamente ordenada (ya sea como único fundamento—auto-fundándose—o como un fundamento que requiere a su vez ser fundado). En este contexto, el no-ser puede denotar la esperanza de un “lugar” más allá de la jerarquía fundada en el ser, señalando la posibilidad de una liberación de la totalidad ontológica opresora. Pero en cuanto esta liberación implica un fundamento ahora entendido como no-ser, la nada termina denotando el origen místico *del* fundamento. Y con esto, si bien se produce una inversión y por tanto un cambio en el contenido de la jerarquía, no es suficiente para superar la lógica de la jerarquía. Dicho en otros términos, la inversión de los términos no implica necesariamente la *subversión* de la estructura.

Pero ¿estamos tan seguros que el sentido del ser es *ser fundamento*? ¿Qué pasa con la jerarquía y de dónde viene su estructura cuando el ser ya no es fundamento? Y si el sentido del *ser* no fuera ser fundamento, entonces, ¿qué sentido tendría la nada?



²⁴ No olvidemos que uno de los autores más importantes del misticismo cristiano y una de las referencias obligadas para Eckhart, a saber, Pseudo Dionisio, es quién, por un lado, llevó al extremo el lenguaje apofático en su *Teología mística*, y, por el otro, justificó y desarrolló la noción de jerarquía en sus libros sobre la *Jerarquía celestial* y la *Jerarquía eclesiástica* (ver Pseudo Dionisio 1987). Es decir, ya en Pseudo Dionisio encontramos un antecedente para la necesidad de darle un origen y fundamento místico a toda jerarquía.

Bibliografía

- Cardenal, Ernesto. *Vida en el amor*. Madrid: Trotta, 2010.
- Carlson, Thomas. *The Indiscrete Image: Infinity and Creation of the Human*. Chicago: The University of Chicago Press, 2008.
- Carlson, Thomas. "Apophatic Analogy: On the Language of Mystical Unknowing and Being-toward-Death". *Rhetorical Invention and Religious Inquiry: New Perspectives*, ed. Walter Jost y Wendy Olmsted. New Haven: Yale University Press, 2000.
- Carlson, Thomas. *Indiscretion: Finitude and the Naming of God*. Chicago: The University of Chicago Press, 1999.
- Castro-Gómez, Santiago. *La Hybris del Punto Cero: Ciencia, raza e ilustración en la Nueva Granada (1750-1816)*. Bogotá: Editorial Pontificia Universidad Javeriana, 2010.
- Dussel, Enrique. *Las Metáforas Teológicas de Marx*. Ciudad de México: Siglo veintiuno editores, 2017.
- . *Filosofía de la liberación*. Ciudad de México: Fondo de Cultura Económica, 2011.
- . *Filosofía ética latinoamericana V: arqueología latinoamericana. Una filosofía de la religión antifetichista*. Bogotá: Universidad Santo Tomás, 1980.
- Eckhart, Maestro. *Meister Eckhart Predigten. Deutsche Werke I*. Frankfurt am Main: Deutscher Klassiker Verlag, 2020.
- . *El fruto de la nada*, ed. Amador Vega. Madrid: Alianza Editorial, 2018.
- . *The Complete Mystical Works of Meister Eckhart*. New York: The Crossroad Publishing Company, 2009.
- Fanon, Frantz. *Piel negra, máscaras blancas*. Madrid: Akal, 2016.
- Givone, Sergio. *La Historia de la nada*. Adriana Hidalgo Editora, 2010.
- Haas, Alois M. *Viento de lo absoluto. ¿Existe una sabiduría mística de la postmodernidad?* Madrid: Ediciones Siruela, 2009.
- Heidegger, Martin. "Was ist Metaphysik?". *GA 9, Wegmarken*, ed. F.-W. von Herrmann. Frankfurt am Main: Klostermann, 1976.
- Laclau, Ernesto. "On the Names of God". *Political Theologies: Public Religions in a Post-Secular World*, ed. Hent de Vries y Lawrence E. Sullivan. New York: Fordham University Press, 2006.
- Lovejoy, Arthur. *The Great Chain of Being: A Study of the History of an Idea*. Cambridge, Massachusetts: Harvard University Press, 1964.
- Maldonado-Torres, Nelson. *Against War: Views from the underside of Modernity*. Durham: Duke University Press, 2008.
- Marion, Jean-Luc. "In the Name: How to Avoid Speaking of 'Negative Theology'". *God, the Gift, and Postmodernism*, ed. John Caputo and Michael Scanlon. Bloomington: Indiana University Press, 1999.
- Marion, Jean-Luc. *L'idole et la distance*. Paris: Editions Bernard Grasset, 1977.
- Moten, Fred. "Blackness and Nothingness (Mysticism in the Flesh)". *South Atlantic Quarterly* 112, 4 (2013): 737-780.
- Platón. *La república, Diálogos IV*. Barcelona: Gredos, 2020.
- Plotino. *The Enneads*. New York: Larson Publications, 1992.
- Pseudo-Dionysius. *Pseudo-Dionysius: The Complete Works*. New York: Paulist Press, 1987.
- San Agustín. *La ciudad de Dios*, traducción de Rosa Marina Sáez. Madrid: Editorial Gredos, 2007.
- Sells, Michael A. *Mystical Languages of Unsayings*. Chicago: The University Press, 1994.
- Tillich, Paul. *The Courage to be*. New Haven: Yale University Press, 2000.
- . *Teología Sistemática*. Barcelona: Ediciones Ariel, 1972.
- Warren, Calvin L. "Black Mysticism: Fred Moten's Phenomenology of (Black) Spirit". *Zeitschrift für Anglistik und Amerikanistik* 65, 2 (2017): 219-229.
- Williams, Duane. "Meister Eckhart and the Varieties of Nothing", *The Journal of the The Eckhart Society* 20, 1 (2011): 85-98.
- Yámblico. Iamblichus: *On the Mysteries*. Atlanta: Society of Biblical Literature, 2003.
- Yountae, An. *The Decolonial Abyss: Mysticism and Cosmopolitics from the Ruins*. New York: Fordham University Press, 2017.



El principio supremo de la moral kantiana como conocimiento práctico

The Supreme Principle of Kantian Morality as Practical Knowledge

FELIPE ZEGERS-QUIROGA¹

Universidad de Chile, Chile
felipe.zegers@ug.uchile.cl

Fecha de recepción: 13/09/2023

Fecha de aceptación: 04/01/2024

Resumen

El presente trabajo se propone esclarecer en qué medida el principio establecido por Kant en la *Fundamentación de la metafísica de las costumbres* constituye un conocimiento. Se mostrará que, al reconstruir la caracterización kantiana de conocimiento práctico, podemos luego examinar de qué manera aquella obra se trata justamente de alcanzar ese tipo de conocimiento. Tal investigación se ve justificada, por un lado, porque ayudaría a comprender el lugar sistemático de la obra en cuanto en ella hay un conocimiento necesario para la ciencia ética y, por otro lado, iluminaría cierto aspecto del debate contemporáneo respecto de si acaso la filosofía moral de Kant es realista o antirrealista en alguna de sus variantes. Para aquello es necesario comenzar por aclarar qué entiende Kant por conocimiento práctico, y luego ver cómo esto se realiza o no en dicha obra, finalmente bosquejaremos los

CÓMO CITAR ESTE ARTÍCULO:

En APA: Zegers-Quiroga, F. (2024). El principio supremo de la moral kantiana como conocimiento práctico. *Resonancias*, (17), 57-78. DOI: 10.5354/0719-790X.2024.72000

En MLA: “El principio supremo de la moral kantiana como conocimiento práctico.” *Resonancias*, n.º 17, julio de 2024, pp. 57-78. DOI: 10.5354/0719-790X.2024.72000

Palabras clave: fundamentación, conocimiento práctico, Kant, ética, moral

Keywords: foundation, practical knowledge, Kant, ethics, moral

¹ Este trabajo ha sido posible gracias al financiamiento de ANID-Subdirección de Capital Humano/Magister Nacional/2022-22220444. ORCID: 0009-0007-6108-8818.

posibles rendimientos filosóficos que esto logra en relación con otras aproximaciones contemporáneas a la filosofía moral de Kant.

Abstract

This paper aims to clarify to what extent the principle established by Kant in the *Foundation for the Metaphysics of Morals* constitutes knowledge. It will be shown that, by reconstructing Kantian characterization of practical knowledge, we can then examine in what way that work is precisely about attaining that kind of knowledge. Such an investigation is justified, on one hand, because it would help to understand the systematic place of the work insofar as it contains a knowledge necessary for ethical science, and, on the other hand, it would illuminate a certain aspect of the contemporary debate as to whether Kant's moral philosophy is realist or anti-realist in any of its variants. To do so, it is necessary to begin by clarifying what Kant means by practical knowledge, and then to see how this is realized or not in said work, and finally to sketch the possible philosophical yields that this achieves in relation to other contemporary approaches to Kant's moral philosophy.

I

En distintas partes de su obra, Kant afirma que la filosofía, y en particular la metafísica, debe constituirse próximamente como una ciencia (por ej. B XXIII o B878). Cabe notar que, en sede práctica, existiría entonces una ciencia de lo moral —lo que Kant suele llamar “Metafísica de las costumbres”² (por ej. Ak. V. 387-388) — que, en tanto ciencia, supone que podemos alcanzar un conocimiento verdadero acerca de lo moral. El presente trabajo intentará investigar en qué medida cierta parte de la filosofía práctica kantiana, específicamente el principio expuesto en su obra *Fundamentación de la metafísica de las costumbres* (en adelante *GMS*), constituye un conocimiento; teniendo en consideración que la *GMS* no es propiamente ni una Crítica ni la exposición de la Ciencia. Para ello debemos, primero, caracterizar adecuadamente qué es — en términos kantianos — el conocimiento, basándonos centralmente en la distinción que él mismo realiza entre conocimiento teórico y conocimiento práctico, y luego analizar la estructura y argumentación de la *GMS* para ver cómo es que coincide cierta concepción de conocimiento con lo allí expuesto.

Esto es relevante principalmente por dos razones; en primer lugar, porque esto contribuye a ganar una mejor comprensión del argumento de la obra, pero,

² No hacemos referencia aquí a la obra con el mismo nombre, sino al proyecto que tiene en mente Kant al escribir la *KrV* [cf. B 878] o lo que dice en el prefacio de la *GMS* que citamos.

sobre todo, del lugar sistemático que tendría la *GMS* en el corpus kantiano. Ya que, si en la *GMS* se establece un conocimiento práctico puro, y el proyecto de una metafísica de las costumbres constituiría la ciencia de lo práctico (conocimiento articulado sistemáticamente [A832/B860]), esto sugiere que este conocimiento será requerido, o al menos supuesto (cf. Ak. V. 8), al constituirse el sistema. Por otro lado, existe una amplia disputa en torno al eventual carácter realista, antirrealista, constructivista, intuicionista etc., de la filosofía moral de Kant (Santos y Schmidt 2018), el cual aborda, entre otras cosas, la relación que existiría entre la idea de conocimiento y los hechos morales. Esto, a su vez, tiene una cantidad importante de problemas de interpretación, por ejemplo, el significado de esos mismos términos (*Id.* VII). Nosotros queremos, por otra parte, alejarnos de tal marco categorial y acercarnos al proyecto y los propios términos que emplea Kant, sobre todo respecto de esa curiosa caracterización de la ética como ciencia, y —de forma más restringida en la *GMS*— de la posibilidad de asentar este conocimiento de lo moral. Debemos recalcar que estas dos razones que justifican la importancia de nuestra investigación quedarán en este trabajo solo sugeridas, ya que, por motivos de extensión, aquí solo cabe enfocarnos en el tratamiento del asunto principal que mencionamos al comienzo.

II

Veamos ahora la distinción kantiana entre conocimiento teórico (en adelante CT) y conocimiento práctico (en adelante CP)³. Comenzaremos entregando maneras preliminares en que se podría entender esta distinción con algunos elementos de la filosofía de Kant; después revisaremos los problemas de estas concepciones; con ello podremos ver luego más claramente las diferencias y posibles continuidades.

Tomemos por ahora la noción de conocimiento general como la concordancia entre un concepto y su objeto, tendríamos entonces conocimiento cuando nuestro concepto respecto de un algo coincide o corresponde con él, esto va en la línea de lo que el propio Kant llama “la definición nominal de la verdad” (A58/B82)⁴. Una primera manera de entender la distinción sería considerar que en el CT hay un determinado objeto dado para el cual debemos hacer coincidir el concepto que tenemos de este; en otras palabras, el concepto debe adecuarse al objeto.

³ Conocimiento aquí siempre será la traducción del término “Erkenntniss” que emplea Kant en alemán.

⁴ Cito acorde a la numeración establecida en la edición académica de las obras de Kant, las traducciones empleadas son aquellas mencionadas en la bibliografía final, excepto por el caso de la *Lógica Jäsche* cuya traducción es propia.

Asimismo, el CP sería aquel en el que tenemos cierto concepto, y debemos hacer que nuestro objeto corresponda con este, de tal manera que el objeto se adecue a lo que *debe ser* según el concepto de nuestro CP; es decir, al revés que el CT, el objeto debe adecuarse al concepto.

Una segunda manera de entender la distinción entre CT y CP, no por ello excluyente, es considerar que refieren objetos de distinta naturaleza; es decir, el conocimiento de ambos tipos no se distinguiría por su dirección de adecuación, ya que en ambos el concepto tendría que corresponder al objeto, sino que se diferenciarían en tanto los objetos a los cuales se intenta conocer tendrían una constitución radicalmente distinta. Bajo este enfoque existirían algo así como objetos teóricos (comúnmente objetos de la naturaleza o de la experiencia sensible) y otros objetos prácticos (generalmente acciones humanas o ideas de la moral, concebidos intelectualmente). Así podría interpretarse el siguiente pasaje en la *Crítica de la razón pura*: “Me contento aquí con definir el conocimiento teórico como uno por el cual conozco lo que *existe*, y el práctico [como un conocimiento] por el cual me represento *lo que debe existir*.” (A633/B661). Aunque debemos resguardarnos de tomar esta caracterización como acabada, ya que, como indica el comienzo de la cita, es solo aplicable al contexto específico de la argumentación de aquellos párrafos sobre teología.

Ahora bien, hay dos objeciones posibles, ligadas entre sí, que surgen a partir de consideraciones del propio Kant. En primer lugar, existe el riesgo de una circularidad al concebir el conocimiento bajo este modelo concepto-objeto, como dice nuestro filósofo

Yo puedo comparar el objeto con mi conocimiento, pero sólo *en cuanto yo lo conozco*. Mi conocimiento entonces debe confirmarse a sí mismo, lo cual está lejos aún de ser suficiente para la verdad. Pues como el objeto está fuera de mí y el conocimiento en mí, siempre puedo sólo juzgar: si mi conocimiento del objeto coincide con mi conocimiento del objeto. Los antiguos llamaban *diallele* a este círculo de explicación (Ak. IX. 50)⁵.

Esto es similar a la preocupación que muestra el mismo Kant en otro lugar — más preponderante— de su obra (A58/B82ss). La circularidad surge de la incapacidad de establecer un criterio firme para establecer la verdad. No obstante,

⁵ Original: “Nun kann ich aber das Object nur mit meinem Erkenntnisse vergleichen, *dadurch daß ich es erkenne*. Meine Erkenntniß soll sich also selbst bestätigen, welches aber zur Wahrheit noch lange nicht hinreichend ist. Denn da das Object außer mir und die Erkenntniß in mir ist, so kann ich immer doch nur beurtheilen: ob meine Erkenntniß vom Object mit meiner Erkenntniß vom Object übereinstimme. Einen solchen Cirkel im Erklären nannten die Alten *Diallele*.”

notemos aquí lo que sostiene al decir que, en este marco, el objeto se encuentra fuera de mí y el conocimiento dentro de mí, y ahora tengamos en consideración la argumentación presente en la *KrV* sobre el eje de la revolución copernicana en filosofía, en la cual “los objetos deben regirse por nuestro conocimiento” (B XVI). Tomando esto en cuenta, no sería correcto hablar de un objeto simplemente dado y completamente externo e independiente al sujeto que lo conoce, por lo cual el modelo de correspondencia ya supondría un marco posible de adecuación que posibilita y limita el conocimiento que pretende tener, así no existiría en ningún caso un acceso directo hacia los supuestos objetos, por lo que al menos en el caso del CT no parece satisfactorio. Sin embargo, podría decirse que en el caso del CP el objeto no sería dado, ya que estaríamos en posesión del concepto, y tendríamos que producir y adecuar el objeto mediante la acción. Pero allí, tanto el concepto como el eventual objeto también tendrían que regirse bajo las condiciones de posibilidad *a priori* en el sujeto, ya que, a pesar de que el pasaje anterior parece referirse al contexto de una metafísica teórica, en principio no hay razones para suponer que no valga lo mismo para el caso de algo así como un CP.

Expuestas ambas objeciones, la posible circularidad al examinar su verdad y la dependencia entre el conocer y lo conocido, no pretendemos desechar la primera aproximación preliminar —el modelo de adecuación— sino afinarlo; así como también examinar la distinción en torno a los tipos de objeto. Revisemos el pasaje principal donde Kant parece hacer la distinción de manera más fiable.

En la medida en que en éstas haya de haber razón, en ellas debe conocerse algo *a priori*, y el conocimiento de ellas puede ser referido a su objeto de dos maneras: o bien meramente [para] *determinarlo* a éste y al concepto de él (que debe ser dado por otra parte), o bien [para], además, *hacerlo efectivamente real*. El primero es el *conocimiento racional teórico* y el otro, *práctico* (B IX-X).

Siguiendo parcialmente lo argumentado por Engstrom (2002 57), podemos encontrar aquí una caracterización genérica de conocimiento, común tanto para el CT como para el CP, así como también una distinción específica entre ambos. Lo primero que observamos es que la diferencia es un añadido, el *además* [auch] que antecede la especificación del CP. Sin embargo, en ambos casos el conocimiento determinaría tanto al objeto como a su concepto, de modo que la pregunta evidente es ¿A qué se refiere aquí con determinación? Una pista la encontramos bastante más adelante en la misma obra: “la *determinación* es un predicado que se añade al concepto del sujeto, y lo aumenta” (A598/B626), es decir, consistiría en un juicio sintético, en tanto no estaría contenido ya en el concepto.

Si existe una determinación que se añade al concepto, entonces este concepto tiene que estar en alguna medida presupuesto, como dice el pasaje *dado* por otra

parte, que actúa así como principio y otorga la regla para el juicio (Ak. XX. 211). Engstrom califica esto como *autodeterminante* (2002 57) e interpreta así la *espontaneidad* del conocimiento kantiano (KrV A51/B75). Pero falta aún aclarar cómo opera la determinación en el caso del objeto; para ello veamos el siguiente pasaje:

Solo son posibles dos casos en los cuales una representación sintética y sus objetos pueden coincidir, [pueden] referirse necesariamente unos a otros, y [pueden], por decirlo así, encontrarse entre sí: o bien cuando sólo el objeto hace posible la representación, o bien cuando sólo ésta hace posible al objeto [...] si es lo segundo, como la representación en sí misma (pues no se trata aquí de la causalidad de ella por medio de la voluntad) no produce su objeto *en lo que respecta a la existencia*, la representación es determinante *a priori* con respecto al objeto, cuando sólo por medio de ella es posible *conocer algo como un objeto* (A92/B124-125).

Engstrom interpreta este pasaje asumiendo lo que sería una contraposición típica en el pensamiento de Kant entre *forma* y *existencia* (2002 58-59). De esta manera, si el segundo caso corresponde al conocimiento *a priori* (i.e. aquel donde la representación hace posible al objeto) donde no produce al objeto respecto de su existencia, entonces él concluye que de lo que se trata en el conocimiento general *a priori* es de producir al objeto respecto de su *forma*, y en ese sentido se entiende también el pasaje antes citado de B XVI. De esta manera, lo común sería tanto del CP como del CT sería la determinación respecto a la *forma* del objeto, y la diferencia sería que en el caso del CP también se vería producida la *existencia* del objeto, y así podemos entender el paréntesis del pasaje antes citado respecto a que *aquí* no se trata de causalidad por medio de la voluntad. Engstrom resume su posición respecto de la diferencia específica de la siguiente manera “mientras que en el caso del conocimiento teórico la realidad del conocimiento depende de la realidad de su objeto, en el caso del conocimiento práctico la realidad del objeto depende de la realidad del conocimiento”⁶ (Engstrom 61). Así, la diferencia relevante es respecto a la dependencia en la existencia del objeto, y no simplemente en relación con la dirección de adecuación.

Además, encontramos que luego de que Kant hable de la definición nominal de la verdad (A58/B82ss), descarta que pueda haber un criterio universal de la verdad de un conocimiento según la materia (i.e. del contenido), pero sugiere que sí sería posible respecto de la forma de, en cuanto se podrían proveer reglas que funcionen como criterio. Esto se comprenderá más adelante.

⁶ Las traducciones de las citas de comentaristas son propias.

Dicho lo anterior, no nos parece que la posición de Engstrom sea suficientemente contundente y persuasiva respecto de la evidencia que presenta, sobre todo en relación con su interpretación posterior del CP como conocimiento eficaz. Revisaremos entonces ahora otros pasajes que pueden iluminar el asunto, que no son considerados explícitamente por el intérprete.

En la primera introducción de la *Crítica de la Facultad de Juzgar*, Kant caracteriza la filosofía como un sistema de conocimiento racional mediante conceptos, y la filosofía crítica (no incluida dentro de aquel sistema) como una investigación en torno a la posibilidad de ese conocimiento. Luego nos dice que la filosofía, en tanto sistema, no puede ser dividida de otra manera sino bajo

la diferencia originaria de sus objetos y a la diversidad esencial —que sobre aquélla reposa— de los principios de una ciencia que los contiene, en filosofía *teórica* y filosofía *práctica*; de esta suerte, una parte tiene que ser la filosofía de la naturaleza, y la otra, la de las costumbres, de las cuales la primera puede contener también principios a priori, pero la segunda (puesto que la libertad no puede ser de ningún modo objeto de la experiencia) nunca otros que principios a priori puros (Ak. XX. 195).

Aquí se parece volver a sugerir que la distinción es respecto de la constitución esencial de los objetos respectivos del CT y el CP. Pero Kant inmediatamente busca descartar lo que considera un gran malentendido, ejemplificado con la política y la economía, entre otros, de que la filosofía práctica *no* trata de aquellos asuntos que puedan ser producidas por nosotros mediante una mera operación —un cálculo de cómo alcanzar determinado fin—, operaciones que podríamos más bien llamar técnico-prácticas, sino que “las proposiciones prácticas [...] únicamente son aquellas que consideran la *libertad* bajo leyes” (Ak. XX. 196), es decir, moral-prácticas. El resto serían algo así como el conocimiento teórico de ciertos asuntos prácticos, pero no filosofía práctica como tal. Nosotros entenderemos aquí el calificativo de *práctico* respecto del conocimiento práctico en el sentido moral predominante en Kant, y no en aquel sentido técnico, siguiendo la terminología especificada por nuestro filósofo en esta primera introducción de la tercera crítica:

todas las proposiciones prácticas que derivan del arbitrio, como causa, aquello que la naturaleza puede contener, pertenecen en su conjunto a la filosofía teórica, como conocimiento de la naturaleza, y sólo las que dan la ley a la libertad son específicamente diferentes de aquéllas en virtud del contenido (Ak. XX. 197).

Solo estas últimas son parte de la filosofía práctica propiamente tal. La filosofía teórica difiere de la práctica tanto en sus principios como en sus consecuencias, pero la posibilidad de sus objetos y, por tanto, del conocimiento como tal, es

distinto en ambas, al regirse tales objetos bajo las leyes de la libertad. Así se entiende, también, que finalmente Kant diga “Práctico es todo lo que es posible por libertad” (A800/B828).

Si nuestra interpretación es correcta, y podemos tomar lo presentado en la *KU* en la misma sintonía de la distinción antes discutida y descrita en la *KrV*, entonces el CP se restringe a un campo bastante estricto. De hecho, al final de la primera crítica, al hablar de *libre albedrío* —que en tal contexto no parece diferenciarse de voluntad como sí lo hace en otros textos— el filósofo declara que “todo lo que esté en conexión con éste, ya sea como fundamento [Grund] o como consecuencia, se denomina práctico” (A802/B830). Notemos aquí la noción de fundamento, ya presente en el propio título de la *GMS*, la cual será relevante en lo que sigue.

En la *Lógica Jäsche*, aunque polémica, encontramos un apéndice titulado justamente “Von dem Unterschiede des theoretischen und des praktischen Erkenntnisses” (Ak. IX. 86); allí el filósofo afirma que el conocimiento práctico consiste en a) *imperativos*, o contienen b) *fundamentos* [Gründe] para imperativos posibles. Aquí se entiende *imperativo* como una proposición que expresa un acto libre que busca realizar algo acorde a un fin, y así constituye un *deber ser* distinto del mero *ser* de un CT, mientras que los fundamentos b) pueden ser tenidos también por teóricos (oponiéndose a conocimiento especulativo) en tanto que de aquellos pueden ser derivados los imperativos.

La idea de que el fundamento es constitutivo del CP es algo que también está presente en la segunda crítica, donde Kant dice explícitamente que este conocimiento práctico es “aquel que sólo tiene que tratar de los fundamentos de determinación [Bestimmungsgründen] de la voluntad” (Ak. V. 20). Si leemos con atención el *sólo* (bloß), unido a lo anterior, es necesario asumir que lo que Kant llama fundamento es parte esencial del CP.

Ahora bien, si vemos el famoso pasaje de la *KrV* donde el filósofo afirma que “Sin sensibilidad no nos sería dado objeto alguno; y sin entendimiento, ninguno sería pensado. Pensamientos sin contenido son vacíos, intuiciones sin conceptos son ciegas [...] el entendimiento no puede intuir nada, y los sentidos no pueden pensar nada. Sólo de su unión puede surgir el conocimiento” (A51/B75ss), parece ser que algo se nos escapa nuevamente, ya que, como sostiene Hutter, es justamente esta unión entre sensibilidad y pensamiento la esencia del conocimiento humano (Hutter 104). Pero debemos de recordar nuevamente que, en este tipo de pasajes, Kant se está refiriendo al conocimiento teórico del entendimiento, propio de los objetos de la experiencia sensible. Mientras que la sugerencia de Kant, al menos en la primera crítica, parece ser que justamente aquellos objetos que podríamos eventualmente conocer fuera de la experiencia sensible son aquellos que conoceríamos mediante la razón, en su uso práctico (A796/824).

Luego de esta exposición, tomaremos como base la distinción de B IX-X con la interpretación antes dada; es decir, entender el conocimiento como una adecuación dentro de un marco de determinación dado, en donde lo que distingue al CP es, por un lado, la dependencia en la existencia del objeto conocido respecto de su concepto, tal que el último tiene la prioridad ontológica y, por otro lado, el que sus objetos se rigen por las leyes de la libertad. Así entendemos también que aquellos conceptos son conceptos normativos que expresan un *deber ser*, un imperativo, pero solo aquellos imperativos que representan la libertad bajo leyes (por tanto, como veremos, incondicionados). A lo que habría que añadir también, y esto es importante, los fundamentos de los cuales se pueden derivar aquellos imperativos, los cuales en cierto sentido también son teóricos — como se explica en la *Lógica Jäsche* y también en las dos primeras críticas— pero por su contenido, de fundamentación de imperativos de la voluntad, son en última instancia también conocimiento práctico. Además, el CP —tomado en un sentido amplio— considera también conceptos morales ligados con lo anterior, lo que Kant llamó “consecuencias”.

Hemos revisado una posible distinción entre CT y CP, ahora debemos ver directamente en el texto de la *GMS* cómo podemos afirmar con sentido que aquí haya conocimiento filosófico de algún tipo, presumiblemente práctico. Mediante este examen podremos también evaluar si acaso nuestra interpretación está en lo correcto, y complementarla de ser necesaria.

III A

Al comienzo de la *GMS*, Kant discute sobre las divisiones de la filosofía. Luego nos dice que la ética sería un conocimiento de la razón *material*, tal que contempla objetos — no es puramente formal — tratando, entonces, acerca de tales objetos y las leyes que los rigen. Lo que distingue particularmente a la ética es que tales leyes son las leyes de la libertad, las cuales determinan todo lo que debe suceder (Ak. IV. 387). Por ello, podemos afirmar que tal ética es un CP, incluso es explícito al decir más adelante que “las leyes morales y sus principios, no sólo se diferencian esencialmente de cualquier otro conocimiento práctico que albergue algún elemento empírico” (Ak. IV. 389). Pareciera así que el CP aborda principios, leyes y objetos morales. Debemos ahora tomar el hilo de argumentación que nos permita observar el procedimiento para obtener tal conocimiento y explicarlo, incluso a sabiendas de que estamos solamente en la fundamentación de una metafísica de las costumbres posible, cuyo único objetivo es la búsqueda (capítulo 1 y 2 de corte analítico) y el establecimiento (capítulo 3 de corte sintético) del principio supremo de la moralidad (Ak. IV. 392).

Kant parte de un hecho dado y es que existe normatividad en el mundo, en particular, existe al menos el *deber* en general y, por consiguiente, lo que él llama la “idea común” de este y sus leyes. Dada esta idea común, afirma que “Cualquiera ha de reconocer que una ley, cuando debe valer moralmente, o sea, como fundamento de una obligación, tendría que conllevar una necesidad absoluta” (Ak. IV. 389), necesidad que supone la validez universal de su planteamiento para todo ser racional, y nos dice que la única manera de establecer una *necesidad absoluta* es mediante una investigación “exclusivamente a priori en los conceptos de la razón pura” (Ak. IV. 389). Como sostiene Scarano, si bien es cierto que del análisis del concepto común de deber se sigue el de su necesidad, no es obvio el paso del requerimiento de establecer tal necesidad a que esto sea *a priori*. En otras palabras, una cuestión es la cualidad de algo de ser necesario, y otra cosa es la cognición o el carácter epistemológico del apriorismo, por tanto, para él esta conexión debe ser argumentada explícitamente (Scarano 12). Más adelante esto se verá con mayor claridad; por ahora resulta suficiente lo postulado por el intérprete “Dado que la *aplicabilidad* de la obligación es válida para todos los seres racionales imaginables, la *cognición* de estas obligaciones o de las leyes que son responsables de ellas no puede depender de los rasgos contingentes que surgen del hecho de que el experimento mental involucra a diferentes seres racionales” (Scarano 14). Lo anterior justifica la importancia de una filosofía moral pura, en tanto permite un conocimiento veraz de los principios morales, despejando así la posibilidad terrible de que, en palabras de Kant, sea una simple quimera, lo que además no solo es un asunto cognitivo, sino práctico.

La *GMS* es una fundamentación de una metafísica de las costumbres futuras, la cual, entre otras cosas, permitiría “explorar la fuente de los principios a priori que subyacen a nuestra razón práctica” (Ak. IV. 390), pero todo indica que la fundamentación consiste ya en parte de esa indagación. ¿Cuál es aquella fuente? ¿Ello sería un descubrimiento o acaso una construcción de tales principios? Luego Kant afirma que tal metafísica de las costumbres “debe indagar la idea y principios de una posible voluntad *pura*” (Ak. IV. 390), lo que es un indicio no solo de la posterior asimilación entre razón práctica y voluntad, sino de que la investigación en torno a los principios de aquella voluntad parece yacer — la fuente — en ella misma, algo que se tornará evidente luego.

III B

El primer capítulo consiste, como indica su título, en el paso del conocimiento común que se tiene generalmente respecto a la moral hacia uno que pueda calificarse de propiamente filosófico. Comienza con la famosa afirmación “No es

posible pensar nada dentro del mundo, ni después de todo tampoco fuera del mismo, que pueda ser tenido por bueno sin restricción alguna, salvo una *buena voluntad*” (Ak. IV. 393). El énfasis aquí debe ser puesto en aquel *no es posible pensar*, es decir, es inconcebible la posibilidad de que esto sea de otra manera (Allison 71), en este y cualquier mundo posible. Si es que desde este conocimiento común nos preguntamos qué es lo bueno, entonces llegaremos a la idea de que solamente lo será la *buena voluntad*, y es solamente su *querer* lo que la hace buena por sí misma (Ak. IV. 394). Este *querer* no debe entenderse como un *deseo*, que exactamente condicionaría a la voluntad, sino como la motivación intrínseca de ella, motivación que no es más que por orden del deber, es decir, una voluntad absolutamente buena es buena por sí misma y por nada más allá de ella, sin limitaciones. Un modo de pensarlo es mediante el contraste con la posible bondad que podrían tener determinadas acciones o sus efectos, ya que la misma acción podría impugnarse dependiendo su contexto, y las consecuencias últimas de cada una son imprevisibles e incontrolables, pero, sobre todo, ambas dependen de su bondad en relación con algo externo a ellas, si no ¿cómo podríamos siquiera decir que son buenas? Su dependencia las condiciona. En cambio, el valor de la *buena voluntad* es como el de una joya que brilla *por sí sola*, como ilustra el propio Kant (Ak. IV 394).

Ahora bien, aún no hemos desentrañado el concepto de buena voluntad, para ello “vamos a examinar antes el concepto del *deber*, el cual entraña la noción de una buena voluntad” (Ak. IV. 397). Para poder explicar el concepto de buena voluntad, debemos primero investigar en qué consiste el concepto de *deber*, en tanto él contiene ya la idea de buena voluntad. Hay una serie de cosas que se hacen o pueden hacer en conformidad con el deber, pero una acción es propiamente buena solamente si es causada estrictamente *por* el deber, “ahí se cifra el valor del carácter, que sin parangón posible representa el supremo valor moral, a saber, que se haga el bien por deber y no por inclinación” (Ak. IV. 398-399).

Kant nos presenta luego tres tesis, aunque la primera de forma implícita, la cual afirma que el valor moral consiste en actuar por el deber mismo y no por alguna inclinación. La segunda tesis sostiene que ese valor tampoco puede consistir en el propósito de ella; es decir, en los efectos o el objeto que se alcanza con su realización. Allí nos dice que aquel valor “no puede residir sino en el *principio de la voluntad*” (Ak. IV. 400); es decir, una voluntad determinada por “el principio formal del querer en general” (Ak. IV. 400). Notemos aquí que, en su formalidad, o sea, la atención exclusiva al *principio* y no a las consecuencias, efectos o inclinaciones varias, es la única manera por la cual tiene sentido que hablemos de una acción propiamente buena. La tercera tesis continúa lo anterior, diciéndonos “*el deber significa que una acción es necesaria por respeto hacia la ley*” (Ak. IV. 400); no quiero detenerme en la explicación que entrega el filósofo acerca del respeto, solo es necesario indicar que para él este es un sentimiento producido

por la misma razón y no por inclinación, como señala en la nota a pie de página (Ak. IV. 401N). Y un poco antes nos dice

Como una acción por deber debe apartar el influjo de la inclinación y con ello todo objeto de la voluntad, a ésta no le queda nada que pueda determinarla objetivamente salvo la *ley* y, subjetivamente, el *puro respeto* hacia esa ley práctica, por consiguiente la máxima de dar cumplimiento a una ley semejante, aun con perjuicio de todas mis inclinaciones (Ak IV. 400-401).

Aquí vemos que la purificación de la noción de deber implica ese *apartar* otras cosas como la inclinación o los objetos de la voluntad, con ello, luego de esta limpieza, lo único que nos queda —*a priori*— es la determinación por una ley a nivel objetivo, y por una máxima acorde a nivel subjetivo. A estas alturas ya contamos con ingredientes esenciales, una voluntad absolutamente buena es aquella que determina la acción conforme al deber, el cual se estructura necesariamente bajo la forma de una ley y, por tanto, la forma de esta ley, el que sea universal y consistente, es lo esencial en la buena voluntad.

Más adelante tenemos los famosos párrafos acerca de la deducción del imperativo categórico, donde queda más claro lo anterior. Si consideráramos otras motivaciones o efectos esperados, entonces siempre se podrían explicar con relación a diversas causas ajenas a la voluntad de un ser racional, pero es precisamente en esa voluntad el “único lugar donde puede ser encontrado el bien supremo e incondicionado” (Ak. IV. 401). Solo en los seres racionales, particularmente en la voluntad que a través de esta misma razón pueden *ellos* determinar, podemos encontrar el bien moral como tal, lo único que pueden controlar por completo y, por tanto, es propiamente suyo. Habiendo una vez despejado todos los factores ajenos a la voluntad que podrían condicionar la acción, como los efectos esperados o incentivos varios, Kant nos dice respecto: “no queda salvo la legitimidad universal de las acciones en general, que debe servir como único principio para la voluntad” (Ak. IV. 402).

Inmediatamente después viene la conocida primera formulación del imperativo categórico (de aquí en adelante IC). Pero me interesa enfocarme aquí en el procedimiento, nos dice *sin tomar en cuenta* aquellas otras cosas que podrían llevarnos a actuar, y *despojando* a la voluntad de otras determinaciones posibles (Ak. IV. 402). En otras palabras, dado el procedimiento de buscar lo único de lo cual se puede decir que es completamente incondicionado, ya que restamos todas las condicionantes posibles, entonces nos queda exclusivamente la *legitimidad universal*, que en último caso no es más que la ley moral y, por tanto, el IC. La forma de la ley a la cual nos referíamos anteriormente es esta universalidad.

III C

En el segundo capítulo de la *GMS*, cuyo tránsito es intrafilosófico, se comienza recalando la importancia de que las ideas o conceptos de la moralidad y el deber no pueden provenir de la experiencia, ya que así estarían condenados al fracaso, como sugerimos anteriormente. Si, por otro lado, sabemos que existe la noción de moralidad, tenemos conciencia de ella, lo cual supone que podemos obrar moralmente, ahora se trata de mostrar sus condiciones de posibilidad. A partir de ejemplos particulares nunca se puede derivar necesidad absoluta, lo cual es requerimiento de la moralidad, por tanto, tales conceptos solamente pueden extraerse *a priori*, buscando en las ideas de la razón que determinen la voluntad. Además, no contamos con ningún caso que podamos decir con completa seguridad que allí la razón determinó por completo a la moralidad, por consiguiente, tal búsqueda sería en vano: “todos los conceptos morales tienen su sede y origen plenamente *a priori* en la razón [...] resulta también que dichos principios no pueden ser abstraídos a partir de un conocimiento empírico y por ello mismo meramente contingente” (Ak. IV. 411).

Más abajo, en ese mismo párrafo, declara que aquellos “conceptos y leyes de la razón pura” que serán expuestos, son “conocimiento racional práctico, pero puro” (Ak. IV. 411). Así, es inútil intentar indagar aquellos principios a partir de la naturaleza humana o incluso de la razón humana, sino que “justamente porque las leyes morales deben valer para cualquier ser racional, se deriven dichas leyes de los conceptos universales de un ser racional en general” (Ak. IV. 412). Esta exposición, en cuanto filosofía pura, sería una metafísica de las costumbres, es decir, una explicación de la moralidad *a priori*.

Ya sabemos cuál debe ser el origen desde el cual debemos extraer los conceptos de la moralidad. Luego encontramos uno de los pasajes más famosos de la *GMS*, donde encontramos la caracterización de la voluntad como razón práctica

Cada cosa de la naturaleza opera [wirken] con arreglo a leyes. Sólo un ser racional posee la capacidad de obrar [handeln] según la representación de las leyes o con arreglo a principios del obrar, esto es posee una *voluntad*. Como para derivar las acciones a partir de leyes se requiere una razón, la voluntad no es otra cosa que razón práctica (Ak. IV. 412).

La idea central de este párrafo es la de la voluntad, entendida como una capacidad que permite distinguir a los seres racionales de la naturaleza. La naturaleza causa efectos de acuerdo con leyes, mientras que la voluntad es la capacidad de obrar, de producir efectos según principios, es decir, no simplemente generar

cambios, sino de hacerlo bajo una cierta regla que se tiene en consideración —la llamada representación de leyes— y, en consecuencia, de tener agencia propiamente tal.

En el caso de que la razón pudiera determinar por completo a la voluntad y no tener otras fuentes de motivación, entonces actuarían los seres humanos de manera completamente racional, pero esa no es la situación, hay otros factores —como las inclinaciones— que nos condicionan. Si fuésemos absolutamente racionales, lo que aparece como objetivamente necesario —actuar bajo principios morales— lo sería también subjetivamente, pero ya que no es el caso, entonces esta representación de leyes aparece como apremio o constricción [nötigung], es decir, como contingente subjetivamente. En una voluntad santa o divina no habría conflictos motivacionales, mas, en contraste, de la humana solamente puede decirse que actúa moralmente cuando lo hace enteramente por deber.

Ahora podemos entender que Kant afirme que “La representación de un principio objetivo, en tanto que resulta apremiante para una voluntad, se llama mandato (de la razón) y la fórmula del mismo se denomina *imperativo*” (Ak. IV. 413). Recordemos que imperativo no se refiere a la formulación lingüística proposicional de un imperativo, sino al contenido de aquella proposición de ser mandataria. Como decíamos, para una voluntad absolutamente buena no se podría decir que está *apremiada* por la ley, ya que siempre está completamente determinada por ella, esta es la razón por la cual el autor nos dice que para los seres divinos no valen los imperativos, no tiene sentido hablar de ellos, ya que el querer coincide siempre con la ley: “De ahí que los imperativos sean tan sólo fórmulas para expresar la relación de las leyes objetivas del querer en general con la imperfección subjetiva de la voluntad de este o aquel ser racional, verbigracia de la voluntad humana” (Ak. IV. 414), así, todos imperativos son expresados por medio de un deber y, por tanto, de un apremio como el antes mencionado.

Esto será relevante para comprender la argumentación del tercer capítulo. Por ahora, es importante notar que la exposición analítica de Kant funciona en gran medida mediante el contraste con la voluntad santa, que sirve como un ideal frente al cual puede compararse la imperfección humana; nosotros podemos conocer directamente solo la voluntad como nos es dada a los humanos, determinable tanto racional como sensiblemente, pero si abstraemos esto último podemos concebir una voluntad puramente racional. Esto nos permite no solo analizar los supuestos que yacen bajo toda idea de moralidad *a priori*, sino comprender cómo es que nos mandan. Aunque esto último aún no queda demostrado en su posibilidad.

El problema es ahora ver la posibilidad de los imperativos categóricos, ya que los hipotéticos no suponen un desafío; entonces el problema es “¿cómo son posibles todos esos imperativos? Esta pregunta no pretende averiguar cómo pueda

pensarse la consumación de la acción que manda el imperativo, sino cómo puede ser pensado el apremio de la voluntad que el imperativo expresa en el problema” (Ak. IV. 417). Notemos que el *cómo puede pensarse* consiste en explicar que mande, no en confirmar que de hecho existen y obligan, ni tampoco en preguntarse por qué obliga. Tal es la importancia de aquel problema que constituye la pregunta central de toda la *GMS*, cuya respuesta final no encontraremos hasta el último capítulo.

Nos dice Kant que dado de que carecemos de la ventaja de que el IC nos sea dado en la experiencia, se debe investigar *a priori*, lo cual implica dar cuenta de su posibilidad. La pregunta por la posibilidad del IC debe tomarse de manera similar a la pregunta planteada en la *KrV* respecto de comprender la posibilidad de juicios sintéticos *a priori*, pero en este caso de un juicio práctico. Reformulando, sería investigar bajo qué condiciones es que podemos tener aquel tipo de juicios, tal que comprendamos como se relacionan sus elementos entre sí, es decir, que sea una proposición no trivial y necesaria, que tenga validez universal. En pocas palabras, consiste en aclarar qué posibilita que este juicio sea sintético y al mismo tiempo *a priori*.

Notemos ahora el siguiente pasaje de la *KrV*, de una nota al pie de página de Kant, a partir de la cual podemos relacionar y entender el pasaje anterior:

Para *conocer* un objeto se requiere que yo pueda demostrar su posibilidad (ya sea por el testimonio de la experiencia, a partir de la realidad efectiva de él, ya sea *a priori*, por la razón) [...] para atribuirle a tal concepto validez objetiva (posibilidad real, ya que la primera era solamente la [posibilidad] lógica), se requiere algo más. Este algo más, empero, no precisa ser buscado en las fuentes teóricas del conocimiento; puede estar también en las prácticas (B XXVI N).

Aunque el contexto es distinto, pareciera que de todas formas estamos ante un caso similar. Si deseamos conocer, de alguna manera, el IC, entonces primero debemos demostrar su posibilidad, lo que puede hacerse *a priori* en la razón, particularmente en la variante práctica. Ya podemos divisar que en la *GMS* el objetivo central del libro, establecer el principio supremo de la moralidad que no es distinto a mostrar cómo es posible, es al menos un primer paso necesario para conocerlo. Kant afirma que no ha demostrado aún la existencia y posibilidad del IC (que fundamenta todo deber posible) y, por tanto, su pretensión de verdad; esto excede las intenciones del capítulo dos, pero quedará establecido luego en el tercero. Esto significa que durante el capítulo dos ha derivado el IC del concepto de una voluntad absolutamente racional, pero solo en el tercero demostrará que es válido para el ser humano.

Continuemos ahora lo anterior; estamos frente a un juicio sintético *a priori* que es práctico justamente porque es una proposición de la moralidad, es moral en tanto manda incondicionadamente a la voluntad, lo que resulta curioso es que ello ocurre independientemente de nuestro querer y de cualquier cosa que se nos pueda mostrar como un fin para la voluntad. Lo cual queda más claro en un comentario a pie de página del autor:

Asocio con la voluntad el acto *a priori* sin presuponer condición alguna por parte de una inclinación y, por tanto, necesariamente (aunque sólo objetivamente, es decir, bajo la idea de una razón que tuviera pleno control, sobre todos los motivos subjetivos). Ésta es una proposición práctica que no deduce analíticamente el querer una acción a partir de otra ya presupuesta (pues nosotros no poseemos una voluntad tan perfecta), sino que asocia inmediatamente dicho querer con el concepto de la voluntad, en cuanto voluntad de un ser racional, como algo que no está contenido en tal concepto (Ak. IV 420N).

Lo interesante es explicar la asociación *a priori* entre la obligación y la voluntad, la cual no está contenida en el concepto de ella, ya que somos seres imperfectos, nuestro querer no coincide siempre con el deber y tenemos distintas fuentes de motivación. Entonces, debemos explicar cómo es que el IC puede mandarnos por sí solo, lo que, nuevamente, implica mostrar una conexión universal y necesaria entre mandato moral y voluntad. Tal enlace será clave, y lo investigaremos en el último capítulo de la obra. Si bien es cierto que solamente en tanto el IC manda para todo ser racional y, por tanto, para el ser humano en cuanto ser racional, debemos aclarar cómo es que manda para el ser humano en específico, sin ser absolutamente racional.

III D

Al inicio del tercer capítulo, Kant vuelve sobre la necesidad de que “ambos conocimientos queden unidos entre sí mediante la vinculación con un tercero en donde se encuentren mutuamente” (Ak. IV. 447); es decir, debemos encontrar el enlace entre voluntad y moralidad, en un elemento en donde se encuentren ya ambos. Para Kant, presuponiendo una voluntad libre, se sigue analíticamente la moralidad y su principio; sin embargo, luego nos dice que el principio como tal (IC) sigue siendo una proposición sintética, que no se puede hallar en el análisis del concepto de buena voluntad. Compartimos con Guyer que resulta confuso el párrafo del autor al respecto (Guyer 149, 152), ya que pareciera ser contradictorio con lo presentado en el capítulo anterior. No obstante, para efectos de esta

investigación creemos que —y es algo que el propio Guyer descarta— la proposición es sintética justamente porque se refiere a una voluntad humana, pues esta no tiene exclusivamente determinaciones racionales, sino también inclinaciones y apetitos, por lo que el análisis de una voluntad absolutamente buena no puede mostrarnos con necesidad la implicación sobre la imperfecta voluntad humana.

Retomando ahora lo anterior, este tercer término sería procurado después por el concepto positivo de libertad, pero para entender esto primero debemos entender que “la *voluntad* es un tipo de causalidad de los seres vivos en tanto que son racionales, y *libertad* sería la propiedad de esta causalidad para poder ser eficiente independientemente de causas ajenas que la *determinen*” (Ak. IV. 446), tal sería una concepción *negativa* de la libertad, en tanto contrasta con las determinaciones causales de la naturaleza, pero es relevante para el concepto positivo, en tanto nos indica que la voluntad es una causa, y en tanto “el concepto de una causalidad conlleva el de *leyes*” (Ak. IV. 446), entonces podemos colegir que la causalidad de la voluntad se rige bajo leyes de un orden distinto a las leyes de la naturaleza. En último término, debido a la deducción de la moralidad como autonomía del capítulo anterior en la cual la voluntad es una ley para sí misma, es que Kant sostiene que “una voluntad libre y una voluntad bajo leyes morales son exactamente lo mismo” (Ak. IV. 447). De lo anterior podemos entender que la libertad positiva, que como indica el subtítulo de la sección es la clave para explicar la autonomía, será el concepto que nos procurará luego el tercer término antes mencionado.

Luego Kant expresa que no es suficiente atribuir libertad a nuestra voluntad sin que podamos, al mismo tiempo, atribuírsela a todo ser racional, en tanto la moralidad es ley para todo ser racional. Dado esto afirma que “todo ser que no puede obrar sino *bajo la idea de libertad* es por eso mismo realmente libre” (Ak. IV. 448); es decir, un agente no puede concebirse a sí mismo como agente —como obrando— si es que al mismo tiempo no se presupone libre, porque de lo contrario sería un mero objeto pasivo dentro de una cadena de causas y efectos, no podría atribuirse autoría y no sería así un sujeto que determina racionalmente su acción, sino simple medio de sus impulsos.

Kant admite haber presupuesto la libertad luego de haber retrotraído el concepto de moralidad hacia ella, pero tal exposición no ha *demostrado* la realidad de tal idea. Y frente a ese vacío, podemos preguntarnos por qué habríamos de seguir el principio de la moralidad, incluso cuando no tengamos un interés que nos impulse a hacerlo. Una alternativa sería la de apreciar el valor de las leyes morales, pero para esto tendríamos que considerarnos ya sometidos a ellas, por lo cual no podemos aún entender “*desde dónde obliga la ley moral*” (Ak IV. 450). Aquí se muestra un aparente círculo vicioso: nos pensamos libres para vernos bajo leyes morales, y pensamos estar sometidos a tales leyes porque somos libres. Pero

tanto la auto-legislación (autonomía) y la libertad de la voluntad son lo mismo, no puede uno fundamentar el otro. Kant propone una posible salida: “indagar si cuando nos pensamos como causas eficientes a priori a merced a la libertad, no adoptamos algún otro punto de vista que cuando nos representamos a nosotros mismos según nuestras acciones como efectos que vemos ante nuestros ojos” (Ak. IV. 450).

Ese otro punto de vista constituye la salida. A partir de la distinción kantiana entre *fenómeno* y *cosa en sí*, es que se puede proporcionar una distinción entre *mundo sensible* y *mundo inteligible* (Ak. IV. 451). Si bien el sujeto puede entenderse a sí mismo como un fenómeno, al recibir los fenómenos externos, pero al mismo tiempo separándose de ellos, es que también se concibe como un yo que obra, que en último caso consiste en adscribirse como tal al *mundo inteligible*. Es la razón aquella capacidad para distinguirse del resto de cosas, es la razón la que distingue entre mundos, en tanto miembro del mundo sensible está sometido a leyes naturales, y como miembro del mundo inteligible es que se considera a sí bajo leyes racionales. Así, como perteneciente al mundo inteligible, piensa la causalidad de su voluntad como libre, es decir, independiente del mundo sensible, y la idea de libertad se liga a la de autonomía que a su vez se asocia con la moralidad. Sin embargo, ya que también nos consideramos como parte del mundo sensible —y no exclusivamente al mundo inteligible— como miembros de ambos nos pensamos sometidos al deber, quedando así descartada la sospecha del círculo vicioso.

Tal mundo inteligible es el fundamento del mundo sensible, en tanto, de lo contrario las acciones de este mundo serían incomprensibles como tales (como acciones realizadas por un agente), por tanto, también las leyes del primero —si bien colisionan con las del segundo— aparecen como imperativos para nosotros. Es así como es posible el imperativo categórico “gracias a que la idea de libertad me convierte en un miembro del mundo inteligible” (Ak. IV. 454), sin por ello ser exclusivamente parte de él —en tal caso las acciones siempre corresponderían a las leyes morales— sino también parte del mundo sensible, transformando un deber conforme a dichas leyes, a la autonomía. Esto incluso se ve confirmado en la razón humana común, como describe el propio autor.

IV

Hemos hecho una revisión general del argumento central de la GMS para nuestros fines, que resulta necesaria para comprender lo que viene. Es momento, entonces, de volver sobre lo dicho al comienzo respecto de la distinción entre

tipos de conocimiento para concluir en nuestra investigación. Como vimos, en distintos pasajes Kant habla explícitamente de conocimiento, y a veces marcadamente de conocimiento práctico. No debería haber duda entonces de que para Kant la *GMS* constituye en sí o indaga en torno a un conocimiento. Ha quedado también claro en qué medida el conocimiento práctico es práctico, es decir, en la medida en que expresa o es fundamento de un imperativo. Ahora falta la pregunta central: ¿En qué sentido es un conocimiento aquel conocimiento —práctico— que está en juego en la *GMS*?

Realicemos una distinción preliminar, una cuestión es si acaso el tener conciencia o saber qué es el IC (o los deberes que fundamenta) es un conocimiento, y otra cosa distinta es si acaso la investigación acerca de la posibilidad del IC —el objetivo del texto de buscarlo y establecerlo— constituye un conocimiento en sí mismo. En primera instancia, parece satisfactorio concluir que podemos conocer el IC, en tanto podemos expresar sus distintas formulaciones, y en la práctica —como hace notar el propio Kant— él ya está integrado de cierta manera en el conocimiento popular acerca de los deberes, aun cuando no se tenga conciencia de su formulación filosófica; y si bien tampoco fue asumida su existencia real, ya que estaba la alternativa de que fuese una simple quimera, finalmente quedó demostrada su posibilidad. Pero el hecho de haber expresado con claridad el IC en la *GMS* y haberlo analizado como base de los deberes no parece ser un conocimiento demasiado novedoso, más bien es una aclaración de algo ya presente, de un conocimiento moral común, y así se entiende el énfasis de Kant respecto a que los dos primeros capítulos son simplemente analíticos. En otras palabras, todas las personas para actuar ya tenemos un conocimiento práctico integrado que nos orienta y, por supuesto, esto puede ser analizado filosóficamente, pero aquí nos interesa más el CP filosófico que el mismo Kant intenta alcanzar, considerando el proyecto de filosofía que parece tener en el conjunto de su obra.

Entonces, cabe suponer que solamente en la tercera parte, teniendo como premisas las dos anteriores, habría algo así como un conocimiento práctico original que la *GMS* asienta. Sin embargo, ya vimos con anterioridad, a raíz de una nota a pie de página de la *KrV* (B XXVI N.), que la indagación en torno a la posibilidad de X, en este caso del IC (lo hecho en el tercer capítulo), es condición necesaria de su conocimiento. Pero ser condición necesaria del conocimiento no es conocimiento de X como tal, y ya dijimos que el IC ya era conocido en cierta manera. Sin embargo, también dijimos que es distinto conocer el objeto como tal que conocer su fundamento, que es a su vez establecer su posibilidad. No obstante, también mostramos en un comienzo que, justamente, para Kant los fundamentos del imperativo son parte del conocimiento práctico; por lo que podemos afirmar que al establecer el fundamento del IC, lo que lo asienta como principio supremo de la moralidad (objetivo de la obra), es que la *GMS* no es solo la investigación *acerca* de un CP tal como es el IC, sino que centralmente es un CP novedoso, al darnos

a conocer su fundamento. Lo que coincidiría también con lo afirmado por Kant tanto en las dos primeras críticas como en la *Lógica Jäsche*, tal como expusimos en un comienzo.

¿Qué es lo conocido? El fundamento del IC, que es la demostración de su posibilidad —lo que nos permite comprender que él mande— que también es la argumentación en torno a la distinción mundo inteligible/sensible, que asimismo tiene como premisa el análisis de la voluntad pura formalmente considerada, que a su vez supone la libertad “el único presupuesto bajo el cual es posible dicho imperativo” (Ak. IV. 461). Libertad que en último término no podemos conocer ni comprender, ya que excede los límites de la razón humana, sino solo suponer si pretendemos hacer sentido de nuestra experiencia moral; tal como tampoco podemos conocer jamás el mundo inteligible. Limitaciones ambas que son ineludibles en la búsqueda que hace la razón de lo incondicionado y, por tanto, limitaciones del conocimiento aquí expuesto.

Si mi interpretación de Kant ha sido correcta hasta aquí, entonces hemos contestado nuestra pregunta inicial.

Por último, podemos decir brevemente que lo anterior —aun considerando que la pregunta acerca del conocimiento dentro de la filosofía moral es distinto a la pregunta por la justificación de la moralidad—, de todas maneras, nos entrega ciertas luces en torno al largo debate en las interpretaciones realistas, versus constructivistas del pensamiento de Kant en sus distintas versiones y gradaciones. Tomaré para estos efectos lo afirmado por Korsgaard en su ensayo acerca de realismo y constructivismo, sin pretensiones de representar con esto el ingente debate anterior y posterior. Allí, ella nos dice que el realismo moral es una posición que concibe los conceptos y proposiciones morales como descriptivos de objetos o hechos intrínsecos del mundo; es decir, refieren a objetos realmente existentes (Korsgaard 302-303). Así, el realismo moral concibe la filosofía práctica como una tarea para descubrir o aclarar tales entidades normativas. Para ella, en último término esto sería no comprender la filosofía práctica, ya que la entenderían como una filosofía teórica que simplemente tendría que aplicar tales conocimientos a las acciones.

Si Kant fuese interpretado como un realista de la manera antes dicha, entonces para él, en la filosofía moral se descubriría la estructura objetiva, realmente existente, de la voluntad que llevaría a ver la constitución real de los conceptos morales que de ella se derivan, y luego podríamos aplicar aquellos conceptos para guiar nuestras acciones. Esto, por supuesto, es difícilmente defendible, en tanto el ejercicio kantiano no equivale a encontrar un hecho u objeto del mundo, sino a demostrar cómo es pensable la moralidad, cuyos fundamentos se muestran al indagar en la forma de la voluntad.

Por otro lado, ella expone la visión constructivista que atribuye a Kant —así como también a Rawls y defendida por ella misma—, en la cual lo que la distingue es que el énfasis de la filosofía práctica debe estar orientado a la solución de problemas prácticos, tal que sus conceptos surgen de aquellos mismos problemas y bajo sus propios términos y marcos condicionantes, siendo los conceptos morales los nombres para aquellas soluciones que el uso de nuestra razón *construye* en cierto sentido. Por ello, “La diferencia importante entre realismo y constructivismo no reside, entonces, en cuál de los dos es verdadero, sino en cómo nos orientan hacia las tareas de la filosofía práctica” (Korsgaard 325); por tanto, tampoco son necesariamente excluyentes, sino enfoques distintos en torno al mismo problema. En cualquier caso, la autora defiende que el problema de Kant es, justamente, el problema práctico “lo que hay que hacer” (Korsgaard 322), y al reflexionar en torno a la naturaleza de tal problema es que se deduciría el imperativo categórico, que nos daría la base para tal solución.

Recordemos que la visión constructivista de la filosofía kantiana proviene de Rawls (Rawls 235ss), y aunque la visión constructivista de Korsgaard es más sofisticada, pensamos que sería más preciso, si adoptáramos tal problema, el término *constitucionalismo* [constitutivism] *transcendental* para caracterizar el planteamiento de Kant, tal como propone Sensen (Sensen 219), como un derivado muy específico del constructivismo que no deje lugar al arbitrio personal sino que sea el reconocimiento de la naturaleza de la razón humana misma.

No obstante, con lo que hemos expuesto, debiese estar claro que ese no es tampoco el problema central de Kant en la *GMS*, sino que ella debe ser entendido como una investigación en torno al conocimiento trascendental⁷ —las condiciones de posibilidad de este— y, por ello, similar al problema de la *KrV*, solo que en clave práctica. Lo que sugiere que el problema central de la *GMS* podría reformularse diciendo *cómo es posible el conocimiento práctico* (cf. B 73), asunto que, como dijimos, es distinto al problema del debate realismo/constructivismo en torno a la justificación de la moralidad, pero de todas formas ayuda a su esclarecimiento. En último caso, no deseamos descartar de plano las interpretaciones realistas o constructivistas, cuyas múltiples variantes permiten flexibilidad al respecto; tampoco intentamos sugerir una suerte de cognitivismo práctico, ya que los problemas abordados son distintos, sino interrogar a Kant en sus propios

⁷ Tenemos a la vista el pasaje de la *KrV* en que Kant caracteriza el conocimiento trascendental, en tanto este “se ocupa, en general, no tanto de objetos, como de nuestra manera de conocer los objetos, en la medida en que ella ha de ser posible *a priori*” (B25), teniendo también en consideración que “no todo conocimiento *a priori* se debe llamar trascendental, sino sólo aquél por el cual conocemos que (y cómo) ciertas representaciones (intuiciones o conceptos) sólo se aplican *a priori* o sólo *a priori* son posibles (es decir, la posibilidad del conocimiento o el uso de él *a priori*)” (B80). Como hemos visto, este trabajo se ha ocupado justamente de mostrar en qué medida la *GMS* se ocupa de la posibilidad *a priori* del IC.

términos; intentando capturar el carácter deberían considerar el carácter *sui generis* de la obra de Kant en su propio tiempo.

Hemos puesto como objetivo de nuestra investigación mostrar en qué medida el principio establecido en la *GMS* constituye un conocimiento; para ello hemos comenzado por desentrañar la distinción entre conocimiento teórico y práctico que realiza el autor, lo que nos dejó con una concepción bastante específica. En segundo lugar, hemos analizado el hilo central a lo largo de la *GMS*, para finalmente mostrar que allí sí tiene sentido hablar del establecimiento de un conocimiento propiamente tal y, por tanto, tal obra puede leerse en aquella clave, la de la persecución de un tipo de conocimiento práctico en los términos especificados. A pesar de que las conclusiones son preliminares, esto nos ayudaría a comprender de mejor manera las posibles conexiones sistemáticas con otras obras, así como a mirar desde otra perspectiva el debate contemporáneo sobre realismo y antirrealismo en la filosofía moral de Kant.



Bibliografía

- Allison, Henry. *Kant's Groundwork for the Metaphysics of Morals: A Commentary*. Oxford: Oxford University Press, 2011.
- Engstrom, Stephen. "Kant's Distinction Between Theoretical and Practical Knowledge". *The Harvard Review of Philosophy* (2002): 49-63.
- Guyer, Paul. *Kant's Groundwork for the Metaphysics of Morals: A Reader's Guide*. Londres: Continuum, 2007.
- Hutter, Axel. *Das Interesse der Vernunft*. Hamburg: Felix Meiner Verlag, 2003.
- Kant, Immanuel. *Kant's gesammelte Schriften*. Herausgegeben von der Preussischen bzw. von der Deutschen Akademie der Wissenschaften. Berlín, 1902 y ss.
- . *Crítica de la facultad de juzgar*. Trad. Pablo Oyarzún. Caracas: Monte Ávila Editores, 2006.
- . *Crítica de la razón práctica*. Trad. Emilio Miñana y Villagrasa y Manuel García Morente. Círculo de Lectores, 1996.
- . *Crítica de la razón pura*. Trad. Mario Caimi. Buenos Aires: Colihue, 2007.
- . *Fundamentación para una metafísica de las costumbres*. Trad. Roberto Aramayo. Madrid: Alianza, 2012.
- Korsgaard, Christine. "Realism and Constructivism in Twentieth-Century Moral Philosophy". *The Constitution of Agency*, ed. Korsgaard, Christine. Oxford: Oxford University Press, 2008. 302-326.
- Rawls, John. *Lectures on the history of moral philosophy*. Cambridge: Harvard University Press, 2000.
- Santos, Robinson dos y Schmidt, Elke Elisabeth. *Realism and Antirealism in Kant's Moral Philosophy: New Essays*. Berlin, Boston: De Gruyter, 2018.
- Scarano, Nico. "Necessity and Apriority in Kant's Moral Philosophy". *Groundwork for the Metaphysics of Morals*, ed. Horn, Christoph y Schönecker, Dieter. Berlin: Walter de Gruyter, 2006. 3-22.
- Sensen, Oliver. "Kant's Constitutivism". *Realism and Antirealism in Kant's Moral Philosophy: New Essays*, ed. Santos, Robinson dos y Schmidt Elke Elisabeth. Berlin, Boston: De Gruyter, 2018. 197-220.



Inteligencia artificial

Editorial: Inteligencia artificial: cuatro planteos filosóficos Artificial Intelligence: Four Philosophical Approaches

TAMARA CHIBEY-RIVAS¹

Universidad Nacional de Córdoba, Córdoba, Argentina
tchibey@gmail.com

ROBERTO RUBIO²

Universidad Alberto Hurtado, Santiago, Chile
rorubio@uahurtado.cl

El presente número recibe su impulso inicial del coloquio “Inteligencia Artificial: desafíos presentes y futuros”. Dicho coloquio, realizado el 7 de diciembre de 2023, fue organizado por Tamara Chibey, Roberto Rubio y Martín Gendler y auspiciado por el Centro de Estudios Mediales de la Universidad Alberto Hurtado. El evento reunió a especialistas nacionales y extranjeros, quienes abordaron filosóficamente las implicancias inmediatas y futuras de los avances recientes en el campo de la inteligencia artificial generativa.

Entre las principales preguntas que se plantearon y desarrollaron en el coloquio, destacan las siguientes: 1) Al interactuar con máquinas operadas por sistemas de inteligencia artificial, ¿corresponde hacerlo desde una actitud intencional, orientada hacia una relación de yo-tú, o más bien hay que orientar nuestra comunicación con ellas a partir de la idea de que son máquinas diseñadas para fines específicos?; 2) ¿De qué manera se pueden limitar los sesgos de género en sistemas de IA que ofrecen asistencia en la toma de decisiones? ¿Cómo asegurar

CÓMO CITAR ESTE ARTÍCULO:

En APA: Chibey-Rivas, T. & Rubio, R. (2024). Inteligencia artificial: cuatro planteos filosóficos. *Resonancias*, (17), 81-84. DOI: 10.5354/0719-790X.2024.75334

En MLA: Chibey-Rivas, T., and R. Rubio. “Inteligencia artificial: cuatro planteos filosóficos.” *Resonancias*, n.º 17, julio de 2024, pp. 81-84. DOI: 10.5354/0719-790X.2024.75334

¹ Cursa un Doctorado en Filosofía en la Facultad de Filosofía y Humanidades de la Universidad Nacional de Córdoba en Argentina y es becaria de Conicet en el Instituto de Humanidades de la misma universidad. También es Investigadora Asociada del Centro de Estudios Mediales de la Universidad Alberto Hurtado en Chile. ORCID 0000-0003-3670-4598.

² Director del Centro de Estudios Mediales. ORCID 000-0002-2482-9062.

allí criterios éticos de transparencia, justicia, privacidad y responsabilidad desde una perspectiva de género?; 3) ¿Qué consecuencias políticas tiene el actual salto de escala en el procesamiento de datos, un salto que va del uso de datos directos e indirectos en smartphones y plataformas digitales a la creación y utilización de bancos de datos para el entrenamiento de grandes modelos de lenguaje y grandes modelos de agencia?; 4) ¿Qué efectos socio-políticos conlleva caracterizar los sistemas de IA mediante rasgos humanos como la cognición, por ejemplo? ¿Cómo pensar los ensamblajes computacionales operantes en los procesos de datificación social y propagación de plataformas y algoritmos?

El carácter filosófico del encuentro quedó plasmado en los siguientes rasgos de las ponencias y de las discusiones que les siguieron: la aproximación sociotécnica al estudio de los sistemas de IA, un profundo interés y preocupación por los alcances éticos y políticos de la implementación masiva de dichos sistemas, y un cuestionamiento de fondo dirigido a la imbricación entre técnica y cultura como clave para el análisis de la así llamada inteligencia artificial.

A partir de las ponencias y debates mencionados se redactaron los cuatro artículos que componen el presente número. El primer artículo, titulado “La IA entre nosotros: ¿qué debemos hacer?”, aborda las formas en las que los seres humanos interactúan con los sistemas con inteligencia artificial actualmente existentes. La autora Diana Pérez analiza las formas de interacción que están a la base de la consideración de tales sistemas como “inteligentes”, tomando en cuenta el rol programático del “test de Turing”. Con vistas a ello, destaca la siguiente tensión: “la IA es un objeto de diseño humano que nos obliga a no considerarlo un artefacto, una máquina, sino a considerarlo uno de nosotros”. Habría que desarrollar, entonces, frente a la IA, una actitud de diseño y a la vez una actitud intencional. Sin embargo, sostiene Pérez: “no podemos hacer las dos cosas simultáneamente”.

La autora divide en dos grandes grupos los avances actuales de la IA: por una parte, la creación de IAs que simulan las conductas de los seres humanos en la interacción con ellos en el entorno cotidiano social; por otra parte, máquinas más inteligentes que las mentes humanas que ayuden de manera eficiente y rápida en tareas y entornos específicos. Con vistas a estos avances, Pérez advierte los siguientes desafíos: en primer lugar, incentivar la regulación de los desarrollos tecnológicos relacionados con la inteligencia artificial; y en segundo lugar, limitar de hecho ciertos desarrollos de esa tecnología, bajo el entendido de que pueden generar un efecto adverso masivo.

En el artículo “Una mirada crítica a la ética de la IA: de preocupaciones emergentes a principios orientadores”, Gabriela Arriagada ofrece una reconstrucción crítica de la interdisciplina denominada “Ética de la inteligencia artificial”. La reconstrucción comienza por las preocupaciones que dieron impulso a los

debates iniciales, luego considera algunas críticas de peso relevantes para el campo de la ética de la IA y, por último, ofrece tres principios orientadores para su desarrollo actual y futuro.

Con respecto a las preocupaciones que dieron inicio a la ética de la IA en la última década, la autora cita el influyente artículo de Mittelstadt *et al.* (2016). En dicho artículo se identifican seis preocupaciones éticas: tres respecto a la evidencia que ofrecen los sistemas de IA (evidencia inconclusa, inescrutable y equivocada) y las otras tres referentes a los resultados injustos, los efectos transformativos y la trazabilidad, respectivamente.

Entre las principales críticas hacia el campo de la ética de la IA, Arriagada destaca los planteamientos acerca de la falacia aditiva y de la inutilidad de los principios. La falacia aditiva consiste en asumir erróneamente que la ética de la IA se alcanza añadiendo sin más enfoques tradicionales de la ética al estudio sobre la inteligencia artificial. Por su parte, la crítica acerca de la inutilidad de los principios fue planteada por Mittelstadt, y consiste en advertir sobre la estrategia de crear principios ambiguos y generales que no se hacen cargo de las correspondientes discusiones normativas y políticas. Teniendo en cuenta esas críticas, la autora ofrece los siguientes principios para orientar el desarrollo una ética aplicada a la inteligencia artificial: una mirada sociotécnica que reconoce la interdependencia entre IA y contextos sociales, estructuras regulatorias y educativas para una integración efectiva de la ética y la necesidad de estándares profesionales que den marco a la práctica ética.

El tercer artículo pertenece a Martín Gendler y se titula “Datificación Social e Inteligencia Artificial: ¿hacia un nuevo ‘salto de escala’?”. El autor analiza los desarrollos de la inteligencia artificial y su relación con los procesos de datificación social. Para ello, ofrece una breve genealogía de tales procesos en conjunto con un análisis de las distintas fases del desarrollo de IA. Entre los principales momentos señalados por dicha genealogía, cabe destacar los siguientes: a) el sistema de recolección, almacenamiento, procesamiento, perfilamiento y aplicación (R.A.P.P.A) como la lógica principal de la operatoria de las principales empresas de Internet y de los objetos técnicos digitales (OTD); b) la industria 4.0 y la correspondiente configuración de objetos técnicos digitales, a los que cabría llamar OTD 4.0. Gendler advierte una relación de doble reciprocidad entre el proceso de datificación y los desarrollos de IA, “dado que el primero opera como condición de base para el entrenamiento y desarrollo de los segundos, mientras que estos, fomentan, optimizan y perfeccionan los diversos mecanismos de recolección, almacenamiento y desarrollo de datos, luego utilizados para los procesos de perfilado y aplicación”.

La propuesta genealógica se despliega en una doble perspectiva: por una parte, Gendler analiza los avances realizados en los sistemas de IA como un elemento clave para potenciar la datificación; por otra parte, nos da a pensar acerca de la

posibilidad de un nuevo “salto de escala” en los actuales procesos de personalización y modulación algorítmica.

El último artículo pertenece a Javier Blanco y lleva por título “Cognición computacional”. Blanco plantea que la noción de cognición computacional, tomada de William Rapaport, es más precisa que la de inteligencia artificial, especialmente para caracterizar los sistemas de *machine learning*.

Blanco advierte acerca de la confusión que suele dominar actualmente las discusiones y análisis acerca de las tecnologías de IA, y propone distinguir entre diversos niveles y tipos de mediaciones computacionales. De ese modo, argumenta, se evitaría “una frecuente falacia de composición que consiste en tomar propiedades específicas de algún subconjunto determinado, en particular a los sistemas producidos por *machine learning*, como aplicables a cualquier sistema computacional”.

Con respecto a la tecnología del *machine learning*, Blanco advierte sus limitaciones, en particular respecto al margen de error en los resultados que se ofrecen, dado el carácter estadístico de los métodos de procesamiento. Sin embargo, no descarta un eventual salto conceptual que traiga consigo otro tipo de programa de aprendizaje maquínico que se automejore. Blanco entiende al *machine learning* como un método de metaprogramación situado en lo que, con Dijkstra, denomina el interjuego entre programación y computación. Este interjuego posibilita un paisaje de oportunidades evolutivas que contribuye a desarrollar y renovar los acoplamientos cognitivos. La propuesta de Blanco apunta a favorecer la coevolución entre seres humanos e IA o, más en general, entre humanos y programas.

A fin de profundizar en la noción cognición computacional, el autor problematiza la distinción entre cognición computacional y cognición humana planteada por Brian Cantwell Smith. Según esto, tal distinción radica en la capacidad de “juicio”, la cual da cuenta de un compromiso existencial humano. Frente a ello, Blanco se pregunta “cómo se establece hoy un compromiso existencial con un mundo poblado de clasificadores estadísticos, cómo estos clasificadores integrados ya a la cotidianidad, a la toma de decisiones, a la cultura, afectan y transforman la capacidad ¿humana? de juicio y sus alcances”.

Para finalizar, deseamos agradecer a las autoras y los autores de los artículos, como así también al Centro de Estudios Mediales, institución que hizo posible el coloquio que dio impulsos para la redacción de este Dossier: *Last but not least*; vaya nuestro sincero agradecimiento a *Resonancias. Revista de Filosofía* por invitarnos a organizar y editar el presente Dossier.



La IA entre nosotros: ¿qué debemos hacer?

AI Among Us: What Should We Do

DIANA INÉS PÉREZ¹

Universidad de Buenos Aires; IIF-SADAF-CONICET

Fecha de Recepción: 08/05/2023

Fecha de Aceptación: 19/10/2023

Resumen

En este trabajo me concentro en las maneras en las que los seres humanos interactuamos con los sistemas con IA actualmente existentes en nuestras sociedades. Muestro que son cierto tipo de formas de interacción las que están en la base de la consideración de tales sistemas como “inteligentes” dado el origen mismo del desarrollo de la IA en base al “Test de Turing”. Distingo dos tipos de desarrollos y sus efectos, tomando como criterios las formas de interacción que estos sistemas demandan. Finalmente, destaco algunos de los desafíos que enfrentamos en nuestros días, el rol central que las humanidades y ciencias sociales tendrán en este ámbito, y sugiero algunas líneas posibles de acción.

CÓMO CITAR ESTE ARTÍCULO:

En APA: Pérez, D. I. (2024). La IA entre nosotros: ¿qué debemos hacer? *Resonancias*, (17), 85-100. DOI: 10.5354/0719-790X.2024.74394

En MLA: Pérez, D. I. “La IA entre nosotros: ¿qué debemos hacer?” *Resonancias*, n.º 17, julio de 2024, pp. 85-100. DOI: 10.5354/0719-790X.2024.74394

¹ Doctora en Filosofía de la Universidad de Buenos Aires. Investigadora Principal del CONICET. Directora del IIF-SADAF-CONICET. Profesora Titular regular de “Metafísica” y Profesora Asociada Regular de “Fundamentos de Filosofía” del Departamento de Filosofía, Universidad de Buenos Aires. Directora del programa de actualización de posgrado “Inteligencia artificial desde una perspectiva humanística” (UBA). Especialista en metafísica de la mente y filosofía de la psicología, con competencia en filosofía del lenguaje y filosofía del arte. <https://orcid.org/0000-0002-6185-7986>

Palabras clave: test de Turing, robots sociales, interacción humano-máquina, regulación, segunda persona

Keywords: Turing Test, social robots, human-machine interaction, regulation, second person

Abstract

In this work I focus on the ways in which human beings interact with AI systems currently existing in our societies. I show that certain types of forms of interaction are the basis for considering such systems as “intelligent” given the very origin of the development of AI based on the “Turing Test”. I distinguish two types of developments and their effects, taking as criteria the forms of interaction that these systems demand. Finally, I highlight some of the challenges we face today, the central role that the humanities and social sciences will have in this area, and I suggest some possible lines of action.

Cuando escuchamos la expresión “inteligencia artificial”, nos vemos llevados a pensar, seguramente por la influencia de las películas de ciencia ficción que consumimos, en robots que conviven con seres humanos como en *Yo Robot*, en grandes máquinas que nos esclavizan y dominan como en *Matrix*, en cyborgs como *Terminator* o en voces que nos seducen como en la película *Her*. Son todos futuros que imaginamos, pero que no están tan cercanos (y, en algunos casos, incluso son tecnológicamente imposibles en este momento). Sin embargo, esto no quiere decir que la IA esté lejos de ser una realidad en nuestras vidas cotidianas; por el contrario, está hoy entre nosotros y es mucho más ubicua -y mucho más imperceptible- de lo que suponemos. Usamos muchos desarrollos tecnológicos que tienen entre sus componentes algoritmos de IA. Para dar solo algunos ejemplos más o menos conocidos: los algoritmos de recomendación de películas o música en *streaming* (como *Netflix*, *Spotify*, etc.), las Apps que nos permiten geolocalizarnos y nos indican qué camino seguir (*Google Maps*, *Waze*, etc.), los asistentes de voz (como *Alexa*, *Siri* y *Cortana*) y el archiconocido *ChatGPT* así como generadores de imágenes como *DALL-E* son desarrollos que involucran IA. También puede haber IA en los sistemas de diagnóstico a través de imágenes médicas, en la posibilidad de mejorar la calidad del foco en la cámara fotográfica de nuestro celular, en los modelos epidemiológicos o meteorológicos que se usan para prevenir catástrofes y desarrollar políticas públicas, en vehículos y drones autónomos y robots de compañía, por mencionar solo algunos de los muchos que nos rodean. En este trabajo me voy a centrar en este tipo de desarrollos actuales de la IA, para reflexionar acerca del lugar que ocupan en nuestras vidas y para discutir la pregunta relativa a qué podemos hacer y qué actitud debemos tomar ante estos desarrollos, nosotros, los seres humanos que habitamos el mundo actual, en las sociedades actuales. Voy a dejar la ciencia ficción y la futurología que conlleva para otra ocasión.

La perspectiva que adoptaré para dar respuesta a esta pregunta es la siguiente: voy a focalizar mis reflexiones en la forma en la que *nosotros*² interactuamos con la IA. Es importante destacar que nuestra forma de interacción con estos dispositivos es, justamente, aquello que le otorga su estatus de máquina inteligente. En efecto, la idea de que ciertas máquinas sean inteligencias artificiales ha quedado asociada al hecho de que puedan interactuar con nosotros de forma humana. Tal es la idea central que desarrollara provocativamente Turing en 1950 cuando, por primera vez, se buscó dar una respuesta a la pregunta “¿Pueden pensar las máquinas?”. En lo que sigue, en primer lugar, voy a presentar el “test de Turing” que sigue siendo una de las canónicas piedras de toque para detectar inteligencia artificial (sección I). Mostraré una tensión que se deriva de este test en lo relativo a nuestras formas de interacción con la IA (sección II). Seguidamente, dividiré en dos grandes grupos los desarrollos de IA actuales, usando como criterio de distinción las formas en las que se espera que nosotros interactuemos con ellas (sección III). Finalmente, extraeré algunas conclusiones acerca de los desafíos que estos dos tipos de desarrollos nos plantean (sección IV).

I. El Test de Turing y nuestra interacción con “máquinas inteligentes”

Para responder a la compleja cuestión acerca de si las máquinas pueden pensar, Turing (1950) propuso reemplazar esta pregunta por otra. La nueva pregunta que formuló era si las máquinas en cuestión podrían inducir a un sujeto humano a creer que estaba interactuando (más específicamente, dialogando) con otro sujeto humano. Es decir, ofreció una “operacionalización” del pensamiento, a través de la conducta (exclusivamente) lingüística de la máquina en el contexto de una interacción dialógica con un ser humano. Si la máquina era capaz de “engañar” al ser humano, se podría considerar que estaba desarrollando una conducta inteligente y que, por lo tanto, era una máquina pensante. Quisiera detenerme en algunos de los detalles de esta propuesta que serán importantes para lo que sigue.

En primer lugar, el test de Turing está basado *exclusivamente* en la conducta lingüística de la máquina.³ Para ello, Turing propuso diseñar un artefacto capaz de entablar un diálogo que no fuera cara a cara, sino mediado tecnológicamente

² Cada vez que use la palabra “nosotros” voy a estar refiriéndome a seres humanos de todos los géneros, etnias, religiones y edades que viven en la actualidad, segunda década del siglo XXI, en las sociedades globalizadas de hoy.

³ Tal vez por eso ha causado tanto revuelo el ChatGPT, y cada vez que alguna versión de este test es pasada exitosamente por alguna computadora, se afirma que las computadoras ya llegaron a ser seres pensantes.

(como los que hoy tenemos cuando chateamos con otra persona o con los *chatbots* de las páginas web de empresas y organismos públicos), con lo cual, no solo se trata de una conducta exclusivamente lingüística, sino además de lenguaje escrito. Asimismo, ese diálogo, para parecer humano, debía incorporar el *timing* propio del procesamiento lingüístico y cognitivo de un humano (por ejemplo, tardar más en responder cuánto es $864 + 1896$ que cuánto es $2+3$), debía estar desprovisto de todo elemento expresivo (tonos de voz y entonaciones específicas) lo que queda garantizado al ser lenguaje escrito y, por supuesto, debía estar programado para “esconder” todos los rasgos no humanos de la máquina, incluyendo su aspecto físico; en síntesis, tenía que *estar diseñada para “engañar” al ser humano en la interacción*.

Como resulta evidente, de acuerdo con el test de Turing, la máquina no es inteligente ni piensa en virtud del material del que está hecha (su *hardware*), ni tampoco en virtud del diseño específico de su *software* (*i.e.* es indiferente si el engaño se logra con un sistema de reglas y representaciones, redes neuronales o cualquier otro algoritmo surgido en algún futuro). Lo relevante para pasar el test no es lo que hay adentro de la caja negra, sino la conducta exhibida por el sistema. Y tampoco es relevante cualquier conducta: lo que importa es la forma de *interacción (lingüística)* que puede llegar a sostener *con los seres humanos*.

Finalmente, la máquina es inteligente porque actúa *como lo haría un ser humano*. No es ni mejor (más inteligente) ni peor (más tonta) que un ser humano. Tenemos que vernos llevados a atribuir intencionalidad, conciencia, pensamiento a la máquina en base a nuestras interacciones con ella para que podamos afirmar (según el test de Turing) que los algoritmos piensan o son inteligentes; es decir, las máquinas deben exhibir el tipo de conducta que usualmente comprendemos adoptando hacia ellos lo que Dennett (1981, 1987) denominó la actitud intencional.

II. ¿Qué actitud adoptar ante la IA? Una discusión dennettiana

Las “maquinas inteligentes” de Turing (1950) son, entonces, artefactos *diseñados* por los seres humanos para que interactúen con nosotros como lo haría otro ser humano induciéndonos, de esta manera, a *atribuirle inteligencia y pensamientos*. La expresión “inteligencia artificial” fue acuñada un poco más tarde, en 1956, cuando se reunieron en el *Dartmouth College* los fundadores de esta disciplina incluyendo, entre otros, a John McCarthy, Marvin Minsky, Claude Shannon, Allen Newell y Herbert Simon. Según Russell y Norvig (2008), dado el tenor de la reunión, hubiera sido preferible el término “racionalidad computacional”

para denominar los desarrollos propuestos. Sin embargo, no solo la racionalidad humana fue objeto de indagación desde los inicios de la IA, sino también otras facultades humanas como “la creatividad, la auto-mejora, y el uso del lenguaje” (Russell y Norvig 2008: 21), además de la capacidad clasificatoria y matemática, el aprendizaje, la detección de patrones y la resolución de problemas, entre otros. En síntesis, el proyecto de la IA, desde sus inicios busca duplicar las facultades psicológicas humanas en una estofa no biológica, sino artificial.⁴ Y la única forma de reconocer capacidades psicológicas humanas en algo es a través de la conducta públicamente observable. Recordemos que todos los proyectos científicos, desde las más diversas disciplinas, que buscaron avanzar en el conocimiento de la mente humana durante el siglo XX tomaron como punto de partida la conducta pública, es decir, la perspectiva de tercera persona o punto de vista objetivo. El acceso científico a la mente está basado en protocolos experimentales en los que la conducta observable resulta ser la base empírica sobre la que la teoría psicológica es validada. De hecho, la visión acerca de la mente humana que dominó la academia anglosajona durante la segunda mitad del siglo XX fue el denominado “funcionalismo” (Putnam 1981, Lewis 1972), que culminó con el desarrollo de la así llamada teoría computacional representacional de la mente humana (Fodor 1987).

Esta visión de la mente tenía una serie de presupuestos que posibilitaron tanto el desarrollo de la IA como el de las ciencias cognitivas en su conjunto.⁵ El funcionalismo sostiene la autonomía del nivel psicológico (*i.e.* la autonomía del estudio de las facultades psicológicas humanas) respecto del nivel biológico. Esto posibilitó el estudio de las facultades cognitivas (por ejemplo, el desarrollo de los estudios chomskianos sobre la facultad del lenguaje) con independencia de los estudios de los mecanismos neurobiológicos involucrados en la *performance* lingüística. Lo mismo ocurrió en otros ámbitos de los estudios cognitivos de la mente, por ejemplo, en la visión, la capacidad conceptual, la teoría de la mente, etc. Esta misma autonomía de lo psicológico respecto de lo biológico está en la base del desarrollo de la IA: se busca diseñar algoritmos que reproduzcan la conducta humana en la resolución de las más variadas tareas “inteligentes”, sin necesidad de preocuparnos por el material en el que estos diseños se realizan.⁶ Así, el

⁴ Russell y Norvig hablan de *duplicar*; otros autores prefieren hablar de *simular* o *modelar* las mentes humanas. Esta discusión terminológica está asociada a la cuestión metafísica de si eso que hacen las máquinas es o no es “realmente” pensamiento. Como estoy enfocándome en la forma en la que nosotros interactuamos con las máquinas, dejo de lado en este trabajo esta cuestión metafísica, y voy a usar indistintamente todos estos términos para referirme a los estados “mentales” de las máquinas; es decir, a los estados mentales atribuidos por nosotros a las máquinas, más allá de la cuestión metafísica de si esta atribución se corresponde con algún proceso que genuinamente podamos considerar un proceso mental, con independencia de nuestras prácticas atributivas.

⁵ Recordemos que las ciencias cognitivas son un proyecto interdisciplinario que incluye la IA entre otras disciplinas, como la psicología, la filosofía, la lingüística, etc.

⁶ En realidad, la base material sí que importa porque determina, por ejemplo, la velocidad de procesamiento, la capacidad de memoria, etc. El desarrollo de la IA no es *solamente* el desarrollo de algoritmos matemáticos, sino que es

desarrollo de duplicados, simulaciones o modelos computacionales capaces de realizar eficientemente tareas cognitivas humanas corrió en paralelo al estudio empírico de las capacidades cognitivas humanas.

Es importante remarcar que la teoría computacional representacional de la mente, tomando como caso paradigmático el lenguaje humano y la hipótesis del lenguaje del pensamiento, fue un paradigma experimental que acumuló mucha evidencia en su favor en los primeros años del desarrollo de las ciencias cognitivas en variadas áreas de investigación, incluyendo facultades de nivel inferior como la visión (Marr 1982). Sin embargo, los desarrollos de la IA que toman como punto de partida las capacidades superiores y los pensamientos proposicionales, rápidamente se encontraron con limitaciones de desarrollo en su aplicación a las capacidades más básicas.⁷ Rápidamente, en el ámbito computacional se comenzó a desarrollar un paradigma alternativo, basado esta vez en las computaciones y tráfico de información que el cerebro humano maneja (en lugar de basarse en la mente humana). Nacían así las redes neuronales y los modelos conexionistas de la cognición. No me voy a detener en esta disputa; simplemente quiero remarcar que, en ambos casos (es decir, las teorías computacionales representacionales proposicionales y los modelos conexionistas), la idea que motoriza el desarrollo es la misma: que se puede diseñar un sistema artificial que duplique o simule las capacidades psicológicas humanas con independencia de lo biológico. La mente/cerebro es así independiente de la materia en el cual se realice (e independiente también de las características del cuerpo en el que se realice). Y, en ambos casos, es la conducta observable lo que resulta la base evidencial de la atribución de intencionalidad o mentalidad a un sistema (biológico o no).

Como señalé al finalizar el apartado anterior, si lo pensamos desde el punto de vista de nuestra interacción con estas máquinas inteligentes, resulta importante señalar que se trata de máquinas diseñadas por nosotros, para que nosotros adoptemos hacia ellas lo que Dennett 1987 denominó la “actitud intencional”. En efecto, estas máquinas están diseñadas de tal manera que nosotros nos veamos obligados a atribuirles deseos, creencias y otros estados psicológicos si queremos tener éxito en nuestra interacción. En palabras de Dennett:

a la par el desarrollo de materiales y tecnologías que ejecutan los algoritmos, almacenan y transmiten información. Pero este desarrollo de la materialidad de la IA no está atada al conocimiento de la materialidad en la que los pensamientos humanos se ejecutan (es decir, en el cerebro), ni busca asemejarse a ella (a diferencia de la mente artificial, el programa, que busca asemejarse a la mente humana). Es en este sentido que la materialidad “no importa”, quiere decir que podría ser cualquiera, la que sea necesaria para que los algoritmos “corran”.

⁷ El problema del marco y el problema de la explosión computacional fueron los primeros en popularizarse en contra de esta versión de la IA de “reglas y representaciones”.

Las movidas futuras de las mejores máquinas computadoras que juegan al ajedrez en la actualidad no pueden predecirse usando la actitud de diseño ni la actitud física; se han vuelto tan complejas que ni sus diseñadores pueden evaluarlas desde la actitud de diseño. La mejor manera de vencerlas es predecir sus respuestas preguntándose cuál sería la movida más racional que la máquina podría hacer, dadas las reglas y los objetivos del ajedrez. Es decir, debemos asumir ... que la computadora “elegirá” la movida más racional. (Dennett 1981)⁸

Pero adoptar la actitud intencional supone rechazar la actitud de diseño, es decir, dejar de considerar a la máquina inteligente como una *máquina diseñada* por alguien para realizar una determinada función o propósito, o sea, para resolver un problema que la persona que la diseñó percibió como tal, y que buscó solucionar con ese diseño.⁹ En este sentido, la IA es un objeto de diseño humano que nos obliga a no considerarlo un artefacto, una máquina, sino a considerarlo uno de nosotros. Y, es importante destacar, no podemos hacer las dos cosas simultáneamente: si bien es cierto que podemos alternar (más o menos voluntariamente) la actitud que adoptemos, no podemos sostener las dos simultáneamente (así como no podemos ver al mismo tiempo un pato y un conejo en una figura ambigua). Más aún, una de esas dos actitudes es la actitud *por default*, es decir, la que adoptamos irreflexivamente en virtud de las conductas que observemos en el sistema. Así, en virtud del diseño, la IA nos lleva a adoptar hacia ella, irreflexivamente, la actitud intencional.¹⁰

En mi opinión, la actitud intencional solo recoge una parte de lo involucrado en nuestras formas de comprensión de las mentes humanas en la interacción cara a cara. La perspectiva de segunda persona (Pérez y Gomila 2022) presenta una visión más acertada de los múltiples aspectos multimodales involucrados en la interacción exitosa. No solo lo expresado lingüísticamente cuenta a la hora de atribuir estados mentales,¹¹ sino también importa el tono de voz, la prosodia, las pausas, la dirección de la mirada, las expresiones faciales y corporales, por

⁸ La traducción es mía.

⁹ Adoptar la actitud de diseño hacia un sistema S, supone aceptar que S tiene un diseñador, un agente intencional que tiene propósitos y buscar crear un objeto con una función específica, para resolver un problema que tiene el diseñador. Por ejemplo, un termostato es un objeto diseñado por seres humanos para regular la temperatura de los ambientes. Adoptando la actitud de diseño, el comportamiento del sistema bajo consideración no es visto como inteligente, sino adecuado o no para cumplir el propósito del agente inteligente, que en este caso es solo el diseñador.

¹⁰ Si alguien tiene alguna duda, pensemos en la sensación que tenemos de que hay otra persona (otra conciencia) del otro lado del ChatGPT cuando fluye la interacción lingüística, sensación frenada *por diseño* cuando ante ciertas preguntas el chat nos responde “No puedo responder esa pregunta, soy un sistema de IA diseñado por la empresa OpenAI...”. Alternativamente pensemos en cómo nos irrita el mal diseño de un *chatbot* cuando buscamos respuesta para un trámite en línea en nuestro banco, compañía de comunicaciones o municipio, y no nos da la opción que le pedimos, nos responde con opciones prediseñadas que no nos sirven, etc. No podemos evitar sentir otra conciencia presente cuando la interacción fluye de manera humana y no podemos evitar irritarnos por el mal diseño cuando no fluye la interacción y deseamos un ser humano del otro lado.

¹¹ Contrariamente a lo que el test de Turing presupone.

mencionar solo algunos elementos obvios. Así, desde esta perspectiva podemos atribuir una gama más amplia de estados psicológicos (no solo los que dependen de las normas de racionalidad central a la actitud intencional dennettiana),¹² dado que parte de todas las expresiones corporales (incluyendo las emocionales) y está fundada en nuestra sensibilidad afectiva (por lo que se trata de una atribución valorativa, no neutra como la propia de la actitud intencional). Desde esta perspectiva, la presencia de elementos lingüísticos es opcional (y no se da necesariamente en el caso canónico del que parte la segunda persona: las interacciones adulto-bebé en el primer año de vida). Tal como sostiene Dennett respecto de la actitud intencional, la adopción de la perspectiva de segunda persona resulta inevitable en las interacciones humanas: se trata de la forma de comprensión humana más básica y perdurable a lo largo de toda nuestra vida, y la aplicamos no solo a los seres humanos con los que interactuamos sino también a mascotas, juguetes, personajes de ficción y, por supuesto, a artefactos con el nivel de sofisticación adecuado, como nuestras computadoras, robots, etc.

III. La IA entre nosotros: ¿personas o herramientas?

Ahora bien, todos somos conscientes de que no es el propósito de los desarrollos más interesantes en IA crear máquinas que sean *exactamente* como los seres humanos. En realidad, el objetivo más publicitado en la industria de la IA es crear artefactos con una inteligencia *superior* a la humana, máquinas que realicen las mismas tareas que los seres humanos, pero que lo hagan de una manera más eficiente, con menos errores, más rápidamente, etc. Es la existencia de este tipo de desarrollos de la IA la que nos lleva a considerar la posibilidad de ser reemplazados por ella. Así, los desarrollos actuales de IA se inclinan en dos direcciones opuestas: (1) la creación de máquinas que simulen las conductas humanas apropiadas para la interacción con seres humanos *in the wild*, es decir, en nuestro entorno social cotidiano; (2) la creación de sistemas *mejores* que los seres humanos, que realicen las tareas asignadas de manera más eficiente y rápida que los seres humanos en entornos específicos para los que fueron especialmente diseñados, aun cuando no interactúen con nosotros como lo hacen

¹² En el fondo, como señalan Russell y Norvig, parece ser más bien la racionalidad humana y no la mente humana lo que la IA busca duplicar. Es cierto que nos autocomprendemos como animales racionales, pero la mente humana no está constituida exclusivamente por la racionalidad lógico-matemática, lingüística, abstracta, sino también por múltiples estrategias inteligentes de supervivencia en la interacción con el entorno físico y social. Todos estos aspectos son los que resultan ajenos al test de Turing y a la IA concebida como modelo de la mente (racional) humana. La mente de un animal racional es mucho más que lenguaje escrito en un *display* y nuestra comprensión de estas mentes depende de mucho más que de la interacción por medio del lenguaje.

los otros seres humanos. Estos dos tipos de desarrollos, como es claro, suponen diferentes formas de interacción con nosotros: en un caso, se busca una interacción social humana típica; en el segundo caso se busca un reemplazo de un humano por una máquina para la realización de una tarea específica (queda abierta la cuestión de cómo habremos de tratar a este tipo de máquinas, dado que resuelven tareas inteligentes humanas, pero no está diseñadas para actuar como un ser humano en el contexto de la interacción, con lo que parecen quedar relegadas al ámbito del diseño artefactual).

Entre aquellos diseños de IA que se orientan a la interacción a la manera humana, se encuentran los robots sociales, que buscan reemplazar a seres humanos que realizan tareas de cuidado (de adultos mayores que viven solos, de infantes). También serían “mejores” que los seres humanos en ciertos aspectos: no se cansan ni necesitan dormir, por lo que realizarían su tarea los 365 días del año, las 24 horas del día y nunca se “olvidan” de acercarle la medicación o alimentar, siempre en el momento apropiado, a sus dueños. Algunos de ellos tienen aspecto humanoide,¹³ otros están diseñados más bien a la manera de animales de compañía. En algunos casos son simples experimentos de laboratorio (todavía no hay robots que puedan interactuar en nuestra vida cotidiana como Robotina, de los Supersónicos -los *Jetsons*- o como los robots que sustituyen a una pareja muerta, como ocurre en algún capítulo de *Black Mirror*). En otros casos han sido probados en contextos terapéuticos,¹⁴ y algunos pueden ser adquiridos hoy para compañía en portales de venta en línea. Estos artefactos incorporan en sus diseños formas de interacción a la manera humana que van más allá de lo lingüístico; por ejemplo, respondiendo al tacto con un sonido amistoso, adaptando sus respuestas a las preferencias de su usuario, interactuando corporalmente a través de movimientos y sonidos similares a los que se dan en interacciones con seres humanos (o mascotas, según corresponda).

Sin duda, la incorporación de estos robots en nuestras vidas abre la pregunta acerca de cómo se verán afectadas, cómo se alterarán las formas actuales de interacción humana y de organización social. Pero me gustaría destacar que el éxito de estos robots sociales parece depender de que efectivamente nos engañen y nos hagan creer que no son robots sino seres humanos (o mascotas). Así surge un dilema para el diseño y la introducción de estos artefactos en las sociedades humanas: o bien se diseñan dejando claro su carácter robótico (no solo en su apariencia, sino también en sus formas de interactuar con seres humanos), en cuyo caso nunca serán realmente otros como nosotros (pero no seremos engañados), o

¹³ Por ejemplo, Myon, desarrollado por el *Neurobotics Research Laboratory* del *Berliner Hochschule für Technik* (“Neurobotics Research Laboratory (NRL) - Myon”).

¹⁴ Por ejemplo, Paro, una foca terapéutica desarrollada por la industria japonesa AIST. “PARO Therapeutic Robot”. *PARO Therapeutic Robot*, www.parorobots.com. Accedido el 16 de junio de 2024

bien se diseñan para que parezcan y actúen como seres humanos (o mascotas) y nos estarán engañando.

El segundo caso que quiero analizar es el de los sistemas con IA orientados a la resolución de una tarea puntal de una manera más eficiente, rápida y desprovista de errores que la humana, tales como muchos de los desarrollos que mencioné al iniciar este trabajo: sistemas que nos orientan en el espacio para viajar por el camino más rápido, con menos congestión de tránsito, que traducen un texto de una lengua que no conocemos a la nuestra, que diagnostican enfermedades o que deciden a quién otorgarle un crédito o un puesto de trabajo, etc. Sin duda, la mayor parte de los sistemas con IA que nos rodean hoy podríamos incluirlos aquí. Respecto de este tipo de desarrollos me gustaría señalar varias cosas.

En primer lugar, cada uno de estos desarrollos sería, en principio, excelente (mejor que los seres humanos) *para una tarea específica*. Pero estamos muy lejos de lograr que un sistema que es muy bueno en algo pueda también ser muy bueno en otra cosa. Es decir, si entrenamos un algoritmo para reconocer tumores en imágenes de ultrasonido, ese mismo algoritmo no estará *ipso facto* calificado para reconocer tumores malignos en fotografías tomadas de la piel de un paciente ni, mucho menos, para reconocer caras de personas en cámara de vigilancia. Para cada propósito específico hay que entrenar un algoritmo diferente. En segundo lugar, cada tarea específica estará orientada por un problema humano que se busca resolver con esta herramienta, por lo que tanto el planteo del problema como las soluciones que se consideran correctas del problema son problemas y soluciones humanos. Así, un algoritmo está entrenado para reconocer el tipo de tumores que conocemos y queremos diagnosticar para tratar, pero no cualquier otra cosa en las imágenes; como sabemos, las formas alternativas de cortar al mundo en partes son muchas, pero solo algunas de ellas son significativas para nosotros (Dupré 1993). En tercer lugar, todo mecanismo clasificatorio tiene falsos positivos y falsos negativos, por lo que hay que tomar decisiones de diseño respecto de qué tasa de casos como estos es aceptable.¹⁵ Esto representa un importante desafío no desprovisto de consecuencias en las vidas humanas (pensamos en los sub- o sobre-diagnósticos y lo que produciría en el sistema de salud y en las vidas de los pacientes). En cuarto lugar, todo sistema clasificatorio tiene sesgos, es decir, prejuicios y presupuestos explícitos o implícitos que forman parte constitutiva del sistema dadas las decisiones de su diseño. Estos sesgos son esenciales para que el algoritmo funcione apropiadamente (por ejemplo, que detecte los tumores que conocemos y sabemos cómo tratar), pero puede resultar un problema si buscamos clasificar personas para entrevistarlas para un trabajo o para otorgarles libertad provisional, porque esas clasificaciones van a depender

¹⁵ Son decisiones que toma el diseñador del sistema; es decir, la empresa que lo diseña y comercializa.

del entrenamiento que haya tenido el sistema en base a los casos anteriores, fundados en los sesgos humanos propios de las prácticas humanas relevantes, y esos ejemplos de entrenamiento del algoritmo pueden estar teñidos de prejuicios y sesgos raciales, de género, etc. que no quisiéramos repetir. En otras palabras, nos basamos en nuestro conocimiento anterior de tipos de tumores y tipos de personas para entrenar los algoritmos, pero mientras que el primer sistema clasificatorio está basado en conocimiento provisto por la comunidad científica, el segundo está basado en prácticas humanas que podemos cuestionar. Más aún, en estos casos las consecuencias sociales son mucho más complejas, profundas e injustas. Un mal diagnóstico (un falso positivo o falso negativo) que nos pueda dar la IA se puede complementar con otros medios diagnósticos, clínicos, análisis de sangre, biopsias, etc., de tal manera que la decisión final de diagnóstico quedará en manos del médico (o médicos, siempre existe la posibilidad de la interconsulta), así como la decisión de qué tratamiento realizar que queda en manos del médico en acuerdo explícito con el paciente. En cambio, la exclusión de alguien de un potencial trabajo, el rechazo al pedido de libertad condicional de alguien o el otorgamiento indebido de libertad condicional a quien volverá a delinquir, o el rechazo de un pedido de crédito bancario para iniciar un emprendimiento, en tanto sean decisiones delegadas a sistemas de IA, causan un impacto irreversible en las personas afectadas y, por lo tanto, en la sociedad en su conjunto, sin que ellas puedan participar de decisión alguna.

Una vez más, la incorporación en nuestras vidas y sociedades de estos sistemas no es inocua. Trae muchos beneficios y puede ser muy útil, pero hay que ser consciente de sus limitaciones y posibles errores (sea por sesgos humanos filtrados en el diseño, sea porque ofrece falsos positivos o falsos negativos). Estos desarrollos no son infalibles y, aunque nos empeñemos en denominarlos “inteligencia” artificial, no son perfectos, cometen errores, como los seres humanos.¹⁶ Abandonar la actitud intencional y volver a la actitud de diseño para interactuar con este tipo de desarrollos parece ser una buena estrategia para hacer visibles estas cuestiones y nos permite realizar otro tipo de preguntas con el fin de decidir si incorporar o no estos diseños en nuestras vidas, en políticas públicas, etc. Tener presente que estamos ante artefactos diseñados por seres humanos para realizar tareas humanas en sociedades humanas es fundamental para nuestra interacción éticamente responsable con estos sistemas.

¹⁶ Obviamente, la estrategia de seguir denominándolos “inteligencia artificial” responde a una cuestión de *marketing*, porque lo que tenga esta etiqueta se vende mejor: está instalada en nuestra sociedad la idea de que estos desarrollos son mejores que nosotros y que no fallan. Nada de eso es verdad: se equivocan y son tan buenos o malos como seamos los seres humanos que los diseñamos.

IV. Y entonces... ¿qué hacemos con la IA?

En este último apartado quisiera hacer dos propuestas concretas ante la pregunta del título: ¿Qué hacemos con la IA? La primera propuesta es promover la regulación de los desarrollos que incluyen IA. En el caso de los robots sociales, las razones son bastante claras: muchos de esos desarrollos (los que salieron del laboratorio y están siendo comercializados) están orientados a población vulnerable (infantes, adultos mayores y personas con capacidades cognitivas limitadas) y en muchos casos se ofrecen como opciones terapéuticas, y todas las opciones terapéuticas están (o deberían estar) fuertemente reguladas tanto en su ejercicio y uso, como en el proceso de producción, y deben ser cuidadosamente testeadas antes de ofrecerse a la venta al público en general. Un muñeco de peluche debe cumplir con ciertas normativas relativas, por ejemplo, a su producción con materiales no tóxicos; pero si se trata de un robot mascota terapéutico debería, además, estar regulada su introducción en las prácticas terapéuticas no supervisadas por personas autorizadas. Asimismo, dado que estos robots buscan reemplazar la compañía humana de cuidadores o amigos (en el caso de aquellos robots que se proponen como compañía para personas que están solas) o parejas (en el caso de los robots sexuales), también se requiere, en mi opinión, algún tipo de estudio acerca del impacto de la introducción de este tipo de artefactos en la calidad de las relaciones interpersonales. Ya está estudiado el efecto de las pantallas en el desarrollo cognitivo y afectivo de los infantes, de tal manera que muchas sociedades de pediatría sugieren limitar el tiempo de exposición en niños y evitar a exposición a pantallas en menores de 2 años.¹⁷ Las redes sociales han contribuido con sus algoritmos de recomendación de *feeds* a ampliar la brecha entre personas con opiniones diferentes, al punto que uno termina leyendo posts de quienes confirman lo que uno piensa, escuchando la música que uno ya conoce, encontrando en Google lo que uno ya buscó, etc. Como modelo de negocios puede ser muy exitoso, pero como modelo del tipo de relaciones interpersonales que consideramos valiosas, la cuestión queda por ser examinada.¹⁸

¹⁷ Véase: Levanta la Cabeza. “La Asociación Española de Pediatría pide restringir el uso de pantallas entre los más pequeños». *Compromiso Atresmedia*, 21 de septiembre de 2023, https://compromiso.atresmedia.com/levanta-la-cabeza/actualidad/asociacion-espanola-pediatria-pide-restringir-uso-pantallas-mas-pequenos_20230921650bffb98383a00012b2f39.html. Accedido el 16 de junio de 2024; también: “Niños y Pantallas | Comunidad SAP”. *Comunidad SAP | El sitio para la comunidad de la Sociedad Argentina de Pediatría*, <http://comunidad.sap.org.ar/index.php/2017/07/31/ninos-y-pantallas/>. Accedido el 16 de junio de 2024; y por último: “Cuánto tiempo puede o debe un niño estar frente a una pantalla”. *HealthyChildren.org*, www.healthychildren.org/Spanish/family-life/Media/Paginas/where-we-stand-tv-viewing-time.aspx. Accedido el 16 de junio de 2024.

¹⁸ Cada vez más voces se levantan en dirección a limitar la exposición de las personas, sobre todo infantes y adolescentes, a las redes sociales. Véase INFOBAE, “Nueva York declaró a las redes sociales como una amenaza para la salud mental de los menores”. *infobae*, 27 de enero de 2024, www.infobae.com/estados-unidos/2024/01/27/nueva-york-declaro-a-las-redes-sociales-como-una-amenaza-para-la-salud-mental-de-los-menores. Accedido el 16 de junio de 2024.

¿Es deseable tener un amigo robot que se acomode a nuestros deseos, nos sugiera cosas que siempre nos gustan y nos acaricie como queremos? ¿O es preferible tener como amigo a alguien que a veces nos contradice, que nos expone otro punto de vista, que nos ofrece una caricia inesperada? ¿Cuál de estos dos tipos de compañía hará nuestra vida más desafiante e interesante, aunque, tal vez, no más cómoda? Yo creo que es preferible el segundo tipo de relación interpersonal, pero más allá de las preferencias personales, hay que evaluar el impacto psicológico y social que estas tecnologías tendrían.

En el caso de aquellos sistemas con IA que no están desarrollados para convivir como otras personas entre nosotros, sino que están orientados a tareas específicas, creo preferible adoptar la estrategia de diseño para su evaluación y, por lo tanto, para pensar sus formas de regulación. Seguramente, esta adoptará formas muy diversas, teniendo en cuenta la tarea que el sistema esté diseñado para realizar. Hay muchas actividades humanas en las que la introducción de tecnología supone una serie de regulaciones propias del ámbito en cuestión, tal como es el ámbito de la medicina. No me voy a detener en estos casos. Me interesa destacar la necesidad de regulación para las tecnologías con IA que se introducen en ámbitos de la vida cotidiana y de políticas públicas. Hoy es relativamente fácil y barato crear una *App* con IA para realizar una amplia variedad de tareas, y estos desarrollos son introducidos al mercado (*i.e.* a las sociedades humanas) sin regulación alguna. Y, como bien sabemos, los “términos y condiciones” son ilegibles e incomprensibles para el usuario típico; nadie sabe exactamente a qué se somete cuando los acepta.¹⁹ Pero muchos de estos desarrollos exacerbaban sesgos y desigualdades, como señalé en el apartado anterior, y generan situaciones conflictivas en varios órdenes de la vida humana. Por ejemplo, recientemente se dio a conocer en la Argentina el caso de un joven que había desarrollado una *App* para “desnudar” gente; es decir, para generar una imagen de una persona desnuda a partir de una foto de esa persona vestida.²⁰ Obviamente las personas “desnudadas” por la *App*, cuyas fotos circularon en redes, no tenían ninguna legislación que las protegiera, dado que eran fotos generadas por IA, o sea, ficticias (la regulación vigente impide la circulación de fotos no autorizadas de las personas, no impide la difusión de imágenes de esas mismas personas que parecen fotos, pero que fueron producidas artificialmente). Los límites pueden ser difíciles de establecer, pero está claro que este desarrollo no regulado generó perjuicios a personas (obviamente, fueron mujeres las perjudicadas; siempre el

¹⁹ Una de las cuestiones sobre las que más se ha trabajado en el ámbito de la ética y regulación de la IA es la de la privacidad y uso de los datos. Dado que esta cuestión no está directamente relacionada con el tema que abordo aquí, es decir, con nuestras formas de interacción con estos desarrollos una vez producidos y ofrecidos al público, no me voy a ocupar del tema de los datos, aunque resulta un tema central para el diseño de regulaciones apropiadas.

²⁰ Segulin, Mariana. “Creó imágenes de sus compañeras sin ropa con inteligencia artificial: ¿hay delito?” Últimas Noticias de Argentina y del Mundo | Todo Noticias, 27 de septiembre de 2023, <https://tn.com.ar/tecnologia/2023/09/27/creo-imagenes-de-sus-companeras-sin-ropa-con-inteligencia-artificial-hay-delito/>. Accedido el 16 de junio de 2024.

más vulnerable es el más perjudicado). Si, por ejemplo, fuera obligatorio para todo sistema que genera imágenes por IA incluir de una manera inviolable (aunque el concepto mismo de inviolable en la era digital es en sí mismo complejo) una especie de marca de agua, algún tipo de identificación visible de la imagen como ficticia, al menos, las personas perjudicadas podrían fácilmente argüir que no son ellas las que están en esa fotografía, sino que es una fotografía trucada. El daño que una imagen puede hacer en la honorabilidad de una persona, sabemos, es muy grande. No parece razonable que sea posible apropiarse de la imagen de alguien y alterarla sin su consentimiento. Pero no hay (en Argentina, y creo que en ningún lado) regulaciones al respecto.

Otro caso conflictivo con la IA generativa es, obviamente, el de los derechos de autor. Podemos pedir a un sistema con IA que genere un cuadro “a lo Van Gogh”. Aun cuando no lo usemos para engañar a nadie, ¿quién puede comercializarlo? ¿El usuario que le indicó a la IA que haga un cuadro de un unicornio a lo Van Gogh? Al fin y al cabo, fue quien tuvo la idea... ¿Y si el artista está vivo y sigue pintando?, ¿quién puede comercializar la imagen? Al fin y al cabo, tal vez a ese artista, en el futuro se le ocurra pintar un cuadro similar.... Una vez más, si no se exige algún tipo de marca de agua que permita distinguir imágenes (o textos) generados por IA de los generados por seres humanos, estamos ante graves problemas que alteran nuestras formas de convivencia actual.²¹

Sé que es difícil pensar el tema regulatorio y es difícil, entre otras razones, porque la tecnología se está desarrollando a una velocidad mucho más grande de lo que pueden desarrollarse y consensuarse nuestros sistemas legales. Sin embargo, cuando es posible avizorar efectos adversos de manera tan clara como en estos casos, tal vez la mejor estrategia sea la precautoria (es decir, “todo está prohibido hasta que se pruebe que no tiene efectos adversos”) e invertir la carga de la prueba: es quien ofrece el desarrollo quien debería probar que no tiene impacto (efectos secundarios) negativos, antes de ofrecer al mercado su desarrollo. Tal como ocurre con la industria farmacéutica: los estudios preclínicos y clínicos para lanzar una nueva droga al mercado corren por cuenta de la empresa que ofrece el medicamento o vacuna, y hay una autoridad regulatoria (como la FDA) que dados estos estudios ya realizados analiza la viabilidad de la introducción de ese producto en el mundo humano.

Por otra parte, si un desarrollo tecnológico resulta tan peligroso *prima facie*, que se estima que sus efectos pueden generar un efecto adverso masivo, siempre es posible limitar el desarrollo en cuestión, tal como ocurrió con la energía

²¹ Véase el manifiesto de los artistas reunidos en el colectivo ArteesEtica: “Arte es Ética - Artivismo y Autorías de Habla Hispana”. Arte es Ética - Artivismo y Autorías de Habla Hispana, <https://arteesetica.org/>. Accedido el 16 de junio de 2024.

atómica, la edición genética y la clonación humana. Si bien no parece ser este el caso con los desarrollos de IA, no es algo que haya que descartar tan rápidamente. Por un lado, varios desarrolladores han realizado llamados públicos para frenar el desarrollo de la IA dado que se está visualizando el peligro del desarrollo exponencial no controlado de estas tecnologías. Pero más allá del escenario postapocalíptico de computadoras tomando el poder, los efectos de las AI actualmente existentes que generan más intolerancia, más segregación y más desigualdades en las sociedades humanas y que alteran nuestras formas humanas de interacción y valoración, deberían ser considerados lo suficientemente importantes como para prender una luz roja. Y esto me lleva a mi segundo punto.

Las personas que nos dedicamos a las humanidades y a las ciencias sociales somos quienes más hemos reflexionado y estudiado las prácticas humanas, nuestras formas de interacción y de organización, nuestras formas de autocomprensión y de organización social y política, somos quienes hemos reflexionado sobre el significado de la vida y sobre los valores y normas que rigen las sociedades humanas.²² Es por eso que creo que la perspectiva humanística debe ser parte de las discusiones y reflexiones alrededor del presente y futuro de la IA. No porque tengamos respuesta a todos los interrogantes que esta tecnología nos plantea, sino porque ya nos hemos formulado las preguntas que estos desarrollos nos plantean. No es la primera vez que una tecnología produce una revolución en el mundo humano; pensemos en la máquina a vapor, la escritura o la imprenta. Estudiar estos casos históricos y los desafíos que plantearon, puede iluminar nuestro presente. No es la primera vez que identificamos conductas discriminatorias y juicios sesgados. Las soluciones que en otras ocasiones se adoptaron pueden servirnos para repensar los desarrollos actuales. Mirar las formas de organización real (que nos provee la antropología) o posible (que nos provee la literatura, el cine y la filosofía política) de las sociedades humanas nos permite replantear los problemas que hoy enfrentamos, que tratamos de resolver con tecnología, pero que también podemos resolver de otras formas. Una vez escuché argumentar que era necesario desarrollar *Apps* de apoyo psicológico porque cada vez hay más gente con trastornos de ansiedad, fobias o depresión. A lo que respondí: “Mejor invitar a esa gente a una clase de Tai Chi o de zumba a la plaza con otra gente!” Si el problema es la soledad, un robot no lo va a solucionar. No es cierto que los problemas que genera la tecnología se resuelven con más tecnología. A veces es tomar distancia, ver dónde estamos parados y a dónde

²² Asimismo, los diversos tipos de estudios sugeridos en el párrafo anterior corresponden a disciplinas sociales y humanísticas, por ejemplo, una evaluación de diversos tipos de terapias psicológicas, incluyendo explícitamente las que quedan en manos de sistemas con IA, el impacto psicológico (afectivo y cognitivo) de las horas ante pantallas de personas de diferentes edades, el impacto educativo en la incorporación de IA en las escuelas, la evaluación sociológica y antropológica del impacto de las redes sociales, el impacto en la personalidad e identidad personal producida por la adopción de avatares en entornos virtuales y el grado comparativo de responsabilidad asociado a las acciones que realizamos en entornos reales versus entornos digitales, por mencionar solo algunos casos.

estamos yendo, y preguntarnos a dónde queremos ir. Y para responder esta pregunta nada mejor que imaginar escenarios alternativos posibles. Para eso están las humanidades y las artes: las alteridades disruptivas, no las que se acomodan a nuestros deseos para mantenernos atados.



Referencias

- "Arte es Ética - Artivismo y Autorías de Habla Hispana". *Arte es Ética - Artivismo y Autorías de Habla Hispana*, <https://arteestetica.org/>. Accedido el 16 de junio de 2024.
- INFOBAE. "Nueva York declaró a las redes sociales como una amenaza para la salud mental de los menores". infobae, 27 de enero de 2024, www.infobae.com/estados-unidos/2024/01/27/nueva-york-declaro-a-las-redes-sociales-como-una-amenaza-para-la-salud-mental-de-los-menores. Accedido el 16 de junio de 2024.
- Dennett, Daniel. *Brainstorms. Philosophical Essays on Mind and Psychology*. Cambridge: MIT Press, 1981.
- Dennett, Daniel. *The Intentional Stance*. Cambridge: MIT Press, 1987.
- Dupré, John. *The Disorder of Things. Metaphysical Foundations of the Disunity of Science*. Cambridge: Harvard University Press, 1993.
- Fodor, Jerry. *Psicosemántica*, Madrid: Técnos, 1994.
- Lewis, David. "Psychophysical and Theoretical Identifications", *Australasian Journal of Philosophy*, 50 (1972): 249-58
- Levanta la Cabeza. "La Asociación Española de Pediatría pide restringir el uso de pantallas entre los más pequeños". *Compromiso Atresmedia*, 21 de septiembre de 2023, https://compromiso.atresmedia.com/levanta-la-cabeza/actualidad/asociacion-espanola-pediatria-pide-restringir-uso-pantallas-mas-pequenos_20230921650bffa98383a00012b2f39.html. Accedido el 16 de junio de 2024.
- Marr, David. *Vision. A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco: W. H. Freeman, 1982.
- "Neurorobotics Research Laboratory (NRL) - Myon." Neurorobotics Research Laboratory, www.neurorobotik.de/robots/myon_en.php. Accedido el 16 de junio de 2024.
- "Niños y Pantallas. Comunidad SAP". *Comunidad SAP. El sitio para la comunidad de la Sociedad Argentina de Pediatría*, <http://comunidad.sap.org.ar/index.php/2017/07/31/ninos-y-pantallas/>. Accedido el 16 de junio de 2024.
- "PARO Therapeutic Robot". *PARO Therapeutic Robot*, www.parorobots.com. Accedido el 16 de junio de 2024
- Pérez, Diana & Gomila, Antoni. *Social cognition and the second person in human interaction*. London: Routledge, 2022.
- Putnam, Hilary. "La naturaleza de los estados mentales". *Cuadernos de Crítica*, vol. 15 México: UNAM, 1981.
- Russell, Stuart & Norvig, Peter. *Inteligencia artificial. Un enfoque moderno*. México: Pearson Prentice Hall, 2008.
- Segulin, Mariana. "Creó imágenes de sus compañeras sin ropa con inteligencia artificial: ¿hay delito?" Últimas Noticias de Argentina y del Mundo | Todo Noticias, 27 de septiembre de 2023, <https://tn.com.ar/tecnologia/2023/09/27/creo-imagenes-de-sus-companeras-sin-ropa-con-inteligencia-artificial-hay-delito/>. Accedido el 16 de junio de 2024.
- Turing, Alan. "Computing Machinery and Intelligence", *Mind*, 1950, 49 (1950): 433-460.



Una mirada crítica a la ética de la IA: de preocupaciones emergentes y principios orientadores a un desvelar ético

A Critical Look at AI Ethics: From Emerging Concerns and Guiding Principles to Ethical Unveiling

GABRIELA ARRIAGADA-BRUNEAU¹

Pontificia Universidad Católica de Chile, Santiago, Chile
gcarriagada@uc.cl

Fecha de recepción: 19/04/2024

Fecha de aceptación: 22/06/2024

Resumen

En este artículo examino el estado actual de la “segunda ola” de la ética en la inteligencia artificial (IA), la cual se centra en la integración de principios éticos fundamentales como la justicia, la privacidad, la transparencia y la explicabilidad en el diseño, uso e implementación de sistemas de IA. Argumento que, aunque esta fase ha sido criticada por su naturaleza abstracta y su falta de contextualización, es imperativo que la emergente “tercera ola” adopte un cambio paradigmático hacia un “desvelar ético”. Propongo que este desvelar ético debería involucrar un profundo y continuo proceso hermenéutico que no solo interprete cómo las tecnologías de IA reconfiguran nuestras estructuras sociales, políticas y personales, sino que también actúe como un medio de liberación del enmarcamiento tecnológico Heideggeriano. Este enfoque sugiere que la ética debe ser considerada no solo como un complemento, sino

CÓMO CITAR ESTE ARTÍCULO:

En APA: Arriagada-Bruneau, G. (2024). Una mirada crítica a la ética de la IA: de preocupaciones emergentes y principios orientadores a un desvelar ético. *Resonancias*, (17), 101-120. DOI: 10.5354/0719-790X.2024.74438

En MLA: Arriagada-Bruneau, G. “Una mirada crítica a la ética de la IA: de preocupaciones emergentes y principios orientadores a un desvelar ético.” *Resonancias*, n.º 17, julio de 2024, pp. 101-120. DOI: 10.5354/0719-790X.2024.74438

Palabras clave: ética de inteligencia artificial, sistemas sociotécnicos, enmarcamiento tecnológico, principios éticos, ética aplicada

Keywords: Ethics of Artificial Intelligence, Sociotechnical Systems, Technological Framing, Ethical Principles, Applied Ethics

¹ Profesora Asistente, Instituto de Éticas Aplicadas e Instituto de Ingeniería Matemática Computacional, Pontificia Universidad Católica de Chile. <https://orcid.org/0000-0002-0006-7024>

como un componente esencial y fundamental en el ciclo de vida del desarrollo de la IA, fomentando así una integración más profunda de consideraciones éticas que guíen tanto la innovación tecnológica como su implementación práctica.

Abstract

In this article, I examine the current state of the “second wave” of ethics in artificial intelligence (AI), which focuses on the integration of fundamental ethical principles such as justice, privacy, transparency, and explicability in the design, use, and implementation of AI systems. I argue that although this phase has been criticized for its abstract nature and lack of contextualization, it is imperative that the emerging “third wave” adopts a paradigmatic change towards an “ethical unveiling.” I propose that this ethical unveiling should involve a deep and ongoing hermeneutic process that not only interprets how AI technologies reconfigure our social, political, and personal structures but also acts as a means of liberation from the technological framing in the Heideggerian sense. This approach suggests that ethics should not only be considered as an add-on but as an essential and fundamental component in the lifecycle of AI development, thus promoting a deeper integration of ethical considerations that guide both technological innovation and its practical implementation.

1. La segunda ola de la ética de IA: preocupaciones emergentes y principios orientadores

La ética en inteligencia artificial se ha consolidado recientemente como una disciplina. En los últimos años, los artículos de investigación dedicados a esta temática han proliferado. Como notan Borenstein *et al.*, los trabajos académicos en Google Scholar de artículos relacionados a búsquedas sobre “ética de inteligencia artificial” van de 21 resultados en 2016 a 341 en 2011. Es más, recientemente actualizamos la búsqueda con los mismos criterios y los resultados superan los 1000 artículos (referencia eliminada para revisión) y, en general, refieren a discusiones centradas en principios éticos aplicados a la ética de la IA. Esta etapa en el desarrollo de la ética de inteligencia artificial ha sido denominada por Aimee Van Wynsberghe como la segunda ola de la ética de IA. La autora describe la primera ola desde el enfoque en preocupaciones de la superinteligencia y escenarios de riesgo existencial asociado a la “sublevación” de las máquinas, haciendo referencias a uno de los primeros y más reconocidos análisis por Nick Bostrom y Eliezer Yudkowsky.

La segunda ola, según Van Wynsberghe, se ha enfocado en abordar las preocupaciones prácticas relacionadas con el uso de técnicas de aprendizaje automático, los desafíos de transparencia y explicabilidad en las cajas negras algorítmicas, la falta de representatividad en los datos de entrenamiento que

ocasionan sesgos en los modelos de IA, también la violación de la privacidad de los ciudadanos a través de sistemas de reconocimiento facial. En términos simples, esta segunda ola discute la aplicación de principios éticos como la justicia, la privacidad, la transparencia y la explicabilidad en el diseño, uso e implementación de los sistemas de IA.

Ahora bien, para visualizar críticamente esta segunda ola, es relevante considerar preocupaciones emergentes relacionadas con cambios en los contextos de toma de decisiones que, necesariamente, influyen en la adopción de sistemas de IA. Me refiero, por ejemplo, al creciente uso de algoritmos (conjunto finito de instrucciones específicas) para la toma de decisiones. Si bien, no todo algoritmo se utiliza como modelo para entrenar un sistema de IA, el estudio ético del uso de algoritmos como mediadores en procesos sociales nos ha informado mucho sobre preocupaciones emergentes que también se presentan en sistemas de IA.²

Recordemos el artículo fundacional relacionado con el uso ético de algoritmos titulado “La ética de los algoritmos: delineando el debate” (Mittelstadt *et al.*), donde se discutió formalmente la creciente influencia de los algoritmos en procesos de decisión que antes dependían únicamente de seres humanos. Ahí, los autores enfatizaban la importancia ética de la mediación algorítmica en procesos sociales, empresariales, gubernamentales, incluso en interacciones personales. Para analizar este fenómeno, reconocieron seis tipos de preocupaciones éticas que se originan en cuestionamientos epistémicos (figura 1). Por ejemplo, la evidencia inconclusa (*inconclusive evidence*) refleja la necesidad de reconocer y gestionar la incertidumbre en sus conclusiones. Asimismo, la evidencia inescrutable (*inscrutable evidence*) establece la importancia de la transparencia y la comprensión del proceso algorítmico que lleva a las conclusiones entregadas por el modelo. Esto se contextualiza en las demandas de una sociedad informada donde la rendición de cuentas (transparencia) y la comprensión pública son pilares para generar confianza en los sistemas tecnológicos. Finalmente, la evidencia equivocada (*misguided evidence*) nos obliga a reflexionar sobre la integridad de los datos de entrada, dado que los errores en las bases de datos pueden llevar a conclusiones erróneas y decisiones injustas.

² Aunque no hay una definición general de inteligencia artificial que sea aceptada en este trabajo cuando me refiera a sistemas de inteligencia artificial (IA), tendré en mente la definición establecida por el AI HLEG (Grupo de expertos de alto nivel sobre inteligencia artificial) de la Unión Europea. Esta definición dice que la IA es “sistemas que muestran comportamiento inteligente al analizar su entorno y tomar acciones —con cierto grado de autonomía— para alcanzar objetivos específicos”. Esta definición, por tanto, se refiere únicamente a lo que se denomina IA estrecha o débil; a saber, que no tiene un espectro completo de habilidades cognitivas en forma de una inteligencia artificial general.

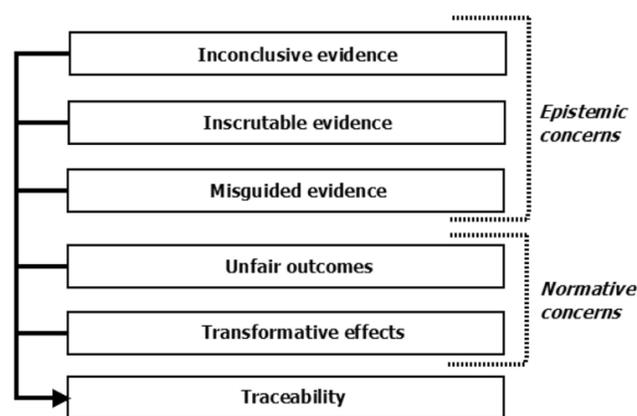


Figure 1. Six types of ethical concerns raised by algorithms.

Figura 1: Seis tipos de preocupaciones éticas sobre uso de algoritmos como mediadores para la toma de decisiones (Mittelstadt et al., 4).

Las siguientes tres preocupaciones se refieren más bien a cuestiones normativas, como los resultados injustos (*unfair outcomes*), donde el enfoque ético requiere una evaluación continua de las acciones que los algoritmos motivan, particularmente cuando podrían reforzar desigualdades existentes o introducir nuevas formas de discriminación. Esto pone de manifiesto la necesidad de criterios éticos robustos que puedan aplicarse al diseño y operación de sistemas algorítmicos para prevenir tales injusticias. En cuanto a los efectos transformativos (*transformative effects*), los algoritmos pueden alterar nuestra percepción del mundo y nuestras estructuras sociales y políticas de maneras significativas y, a menudo, imperceptibles. Por tanto, debemos reflexionar sobre cómo estos cambios influyen en nuestra realidad y los valores que sustentan nuestras sociedades. Finalmente, la trazabilidad (*traceability*) es crucial para poder atribuir responsabilidades cuando se utilizan algoritmos.

Si observamos detenidamente estas categorías de preocupaciones normativas y epistémicas sobre el uso ético de algoritmos, es posible identificar que estas no solo señalaron la urgencia de un enfoque interdisciplinario para abordar los desafíos éticos relacionados con el creciente uso de algoritmos, en particular en contextos de decisión de alto impacto social, sino que también configuran varios de los grandes desarrollos que hemos visto en el avance de la ética de la IA como disciplina. Así, podemos reconocer cómo los avances en esta área se benefician de debates y cuestionamientos en áreas colindantes, como la ética de algoritmos. Por ejemplo, reconocer la incertidumbre inherente en el conocimiento producido por los algoritmos se puede relacionar con la percepción de la tecnología y la IA

como falible (escepticismo epistémico). La evidencia inconclusa generada por los algoritmos nos desafía a ver los avances e incorporaciones de la tecnología como una práctica contextualizada, desde la cual se consideren los valores y prejuicios subyacentes que podrían influir en los resultados de los algoritmos y otras tecnologías, como los sistemas de IA.

En esta línea, autoras como Meredith Broussard subrayan que la IA no es una solución universal ni objetiva y que, por tanto, no debemos caer en un “tecnochovinismo”, el cual se define como la creencia de que “la tecnología es siempre la solución” (Broussard, 4). Esta caracterización retrata un optimismo ciego que incita a asumir que la tecnología es menos sesgada que una decisión humana, influenciando no solo las expectativas sobre el funcionamiento de la IA, sino también la evaluación ética que hacemos de esta. Si a esta advertencia le sumamos el factor de “escrutabilidad” y el concepto de “evidencia equivocada” del artículo fundacional sobre la ética de algoritmos, podemos ver cómo en la ética de IA también se nos interpela a cuestionar la validez de los inputs (entradas) y outputs (salidas) de un sistema.

Por esto, una de las preocupaciones epistémico-éticas fundamentales del desarrollo de la inteligencia artificial se ha transformado en discusiones sobre transparencia y explicabilidad respecto a la opacidad y falta de conocimiento sobre qué hace la IA, así como los sesgos y patrones de discriminación que provienen de datos de entrenamiento que reflejan una sociedad altamente sesgada e injusta. Estas son las mismas características que Van Wynsberghe destaca como parte de la segunda ola de la ética de la IA. Como consecuencia, las preocupaciones normativas planteadas por Mittelstadt *et al.* al inicio del debate sobre la ética de algoritmos, lograron permear no solo en directrices para el uso de algoritmos, sino que también se extendieron a la IA, en parte, debido al gran uso de algoritmos como base para modelar sistemas de IA. Así, a medida que la discusión comenzó a proliferar, el mundo académico,³ que incipientemente se acercaba a estos nuevos desafíos, y el desarrollo tecnológico aplicado a la sociedad, que avanzaba sin marcos, directrices ni estándares que generaran alertas sobre qué se estaba haciendo y cómo, se encontraron en una tormenta perfecta. Este escenario propició el surgimiento de una serie de principios de ética de la IA como respuesta para enfrentar estos avances. No obstante, este pilar fundamental de la segunda ola de la ética de la IA no tardaría en ser criticado por su “inutilidad”.

³ Véase Binns; Boddington; Burrell; Chouldechova; Coeckelbergh; Danks and London; Dencik *et al.*; Dennis *et al.*; Dignum; Etzioni and Etzioni; Floridi *et al.*; Narayanan; ‘Robotics’.

1.1 La (in)utilidad de los principios éticos de IA

Uno de los artículos críticos más comentados sobre la proliferación de principios éticos en IA es el de Brent Mittelstadt, titulado “Los principios por sí solos no pueden garantizar una IA ética”. Aquí, Mittelstadt señala que muchas de las iniciativas éticas desarrolladas hasta esa fecha se basaban en el uso de estos principios de IA desarrollados durante la segunda ola. Su uso era predominante en la industria, donde, sin embargo, se traducían en un simple caso de “señalización de virtud” (*virtue signaling*) (Mittelstadt 501), en el cual las prácticas éticas se interpretan como tácticas para posponer la regulación (o simular una autorregulación), cambiando el foco de las discusiones a problemas teóricos y respuestas tecnológicas, lo que genera una apariencia de consciencia y enfoque ético. En la práctica, nos dice Mittelstadt, esto implica que los principios éticos de IA son ambiguos y generales, ya que no abordan las discusiones normativas y políticas que subyacen en conceptos clave como la justicia y la privacidad, y pueden, por lo tanto, continuar alimentando un desarrollo de la IA que no profesionaliza buenas prácticas ni logra integrar la ética en sus procesos productivos.

En respuesta a este contexto, Mittelstadt compara la incipiente ética en inteligencia artificial con la ética médica, identificando que hay principios base comunes que podrían estimular un principialismo robusto en IA. Con el respaldo de la OCDE y del Grupo de Expertos de Alto Nivel sobre Inteligencia Artificial de la Comisión Europea (AI HLEG), se propusieron cuatro principios para guiar el desarrollo de una IA confiable: respeto por la autonomía humana, prevención de daños, equidad y explicabilidad (High-Level Expert Group on AI). Sin embargo, como bien analiza Mittelstadt, el principialismo de la ética médica surgió en un contexto profesional diametralmente diferente. La ética médica se desarrolla en contextos con una institucionalización y responsabilidades morales profesionales arraigadas en el quehacer médico, con comités de ética e investigación, instituciones médicas fiscalizadoras, cuerpos profesionales y la base transversal que es el juramento hipocrático. Esa robustez permitió que los principios base de la bioética gestaran un lenguaje común en la ética médica (Beauchamp and DeGrazia; Bosk) y que permitiera identificar y conceptualizar los diversos desafíos éticos. El contexto de creación e implementación de la IA es diferente.

Mittelstadt identifica cuatro grandes aspectos que generan las condiciones necesarias para que el principialismo funcione y que, en comparación con la medicina, el desarrollo de la IA carece (502-508): (1) objetivos comunes y deberes fiduciarios, (2) una historia y normas profesionales consolidadas, (3) métodos probados para traducir principios en prácticas y (4) mecanismos robustos de rendición de cuentas legales y profesionales. En comparación con la medicina, que está orientada por un objetivo unificado de promover el bienestar del paciente y se apoya en una estructura profesional bien establecida, normativas claras y

mecanismos de responsabilidad sólidos, beneficiándose de relaciones fiduciarias y un marco regulatorio riguroso que fomenta la confianza y la colaboración. El sector de la IA se enfrenta a un entorno en el cual los intereses de desarrolladores, usuarios y afectados no siempre están alineados, lo que complica la toma de decisiones éticas en un contexto más competitivo que colaborativo (y altamente privatizado), exacerbado por una regulación que varía regionalmente sin establecer deberes fiduciarios claros hacia los sujetos de datos y los usuarios.

En la misma línea, un año después, Thilo Hagendorff criticó las directrices éticas desarrolladas hasta ese entonces, cuestionando que estas no tienen un impacto real transversal en la toma de decisiones humanas en el campo de la IA. Su crítica es clara: “la ética carece de un mecanismo de refuerzo. Las desviaciones de los diversos códigos de ética no tienen consecuencias” (Hagendorff 113). Analizando 22 de las principales directrices de ética de la IA publicadas hasta 2020, Hagendorff argumenta que estas carecen de mecanismos para reforzar sus propias afirmaciones normativas. Así, la ética sirve más para calmar voces críticas del público que para robustecer prácticas en la disciplina.

En relación a este punto, aunque considero que Hagendorff puede tener un argumento sólido sobre cómo estos principios se traducen en prácticas robustas adoptadas por instituciones y profesionales, difiero en su diagnóstico crítico hacia la ética por no tener mecanismos de refuerzo (*reinforcement mechanisms*). El propósito fundamental de la ética no es ejercer coerción o imponer sanciones (en este aspecto nos podemos valer de la regulación), sino más bien guiar el comportamiento humano a través de una reflexión moral y el fomento de la integridad personal y colectiva. Para esto, los principios pueden servir como traductores para definir ciertas preocupaciones éticas dentro del contexto de la IA. La tarea pendiente, no obstante, es integrar la ética como una brújula moral para la acción, que nos permita justificar las decisiones y acciones que tomamos respecto al diseño, uso e implementación de sistemas de IA.

A pesar de esta distinción, comparto el punto crítico de Hagendorff: debido a que las desviaciones de los códigos éticos establecidos no conllevan consecuencias inmediatas, la ética de la IA, a menudo, es percibida como algo externo o adicional a las preocupaciones técnicas, como un marco no vinculante impuesto por instituciones externas a la comunidad técnica. Esta falta de responsabilidad vinculante que apele a códigos y estándares profesionales provoca que, en palabras de Hagendorff, los desarrolladores carezcan de una sensación de responsabilidad o de una percepción de la importancia moral de su trabajo, siendo los incentivos económicos un factor que frecuentemente sobrepasa el compromiso con los principios y valores éticos. Me parece, por tanto, que lo planteado por Hagendorff más bien enfatiza que la ética de la IA necesita lograr integrar estos

principios como parte de las metodologías de trabajo de los desarrolladores de IA y de las instituciones que los financian.

Hoy por hoy, aunque estas críticas han influido en los debates académicos, no hemos logrado superar estas dificultades. Un artículo más reciente de Luke Munn califica nuevamente a la ética de la IA como “inútil”, argumentando que la ineficacia de los principios éticos se debe, en gran parte, a que son abstractos, aislados (descontextualizados) y carecen de fuerza (para demandar su cumplimiento). Como resultado, dice Munn, existe una brecha entre los principios éticos y la práctica. Para él, en un mundo de suma cero, la obsesión con los principios de la IA no solo es inútil, sino también peligrosa, ya que desvía recursos humanos y financieros de enfoques más productivos. Un ejemplo de esto es cuando Google anunció la creación de un cuerpo independiente para revisar las prácticas éticas de IA de la compañía, pero este grupo no tendría poder real para vetar proyectos o detenerlos de manera significativa.

El llamado de Munn, en respuesta a estas limitaciones, es que se requiere un cambio de un enfoque basado en la adhesión universal a principios y reglas hacia un enfoque ético sensible al contexto. Esto implica un enfoque menos disciplinario y más emancipador, que permita a los actores morales actuar de manera autónoma y responsable, basándose en un conocimiento integral y en la empatía necesaria para tomar decisiones moralmente relevantes. Es decir, empoderar a quienes impulsan el desarrollo de la IA a tener estándares profesionales. En esto último concuerdo con Munn: el aspecto crucial de las críticas al desarrollo de la segunda ola de ética de la IA es que, aunque los principios han logrado cierta transversalidad en su adopción simbólica, no se ha logrado una apropiación de facto. Esta apropiación en la práctica es crucial, considerando que la legislación y regulación en materias de IA siempre estará incluso más retrasada que las consideraciones éticas.

De estas críticas consistentes en el tiempo, destaco que el paso a una tercera ola de la ética de inteligencia artificial debe exigir que los avances y desarrollos técnicos en el área se planteen desde la integración de la ética como parte del desarrollo de la IA, fortaleciendo estándares profesionales y estructuras reguladoras que enmarquen este proceso. Esto, sin embargo, implica necesariamente enseñar una ética aplicada a la inteligencia artificial para formar profesionales que usen los principios éticos, los cuales, aunque por sí solos pueden ser ineficientes, contextualizados y traducidos a prácticas concretas, pueden guiar mejores prácticas.

En respuesta a este escenario en el que aún nos encontramos, en la siguiente sección propongo como podemos navegar hacia una tercera ola de desarrollo de la ética de la IA desde la contextualización de la IA como sistema sociotécnico, comprendiendo esta transición desde el enmarcamiento (*Gestell*) heideggeriano.

2. Navegando hacia una tercera ola de la ética de IA: contextualización y el desvelar ético

Lo discutido anteriormente refleja la complejidad de avanzar en criterios éticos para el desarrollo e implementación de la IA. No basta con, simplemente, añadir la ética a la IA; una simple añadidura sería caer en lo que Johnson y Verdicchio llaman una falacia aditiva. Los autores han desafiado la noción simplista de “incrustar” (*embedding*) la ética en la IA, argumentando que, al añadir la ética a la tecnología; es decir, “Ética + IA = IA Ética,” se comete una falacia al asumir erróneamente que los principios éticos pueden ser codificados directamente en la IA. Su argumento se basa en que, para lograr una adición de ese tipo, los campos de la IA y la ética deberían compartir características ontológicas, o sea, deben tener naturalezas compatibles, cosa que no tienen.

El origen ontológico de la IA tiene una base computacional, lo que implica que, para lograr dicha adición, la ética también debe ser “computacionalizada”, capturando principios éticos en formulaciones matemáticas. Esta discrepancia pone énfasis en que cuando hablemos de ética en IA no debemos entenderlo como un proceso que implique incluir elementos o consideraciones éticas en la tecnología misma, sino que nos llama a ampliar nuestra manera de entender la relación que puede haber entre la ética y la tecnología. Desde esta limitación ontológica, una comprensión amplia de la IA como un sistema sociotécnico abre una dimensión de convergencia al considerar que las inteligencias artificiales no son solo un conjunto de herramientas computacionales, sino que operan dentro de redes complejas de relaciones humanas, normas sociales y prácticas organizativas, y es en esa dimensión donde puede haber una simbiosis.

Por esto, una de las tendencias que han surgido y a las cuales adscribo es definir la IA como un sistema sociotécnico, lo que implica concebirlo como una entidad que no opera de manera aislada, sino en contextos específicos; está inmersa en la sociedad, sujeta a sus normas y valores, y afectada por sus prácticas. Por lo tanto, cualquier consideración ética en la IA debe tener en cuenta el impacto y la interacción que la tecnología adopta y ejerce sobre el tejido social.

En la literatura, se han propuesto diversas definiciones y caracterizaciones de los sistemas sociotécnicos. En general, estos se comprenden como sistemas que dependen no solo del hardware técnico, sino también del comportamiento humano y de las instituciones sociales para su correcto funcionamiento (Davis *et al.*; Nickel). Según esta comprensión, Van de Poel (391) identifica que los sistemas sociotécnicos consisten en combinaciones de tres bloques de construcción básicos:

- (i) Artefactos técnicos: objetos físicos que pueden cumplir y han sido diseñados para una cierta función técnica —tienen una naturaleza física^{3/4}.
- (ii) Agentes humanos: personas que intencionan acciones e interactúan con artefactos técnicos.
- (iii) Instituciones: reglas o normas sociales que deben ser seguidas por los agentes—prescriben expectativas de comportamiento moral^{3/4}.

A esta base fundacional, Van de Poel agrega dos características propias de los sistemas de IA: (iv) agentes artificiales y (v) normas técnicas. La idea subyacente aquí es que los agentes artificiales tienen propiedades que los distinguen de los artefactos técnicos tradicionales. Ahora bien, dado que las instituciones son construcciones sociales, nos dice Van de Poel, estas no pueden ser percibidas y seguidas directamente por agentes artificiales. Lo que sí pueden hacer es seguir un componente análogo que aplica a los sistemas de IA, a saber, las normas técnicas, que propician efectos causales-físicos en vez de en términos intencionales.

Basado en esto, Van de Poel articula que los sistemas de inteligencia artificial se definen como sistemas sociotécnicos que exhiben una complejidad aumentada en relación con los sistemas sociotécnicos convencionales, atribuible a su aptitud inherente para el aprendizaje y la adaptación. Esta propiedad distintiva tiene el potencial de inducir resultados imprevistos, los cuales plantean desafíos significativos en términos de previsión y regulación en comparación con sistemas tradicionales.

De ahí que la concepción de la inteligencia artificial como un sistema sociotécnico subraya la necesidad de examinar minuciosamente las interacciones que surgen entre los componentes tecnológicos (tales como agentes artificiales y normas técnicas) y los elementos sociales (incluyendo normativas, relaciones interpersonales y estructuras organizativas) en el proceso de diseño e implementación de sistemas basados en IA. Este análisis sociotécnico profundiza en la comprensión de los sistemas de IA más allá de su función computacional, adentrándose en su rol como configuradores de espacios sociales. Tal entrelazamiento resalta la importancia de considerar la tecnología de IA no solo como un ente de ejecución de tareas, sino como un agente activo en la reconfiguración de la interacción social y el conocimiento.

La consideración de la IA desde una perspectiva sociotécnica, en consecuencia, aboga por una comprensión holística de la tecnología, que incorpora a la IA en el tejido de las relaciones humanas y los marcos institucionales. Al reconocer que los sistemas de IA pueden tanto perpetuar como alterar las desigualdades

sociales, se pone de manifiesto la necesidad de desarrollar marcos de gobernanza tecnológica que sean éticamente robustos y eviten la falacia aditiva.

Estas consideraciones, no obstante, reflejan algo más que la mera necesidad de comprender la simbiosis y retroalimentación que hay entre técnica y sociedad para poder prevenir impactos negativos como un trato injusto o una violación a la privacidad. Una tecnología como la inteligencia artificial genera lo que Swierstra y Molder llaman “impactos suaves”. Estos impactos suaves se refieren a las modificaciones sociales y personales que las tecnologías tienen en las personas, tales como la reevaluación de normas sociales y valores, la influencia en identidades sociales, incluso la reconfiguración de las expectativas sociales sobre bienestar. Por ejemplo, los sistemas de reconocimiento facial plantean nuevas expectativas sobre la privacidad y los límites de la vigilancia. También vemos cómo diagnósticos médicos asistidos por sistemas de IA pueden cambiar el valor que se le da a la tecnología y al rol humano en contextos de cuidado. Esto sugiere que una evaluación ética de la tecnología no debe limitarse a los impactos directos y evidentes, sino que también debe considerar estos efectos más sutiles y a largo plazo.

Para lograr un entendimiento más profundo de estos impactos suaves, propongo entender el paso a la tercera ola de la inteligencia artificial como una emancipación de los condicionamientos que la IA puede tener en nuestra sociedad, contextualizando su desarrollo e implementación para que nos sirva como una tecnología reveladora de las nuevas dinámicas éticas y epistémicas que implica la evolución de una sociedad que adopta y se tecnifica activamente con la IA. Esto implica usar los principios éticos que hemos construido como cimientos de una edificación que tenga ciertos mínimos éticos comunes, pero que se levante a través de un proceso hermenéutico que cuestione a la tecnología y nos permita tener un conocimiento y agencia sobre las verdades a las que accedemos gracias a ella.

2.1. Un proceso hermenéutico y un desvelar ético

Para embarcarnos en este proceso hermenéutico, primero debemos entender qué implica. Grunwald realiza una crítica a los métodos de evaluación tecnológica que se centran exclusivamente en las consecuencias de las tecnologías, argumentando que estos enfoques son insuficientes. Propone, en respuesta a su crítica, que también deberíamos examinar el conocimiento hermenéutico que tenemos disponible, es decir, la comprensión profunda e interpretativa sobre cómo las tecnologías son entendidas y significadas socialmente. Tomando esta definición de conocimiento hermenéutico, sugiero que parte de nuestras inquietudes epistemológicas respecto a una ética de la IA debe basarse en un proceso

iterativo para entender cómo significamos la tecnología, o sea, el significado social de la tecnología de IA. Así, al formar un espiral de entendimiento donde cada nuevo descubrimiento o interpretación modifica y enriquece nuestra comprensión inicial, podemos contextualizar los principios éticos que hemos generado. Ya no como simples directrices abstractas que han de “guiar” el actuar ético, sino como estructuras que cimientan nuestro proceso constante de cuestionar el rol de la tecnología en nuestra sociedad.

Esta aproximación hermenéutica, además, se complementa si la desenmarcamos en un sentido Heideggeriano. Para Heidegger, la esencia de la tecnología no es nada tecnológico, ya que la técnica no es un simple conjunto de herramientas; es más bien una forma de entender y relacionarse con el mundo (Heidegger). La técnica moderna, nos dice Heidegger, configura nuestra relación con el mundo de manera que todo se presenta como un recurso disponible para ser explotado, ayudándonos a desvelar (*Entbergen*) verdades. El desafío clave es no dejar la técnica en el dominio únicamente de lo técnico y, en cambio, cuestionarnos cómo podemos relacionarnos con la técnica de manera que nos permita acceder a una comprensión más rica de la verdad y del ser, en lugar de quedar atrapados en una visión meramente utilitaria del mundo.

Pero para alcanzar ese vínculo, debemos emanciparnos del enmarcamiento (*Gestell*) que la tecnología moderna tiene sobre nosotros, condicionándonos a un modo de comprensión del mundo y, por tanto, nuestra interacción con él. Según Heidegger, el enmarcamiento convierte todo en recursos disponibles para ser explotados. Si aplicamos esto al contexto de la IA, puede interpretarse como las tendencias a caer en ese tecnochovinismo o solucionismo tecnológico,⁴ donde la tecnología de IA se ve como maximizadora y optimizadora de procesos y decisiones que impactan a los humanos, y no como una tecnología facilitadora para el bienestar humano. Esto también puede entorpecer nuestro acceso a diferentes conocimientos y verdades, ya que la IA, cuando no se contextualiza y no se entiende desde su dinámica inherentemente sociotécnica, puede caer fácilmente en vicios como la señalización virtuosa o la apariencia de autorregulación mencionados anteriormente. Un paso hacia la integración de la ética en el desarrollo de la tecnología de IA debe basarse en facilitar una relación auténtica de conocimiento, donde no nos condicionemos a la estructura tecnológica, sino que desvelemos sus dinámicas sociotécnicas para significarlas.

⁴ El solucionismo tecnológico, como lo presenta Morozov tiende a simplificar problemas complejos, reduciéndolos a cuestiones que pueden ser resueltas mediante la tecnología, con aplicaciones, software o dispositivos. Esta simplificación, a menudo, pasa por alto factores socioculturales, históricos y políticos que son cruciales para entender y abordar dichos problemas de manera efectiva. Así, este solucionismo se caracteriza por un optimismo desmedido en la capacidad de la tecnología para mejorar la sociedad. Esto puede llevar a una dependencia excesiva de soluciones digitales y a la promoción de tecnologías como panaceas universales, tal y como ocurre con ciertas expectativas de integrar la IA indiscriminadamente.

Así, un proceso hermenéutico enfatiza cómo interpretamos y reinterpretamos tecnologías como la IA. A través de un “desvelar” continuo, empezamos a ver no solo lo que la IA puede hacer (su utilidad funcional) sino también lo que significa para la sociedad (su significado social y ético). Este “desvelar ético” lo defino así: entender cómo la IA puede cambiar nuestras relaciones sociales, nuestras estructuras de poder y nuestras concepciones de los diferentes principios que hemos establecido como su base normativa.

Esta visión, por tanto, invita a un cuestionamiento continuo sobre cómo la tecnología de IA puede ser diseñada y desplegada de manera que realmente beneficie y enriquezca la sociedad, no solo en términos de eficiencia y capacidad, sino también para fomentar una comprensión más profunda y crítica de nosotros mismos y del mundo que habitamos. Este paso a una tercera ola de la ética de inteligencia artificial debe, entonces, articular los principios existentes para que apoyen una reflexión más profunda sobre el desarrollo tecnológico, evitando integraciones superficiales que promuevan prácticas como la señalización virtuosa o la apariencia de autorregulación, mencionadas anteriormente. Esto implica comprometerse con un proceso continuo de reflexión y revisión sobre cómo la tecnología revela o encubre la verdad de nuestro ser en el mundo y cómo su rol en la sociedad influye en estos procesos epistémicos y éticos que determinan la coevolución de una sociedad enmarcada en esta tecnología.

Pero ¿qué implica, en la práctica, integrar un proceso de conocimiento hermenéutico que nos permita avanzar hacia una desvelación ética en el desarrollo de la IA? Primero, esto implica educar a los desarrolladores y usuarios de IA en un pensamiento crítico que contemple la tecnología como un actor que configure activamente nuestras realidades sociales y personales, demandando así una responsabilidad ética más profunda en su creación y uso. Uno de los ejes de desarrollo de esta nueva ola debe centrarse en la educación en la ética de la IA, que no debería limitarse a inculcar principios normativos prefijados, sino que debe fomentar un entendimiento dinámico de cómo las tecnologías revelan y ocultan aspectos del mundo que habitamos.

Además, es crucial integrar estos principios éticos en la práctica diaria del desarrollo de la IA. Esto no solo se logra a través de directrices y políticas, sino también mediante la implementación de mecanismos de revisión y ajuste constantes que permitan a las tecnologías adaptarse a nuevos entendimientos y desafíos éticos que surjan a lo largo de su ciclo de vida. Estos mecanismos deben ser transparentes y accesibles para los usuarios y desarrolladores, permitiendo un diálogo abierto sobre las decisiones éticas en el diseño y uso de la IA. Este diálogo, por cierto, debe surgir en un contexto cultural de reflexión continua. Esto incluye espacios de diálogo y crítica donde los implicados puedan discutir y revisar no solo los aspectos técnicos, sino también las implicancias sociales y

éticas de su trabajo. Estos foros deben ser inclusivos, incorporando una variedad de perspectivas, incluidas aquellas de las comunidades afectadas por estas tecnologías, para garantizar que la IA se desarrolle de manera justa.

Este paso hacia una tercera ola apunta a una contextualización de la IA como un sistema sociotécnico que influye y es influido por la sociedad y sus dinámicas. Es volver a responder esa pregunta por la técnica que nos planteaba Heidegger: al entender la esencia de la IA, podemos comprender cómo influye en nuestra forma de ser en sociedad y, por lo tanto, nos permite desvelar y emanciparnos de su influencia, transformándola en un medio para conocer y significar nuestra experiencia vital. Así, avanzar en este conocimiento hermenéutico de la IA nos permite desvelar la ética de la IA, al establecer una relación entre las expectativas normativas, las prácticas mediadoras y la tecnología misma, que enriquece nuestra condición humana y nuestro proceso de coevolución en una sociedad tecnificada por la IA.

Este enfoque hermenéutico no solo desafía la aplicación superficial de la tecnología, sino que nos exige repensar cómo los principios éticos están fundamentalmente entrelazados con el avance tecnológico. En la práctica, esto significa que las consideraciones éticas de la IA deben evolucionar de ser reactivas y situacionales a convertirse en proactivas e integrales en cada fase del desarrollo tecnológico. Tal enfoque exige un cambio de ver la ética de la IA meramente como un cumplimiento regulatorio o una estrategia de gestión de riesgos, para abrazarla como un pilar fundamental de la innovación y la filosofía de diseño. Esta transformación implica incorporar deliberaciones éticas desde el inicio de la conceptualización tecnológica, asegurando que cada paso del desarrollo esté alineado con la previsión ética y los valores sociales.

Además, la noción de desvelar éticamente dentro de la IA requiere un diálogo continuo entre múltiples partes interesadas, incluidos eticistas, tecnólogos, legisladores y el público. Este diálogo no debería limitarse solo a círculos académicos o de expertos, sino extenderse a la comunidad más amplia para democratizar el desarrollo de las tecnologías de IA. Al involucrar una variedad diversa de voces en la conversación, podemos anticipar y mitigar mejor los desafíos éticos que plantea la IA, asegurando que la tecnología avance de una manera no solo técnicamente competente sino también socialmente responsable y culturalmente sensible.

Por último, la aplicación de un enfoque hermenéutico en la ética de la IA también debería involucrar una evaluación y adaptación continuas de las directrices éticas a medida que la tecnología y sus implicaciones sociales evolucionen. Este proceso dinámico de calibración ética nos permite responder a nueva información, percepciones y desafíos, asegurando que la ética de la IA permanezca relevante y robusta. Al fomentar un marco ético adaptable, podemos facilitar una

integración más reflexiva de la IA en la sociedad, mejorando el potencial de estas tecnologías para servir como verdaderos instrumentos para el beneficio humano y social.

3. Comienzos del desvelar ético: desmitificar, enseñar e integrar

La combinación de un proceso de conocimiento hermenéutico y una revelación emancipadora del enmarcamiento de la tecnología nos invita a reconsiderar no solo cómo construimos tecnología, sino también para qué y para quién la construimos. Este marco no solo amplía nuestra comprensión de la tecnología como fenómeno social y ético, sino que también nos desafía a desarrollar tecnologías que verdaderamente enriquezcan la condición humana, reconociendo y respetando nuestra integridad y valores compartidos.

Así, al fomentar una ética de la IA sociotécnica que reconozca y respete la complejidad de la existencia humana, se invita a repensar el desarrollo tecnológico de la IA. Esta perspectiva sociotécnica requiere un enfoque de ética de IA que sea dinámico y adaptativo, capaz de responder a los cambios tecnológicos y sociales continuos. Tal enfoque no solo mejora la capacidad de la IA para servir al bien público, sino que también fomenta un desarrollo tecnológico más consciente y responsable en un sentido existencial, entendiendo que la tecnología tiene impactos duraderos y profundos en la sociedad.

En consecuencia, aunque las críticas a los principios éticos de IA sugieren su inutilidad por la falta de consecuencias tangibles, es crucial reconocer el papel significativo que han jugado en fomentar el debate y la conciencia ética en este ámbito. Los principios éticos en esta segunda ola de la ética de IA han funcionado como un catalizador para que la comunidad de la IA y la sociedad en general presten atención a las implicaciones morales y sociales del desarrollo de la IA. Si buscamos establecer un lenguaje común y objetivos éticos claros, estos principios son la base para que diferentes partes interesadas, desde desarrolladores hasta legisladores, tengan un punto de partida para discutir y evaluar las tecnologías emergentes. Por lo tanto, aunque puedan ser vistos como ineficaces en términos de ejecución directa, considero relevante enfatizar que su valor reside en elevar la ética de la IA a un tema de interés y discusión global, con directrices que nos permitan alcanzar ciertos mínimos comunes.

Además, las críticas a la ética de la IA, tanto por la falta de eficacia de los principios éticos como por las falacias aditivas cometidas, apuntan a la necesidad de contextualizar la IA como una tecnología sociotécnica, motivando el

desarrollo de metodologías interdisciplinarias que integren estos principios traduciéndolos a la práctica tecnológica. Operacionalizar intervenciones éticas, sin embargo, no puede alcanzarse solo desde formalizaciones computacionales (*e.g.*, métricas de *fairness*) o herramientas técnicas (*e.g.*, para auditorías de datos) ni solo desde principios o directrices (*e.g.*, disminuir la discriminación y aumentar la representatividad). Incorporar la ética activamente en el diseño y desarrollo de sistemas de IA, tal y como he sugerido aquí, implica una colaboración más estrecha entre eticistas, ingenieros, desarrolladores, científicos sociales y otros incumbentes (*stakeholders*) para crear un marco de trabajo que no solo sea teóricamente robusto, sino aplicable en contextos reales y que pueda desarrollarse desde un proceso hermenéutico fructífero. Hacer ética de inteligencia artificial que no caiga en la ineficiencia implica, entonces, el desarrollo de visiones más holísticas y colaborativas.

Por otro lado, es esencial fortalecer los currículos educativos en el campo de la IA para asegurar que los futuros profesionales comprendan que la ética no es un añadido superficial, sino una parte integral de la práctica profesional. Las percepciones de los desarrolladores respecto a la ética tienden a ser negativas o escépticas. Estudios exploratorios han evidenciado que, en algunos casos, se reconoce como una necesidad, pero que su experiencia con la ética es restrictiva, ya que impide avanzar en innovación, además no tiene directrices claras de cómo aplicarla a su trabajo (Arriagada *et al.*). En otros casos, los desarrolladores de IA entienden su trabajo desde una neutralidad; es decir, que ser un desarrollador no es ni ético ni no-ético, ya que sus decisiones son sobre eficacia, optimización y decisiones técnicas (Griffin *et al.*). En ambos estudios, se evidencia la adopción de una neutralidad peligrosa, que descontextualiza a la IA como una herramienta y no la entiende en su dimensión sociotécnica. En otras palabras, refleja posturas, perspectivas y prácticas en los desarrolladores de IA que los mantienen en un enmarcamiento que limita su entendimiento de esta tecnología y, por lo tanto, les impide identificar y significar su rol y experiencia en este camino de coevolución con la tecnología.

El llamado, por ende, al dar un paso a la tercera ola de la ética de IA es contextualizar y situar el conocimiento ético desde un proceso de conocimiento hermenéutico que tiene un rol emancipador y busca crear una cultura ética y de responsabilidad colectiva. Además, este enfoque hermenéutico implica entender la tecnología no como una entidad aislada, sino como parte de un tejido social y cultural más amplio, donde su desarrollo y aplicación están inevitablemente influidos por las estructuras de poder y las normas sociales existentes. Por lo tanto, es crucial que los desarrolladores de IA trabajen de cerca con comunidades, grupos de interés y expertos en ética para interpretar y entender las implicaciones éticas de la tecnología en diferentes contextos.

4. Conclusiones

En este artículo he presentado una serie de elementos que contribuyen a dilucidar el cambio paradigmático en el rol que tiene la ética de la IA en el desarrollo de esta tecnología para su tercera ola. La segunda ola de la ética de la IA, como se ha discutido, centra su enfoque en abordar preocupaciones prácticas emergentes vinculadas al uso cotidiano de tecnologías, a través de principios y definiciones como la transparencia, la privacidad y la justicia. Este enfoque, si bien es un avance respecto a las preocupaciones más teóricas de la primera ola, todavía puede estar impregnado de cierto solucionismo tecnológico, al intentar resolver problemas éticos complejos mediante soluciones tecnológicas “guiadas” por principios, como métricas de justicia para disminuir sesgos o sistemas de IA explicativos. Sin embargo, estas soluciones a menudo no consideran la totalidad de las estructuras sociales y políticas subyacentes, lo que puede llevar a la implementación de soluciones que no abordan las raíces de los problemas.

Conuerdo así, parcialmente, con las críticas a la segunda ola de la ética de IA, que apuntan a la insuficiencia de principios éticos abstractos y descontextualizados para abordar los retos prácticos y morales que presenta esta tecnología. Sin embargo, me parece que esta crítica no apunta a la brecha de fondo. Y es que los principios éticos no tienen como función última regular acciones y comportamientos. Para poder habilitar un contexto en el cual los principios éticos tengan una bajada aplicada, se requiere una comprensión profunda de las condiciones sociotécnicas en las que se despliegan los sistemas de IA y las implicancias éticas y epistémicas que esto tiene en contextos de desarrollo científico y social. De lo contrario, los principios éticos corren el riesgo de permanecer como ideales teóricos sin aplicabilidad real, incluso si hay regulaciones que exijan su uso.

Para esto, sostengo que la ética de la IA debe ser considerada como parte integral del diseño y la operación de sistemas tecnológicos, implicando una interacción constante con las normativas sociales y las dinámicas de poder. En el análisis sugiero que una aproximación hermenéutica, que contextualice y sitúe el conocimiento ético dentro de procesos dinámicos y cambiantes, es esencial para desarrollar prácticas que sean genuinamente éticas y efectivas. Esto resuena con la crítica hecha por Heidegger sobre el “*Gestell*” o enmarcamiento, donde la tecnología no solo se ve como un conjunto de herramientas, sino como un sistema que configura nuestra relación con el mundo, reduciendo todo a recursos para ser explotados. En este sentido, el solucionismo tecnológico puede ser visto como una manifestación del enmarcamiento, donde la tecnología se promueve como la solución universal, ignorando las dimensiones éticas, políticas y sociales más profundas.

Sugiero, por tanto, que la tercera ola de la ética de la IA puede ofrecer un camino de integración y un desvelar de la ética, al enfocarse en cómo la tecnología puede ser diseñada y utilizada de manera que respete y mejore la condición humana sin reducir a las personas y sus culturas a meros datos o recursos. Este enfoque más matizado y crítico permite una integración de la IA en la sociedad que es consciente de las complejidades humanas y éticas, y que busca soluciones que sean no solo efectivas desde el punto de vista tecnológico, sino también justas y sostenibles desde una perspectiva social y cultural.

Este cambio de paradigma, sin embargo, tiene también desafíos significativos en la estandarización y aplicación global de prácticas éticas en IA debido a variaciones culturales, legales y económicas que pueden influenciar esas devoluciones contextuales. No obstante, esto puede ser también entendido como una oportunidad para desarrollar marcos éticos y metodologías éticas de trabajo que se ajusten a contextos locales mientras se mantienen alineadas con principios éticos transversales de la segunda ola.

A raíz del análisis propuesto en este trabajo, se llama a que investigaciones futuras se enfoquen en desarrollar métodos para la integración operativa de la ética en el ciclo de vida de los sistemas de IA, desde la concepción hasta la implementación y revisión desde las consideraciones sugeridas, para una integración más efectiva de los principios éticos en la práctica tecnológica. A través de un enfoque hermenéutico y la colaboración interdisciplinaria, podemos aspirar a desarrollar tecnologías de IA que sean éticas en un sentido robusto y no meramente por adición.



Referencias

- Arriagada, Gabriela; López, Claudia y Mendoza, Marcelo. *Ethics of AI and IT*. CRC Press, Taylor and Francis, próxima publicación.
- Beauchamp, Tom, y David DeGrazia. "Principles and Principlism". *ResearchGate*, 2004, pp. 55-74, https://doi.org/10.1007/1-4020-2127-5_3
- Binns, Reuben. "Fairness in Machine Learning: Lessons from Political Philosophy." *Proceedings of Machine Learning Research*, vol. 81, 2018, pp. 1-11.
- Boddington, Paula. *Towards a Code of Ethics for Artificial Intelligence*. 1st ed., Springer Cham, 2017. <https://doi.org/10.1007/978-3-319-60648-4>
- Borenstein, Jason, et al. "AI Ethics: A Long History and a Recent Burst of Attention." *Computer*, vol. 54, n.º 01, Jan. 2021, pp. 96-102. <https://doi.org/10.1109/MC.2020.3034950>.

- Bosk, Charles L. "Bioethics, Raw and Cooked: Extraordinary Conflict and Everyday Practice". *Journal of Health and Social Behavior*, vol. 51 Suppl, 2010, pp. S133-146. *PubMed*, <https://doi.org/10.1177/0022146510383839>.
- Bostrom, Nick, y Eliezer Yudkowsky. "The Ethics of Artificial Intelligence." 2014, pp. 316-34.
- Broussard, Meredith. *Artificial Unintelligence: How Computers Misunderstand the World*. MIT Press, 2018.
- Burrell, Jenna. "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms." *Big Data & Society*, vol. 3, n.º 1, June 2016, p. 2053951715622512. <https://doi.org/10.1177/2053951715622512>
- Chouldechova, Alexandra. "Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments". *Big Data*, vol. 5, n.º 2, 2017, pp. 153-63, <https://doi.org/10.1089/big.2016.0047>
- Coeckelbergh, Mark. "Responsibility and the Moral Phenomenology of Using Self-Driving Cars". *Applied Artificial Intelligence*, vol. 30, n.º 8, Sept. 2016, pp. 748-57. <https://doi.org/10.1080/08839514.2016.1229759>
- Danks, David, y Alex John London. "Algorithmic Bias in Autonomous Systems". *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, AAAI Press, 2017, pp. 4691-97, <http://dl.acm.org/citation.cfm?id=3171837.3171944>
- Davis, Matthew C., et al. "Advancing Socio-Technical Systems Thinking: A Call for Bravery". *Applied Ergonomics*, vol. 45, n.º 2, Mar. 2014, pp. 171-80. *PubMed*, <https://doi.org/10.1016/j.apergo.2013.02.009>
- Dencik, Lina, et al. "Towards Data Justice? The Ambiguity of Anti-Surveillance Resistance in Political Activism". *Big Data & Society*, vol. 3, n.º 2, Nov. 2016, p. 2053951716679678, <https://doi.org/10.1177/2053951716679678>.
- Dennis, Louise, et al. "Formal Verification of Ethical Choices in Autonomous Systems". *Robotics and Autonomous Systems*, vol. 77, Mar. 2016, pp. 1-14, <https://doi.org/10.1016/j.robot.2015.11.012>.
- Dignum, Virginia. "Ethics in Artificial Intelligence: Introduction to the Special Issue". *Ethics and Information Technology*, vol. 20, n.º 1, Mar. 2018, pp. 1-3. *Springer Link*, <https://doi.org/10.1007/s10676-018-9450-z>.
- Etzioni, Amitai, and Oren Etzioni. "Incorporating Ethics into Artificial Intelligence". *The Journal of Ethics*, vol. 21, n.º 4, Dec. 2017, pp. 403-18. <https://doi.org/10.1007/s10892-017-9252-2>.
- Floridi, Luciano, et al. "AI4People-An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations". *Minds and Machines*, vol. 28, n.º 4, Dec. 2018, pp. 689-707. <https://doi.org/10.1007/s11023-018-9482-5>.
- Griffin, Tricia A., et al. "The Ethical Agency of AI Developers". *AI and Ethics*, Jan. 2023, <https://doi.org/10.1007/s43681-022-00256-3>.
- Grunwald, Armin. "The Objects of Technology Assessment. Hermeneutic Extension of Consequentialist Reasoning". *Journal of Responsible Innovation*, vol. 7, n.º 1, jan. 2020, pp. 96-112. <https://doi.org/10.1080/23299460.2019.1647086>.
- Hagendorff, Thilo. "The Ethics of AI Ethics: An Evaluation of Guidelines". *Minds and Machines*, vol. 30, n.º 1, mar. 2020, pp. 99-120. <https://doi.org/10.1007/s11023-020-09517-8>.
- Heidegger, Martin. *THE QUESTION CONCERNING TECHNOLOGY AND OTHER ESSAYS*. Translated by William Lovitt, Garland Publishing, Inc., 1977.
- High-Level Expert Group on AI. *Ethics Guidelines for Trustworthy AI*. Publications Office, 2019, <https://doi.org/10.2759/346720>.
- Johnson, Deborah, y Mario Verdicchio. "Ethical AI Is Not about AI - Communications of the ACM." 1 feb. 2023, <https://cacm.acm.org/opinion/ethical-ai-is-not-about-ai/>.
- . "Ethics+AI Does not=Ethical AI:The Additivity Fallacy." University of Vienna.
- Mittelstadt, Brent. "Principles Alone Cannot Guarantee Ethical AI". *Nature Machine Intelligence*, vol. 1, n.º 11, Nov. 2019, pp. 501-07. *arXiv.org*, <https://doi.org/10.1038/s42256-019-0114-4>.
- Mittelstadt, Brent Daniel, et al. 'The Ethics of Algorithms: Mapping the Debate'. *Big Data & Society*, vol. 3, n.º 2, Dec. 2016, p. 2053951716679679, <https://doi.org/10.1177/2053951716679679>.
- Morozov, Evgeny. *To Save Everything, Click Here. Technology, Solutionism, and the Urge to Fix Problems That Don't Exist*. Penguin Random House, 2014, <https://www.penguin.co.uk/books/183571/to-save-everything-click-here-by-morozov-evgeny/9780241957707>.
- Munn, Luke. "The Uselessness of AI Ethics". *AI and Ethics*, vol. 3, n.º 3, Aug. 2023, pp. 869-77. <https://doi.org/10.1007/s43681-022-00209-w>.
- Narayanan, Arvind. "Tutorial: 21 Fairness Definitions and Their Politics". <https://www.youtube.com/watch?v=jlXluYdn-yyk>. ACM Conference on Fairness, Accountability, and Transparency (ACM FAccT).
- Nickel, Philip J. "Trust in Technological Systems". *Norms in Technology: Philosophy of Engineering and Technology*, vol. 9, editado por M. J. de Vries et al., Springer, 2013, <https://philarchive.org/rec/NICTIT>.
- "Robotics: Ethics of Artificial Intelligence". *Nature*, vol. 521, n.º 7553, May 2015, pp. 415-18. <https://doi.org/10.1038/521415a>.

- Swierstra, Tsjalling, y Hedwig te Molder. "Risk and Soft Impacts." *Handbook of Risk Theory: Epistemology, Decision Theory, Ethics, and Social Implications of Risk*, edited by Sabine Roeser et al., Springer Netherlands, 2012, pp. 1049-66. https://doi.org/10.1007/978-94-007-1433-5_42.
- van de Poel, Ibo. "Embedding Values in Artificial Intelligence (AI) Systems." *Minds and Machines*, vol. 30, n.º 3, Sept. 2020, pp. 385-409. <https://doi.org/10.1007/s11023-020-09537-4>.
- van Wynsberghe, Aimee. "Sustainable AI: AI for Sustainability and the Sustainability of AI". *AI and Ethics*, vol. 1, n.º 3, Aug. 2021, pp. 213-18. <https://doi.org/10.1007/s43681-021-00043-6>.



Datificación social e inteligencia artificial: ¿hacia un nuevo “salto de escala”?

Social Datafication and Artificial Intelligence: Towards a New “Scale Leap”?

MARTÍN GENDLER¹

IIGG-CONICET/IDEPI-UNPAZ, Argentina

martin.gendler@gmail.com

Fecha de recepción: 28/04/2024

Fecha de aceptación: 30/06/2024

Resumen

Desde la década de 1970 se asiste a un proceso de fuerte penetración de las tecnologías digitales en todos los ámbitos de la vida. Actualmente, su proceso de masificación y concretización ha permitido entrever una nueva fase del informacionalismo, íntimamente vinculado con el desarrollo de las tecnologías 4.0, donde el avance de la datificación social se presenta como un factor fundamental.

El presente artículo busca analizar los desarrollos actuales de inteligencia artificial dentro de las lógicas del proceso de datificación social, entendiéndolos como un elemento clave tanto para potenciar el mismo, como también abriendo la puerta a pensar un nuevo “salto de escala” en los actuales procesos de personalización y modulación algorítmica.

¹ Doctor en Ciencias Sociales, Licenciado y Profesor en Sociología (FSOC-UBA), Diplomado Superior en Gobernanza de Internet (UFRGS). Becario posdoctoral del CONICET. Miembro del Instituto de Investigaciones Gino Germani (IIGG-UBA) y del Instituto de Estudios para el Desarrollo Productivo y la Innovación (IDEPI-UNPAZ). Docente UBA y UNPAZ.
ORCID: <https://orcid.org/0000-0001-5781-6367>

CÓMO CITAR ESTE ARTÍCULO:

En APA: Gendler, M. (2024). Datificación social e inteligencia artificial: ¿hacia un nuevo “salto de escala”? *Resonancias*, (17), 121-141. DOI: 10.5354/0719-790X.2024.74503

En MLA: Gendler, M. “Datificación social e inteligencia artificial: ¿hacia un nuevo “salto de escala”?” *Resonancias*, n.º 17, julio de 2024, pp. 121-141. DOI: 10.5354/0719-790X.2024.74503

Palabras clave: inteligencia artificial, plataformas, datificación, tecnologías 4.0, internet

Keywords: Artificial Intelligence, Platforms, Datafication, Technologies 4.0, Internet

Para ello se despliega una breve genealogía de los procesos de datificación social, en conjunto con un análisis de las distintas fases del desarrollo de IA hasta el momento para ahondar en su entendimiento, características y efectos sociotécnicos.

Abstract

Since the 1970s, there has been a process of strong penetration of digital technologies in all areas of social life. Currently, its process of massification and concretization allowed to glimpse a new phase of informationalism, intricately linked to the development of 4.0 technologies, where the advance of social datafication is presented as fundamental.

This article seeks to analyze the current developments of Artificial Intelligence within the logic of the social datafication process, understanding them as a key element both to enhance it, as well as opening the door to thinking about a new “scale leap” in current processes. of personalization and algorithmic modulation.

To this end, a brief genealogy of the social datafication processes is displayed, linked to an analysis of the different phases of AI development so far to delve into the understanding, characteristics, and sociotechnical effects.

1. Introducción

A la hora de pensar en los procesos de datificación social, es común que se visualicen enormes bases de datos, sistemas algorítmicos o desarrollos cuasi salidos de la ciencia ficción. Lo cierto es que, si bien, actualmente nos encontramos en un período donde la datificación traspasa capilarmente casi todas las esferas de la vida social, política, económica y cultural, principalmente, a galope de los desarrollos y masificaciones de las tecnologías digitales (Mayer-Schönberger y Cukier 2013), es posible comprender que su genealogía viene de larga data (Harari 2018). Por ejemplo, uno de sus elementos clave, hoy día de moda, los algoritmos, datan de la antigua Grecia debiendo su nombre a una creación de reglas para ejercicios aritméticos creados en el siglo IX (Finn 2017).

Siguiendo a Mejías y Couldry (2022), la datificación comprende los procesos donde un elemento de la realidad concreta o simbólica *se convierte en datos*. Esto implica procesos de recolección, medida, almacenamiento de esos datos para distintos usos teóricos o aplicados, desde los sistemas de medición del tiempo y el espacio, las cantidades, el clima, los tamaños, hasta su aplicación a gran escala para la gestión de las poblaciones humanas. Para Deleuze (2014) y Foucault (2007), esta última aplicación comenzó a perfeccionarse y masificarse a partir del despliegue del diagrama de poder biopolítico, apoyado mayormente en los desarrollos de los saberes expertos de la medicina, la biología y la economía política. En este sentido, los datos empezaron a ser trabajados en pos de la confección de diversas *curvas de normalidad* poblacionales que permitieran establecer

indicadores para moldear subjetividades, cuerpos, edificaciones y gestiones de acciones y comportamientos, intentando *llevar hacia la media* a individuos y poblaciones. Cabe destacar que estas fueron acciones motorizadas y cuasi monopolizadas por el Estado y sus distintos organismos, adaptando los dispositivos disciplinarios que apuntaban al moldeo de cuerpos en espacios cerrados o delimitados hacia dispositivos biopolíticos de seguridad, los cuales enfocaban en los movimientos y comportamientos en circulación.

Avanzando en el tiempo, los procesos de datificación fueron perfeccionándose poco a poco, atravesados por los procesos bélicos de la primera mitad del siglo XX. En este sentido, los diversos mecanismos y estrategias para relevar, medir y procesar datos respecto de cantidad de efectivos, logística de alimentos, recursos bélicos y armamento, posiciones de ataque y defensa, como también el desarrollo de la criptografía, permitieron no solo brindar un salto de escala a los elementos de medición y aplicación de datos, sino también a sus lógicas (Kitchin 2014).

Es por este motivo que el final de la Segunda Guerra Mundial es el punto clave que Deleuze (1999) identifica como el inicio del cambio de diagrama de poder y transformación societal hacia una Sociedad de Control, donde los objetos técnicos, especialmente los digitales (Gendler 2023b), serían sus elementos clave (Rodríguez 2008).

Es posible signar estos movimientos en tres acontecimientos. En primer lugar, las Conferencias Macy de 1946 y el nacimiento de la cibernética (Rodríguez 2019), cuyos postulados apuntaban, principalmente, a la construcción de un paradigma transdisciplinar en el cual la recolección y medida de datos en pos sería vital para la construcción de un sistema capaz de predecir las conductas humanas, en función de modularlas para ordenar y conducir la vida social, previniendo la entropía (Tiqqun 2015). En segundo lugar, y a galope de los desarrollos cibernéticos, durante la década de 1960 se inició el trabajo de investigación, diseño y creación de distintos desarrollos técnicos en el sector militar de los EEUU, principalmente, en función de disponer de un sistema de distribución de información descentralizado capaz de potenciar las distintas carreras que tenían lugar en el marco de la Guerra Fría (Zuckerfeld 2014). Dentro de estos desarrollos, es posible identificar a ARPANET, la predecesora de Internet, pero también otras investigaciones incipientes, como la inteligencia artificial (IA), la realidad aumentada (RA) o la realidad virtual (RV).

En tercer lugar, se destaca el punto de quiebre que significó la crisis del petróleo de 1973 (Castells 1999), principalmente sus procesos posteriores. En pocas palabras, el jaque al sistema de producción capitalista basado en la materia y energía, generado por el boicot de los países productores de petróleo hacia el mundo occidental, desencadenó no solo una fuerte crisis económica,

sino también la planificación de alternativas posibles para salir de dicha crisis (Zukerfeld 2010). Se destaca que gran parte de estas opciones comenzaron a considerar a las tecnologías digitales como uno de los factores clave para la reorganización económica y social, considerando que la informatización de la sociedad podría traer respuestas para dar nuevos aires al capitalismo, generando un cambio en su modo de desarrollo productivo, pasando del industrialismo al informacionalismo (Castells 1999). Ejemplos de estos son el Plan Jacudi de Japón y el Informe Nora-Minc francés, continuados luego en la década de 1990 por el proyecto Infraestructura Global de la Información de los EEUU y el Informe Bangemann de la Unión Europea, teniendo su mayor hito en las Cumbres Mundiales de la Sociedad de la Información de la ONU y la UIT celebradas entre 2003 y 2005 (Mattelart 2006). Todas estas instancias, desplegadas en sintonía con la grilla de inteligibilidad neoliberal (Foucault 2007) en su paulatina conjunción con la cibernética como saber experto, dieron pie a un nuevo modo de configuración social (Eliás 1994) vinculado con las tecnologías digitales, los procesos de informacionalización y, por supuesto, transformaciones y nuevos impulsos a los procesos de datificación social.

Es a inicios de la década de 2010, tras los eventos ocasionados en la “crisis de las .com” de 2002, la crisis económica mundial de 2008 (Srnicsek 2016) y, especialmente, el nacimiento y masificación del *smartphone* (Sadin, 2017), que es posible signar el inicio de una nueva fase (Zukerfeld 2020) u oleada (Girolimo y Feldman 2020) del informacionalismo, uno principalmente vinculado con las tecnologías 4.0, las cuales tienen a la datificación como uno de sus factores clave (Gendler 2023a).

El presente artículo busca analizar los desarrollos actuales de inteligencia artificial, uno de los principales desarrollos de tecnologías 4.0, dentro de las lógicas del proceso de datificación social, entendiéndolos como un elemento clave tanto para potenciar el mismo, para la consolidación del ecosistema de tecnologías 4.0, como también abriendo la puerta a pensar un nuevo “salto de escala” en los actuales procesos de personalización y modulación algorítmica.

Para ello se despliega una breve genealogía de los procesos de datificación social en ambas etapas del informacionalismo, en conjunto con un análisis de las distintas fases de la IA hasta el momento, entendiendo a este desarrollo sociotécnico como uno de los principales en la nueva etapa inaugurada. Asimismo, se aborda el ecosistema presente de la IA para ahondar en el entendimiento, características y efectos que la masificación de este desarrollo tiene actualmente.

2. Del proceso de datificación social en el informacionalismo

Siguiendo a Castells (1999), el informacionalismo, como nuevo modo de desarrollo dentro del modo de producción capitalista, trajo aparejados varios cambios y transformaciones, principalmente en torno al principal insumo de los bienes, servicios y mercancías en general. Dentro de estos, es posible identificar el mayor peso que va asumiendo la información y el conocimiento en detrimento de la materia y la energía (Zukerfeld 2010), elementos claves durante el industrialismo. Esta cuestión no es menor dado que los principales bienes de este nuevo modo de desarrollo, los bienes informacionales, especialmente los primarios (Zukerfeld 2010), son elementos cuya materialidad está centrada en los bits que los componen. La ontología de los bits permite apreciar que estos son fácilmente replicables con una pérdida casi imperceptible de calidad, lo cual tiende su replicabilidad a un coste de cero (Rullani 2000); son exactamente iguales unos a otros y, a diferencia de interpretaciones que hablan de lo digital como sinónimo de “inmaterial”, los bits son un soporte material, aunque intangible (Cafassi 1998). Asimismo, se identifica que los objetos técnicos digitales (OTD) (Gendler 2023b) pueden integrar en los mismos artefactos todas las funciones de almacenamiento, procesamiento, reproducción, transmisión y conversión que las tecnologías analógicas de la información hacían por separado, teniendo su clave en la integración, pero también en su flexibilidad.

Sumado a estas características, los OTD también conllevan en su diseño o despliegan en su acción procesos de interpretación (Blanco y Rodríguez 2015) de otros OTD para su funcionalidad, ya sea entre softwares o en su vinculación con un hardware, siendo la interconexión y *diálogo* constante entre los mismos un factor fundamental de su funcionamiento y tecnicidad (Simondon 2016).

Ahora bien, el desarrollo del informacionalismo y la masificación de los OTD en todos los intersticios de la vida social trajo aparejado la posibilidad de que su materialidad, basada en bits, no solo sirva para transmitir información digital y permitir su funcionamiento e interpretación conjunta, sino también para su *almacenamiento*; es decir, para *convertir a esos bits en datos medibles y reapropiables*. Esto abriría una nueva puerta al impulso de la datificación social, incrementando fuertemente la cantidad de datos disponibles al facilitar su registro, dado que cada flujo de información y actividad de (y en) los OTD pasaba a ser plausible de ser convertida en datos. Sin embargo, como bien menciona Lessig (2006), durante los primeros pasos de la “red de redes”, especialmente en su etapa como ARPANET y como NSFNET, esta no permitía registrar e identificar

datos e información digital por las características de diseño de sus protocolos principales, especialmente el TCP/IP y el *packet switching*².

Sería la invención y desarrollo de las *cookies* en 1994, pensadas para mejorar las funcionalidades del primitivo comercio electrónico, aquellas que sentarían las bases de un OTD capaz de almacenar los registros de actividad y características provenientes de individuos o colectivos que usaran otros OTD (Rossi 2018; Gendler 2023a), sentando un precedente que luego fue apropiado cooptativamente (Lago Martínez, Méndez y Gendler 2017) por las principales agencias de seguridad nacionales tras los atentados del 11 de septiembre de 2001³. En breve, estas agencias requirieron a los principales proveedores de servicio de internet de los EEUU a almacenar y registrar no solo los datos personales de los ciudadanos que se conectaban a Internet, sino también sus registros de actividad⁴ (Sadin, 2018), dando paso a nuevos OTD especializados en dicha labor.

En relación, la “crisis de las .com” del año 2001/2002 (Srnicek 2016) permite visualizar otro punto de inflexión en estos procesos dado que, tras esta, el ecosistema de Internet y su modelo de negocios experimentaría una fuerte mutación, normalmente identificada en el paso de la web 1.0 a la 2.0 con la emergencia y despliegue de los nuevos medios conectivos (Van Dijck 2016), las llamadas “redes sociales”⁵. En pocas palabras, la relación acerca que la crisis había sido generada por la falta de un modelo de negocio y una rentabilidad claros en las empresas

² Por las características originarias de la arquitectura de la red de redes, el diseño de estos protocolos se centró en intentar ofrecer simplicidad, robustez, modularidad, rendimiento y eficiencia a los intercambios de flujos de datos, texto y contenidos entre los internautas. Lessig (2006) argumenta que fue debido a esto que no hubo una preocupación o interés inicial en que dichos protocolos contuvieran alguna funcionalidad técnica de almacenamiento o registro de la información, la cual podía llegar a generar dificultades en la primitiva operatividad de la red. A esto podemos añadir que, en estos primeros tiempos, los diseñadores de la red eran también, en su gran mayoría, sus mismos usuarios, quienes se movían mayormente bajo una lógica colaborativa, considerándose innecesario el elaborar mecanismos de registro de información dado que esta era, cuasi en su totalidad, compartida libremente en pos de la mejora continua de procesos y desarrollos (Gendler 2023a).

³ Una *cookie* informática es un pequeño archivo de datos diseñado para almacenar información sobre el usuario en un sitio web. Como se ha mencionado, originalmente estos OTD tenían como objetivo guardar información temporalmente en la sesión del navegador del usuario para facilitar tareas como las compras en línea. Hoy en día, existen distintos tipos de *cookies*, algunas que siguen cumpliendo el rol de las primeras, otras que se almacenan en el navegador o en el dispositivo del usuario hasta que este las elimine. Se destaca que estas últimas, mayormente, se enfocan en almacenar preferencias o en hacer seguimiento de las acciones del usuario transmitiendo información hacia quienes las han creado o movilizado. Es importante destacar que, en la actualidad, múltiples actores pueden utilizar *cookies* con fines de registro y almacenamiento (no solo temporal) de datos directos e indirectos.

⁴ Cabe destacar que entre estos dos conjuntos hay una diferenciación clave. Por un lado, los datos directos, otorgados de forma explícita por los usuarios, como ser aquellos necesarios para configurar un perfil o un formulario. Por otro lado, los datos indirectos, es decir, todos aquellos que pasan a ser almacenados fruto de la actividad de los individuos sin que sean brindados por ellos de forma explícita (Gendler 2019).

⁵ Siguiendo a Srnicek (2016), la llamada “crisis de las .com” fue una crisis económica que tuvo lugar a inicios del nuevo milenio, principalmente, debido a la sobreestimación de las posibilidades de ganancia de las empresas dedicadas a Internet, materializado en el valor de sus acciones en el mercado bursátil. En pocas palabras, cuando las expectativas empezaron a desinflarse debido al hecho concreto de que estas empresas no reportaban márgenes de ganancias significativos, el valor de sus acciones se derrumbó.

de Internet y sus principales OTD en la última década del siglo XX, generó que se empiecen a pensar nuevos mecanismos de rentabilidad para ellas, como el cobro por el uso (Rullani 2000), o el que finalmente terminó primando: la publicidad. Cabe aclarar que esta sería distinta de las presentes en los medios de comunicación masivos como también de las que se podían encontrar en las webs primitivas, dado que el nuevo desarrollo sería una *publicidad orientada* a ser mostrada a aquellos usuarios que puedan compatibilizar con ese consumo/producto. Es decir, las nacientes redes sociales y otras webs sobrevivientes de la crisis -como Google, Ebay o Amazon- pasaron a aplicar e ir perfeccionando los distintos mecanismos y OTD disponibles para almacenar datos directos e indirectos de los individuos en función de procesarlos, generando como resultado una publicidad que se mostraría solo a aquellos que “hubieran demostrado que le interesaba” (Samela 2016).

Sería el nacimiento del smartphone en 2007 y su posterior difusión, con la posibilidad de establecer una conexión virtualmente ininterrumpida, geolocalización constante y nuevos desarrollos y aplicaciones orientados al mercado móvil (Sadin 2017), lo que abriría la puerta a un nuevo modelo de negocios y a nuevas formas de organización de los OTD: el modelo de las plataformas digitales.

Siguiendo a Van Dijck, Poell y Waal (2018) es posible diferenciar a las plataformas en dos grandes grupos. En primer lugar, las plataformas infraestructurales, entre las cuales se encuentran las GAFAM⁶, que consisten en los principales sistemas sociotécnicos para el intercambio de recursos, mensajes y contenidos; para la coordinación de prácticas, movimientos y, especialmente, de los flujos de datos y comunicaciones. Cabe agregar que estas plataformas son actores hegemónicos de varios segmentos particulares dentro del ecosistema de medios conectivos, cumpliendo cada una de ellas el rol de brindar su infraestructura para un enorme abanico de plataformas específicas, aquellas llamadas “plataformas sectoriales”, las cuales no podrían operar sin esta base brindada por las infraestructurales.

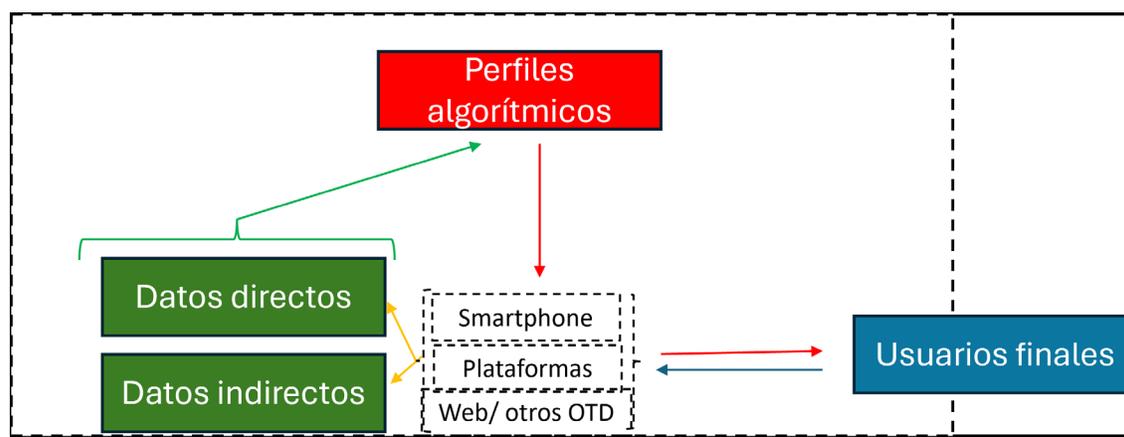
Es necesario destacar que estos procesos generarían un fuerte salto en la masividad de la cantidad de datos almacenados, enriqueciendo y complejizando las bases de datos disponibles para su procesamiento y aplicación (Crawford 2023), consolidando el proceso de datificación a la vez de *llevarlo a otro nivel*. En otras palabras, al irse arraigando las plataformas como modelo y como espacio de actividad y flujos de atención en Internet (Celis Bueno 2017), en gran parte gracias a las posibilidades abiertas por el smartphone y el desarrollo del 4G, y al tener estas como principal funcionalidad el registro de datos directos e indirectos

⁶ Google (Alphabet), Amazon, Facebook (Meta), Apple y Microsoft.

de los usuarios para orientar publicidad, la cantidad de datos digitales en juego pasó a verse incrementada exponencialmente, volviéndose su almacenamiento, recolección y procesamiento cuasi una *lógica* del diseño y operatoria de los OTD.

El enorme volumen de datos disponibles, como también la optimización y mejora de las formas de procesamiento, pronto dio pie a nuevos dispositivos y OTD que fueron mucho más allá de la lógica de la publicidad, generándose mecanismos de personalización no solo publicitarios, sino de toda la actividad de las personas en su relación con OTD conectados a Internet (Gendler 2023a). En otras palabras, generando *perfiles generales personalizados que se aplican a los individuos* para modular su actividad en vínculo con los OTD, iniciando un acompañamiento algorítmico de la vida (Sadin 2017) y dando a paso a un nuevo tipo de gubernamentalidad: la gubernamentalidad algorítmica (Rodríguez 2019; Rouvroy y Berns 2015), la cual, a diferencia del período biopolítico, sería también ejercida por el sector privado y no solo por el Estado. De esta manera, podemos pensar al sistema “recolección-almacenamiento-procesamiento-perfilamiento y aplicación” (RAPPA) como la lógica principal de la operatoria de las principales empresas de Internet y OTD, distinguiendo los procesos de datificación social, especialmente durante mediados de la década de 2010.

Figura n.º1. Sistema RAPPA



Fuente: Elaboración propia

2.1. Del paradigma 4.0

Si bien el nacimiento y masificación del modelo plataformas, junto a la paulatina introducción del sistema RAPPa como lógica de la tecnicidad de los OTD interconectados a Internet, son los elementos que permiten entender al pasaje a una segunda “fase” (Zukerfeld 2020) u “oleada” (Girolimo y Feldman 2020) informacional, principalmente por las tendencias de automatización, digitalización y plataformización que conllevan, cabe destacar otros procesos que se desplegaron de forma paralela a lo previamente mencionado en el aparato anterior. En breve, entre 2009 y 2011 empezó a ser motorizado el plan Industria 4.0, el cual apuntaría a una “próxima generación” de OTD (Schwab 2016; Benanav 2022): los OTD 4.0. De este modo, se propició la inversión y el desarrollo de tecnologías digitales de larga data como la IA, la RV, la RA, la biología sintética, la impresión 3D, a la vez de potenciar innovaciones más recientes como la Internet de las cosas (IoT) o el *blockchain*. Siguiendo a Schwab (2016), los OTD propios del ecosistema 4.0 tendrían incumbencia en los dominios físicos, digitales y biológicos (y en sus intersecciones) generando nuevos saltos de escala en la innovación y posibilidades de mejora de las sociedades en conjunto en lo que respecta a la esfera económica, pero también a la política, social y cultural, auspiciando una cuarta revolución industrial, una “sociedad de la industria 4.0” (Finquelievich et al. 2024) o, incluso, un capitalismo 4.0 (Galliano 2020).

Cabe destacar una cuestión fundamental: el elemento que opera como condición vital para el funcionamiento, para la innovación y mejora continua, para la potenciación de las posibilidades de todos los OTD 4.0, y que, a la vez, obra como elemento común conjugándolos en un mismo ecosistema, no es otro que el proceso de datificación social. Es decir, todos los desarrollos e innovaciones de este nuevo paradigma al interior del informacionalismo funcionan en base a su vínculo con los datos digitales, ya sea por su entrenamiento en ellos (IA), por su forma de visualización y despliegue (RA y RV), por consistir el registro, traslado y operatoria de ellos en su cuestión fundamental (IoT y sistemas ciberfísicos), entre otros casos, a la vez de también posibilitar su recolección, almacenamiento y procesamiento. De este modo, es posible advertir que los procesos e hitos destacados en el apartado previo en torno a la consolidación y masividad del proceso de datificación social actual *obraron de condición de posibilidad*, tanto para el desarrollo y masificación de las nuevas tecnologías 4.0 como también para esta nueva etapa del informacionalismo.

Figura n.º2. Tecnologías propias del paradigma 4.0 y dominios de incumbencia

Dominio	Tecnología	Año de aparición
Físico	Impresión 3D	1980
	Nuevos materiales	2004
Físico-Digital	Sistemas ciberfísicos	década de los 2000
	Vehículos autónomos	1994
	Robótica Avanzada	1980
	Robótica Colaborativa	1996
	IoT	2008
	Realidad Virtual	década de 1960
	Realidad Aumentada	década de 1960
	Realidad mixta	década de 1990
	Nanotecnología	década de 1960
Digital	Gemelo Digital	década de 1990
	Computación en la Nube	entre las décadas de 1960 y los 2000
	Fog Computing	2014
	Blockchain	2008
	Plataformas	década de los 2000
Biológico	Biología sintética	2000
	Ingeniería genética	1973
Transversal (Físico-Digital-Biológico)	IA	década de 1950
	Big Data Analytics	década de los 2000
	Bioteecnología	década de 1920
	Neurotecnología	década de los 2000

Datificación

Fuente: Elaboración propia

3. De la inteligencia artificial

Siguiendo a Tegmark (2017), Crawford (2023) y Prodnick (2022), los distintos desarrollos relacionados con la inteligencia artificial tienen una larga trayectoria con varias fases y etapas. En primer lugar, es necesario mencionar que el término IA funciona, en parte, como “concepto paraguas” de múltiples desarrollos distintos, como son el *machine learning*, el *deep learning*, las redes neuronales, la inteligencia artificial generativa, el procesamiento de lenguaje natural, la robótica avanzada, entre muchos otros. En segundo lugar, que muchos de los anuncios y discursos respecto de sus potencialidades responden más bien a una fase “general” de la IA que, al menos de momento, pareciera lejos de estar por llegar, teniendo actualmente un estadio “estrecho” (Madakam et al. 2020). En tercer lugar, sus umbrales de acción y sus potencialidades se vinculan no solo con los avances teóricos y técnicos en torno a capacidades y posibilidades, sino también con el *tipo, cantidad y formas de procesamiento de datos de entrenamiento* utilizados para su programación, optimización y mejora.

Cabe aclarar que los datos de entrenamiento son diversos conjuntos y bases de informaciones recolectadas que pueden ser de distintos tipos: sonidos, textos, imágenes, videos, datos geoespaciales, datos biométricos, series temporales, entre un enorme etcétera. En este sentido, el funcionamiento de la IA opera bajo la lógica *input-output*, pero con algunos condimentos, dado que los procesamientos que se realizan en el ínterin están mayormente condicionados por algoritmos de aprendizaje y de clasificación de los datos de entrenamiento, los cuales generan un *output* en función de las características de los datos, de su etiquetado y de las reglas y lógicas de estos algoritmos de procesamiento.

Nacida oficialmente en 1956 en la Conferencia de Dartmouth, la IA, sus lógicas y los tipos de datos de entrenamiento utilizados tuvieron que atravesar distintas etapas hasta llegar al día de hoy. Siguiendo a Crawford (2023), los primeros datos de entrenamiento consistían en datos lógicos, principalmente figuras geométricas, funciones aritméticas, cuestiones vinculadas a la matemática, la lógica y los juegos de mesa. Cabe destacar que estos primeros desarrollos de IA, cuyo enfoque es conocido como “Good Old Fashioned AI” (GOF AI), utilizaban estructuras formales de lógica y normas de inferencia teniendo un rígido sistema de reglas introducido para tomar decisiones y realizar tareas. Asimismo, trabajaban con datos guardados en la memoria de los sistemas bajo procesos seriales o secuenciales; es decir, bajo un enfoque lineal “uno a uno”, procesando cada dato individualmente, siguiendo una secuencia predefinida de instrucciones (McCarthy 1979)⁷. Ya entrada la década de 1970, se pasó a desplegar sistemas de reglas que intentaban emular el razonamiento lógico, teniendo algunos inconvenientes a la hora de experimentar con algunas imágenes de la vida real, especialmente por su falta de flexibilidad y capacidad para adaptarse a nuevas situaciones y a datos no estructurados o mal definidos en las reglas existentes (Crawford, 2023). Siguiendo a Toosie *et al.* (2021), fue en esta década que, a raíz de los informes ALPAC de los EEUU del año 1966 y Lighthill del Reino Unido del 1973, inició una etapa denominada como “el primer invierno de la IA”. En pocas palabras, en estos informes se destacaba que los avances y desarrollos realizados “no estaban a la altura de las expectativas”, generando una abrupta disminución del financiamiento y del interés en investigaciones y desarrollos de IA.

Fue debido a estos inconvenientes técnicos y financieros que en la década de 1980 se pasó a aplicar modelos probabilísticos y estadísticos, intentando calcular el mayor número de opciones hasta encontrar resultados óptimos, iniciando una nueva ola de interés por estos desarrollos. Es importante resaltar aquí un cambio fundamental: el pasaje de los modelos secuenciales a un tipo de paradigma de procesamiento paralelo, también llamado “modelo conexionista”, el cual

⁷ Uno de sus ejemplos paradigmáticos fue el lenguaje de programación LISP, el cual se centraba en el procesamiento de listas de datos y la manipulación e interpretación de símbolos (McCarthy, 1979).

pasó a ejecutar múltiples operaciones en simultáneo, habilitando procesamientos de manera distribuida (McClelland y Rumelhart 1981). Este modelo, mucho más flexible y adaptable que el secuencial, especialmente en cuanto a sus reglas y lógicas de funcionamiento, habilitó el poder trabajar con grandes volúmenes de datos en función de la frecuencia de los *inputs*, reconociendo patrones complejos y utilizando etiquetas manuales para guiar el aprendizaje de características relevantes de los (conjuntos de) datos.

Dentro de las múltiples operaciones que se empezaron a llevar a cabo, Crawford (2023) destaca una en particular: el trabajo con las frecuencias de relación entre palabras, “reduciendo el habla a datos” (p. 156) y optimizando sistemas de etiquetado que permitieran detectar relaciones y establecer inferencias. Cabe destacar que las principales fuentes elegidas fueron los registros digitalizados de una demanda antimonopolio contra IBM de finales de la década de 1960, los emails de los empleados de la empresa Enron, acusada por fraude, y del proyecto Penn Treebank, una base de texto de cuatro millones de palabras del inglés estadounidense, principalmente reportes policiales y cables ministeriales. Esto no es menor para la autora, ya que indica que la procedencia de estos primeros datos de entrenamiento sentaría las bases para varias lógicas y sesgos de aprendizaje que tendrían los desarrollos de IA hasta nuestros días⁸.

Adentrándose en la década de 1990, se pasó a trabajar con datos de entrenamiento basados en imágenes, especialmente aquellos provenientes del proyecto Face Recognition Technology (FERET), los cuales incluían fotos de prontuarios policiales, licencias de conducir, entre otros, en función de permitir el etiquetado, clasificación y posibles vinculaciones por género, edad y raza. También de esta década constan los primeros registros de datos de entrenamiento de audio y de video. Cabe destacar que en este período se advirtió un “segundo invierno de la IA”, principalmente representado en las múltiples limitaciones que presentaban sus desarrollos en contraposición a la euforia abierta por la apertura y masificación de Internet, incluso al punto de ser renombrados muchos de sus proyectos como “informáticos” o “analíticos”, evitando la sigla IA para no perder su financiamiento (Toosie *et al.* 2021).

Llegada la década de los 2000, y con una Internet cada vez más masiva, los datos de entrenamiento pasaron a constituirse de las diversas producciones y contenidos que los usuarios desarrollaban en la web, en coincidencia con las apropiaciones del uso de las *cookies* que se mencionaron previamente en este artículo, popularizando el término “Big Data” (Magnani 2019). En este sentido, Crawford (2023) y Tegmark (2017) destacan que, incluso previamente al modelo

⁸ De todos modos, también es importante señalar otros desarrollos en la misma línea que no necesariamente utilizaban bases sesgadas, como el sistema de redes neuronales NETtalk (Sejnowski y Rosenberg 1987).

de plataformas, la recolección y almacenamiento de datos provenientes de la web para utilizarlos como base de entrenamiento del procesamiento y generación de *outputs* de IA pasó a ser la forma hegemónica de operación. Asimismo, los autores relatan como clave en esta década el caso de ImageNet, un proyecto que intentó ser la principal base de datos de objetos del mundo, el cual requirió de una enorme cantidad de etiquetado humano inicial para permitir luego que los desarrollos de IA puedan iterar y procesar en función de aprender a clasificar automáticamente. Sin embargo, Crawford (2023) menciona que las etiquetas posibles para cada imagen venían ya prefijadas desde las bases de datos de palabras provenientes de la década de 1980, las cuales ya contaban con fuertes problemas en cuanto a sesgos. De este modo, el clasificador humano quedaba mayormente supeditado a utilizar algunas de las opciones ya prefijadas, estén sesgadas o no.

Cabe destacar asimismo que las posibilidades abiertas por el exponencial incremento cuantitativo de datos de entrenamiento, en conjunción con las optimizaciones del entrenamiento reforzado y la incorporación de las redes bayesianas a los desarrollos de IA, optimizando el conocimiento sobre la interdependencia de variables para hacer inferencias y predicciones (Pearl 2009), hicieron que pronto el “segundo invierno de la IA” diera paso a un nuevo período de entusiasmo (Toosie *et al.* 2021).

Ahora bien, como ya se ha mencionado, la masificación del *smartphone* incrementó exponencialmente el volumen de datos disponibles, mientras que el modelo plataforma permitió a un pequeño puñado de empresas (y las que deseen pagar por ellos, los famosos “terceros”) disponer de los datos que se generasen en el “terreno” de su infraestructura. Esta fue la condición de base para pasar del modelo de la orientación al de la personalización, creando los perfiles personalizados que luego se aplicarían a los individuos para modular su acción, como también para obtener información constante y en tiempo real útil para mejorar su plataforma y sus procesos de recolección de datos, para poder vender o intercambiar con otras empresas, incluso, para disponer de un cierto poder de lobby político (Zuazo 2018). En otras palabras, para acentuar cierto oligopolio sobre el principal insumo del informacionalismo: la información y el conocimiento derivado de la misma. Asimismo, el enorme nuevo volumen de datos generó una optimización de los desarrollos de aprendizaje no supervisado; es decir, aquellos que trabajan sobre datos no etiquetados explorando sus estructuras internas para intentar encontrar patrones que permitan generar *outputs* de agrupamientos, *clusterización* y detección de anomalías (Pandel y Patel 2023).

4. ¿Un nuevo salto de escala?

4.1. Aceleración y nuevos elementos

Siguiendo a Prodnik (2022), desde inicios de la década de 2010 las empresas de plataformas empezaron a incorporar diversas aplicaciones de IA en las lógicas de diseño y funcionamiento de sus desarrollos, especialmente, en torno a ir optimizando el funcionamiento de sus perfiles, búsquedas y recomendaciones algorítmicas. Ejemplos de esto fueron la paulatina incorporación de desarrollos de aprendizaje automático para personalizar el *news feed* en Facebook, las búsquedas en Google o la identificación de tendencias en Twitter. Sin embargo, es a partir de 2016 cuando es posible identificar una mayor velocidad y profundidad en la incorporación de OTD de IA, llegando al punto que para 2019, cerca del 90% de la actividad en Internet se encontraba mediada por la lógica del sistema RAPP (Gendler 2019). Asimismo, el desarrollo e incorporación de nuevas técnicas de procesamiento de información como el *aprendizaje federado* desarrollado por Google en 2016, el cual permite que los datos sean procesados por modelos de aprendizaje en varios dispositivos dispersos (entre ellos, los *smartphones* de los usuarios) en lugar de centralizarse en un único servidor (Li *et al.* 2020), también permiten apreciar una mayor *sinergia creciente* entre la datificación y el desarrollo de los OTD IA.

En este sentido es posible establecer una relación de *doble reciprocidad* entre el proceso de datificación y los desarrollos de IA, dado que el primero, cada vez más, opera como condición de base para el entrenamiento y desarrollo de los segundos, mientras que estos fomentan, optimizan y perfeccionan los mecanismos de recolección, almacenamiento y procesamiento de datos, luego utilizados para los procesos de perfilado y aplicación.

Cabe destacar que la pandemia global ocasionada por el COVID-19, la emergencia sanitaria y los diversos procesos de gestión de las poblaciones desplegados intentando minimizar los efectos del virus, implicaron un mayor uso y despliegue de los OTD intentando sostener los espacios de trabajo, educación y sociabilidad otrora presenciales (Gendler y Andonegui 2021). Asimismo, diversos mecanismos digitales de geolocalización e identificación fueron desplegados en varios países y regiones en búsqueda de controlar las acciones y disminuir posibles riesgos de contacto. Todos estos aspectos implicaron un crecimiento exponencial de la cantidad de datos generados y distribuidos en Internet vía OTD, especialmente dentro de las plataformas hegemónicas, incrementando las cantidades disponibles para ser almacenados, procesados e incluidos en nuevas bases de entrenamiento en pos de mejorar los desarrollos de IA. Esto permitió generar una aceleración de los desarrollos en marcha al disponer de un enorme

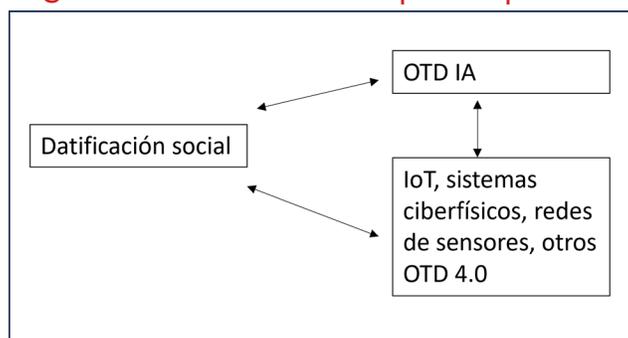
volumen nuevo de datos para entrenamiento en un período mucho menor de lo esperado, por ejemplo, en lo que respecta a la IA generativa⁹, pero también en la optimización de los asistentes de voz, los vehículos autónomos, los sistemas de traducción automática, el diagnóstico de imágenes, entre otros.

Pronto no solo fue posible visualizar los despliegues y desarrollos de nuevos proyectos y avances, como el Metaverso/Omniverso (Ball 2022) y la inteligencia artificial generativa, sino también el perfeccionamiento de otros componentes del ecosistema de tecnologías 4.0 como son la IoT, los sistemas ciberfísicos, la red de sensores, entre otros. Como se ha mencionado previamente, estos OTD ya existían, pero principalmente en fase experimental o de aplicación controlada. En este sentido, los avances en los desarrollos de IA también permitieron y aceleraron su incorporación a estas tecnologías 4.0, permitiendo, a su vez, que estas empezaran a tomar mayor relevancia en la escena.

Respecto del IoT, los sistemas ciberfísicos y las redes de sensores, cabe destacar que su mayor despliegue incorporaría una nueva dimensión a la cuestión, dado que, además de la posibilidad de recolección de datos directos e indirectos de los individuos, se suma la posibilidad del registro de *datos del entorno* (ya sea de la naturaleza, de las casas, oficinas, elementos de la ciudad, entre otros), abriendo la puerta a la cosecha constante y en tiempo real de nuevas fuentes de información previamente poco exploradas. De esta forma es posible evidenciar un tercer factor que se suma a la relación de reciprocidad entre los OTD de IA y el proceso de datificación social, dado que los datos almacenados, procesados y transmitidos por los OTD 4.0 permiten también reforzar y dar un mayor nivel de complejidad para el entrenamiento de los desarrollos de IA, profundizando, a su vez, el impulso de la datificación social en su fase actual, a la vez que estos optimizan y mejoran la operatoria de los mecanismos de recolección y tratamiento de datos de los OTD 4.0.

⁹ Ejemplos posibles son los ya célebres Chat-GPT3 y Dall-E de OpenIA, Deep Dream y Bard/Gemini de Google, Co-Pilot de Microsoft, entre un enorme etcétera.

Figura n.º3. Relación de triple reciprocidad



Fuente: Elaboración propia

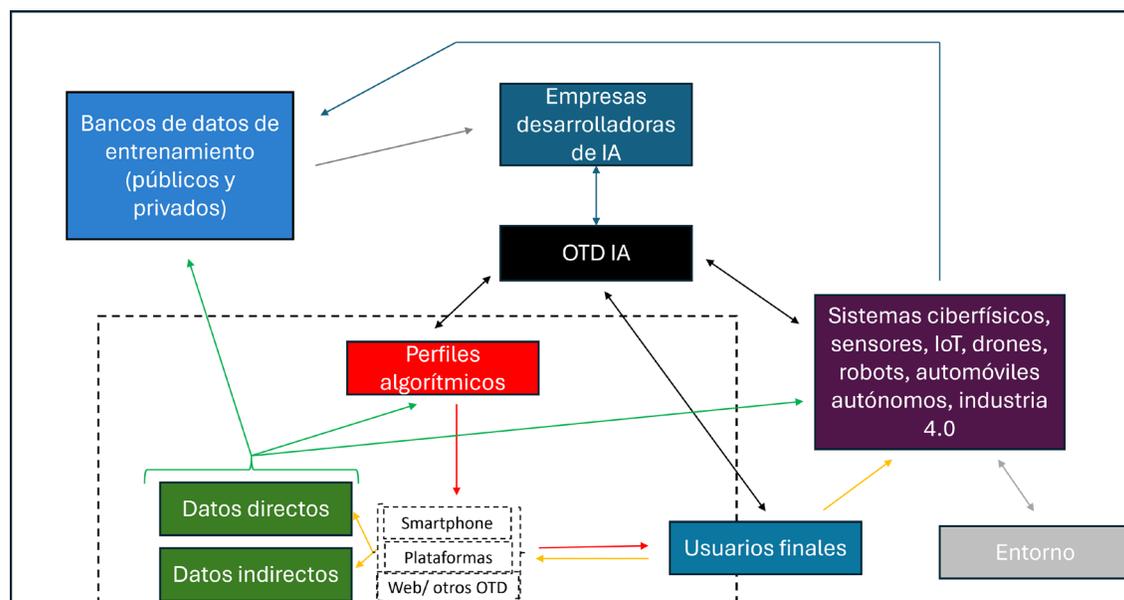
4.2. Los actores

Actualmente, las empresas que generan y despliegan los diversos desarrollos de OTD IA cuentan en sus filas con grandes firmas ya establecidas; entre ellas, muchas de las empresas de plataformas como Google, Amazon, Facebook, Apple y Microsoft, junto a otras emergentes específicas como OpenIA, Anthropic, Databricks, MistralAI, entre otras. Respecto de las GAFAM, cabe destacar que cada una de ellas se encuentra invirtiendo fuertemente tanto en desarrollos de OTD IA que ayuden a optimizar la operatoria RAPPa en los segmentos del ecosistema de Internet donde son líderes, como también como puerta al ingreso a competir en otros segmentos, muchas veces en asociación con las emergentes específicas (Molina 2024). Cabe también mencionar a un tercer grupo de empresas: un enorme abanico de *startups*, las cuales muchas veces se encuentran relacionadas directa o indirectamente con las empresas anteriores al no disponer de los recursos técnicos o económicos para generar desarrollos de forma autónoma. De esta forma, es posible plantear que las *startups* de IA podrían considerarse como empresas *sectoriales* que dependen necesariamente de los desarrollos llevados a cabo por las *infraestructurales* de la IA, ya sean las GAFAM o las nuevas emergentes consolidadas. Ejemplos de esto pueden ser la asociación de Microsoft con *startups* como MistralAI, Alt.ai y Bonsai o la recurrencia de múltiples *startups* a la infraestructura provista por OpenIA o Gemini (de Alphabet/Google), para sus desarrollos y mejoras (Widder, West y Whittaker, 2023). Esto podría permitir augurar que las mismas divisiones que se generaron dentro del ecosistema de las plataformas parecerían extenderse a estos nuevos desarrollos.

4.3. Salto de escala

Todos los elementos antes mencionados permiten conducirnos al tratamiento de la principal problemática planteada en este artículo. Teniendo en cuenta todos los avances, despliegues y masificaciones de OTD IA en los últimos años en estrecha reciprocidad con los procesos de datificación social, sumando a su vez al factor de los OTD 4.0, es posible considerar una *ampliación del sistema RAPP*, tanto en lo que respecta a sus elementos y actores, como también a su operatoria.

Figura n.º4. Sistema RAPP ampliado



Fuente: Elaboración propia

La figura n.º4 nos permite entrever como los datos directos e indirectos de individuos y colectivos ya no son el único recurso utilizado para generar los procesos de perfilamiento, recomendación y gubernamentalidad algorítmica a través de distintos OTD como sucedía, principalmente en las plataformas, hasta, al menos, 2016. En cambio, estas fuentes se amplían tanto por los bancos de datos de entrenamiento disponibles, alimentados constantemente, como por las nuevas

dimensiones de registro y recolección de datos habilitadas por la introducción y desarrollo de los OTD 4.0. De este modo, los perfiles algorítmicos aplicados para modular las distintas esferas de la vida, en la actualidad no solo disponen de una amplificación exponencial y en tiempo real de los datos que permiten ir entrenándolos y perfeccionándolos, sino también de un mayor poder de cómputo y procesamiento para la modificación constante de estos perfiles y su aplicación brindado por los OTD IA, recordando que estos, a la vez, perfeccionan estas capacidades gracias a su entrenamiento constante con estas masas incrementales de datos y ayudan a optimizar las operatorias de los OTD 4.0.

Esto permite vislumbrar el estado alcanzado actualmente por el proceso de datificación social, donde este es tanto condición necesaria como también resultado del despliegue del sistema RAPPa ampliado vía OTD IA y OTD 4.0. Asimismo, es importante destacar que este estado de situación hoy día desborda los “terrenos” de las plataformas y sus satélites, apuntando también a modelos más amplios como el de la Industria 4.0, los asistentes personales, las *Smart Cities*, las *Smart Houses*, el turismo inteligente, entre muchos otros. En otras palabras, los OTD IA y los OTD 4.0 no solo pueden identificarse como la pieza clave para la mejora y potencia del sistema RAPPa sino también como condición de posibilidad para su ampliación y, con ello, también la potencial extensión del proceso de datificación social a múltiples espacios de la vida previamente poco alcanzados.

5. Reflexiones abiertas

El presente artículo ha iniciado con un recorrido específico por la genealogía del proceso de datificación social, contemplando las instancias donde las cuestiones de lo existente y lo simbólico han sido paulatinamente traducidas en torno a datos, y como ello ha sido utilizado de distintas formas en torno al entendimiento y la gestión de la realidad, de las poblaciones y los individuos. Asimismo, se ha trazado una breve genealogía de los desarrollos de IA para comprender sus distintas etapas, instancias, lógicas de funcionamiento sociotécnicas históricas, “inviernos y veranos”, como también, sus concretizaciones. En este sentido, ha sido posible identificar que la actual etapa de los desarrollos de OTD IA tiene al proceso de datificación social como condición de posibilidad, permeando las principales directrices en torno a su diseño y aplicación. Asimismo, ha sido la optimización y masificación de los OTD IA, junto con otros factores aquí relevados, lo que permite entrever algunos de los motivos de la consolidación y fortalecimiento del proceso de datificación social, junto también con la introducción en la ecuación de otros nuevos OTD, como los vinculados al paradigma 4.0

En este sentido, se ha revisitado aquí cómo las lógicas de perfilamiento, recomendación y gubernamentalidad algorítmica encabalgadas al sistema RAPPa encuentran nuevos horizontes, ampliaciones y mejoras con la introducción de estos nuevos factores, tanto para las conducciones de conductas ya efectuadas actualmente como para nuevos niveles, dimensiones y elementos en la gestión de las poblaciones y de la vida en general.

De todos modos, cabe destacar que no solo las empresas mencionadas previamente son parte del ecosistema de la IA, sino que este también se encuentra conformado por otros actores de peso que, poco a poco, empiezan a reclamar protagonismo. Nos referimos a las empresas proveedoras de infraestructura, los creadores de *hardware*, las organizaciones de estándares, la academia y las universidades, los comités de ética, las ONG y distintos organismos del Estado involucrados. Es posible anticipar aquí que el futuro y la concretización de los desarrollos de OTD IA, como también la profundización, recaudos o interpelaciones respecto de la lógica RAPPa ampliada que estos despliegan resulten de las distintas estrategias de saber-poder como también de las alianzas, rivalidades y consensos que puedan tenderse entre estos actores en conjunto con las principales empresas, especialmente reflejados en inversiones, políticas públicas y regulaciones.

Similar será el caso respecto del 5G y potencialmente el 6G -los cuales se espera que sean la piedra angular para incrementar la masificación de todo el funcionamiento del ecosistema 4.0, incluida la IA (Scasserra *et al.* 2021) y, por tanto, puedan abrir paso a nuevos horizontes de la datificación social- cuyos procesos de diseño e implementación también se encuentran en disputa entre los actores del ecosistema por intentar modular su tecnicidad en función de sus propios intereses (Gendler 2024).

Se espera que lo aquí expuesto haya contribuido a una profundización del conocimiento sobre la datificación social, los desarrollos de IA y los OTD propios del ecosistema 4.0, y sobre su estado de situación; y a intentar acercar el entendimiento respecto de lógicas y modos de operatoria de varios de los sistemas estratégicos de saber-poder vinculados y, cada vez más, cuasi simbióticamente entrelazados. Es posible augurar que el tiempo y las investigaciones futuras permitan dictaminar respecto del acercamiento o el alejamiento a la realidad empírica de varias de las prospectivas aquí anunciadas.



6. Bibliografía

- Ball, Matthew. *The Metaverse: And How It Will Revolutionize Everything*. Nueva York: Liveright, 2022.
- Benanav, Aaron. *La automatización y el futuro del trabajo*. Madrid: Traficantes de Sueños, 2022.
- Blanco, Javier y Rodríguez, Pablo Esteban. "Sobre la fuerza y la actualidad de la teoría simondoneana de la información". *Amar a las máquinas Cultura y técnica en Gilbert Simondon*. Javier Blanco, Diego Parente, Pablo Esteban Rodríguez, y Andrés Vaccari, editores. Buenos Aires: Prometeo, 2015.
- Castells, Manuel. *La era de la información*. México: Siglo XXI Editores, 1999.
- Cafassi, Emilio. "Bits moléculas y mercancías (breves anotaciones sobre los cambios en el submundo de las mercancías digitalizadas)". *La ciudad y sus TICs: tecnologías de información y Comunicación*. en Susana Fínquelevich y Esther Schiavo, compiladoras. Buenos Aires: Universidad Nacional de Quilmes, 1998.
- Celis Bueno, Claudio. "Economía de la atención y visión maquina: hacia una semiótica asinificante de la imagen", *Hipertextos 7/5* (2017): 41-53.
- Crawford, Kate. *Atlas de Inteligencia Artificial. Poder, política y costos planetarios*. Buenos Aires: FCE, 2023.
- Deleuze, Gilles. *Posdata a las sociedades de control. Conversaciones*. Valencia: Pretextos, 1999.
- . *El Poder. Curso sobre Foucault. Tomo II*. Buenos Aires: Cactus, 2014.
- Elias, Norbert. *Conocimiento y poder*. Madrid: La Piqueta, 1994.
- Finn, Edward. *What Algorithms Want: Imagination in the age of computing*. Cambridge: The MIT Press, 2017.
- Fínquelevich, Susana; Odena, Belén; Girolimo, Ulises; Gendler, Martín; Feldman, Patricio. y Cortez Oviedo, Sebastián. *Plataformas digitales en la ciudad: transformación y rediseño del espacio urbano y la vida cotidiana*. Buenos Aires: Café de las Ciudades, 2024.
- Foucault, Michel. *Nacimiento de la biopolítica. Curso en el Collège de France (1978-1979)*. Buenos Aires: FCE, 2007.
- Galliano, Alejandro. *¿Por qué el capitalismo puede soñar y nosotros no? Breve manual de las ideas de izquierda para pensar el futuro*. Buenos Aires: Siglo XXI, 2020.
- Gendler, Martín. "Personalización algorítmica y apropiación social de tecnologías. Desafíos y Problemáticas". *Tecnologías Digitales: Miradas críticas de la apropiación en América Latina*. Ana Rivoir y María Julia Morales, coordinadoras. Buenos Aires: CLACSO, 2019.
- Gendler, Martín. "De la cibernética al metaverso: una genealogía de características, transparencias y opacidades algorítmicas", *Disparidades. Revista De Antropología 78/1* (2023a).
- . "Del Objeto Técnico al Objeto Técnico Digital: apuntes y aportes desde (y a las) elaboraciones de Gilbert Simondon en un marco algorítmico", *Anacronismo e Irrupción 13/25* (2023b).
- . "De la Neutralidad de la Red al 5G. Cambios en las dinámicas de poder del ecosistema de Internet". *InMediaciones de la Comunicación 19/1* (2024), 81-105.
- Gendler, Martín y Andonegui, Fernando. "El COVID-19 y las regulaciones digitales en Argentina", *Controversias y Concurrencias Latinoamericanas 12/22* (2021): 175-202.
- Girolimo, Ulises y Feldman, Patricio. "Incorporação e desenvolvimento de tecnologias de informação em cidades Argentinas: uma análise de caso", *DRd - Desenvolvimento Regional em debate, 10* (2020): 141-170.
- Harari, Yuval Noah. *Homo Deus. Breve historia del mañana*. Buenos Aires: Debate, 2018.
- Kitchin, Rob. *The data revolution: big data, open data, data infrastructures & their consequences*. Londres: Sage, 2014.
- Lago Martínez, Silvia; Gendler, Martín; Méndez, Anahí. "Teoría, debates y nuevas perspectivas sobre la apropiación de tecnologías digitales". *Contribuciones al estudio de procesos de apropiación de tecnologías*. Roxana Cabello y Adrián López, editores. Rada Tilly: Ediciones del Gato Gris, 2017.
- Lessig, Lawrence. *The Code version 2.0*. Cambridge: Basic Books, 2006.
- Li, Li; Fan, Yuxi; Tse, Mike; y Kuo-Yi, Lin. "A review of applications in federated learning". *Computers & Industrial Engineering, 149* (2020) <https://doi.org/10.1016/j.cie.2020.106854>
- Madakam, Somayya; Holmukhe, Rajesh y Revulagadda, Rajeev. "The next generation intelligent automation: Hyperautomation". *JISTEM-J. Inf. Syst. Technol. Manag. 19* (2022).
- Magnani, Esteban. *La Jaula del Confort*. Buenos Aires: Sherpa, 2019.
- Mattelart, Armand. "Pasado y presente de la Sociedad de la Información. Entre el Nuevo Orden Mundial de la Información y la Comunicación y la "Cumbre Mundial sobre la Sociedad de la Información", *TELOS 67* (2006).
- Mayer-Schönberger, Viktor. y Cukier, Kenneth. *Big Data: A Revolution That Will Transform How We Live, Work and Think*. Londres: John Murray, 2013.
- McClelland, James y Rumelhart, David. "An interactive activation model of context effects in letter perception: I. An account of basic findings". *Psychological review 88/5* (1981).

- Mejias, Ulises y Couldry, Nick. “Datificación”, *Revista Latinoamericana de Economía Y Sociedad Digital*, issue especial (2022). <https://doi.org/10.53857/hced6196>
- Molina, Carlos. “La clave de la estrategia de IA de cada gigante digital”. *MultiVersial*. 14 de feb de 2024. <https://www.multiversial.es/p/la-clave-de-la-estrategia-de-ia-de>
- Pandey, Kritika y Patel, Sanskruti. “Deep Learning with Convolutional Neural Networks: from Theory to Practice”. 7th International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli: India, 2023: 1217-1224. <https://ieeexplore.ieee.org/document/10125854>
- Pearl, Judea. *Causality*. Cambridge: Cambridge University Press, 2009.
- Prodnik, Jernej. “La lógica algorítmica del capitalismo digital”, *Hipertextos* 10/18 (2022).
- Rodríguez, Pablo. “¿Qué son las sociedades de control?”. *Revista Sociedad* 27 (2008).
- . *Las palabras en las cosas: saber, poder y subjetivación entre algoritmos y biomoléculas*. Buenos Aires: Cactus, 2019.
- Rossi, Aníbal. “¿Burbujas de filtro? Hacia una fenomenología algorítmica”, *InMediaciones De La Comunicación* 13/1 (2018): 263-281.
- Rouvroy, Antoinette y Berns, Thomas. “Governamentalidade algorítmica E Perspectivas De emancipação: O díspar Como condição De individuação Pela relação?”, *Revista Eco-Pós* 18/2 (2015): 36-56.
- Rullani, Enzo. “El capitalismo cognitivo, ¿un déjà-vu?”. *Capitalismo cognitivo, propiedad intelectual, y creación colectiva*. AA.VV, editores. Madrid: Traficantes de sueños, 2000.
- Sadin, Éric. *La Humanidad Aumentada: la administración digital del mundo*. Buenos Aires: Caja Negra, 2017.
- . *La Silicolonización del mundo: la irresistible expansión del liberalismo digital*. Buenos Aires: Caja Negra, 2018.
- Samela, Gabriela. “La lógica algorítmica en el acceso a los contenidos y a las interacciones personales en Internet”, *XI Jornadas Latinoamericanas de Estudios de Ciencia y Tecnología*, Universidad Tecnológica Federal de Paraná. Curitiba: Brasil, 2016.
- Scasserra, Sofía; De Marco, Celeste; Pereira, Mariano; Jozami, Mora y Mora, Carolina. *Nuevos servicios exportables a partir de la red 5G: ¿Cómo aprovecharlos para reducir la brecha de género?* Buenos Aires: BID, 2021.
- Schwab, Klaus. *La cuarta revolución industrial*. Barcelona: Debate, 2016.
- Sejnowski, Terrence y Rosenberg, Charles. “Parallel networks that learn to pronounce English text”. *Complex systems* 1/1 (1987): 145-168.
- Simondon, Gilbert. *Comunicación e información*. Buenos Aires: Cactus, 2016.
- Srnicek, Nick. *Capitalismo de Plataformas*. Buenos Aires: Caja Negra, 2016.
- Tegmark, Max. *Life 3.0 being human in the age of artificial intelligence*. Vintage Books, 2017.
- Tiqqun. *La hipótesis cibernética*. Buenos Aires: Hekht, 2015.
- Toosi, Amirhosein; Bottino, Andrea; Saboury, Babak; Siegel, Eliot y Rahmim, Arman. “A brief history of AI: how to prevent another winter (a critical review)”. *PET clinics* 16/4 (2021): 449-469.
- Van Dijck, Jose. *La cultura de la conectividad: una historia crítica de las redes sociales*. Buenos Aires: Siglo XXI, 2016.
- Van Dijck, Jose, Poell, Thomas. & Waal, Martijn. *The Platform Society. Public Values in a Connective World*. Oxford: Oxford University Press, 2018.
- Widder, David Gray, West, Sarah y Whittaker, Meredith. “Open (For Business): Big Tech, Concentrated Power, and the Political Economy of Open AI”. *SRRN* (2023). <http://dx.doi.org/10.2139/ssrn.4543807>
- Zuazo, Natalia. *Los dueños de Internet*. Buenos Aires: Planeta, 2018.
- Zukerfeld, Mariano. *Capitalismo y Conocimiento: Materialismo Cognitivo, Propiedad Intelectual y Capitalismo Informativo*. Tesis Doctoral. Buenos Aires: FLACSO, 2010.
- . “Todo lo que usted quiso saber sobre Internet pero nunca se atrevió a googlear”, *Hipertextos* 2/1 (2014): 64-103.
- . “Bits, plataformas y autómatas: las tendencias del trabajo en el capitalismo informacional”, *Revista Latinoamericana de Antropología del Trabajo* 4/7 (2020): 1-50.



Cognición computacional

Computational Cognition

JAVIER BLANCO¹

Universidad Nacional de Córdoba, Argentina

javier.blanco@unc.edu.ar

Fecha de recepción: 29/04/2024

Fecha de aceptación: 27/06/2024

Resumen

En este trabajo sostendremos una interrogación a los sistemas computacionales y a la noción misma de computación desde perspectivas conceptuales, políticas y culturales. La ubicuidad y la insoslayable y creciente importancia que estos sistemas tienen en todos los ámbitos de la vida suele llevar a análisis apresurados y distópicos de sus alcances y límites, considerando algunas consecuencias propias de un estadio particular de su desarrollo como constitutivas o inevitables. Tratando de desenredar una vieja, pero renovada trama de confusiones, tomaremos algunas ideas y descubrimientos fundacionales de la computación como la efectividad, la programabilidad y la abstracción, los programas como datos y el interjuego entre manipulación humana y mecánica de símbolos propuesta por Dijkstra, como claves

¹ PhD. en ciencias de la computación, otorgado por la Universidad de Eindhoven, Países Bajos. Profesor Titular con dedicación exclusiva en la Facultad de Matemática, Astronomía, Física y Computación. Director de la Maestría en Tecnología, Políticas y Culturas, dependiente de la Facultad de Ciencias Sociales, de la Facultad de Filosofía y Humanidades y de la Facultad de Artes. Director Alternativo del Doctorado en Ciencias de la Computación, Facultad de Matemática, Astronomía, Física y Computación. <https://orcid.org/0000-0003-4416-7892>

CÓMO CITAR ESTE ARTÍCULO:

En APA: Blanco, J. (2024). Cognición computacional. *Resonancias*, (17), 143-155. DOI: 10.5354/0719-790X.2024.74507

En MLA: Blanco, J. "Cognición computacional". *Resonancias*, n.º 17, julio de 2024, pp. 143-155. DOI: 10.5354/0719-790X.2024.74507

Palabras clave: computación, inteligencia artificial, metaprogramación, cognición, Coevolución

Keywords: Computation, Artificial, Intelligence, Metaprogramming, Cognition, Co-evolution

para construir una mirada que se sustraiga a los frecuentes prejuicios antropomórficos y que habilite pensar un horizonte de mejores entramados entre humanos y algoritmos. Para ello presentaremos tres niveles de análisis de sistemas computacionales que dan cuenta de propiedades diferentes en cada nivel de manera de poder sopesar mejor no solo la actualidad de la IA, sino también las posibilidades de evolución de las mediaciones computacionales en general.

Abstract

In this work, we interrogate computational systems and the very notion of computing from conceptual, political and cultural perspectives. The ubiquity and the unavoidable growing importance of these systems in all areas of life, usually lead to hasty and dystopian analyzes of their scope and limits. Some consequences derived from this particular stage of development are frequently considered ineluctable. In order to unravel an old but renewed web of confusions, we will consider some founding ideas and discoveries from computer science, such as effectiveness, programability and abstraction, programs as data and the interplay between human and mechanic symbol manipulation proposed by Dijkstra as keys to build a perspective that avoids anthropomorphic prejudices and enables us to think about an horizon with better couplings between humans and algorithms. To do this, we will present three levels of analysis of computational systems that account for different properties at each level in order to assess not only the current state of the art of AI but also the possibilities of evolution of computational mediations in general.

“For anyone wanting a precise analysis of what we refer to today as artificial intelligence, which seems now to have become the horizon of everyday life (and I will return to this question), it is necessary to begin from the following postulate: all noetic intelligence is artificial.” (Stiegler 2018)

1. La inteligencia artificial, no (siempre) inteligente, no (solo) artificial

El sintagma “inteligencia artificial”, antes que una definición, tiene una historia no tan larga, pero sí muy intensa. Consecuentemente, no aspiramos aquí a formular una definición precisa, sino que abordaremos de manera crítica algunas de sus caracterizaciones, sobre todo, en los campos de las humanidades y de las ciencias sociales, en busca de formular indicaciones conceptuales y terminológicas que faciliten la comprensión del alcance de estas nuevas entidades en la vida social y cultural. Una de las dimensiones que consideraremos es entenderla como una nueva forma de desarrollo de software, con diferencias importantes con la mayoría de los programas que se desarrollaron en los últimos 70 años. Otra tiene que ver con las maneras de comprender los vínculos diversos entre

humanos y sistemas computacionales y sus posibilidades actuales de evolución. En este sentido, consideramos que las metáforas antropomórficas son confusas e inmotivadas y que obturan un análisis certero de estas dos dimensiones. Tratar de establecer paralelismos entre la inteligencia artificial y la humana (que como nos recuerda Stiegler siempre fue artificial, al menos en su versión noética) aparece como un camino poco promisorio. Por un lado, no es tan claro cómo funciona la inteligencia humana, por lo que parece poco propicio considerarla como parámetro con el cual contrastar los desarrollos de software de IA. En cualquier caso, es casi una obviedad que los funcionamientos son disímiles, al punto que los fundadores de la IA han sido muy cautos, hablando, por ejemplo, de programas que realizan tareas para las que un humano necesitaría inteligencia para poder realizar.

En el interesante libro *Atlas of AI*, Kate Crawford (2021) instala una celebrada frase: “la inteligencia artificial no es artificial ni es inteligente”. Crawford argumenta que los sistemas de IA no son inteligentes ya que carecen de autonomía, racionalidad y discernimiento sin que medie un entrenamiento computacional extensivo e intensivo con grandes bases de datos o reglas y recompensas predefinidas. Todos estos argumentos son válidos, sobre todo, si nos referimos a este momento de la IA, donde los programas producidos a través de alguna variante de *machine learning* son reconocedores estadísticos de patrones, más asociados a una extensión de la percepción (ver más abajo la idea del nooscopio) que a lo que en los humanos es considerado inteligente.

Todos estos argumentos son correctos y, en algún sentido, dan cuenta de una obviedad. Es también notorio cómo esta frase se repite en los debates sobre IA como si fuera una revelación y refleja cierto alivio. Aparentemente, reaparece aquí un recurrente prejuicio humanista que se asienta en suponer que la inteligencia humana es, en cierto sentido, paradigmática y que, por lo tanto, el grado de inteligencia de cualquier entidad se mediría por su cercanía o parecido con ella. No hay buenos argumentos para sustentar esto, y parece reflejar una mirada autocomplaciente de los humanos, poco crítica con los obvios límites y dificultades del pensamiento humano.

Para negar la artificialidad de la IA, Crawford afirma que esta es corporizada (*embodied*) y material. Se argumenta también (al igual que Pasquinelli (2023)) que gran parte del esfuerzo para producir los programas usando machine learning proviene de trabajo humano de clasificación y etiquetado de datos, en general mal remunerado, indispensable para llevar a cabo el proceso de entrenamiento. También eso es cierto y frecuente, aunque no es necesario, y hay muchos sistemas que se desarrollan con datos sintéticos. De cualquier manera, tanto el trabajo humano como la explotación de recursos materiales están, desde siempre, asociados a la producción de lo artificial.

Por diferentes motivos, William Rapaport (2023) también critica, por confuso, el sintagma “inteligencia artificial”. Por un lado, plantea que la idea de inteligencia no puede ser definida con precisión y que en muchos trabajos fundacionales (por ejemplo, de Minsky y Simon) se evitaba cualquier definición, asumiendo que el proyecto era hacer máquinas (o programas) que pudieran hacer cosas que requerirían inteligencia si fueran hechas por humanos. Estas tareas, argumenta Rapaport, quedan mejor enmarcadas bajo el término de ‘cognición’ en lugar de ‘inteligencia’, ya que incluyen estados mentales y procesos como creencias, conciencia, emociones, lenguaje, memoria, percepción, planificación y resolución de problemas.

Por otro lado, la idea de “artificial” tiene también para él varios problemas. El que enfatiza es que la palabra remite a algo poco legítimo, sin realidad propia, a un simulacro de, en este caso, la inteligencia real. En cualquier caso, también puede considerarse que la inteligencia humana es esencialmente artificial, constituida además gracias a la multiplicidad de andamiajes externos que la hacen posible, siendo el lenguaje uno de los fundamentales. Esta artificialidad constitutiva de la inteligencia humana habilitaría, siguiendo a Stiegler, también la producción de estupidez, a partir de la industrialización de órganos exosomáticos constitutivos de los andamiajes cognitivos. Estamos hoy en medio de ese proceso cuya futura evolución es aún incierta y en disputa.

Ahora bien, cuando nos referimos a los sistemas computacionales que hoy constituyen la llamada IA, Rapaport prefiere hablar de “inteligencia (o cognición) sintética” o, más específicamente indicando la naturaleza de esos procesos de síntesis, “cognición computacional” (Rapaport 1354). La mayor generalidad de la palabra “cognición” y la mayor especificidad de “computacional” parecen delimitar con mucha más precisión esta noción elusiva. En lo que sigue, seguiremos usando el acrónimo IA, pero lo entenderemos como refiriendo a la cognición computacional. También esta especificidad permitiría entender mejor la crítica de Crawford esgrimida contra la “artificialidad”; tiene más sentido si se quiere indicar que al IA no es solo computacional.

2. Multiescalaridad de las tecnologías computacionales

“From a bit to a few hundred megabytes, from a microsecond to a half an hour of computing confronts us with completely baffling ratio of 10^9 ! The programmer is in the unique position that his is the only discipline and profession in which such a gigantic ratio, which totally baffles our imagination, has to be bridged by a single technology. He has to be able to think in terms of conceptual hierarchies that are much deeper than a single mind ever needed to face before. Compared to that number of semantic levels, the average mathematical theory is almost flat. By evoking the need for deep conceptual hierarchies, the automatic computer confronts us with a radically new intellectual challenge that has no precedent in our history.” (Dijkstra 1988)

Entender las características de los sistemas de lo que se llama hoy IA requiere distinguir, al menos, tres tipos de registros con grados de generalidad decreciente que suelen confundirse en algunos análisis. En esta sección abordaremos el registro más general desde el que daremos cuenta, sucintamente, de toda mediación computacional.

En el registro más general nos encontramos con las propiedades compartidas por todos los sistemas computacionales, que son los que han transformado drásticamente, durante ya casi un siglo, la escena intelectual, presentando desafíos inéditos, nuevas formas de conocimiento, y han hecho posible formas de acción diferida de complejidad creciente. En los últimos años, los indiscutibles logros de la tecnología de machine learning han revitalizado el rol de las mediaciones computacionales en la estructuración del mundo. Sin embargo, en la mayoría de los trabajos y proyectos actuales suele asumirse que es el desarrollo de la IA la causa principal de esa renovada relevancia, y que los programas generados por machine learning representan el apogeo de las tecnologías computacionales. Contrario a esto, consideramos que la profunda transformación en curso de la condición humana está impulsada por la novedad radical asociada a la capacidad de programar comportamientos de manera precisa y efectiva, y por la -literalmente- infinita variedad y versatilidad del mundo algorítmico, variedad que no necesariamente tienen los programas producidos por machine learning.

Como bien indica Benjamin Bratton (2015), el mundo actual está constituido por una pila de capas digitales que habilitan una multiplicidad de operaciones. Extrapolando ideas de la cibernética, consideramos estas capas como nuevas dimensiones de la materialidad del mundo, que permiten múltiples formas de percepción, de cognición y de acción. La llamada “digitalización” no es una reducción o simplificación del inabarcable flujo de la realidad, sino que constituye una serie de capas superpuestas sobre este *continuum*, y cuya razón de exis-

tencia es habilitar nuevas formas de operación computacional, que, desafiando ciertas intuiciones, multiplican las capacidades y posibilidades. Es una obviedad que las mediaciones computacionales están transformando drásticamente la existencia, aumentando su complejidad, viendo a diario la aparición de nuevas herramientas o procesos de transformación en todas las esferas del mundo de la vida. Pese a esto, muchos autores siguen insistiendo en que lo digital implica una reducción de la realidad o de las personas a “meros datos”, que pierden su singularidad convirtiéndose en objetos de cálculo. Como veremos más adelante, es mucho más fértil conceptualmente asumir esta “datificación” y las consecuentes transformaciones de lo real como una enorme oportunidad evolutiva.

Es importante poner en foco este registro general de análisis de las mediaciones computacionales para evitar una frecuente falacia de composición que consiste en tomar propiedades específicas de algún subconjunto determinado; en particular a los sistemas producidos por machine learning, como aplicables a cualquier sistema computacional. Esta falacia tiene consecuencias conceptuales bastantes insidiosas, ya que este tipo de tecnología produce programas muy diferentes a los usuales, al punto que algunos autores no los reconocen *a priori* como tales (Janlert 2008), que no tienen nada que ver con la mayoría de los sistemas existentes o posibles. Para sumar a la confusión, la palabra “algoritmo” ha tomado cierto uso social asociado a los programas producidos por machine learning que, en esencia, son clasificadores (o generadores) estadísticos, muy diferentes de lo que se considera paradigmáticamente un algoritmo. Vamos a remarcar esto con cierta insistencia en aras de escapar a lo que termina siendo una mirada determinista o incluso tecnofóbica frente a la acelerada evolución de las mediaciones algorítmicas.

Otra característica fundamental de los sistemas computacionales se deriva de un hecho conocido desde los trabajos fundacionales de Turing y que fue consistentemente desarrollado y explotado por la teoría de la computación: los programas son también datos. Esta propiedad tiene muchísimas consecuencias, siendo fundamental también en el desarrollo de la práctica de la programación. Existen múltiples tipos de metaprogramas (es decir, programas que toman a otros programas como datos) indispensables para el desarrollo de las tecnologías computacionales. Compiladores, intérpretes, máquinas virtuales, sistemas operativos, *shells* de comandos, macroexpansiones, super-compiladores, son algunos de los ejemplos de metaprogramas de uso cotidiano. También los clasificadores estadísticos producidos por la tecnología de machine learning son el *output* de un meta-programa de entrenamiento. La meta-programación es la regla, no la excepción.

Esto permite comprender a la tecnología computacional como una metatecnología, es decir, como un espacio de producción de diversas tecnologías, cada una

con sus especificidades. De hecho, casi todas las formas técnicas previamente existentes han sido transformadas por la inclusión de subsistemas computacionales como parte de los objetos o ambientes técnicos. Un auto, un televisor, una lavadora o una afeitadora eléctrica, cada vez más, funcionan gracias a crecientes cantidades de software integrado en ellos.

Otra de las características salientes de la tecnología computacional es la marcada por Dijkstra en el epígrafe: la cantidad, profundidad y complejidad de los niveles de abstracción involucrados. Las relaciones intranivel e internivel son múltiples y generalmente se realizan también como programas. Desde esta perspectiva puede apreciarse otra de las características novedosas de los sistemas computacionales: la capacidad de vincular diferentes órdenes de magnitud.

3. ¿Qué podemos aprender de las máquinas que aprenden?

“The Fathers of the field had been pretty confusing: John von Neumann speculated about computers and the human brain in analogies sufficiently wild to be worthy of a medieval thinker and Alan M. Turing thought about criteria to settle the question of whether Machines Can Think, a question of which we now know that it is about as relevant as the question of whether Submarines Can Swim.” (Dijkstra 1984)

En un segundo término, nos encontramos con los sistemas de machine learning, una tecnología que adquirió un desarrollo acelerado, sobre todo, en las últimas dos décadas. Esta tecnología involucra una nueva forma de programación, basada en el reconocimiento automático de patrones en los más diversos espacios de datos, generando clasificadores estadísticos capaces de comprimir y operativizar la información —considerada en un sentido laxo— contenida en esas grandes masas de datos.

En la actualidad, las tecnologías de machine learning plantean problemas específicos que hace falta enfocar. En un trabajo pionero (Joler y Pasquinelli 2021), los autores muestran a este tipo de tecnología como un nooscopio; es decir, un instrumento de magnificación del conocimiento, análoga a los microscopios o telescopios. Como aquellos, el nooscopio permite ver lo invisible a costa de producir ciertas distorsiones. En general, los sistemas producidos por machine learning permiten establecer correlaciones que serían imposibles de encontrar por métodos “humanos”. Los programas que emergen del proceso de entrenamiento son clasificadores estadísticos capaces de determinar con una alta probabilidad

si algún dato nuevo satisface el mismo criterio implícito que los elementos del conjunto de datos de entrenamiento.

Además de los clasificadores, y dependiendo de estos, en el presente hay una proliferación de sistemas generativos capaces de crear diferentes expresiones cognitivas y culturales. Entre estos programas son ejemplares, por su multiplicidad de uso, los grandes modelos de lenguaje, los cuales son capaces de producir textos de gran complejidad sobre un tema específico, entablar conversaciones amigables o incluso construir programas que resuelvan algún problema propuesto. Cabe preguntarse si estamos ante un verdadero salto cualitativo en las capacidades de los programas o, por ejemplo, si esta capacidad de programar permitirá a estos sistemas el automejoramiento -anhelado o temido- de las IA.

Hay una opacidad intrínseca en los clasificadores estadísticos. El código no es comprensible en su totalidad y la forma de representación de los datos queda diseminada en los pesos de una red, no es reconocible ni siquiera para quienes programaron el sistema. Los programas desarrollados por técnicas de programación usuales son artefactos con propiedades lógicas deducibles *a priori*; por lo tanto, analizables y demostrables matemáticamente. En cambio, los clasificadores estadísticos, sobre todo los producidos por *deep learning*, no tienen propiedades lógicas reconocibles y, en general, solo se enuncian propiedades empíricas y con credenciales epistémicas más dudosas. Se conoce el proceso de entrenamiento, y pueden calcularse ciertos parámetros de adecuación o robustez del modelo producido, aunque son usualmente *ad hoc*, por lo tanto, no es siempre claro el significado del valor de dichos parámetros.

Los sistemas de IA generativa aprenden a producir expresiones que son reconocidas como válidas por el clasificador asociado. La producción es aleatoria. Por ejemplo, en los sistemas de producción de lenguaje, las palabras se van produciendo secuencialmente usando funciones probabilísticas ponderadas por la subred semántica asociada al prompt introducido. No son grandes productores de novedades y no tienen ninguna aspiración de verdad. Producen textos probables basados en un entrenamiento con billones de entradas lingüísticas. Son capaces de cometer errores garrafales, pero, en general, darán respuestas estadísticamente aceptables. No producirán nunca grandes novedades y no hay ninguna razón para confiar en su adecuación a los hechos. A veces lo harán, otras no, y no hay ni habrá ningún criterio intrínseco para distinguirlas. Esto es válido también para la extraña habilidad que tienen de “programar”. Los programas escritos en lenguajes de programación usuales son también textos, asociados semánticamente a su contexto de producción. Si se formula con claridad el tipo de problema a resolver, los generadores pueden producir programas asociados estadísticamente a ese problema. Es difícil o, literalmente, improbable que se generen programas que incluyan alguna novedad interesante, por lo que parece

igualmente improbable que vayan a producir un automejoramiento, más aún dado que el programa generado mismo no posee un código comprensible. A riesgo de remarcar lo obvio, esta improbabilidad solo se aplica a los sistemas producidos por la actual tecnología de machine learning; esto no invalida la posibilidad de otro tipo de programa que se automejore, como ya presuponía Turing en 1947 (Turing 1947).

Es importante conocer todas estas limitaciones para saber qué esperar y qué no de estos sistemas y de sus posibles evoluciones. Seguramente seguirán mejorando en base a entrenamientos más exhaustivos y correcciones de algunos errores comunes. De hecho, los grandes procesadores de lenguaje más populares pasaron por una cantidad de arreglos y ajustes realizadas por miles de trabajadores para reducir comportamientos lingüísticos ofensivos o meramente para mejorar las interacciones. De cualquier manera, a menos que haya algún nuevo salto conceptual, las limitaciones enunciadas seguirán estando.

Si bien no hay ninguna pretensión de verdad en los generadores de lenguaje, son sistemas de mucha utilidad. Y aunque no puedan revelar de manera confiable conocimiento del mundo, sí constituyen un acceso privilegiado a lo que en la teoría política se conoce como “ideología”; es decir, las creencias, ideas, significados que estructuran la vida social. No sabemos si los textos producidos son verdaderos, pero sí sabemos que reflejan lo que se ha dicho o escrito sobre un tema dado (Weatherby 2023).

4. Ensamblajes computacionales

Para poder analizar en contexto las formas de existencia de los objetos digitales o computacionales, son necesarias herramientas conceptuales que den cuenta de su pluralidad y sus ambivalencias. John Johnston introduce el concepto de ensamblaje computacional para referir a la conjunción entre un mecanismo computacional dado y sus correlatos discursivos. Es decir, el conjunto de los programas, aplicaciones y sistemas computacionales y los marcos conceptuales, nociones y enunciados que conjuntamente construyen sentidos y condiciones de inteligibilidad de dichos sistemas

Este registro que refiere a lo que podríamos llamar la computación “*in the wild*”, permite analizar las características que provienen tanto del tipo de programación como del contexto de desarrollo y de aplicación de los sistemas en cuestión. En el presente, abundan ensamblajes computacionales basados en sistemas de machine learning, que son, en general, quienes detentan (o padecen) las

características que se discuten en muchos trabajos filosóficos, políticos o culturales. Sesgos, reproducción del pasado, predicción o alteración de las preferencias o las conductas, apropiación de la atención, desinformación y *fake news* son algunos de los males que se pretenden conjurar con la urgencia de los tiempos políticos. La buena noticia es que estas características no son inherentes a los sistemas algorítmicos en general y, en cierta medida, ni siquiera a los sistemas de machine learning, sino a los contingentes ensamblajes computacionales. Como ejemplo de esto, la noción de *gubernamentalidad algorítmica* planteada por Rouvroy y Berns (2013) cae en este registro, pero el uso de la palabra “algorítmica” sugiere una generalidad que no tiene. Para ser precisos habría que hablar de “gubernamentalidad basada en machine learning en el capitaloceno”, pero claramente el sintagma propuesto es mucho más atractivo y memorable.

Volviendo al nooscopio, el análisis realizado por los autores, y continuado en (Pasquinelli 2023), considera el contexto de producción de diferentes sistemas de machine learning mostrando que, en esos casos, los programas se construyen esencialmente a través de trabajo humano. Desde esta perspectiva buscan, también, dilucidar las profundas transformaciones en el mundo del trabajo.

Quizá una de las diferencias actuales esenciales entre cognición computacional y cognición humana está dada por la “falta de juicio” del lado computacional. Brian Cantwell Smith (2019) plantea la idea de “juicio” (*judgement*) como capacidad que, en el presente, sería distintivamente humana y que da cuenta de un compromiso existencial con el mundo. Más allá de la pregunta acerca de qué sería necesario para que sistemas de cognición computacional adquieran capacidad de juicio, cabe asumir una nueva paradoja reflexiva y preguntarse cómo se establece hoy un compromiso existencial con un mundo poblado de clasificadores estadísticos, cómo estos clasificadores integrados ya a la cotidianidad, a la toma de decisiones, a la cultura afectan y transforman la capacidad ¿humana? de juicio y sus alcances.

5. El interjuego: Machine learning como método de (meta)programación

“Hence, computing science is --and will always be-- concerned with the interplay between mechanized and human symbol manipulation, usually referred to as “computing” and “programming” respectively.” (Dijkstra 1988)

Programar es un proceso sofisticado de manipulación humana en determinados sistemas formales, basados en propiedades lógicas de los lenguajes de

programación y de las especificaciones del problema a resolver. Muchas herramientas conceptuales y lógicas están involucradas en este proceso. Se espera que quien programa pueda proveer no solo el programa funcionando, sino buenos argumentos de por qué resuelve el problema propuesto.

Los procesos de entrenamiento por machine learning no son tan generosos. Si bien, al final del proceso se dispone de un clasificador o un generador, la justificación de corrección o, al menos, de adecuación de la solución rara vez puede producirse. A lo sumo se puede dar cuenta del proceso de producción. Es por ello que estos clasificadores no poseen las certificaciones de calidad que sí tienen otro tipo de programas. Janlert (2008) duda de que pueda llamarse programas a estos sistemas, ya que estos deberían no solo establecer un sistema causal del proceso de ejecución y una descripción del proceso mismo, sino también una serie de justificaciones, de explicaciones racionales que permitan confiar en el sistema producido. Sin duda, estos sistemas son capaces de causar los procesos requeridos, pero ya su descripción es poco comprensible y la justificación suele ser inexistente (e imposible de construir). Estamos frente a lo que Janlert llama “*dark programming*”. Esta manera de programar da lugar a sistemas opacos en los que la forma de representación suele no ser explícita.

Desde sus inicios, la computación se basó en un mito fundacional: la posibilidad de liberar la mente de las tareas repetitivas. Paradójicamente, el vínculo que suele establecerse hoy con los sistemas de IA suele fomentar la repetición de conductas; disminuir la capacidad de atención necesaria para desarrollar nuevas ideas; aumentar, en resumen, la entropía cognitiva. Creemos que ese estado de cosas no es necesario. Las múltiples formas de programación, los interjuegos presentes y posibles entre programación y computación, abren un paisaje de oportunidades evolutivas que pueden renovar los acoplamientos cognitivos. Como intentó Hilbert a inicios del siglo pasado con la matemática, también podemos reformular las bases de la programación de sistemas de IA hoy, sin que nadie nos expulse del paraíso que Turing ha creado para nosotros.

6. La coevolución

“In the long run I expect computing science to transcend its parent disciplines, mathematics and logic, by effectively realizing a significant part of Leibniz’s Dream of providing symbolic calculation as an alternative to human reasoning. (Please note the difference between “mimicking” and “providing an alternative to”: alternatives are allowed to be better.)” (Dijkstra 1988)

Periódicamente vuelve a enunciarse la cuestión del “peligro existencial” planteado por la evolución de la IA. El riesgo de extinción de la humanidad estaría dado cuando la IA “supere” la inteligencia humana, alcance lo que se denomina “inteligencia artificial general” o super-inteligencia que sería incontrolable para humanos y que tomaría el control del mundo. Si bien eso está lejos de los actuales desarrollos, otro factor de preocupación es que las mismas IA sean capaces de automejorarse, adelantando el momento de la “singularidad”.

No abordaremos aquí esta pregunta, en particular, porque parece estar basada en suposiciones conceptuales erróneas, más allá de lo sugestivas que puedan resultar para la imaginación. Las diferentes formas de cognición computacional tienen características propias que no necesariamente coinciden con las formas cognitivas usuales en humanos. Y por supuesto que los sistemas de IA actuales no incorporan deseos ni emociones humanas, y no hay ninguna razón para que los tengan.

La mirada antropomórfica sobre la IA -y los usos lingüísticos asociados- es de alguna manera comprensible, pero oculta más de lo que muestra. Una de las maneras más insidiosas de este sesgo conceptual se manifiesta cuando se critica la IA para repetir que los sistemas computacionales no poseen real inteligencia en comparación con la humana, volviendo a situarse, anacrónicamente, en cierto dualismo cartesiano. Por el contrario, como afirma Dijkstra, la posibilidad de construir programas que permitan formas cada vez más sofisticadas de procesamiento simbólico es un camino para mejorar las capacidades de razonamiento humano o, lo que desde nuestra perspectiva sería lo mismo, de razonamiento de los acoplamientos cognitivos entre mentes y programas. Y no hace falta para esto tener sistemas computacionales autónomos (tampoco los humanos lo son), sino mejores integraciones entre dispositivos cognitivos, humanos o no. La autonomía siempre es relativa y gradual, y ese grado es mayor, de acuerdo con Simondon, cuando incrementa las posibilidades de individuación.

Parafraseando a Hansen, vemos como la distribución de la cognición, la percepción y la acción en el mundo se encuentra en un proceso de rápida evolución. La IA actual, en tanto reconocimiento estadístico de patrones, es una parte

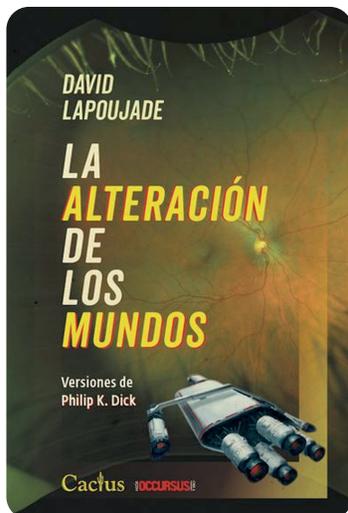
importante de este proceso de redistribución, pero sus límites son ineluctables. Dos caminos posibles y compatibles entre sí se siguen de acá. Uno consiste en mejorar los ensamblajes computacionales basados en este momento tecnológico. Es posible (y necesario) constituir ensamblajes virtuosos que combinen tecnologías libres, vínculos cognitivos más transparentes y valores políticos progresivos. El otro camino, en la línea de la coevolución propuesta por Ashford Lee (2020), es más complejo, pero más transformador, y tiene que ver con encontrar nuevas bases conceptuales para el desarrollo de mediaciones computacionales (Marcus y Davis). Sistemas que superen la opacidad de los clasificadores estadísticos, que puedan acceder a su propio código y que habiliten ensamblajes donde, tanto humanos como ambientes tecnológicos se integren en procesos mutuamente reflexivos, entendiéndose por esto la posibilidad de transformar el ambiente mismo “desde dentro”. Quizá haya que reconstruir muchos sistemas actuales basados en una tecnología más transparente y confiable, aunque esto implique cierta ralentización del desarrollo. Es un camino difícil, pero, como decía Rilke, que algo sea difícil es una razón más para elegirlo.



7. Bibliografía

- Ashford Lee, Edward. *The Co-evolution. The Entwined Futures of Human and Machines*. The MIT Press, 2020.
- Bratton, Benjamin. *The Stack*. MIT Press, 2015.
- Crawford, Kate. *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press, 2021.
- Dijkstra, Edsger. “The threats to computing science.” EWD 898 Texas: Universidad de Austin, 1982.
- . “On the cruelty of really teaching computing science.” EWD 1036 Texas: Universidad de Austin, 1988.
- Hansen, Mark. “Engineering Pre-individual Potentiality: Technics, Transindividuation, and 21st-Century Media”. *Sub-Stance*, vol. 41, no. 3, 2012, pp. 32-59.
- Janlert, Lars-Erik. “Dark programming and the case for the rationality of programs”. *Journal of Applied Logic*, vol. 6, 2008, pp. 545-552.
- Joler, Vladan y Matteo Pasquinelli. “El Nooscopio de manifiesto”. *La Fuga*, no. 25, 2021. Disponible en <https://lafuga.cl/el-nooscopio-de-manifiesto/1053>.
- Marcus, Gary y Davis, Ernest. *Rebooting AI: Building Artificial Intelligence we can trust*. Pantheon Books, 2019.
- Pasquinelli, Matteo. *The Eye of the Master: A Social History of Artificial Intelligence*. Verso, 2023.
- Rapaport, William. *Philosophy of Computer Science: An Introduction to the Issues and the Literature*. John Wiley and Sons, Inc. 2023.
- Smith, Brian Cantwell. *The Promise of Artificial Intelligence. Reckoning and Judgment*. MIT Press, 2019.
- Stiegler, Bernard. “Artificial Stupidity and Artificial Intelligence in the Anthropocene”. Conferencia del 23 de noviembre de 2018, Institute of Ereignis, Shanghai.
- Turing, A.: “The Automatic Computing Engine”. Conferencia dictada en la London Mathematical Society, 20 February 1947.
- Weatherby, Leif. “ChatGPT Is an Ideology Machine”. *Jacobin*. 17 de abril de 2023. Science and Technology / Theory.

BR ESEÑAS



David Lapoujade. *La alteración de los mundos, Versiones de Philip K. Dick*
Ciudad Autónoma de Buenos Aires, Argentina:
Cactus, 2021

NICOLÁS ROJAS-CORTÉS¹

Universidad de Chile, Santiago, Chile
nicolas.rojas.c@ug.uchile.cl

«Jamás tuve una opinión muy elevada de lo que nos hemos habituado a llamar la “realidad”. La realidad, para mí, no es tanto algo que se percibe sino algo que se hace. *Hay que crearla más rápido de lo que ella nos crea*».²

Philip K. Dick

Alguna vez Friedrich Nietzsche dijo que el dedicarse a la filosofía implicaba enseñorearse del mundo o, en otras palabras, imponer el ser sobre el devenir. Pero ¿cómo crear mundos desde una existencia delirante? Cuando hablamos de Dick no hablamos de alguien del que podría predicarse con facilidad la “normalidad”³. De hecho, hablamos de alguien que enfrentaba a la religión con sus propias visiones, y he ahí donde David Lapoujade detecta el punto de partida para hablar

CÓMO CITAR ESTE ARTÍCULO:

En APA: Rojas-Cortés, N. A. (2024). David Lapoujade. *La alteración de los mundos, Versiones de Philip K. Dick*. Ciudad Autónoma de Buenos Aires, Argentina: Cactus, 2021. *Resonancias*, (17), 159-165. DOI: 10.5354/0719-790X.2024.73338

En MLA: Rojas-Cortés, N. A. “David Lapoujade. La alteración de los mundos, Versiones de Philip K. Dick. Ciudad Autónoma de Buenos Aires, Argentina: Cactus, 2021.” *Resonancias*, n.º 17, julio de 2024, pp. 159-165. DOI: 10.5354/0719-790X.2024.73338

¹ Agradecimientos: Agencia Nacional de Investigación y Desarrollo / Subdirección de Capital Humano / Beca de Doctorado Nacional 21210804. <https://orcid.org/0000-0003-4969-6830>.

² Philip K. Dick en Lapoujade 2022 147. Las cursivas son de Lapoujade, y también sirven a nuestro interés.

³ Sea lo que sea que quiera decir esto en un contexto psicológico clínico.

del ciudadano de Disneyland como un notable exponente de la ciencia ficción (CF). Con tal puntapié inicial, la *Introducción: el delirio*, nos invita a considerar a Philip K. Dick (PKD) como un peligroso pensador que es capaz de hacer vacilar la noción misma de realidad (2021 17).

Así, con el delirio deslizándose por todas partes como un elemento fundamental para estructurar el pensamiento de Dick, *La alteración de los mundos* se divide en doce capítulos⁴ cuya organización interna presupone los logros de cada capítulo anterior. Podría sonar una nimiedad destacar esto, ya que se supone que los libros deben leerse por completo, pero la estructura que Lapoujade le otorga a su obra evidencia que, incluso de un *corpus* tan poco sistemático⁵ como el de PKD, puede realizarse una lectura fructífera cuando los textos son enfrentados y analizados con las categorías correctas.

El primer capítulo del ensayo, *Los mundos*, disputa⁶ la distinción entre mundo “objetivo” y “subjetivo” (2022 25). En la narrativa de Dick no parece haber un criterio mínimo con el que esa distinción tenga sentido. Sin embargo, tampoco se quiere decir que los mundos creados por él sean solamente mundos imaginados dentro del psiquismo del delirante. Más bien, los mundos pueden colapsar como decíamos antes, porque ellos entran en guerra. Eso lo muestra la novela *Eye in the Sky* (1957) o *Flow My Tears, the Policeman Said* (1974) en los que la realidad común es asediada violentamente por las alucinaciones de los personajes.

Sin embargo, la razón de lo anterior no se agota en afirmar que “simplemente son obras de CF”. Lapoujade se atreve a afirmar en *La causalidad* (2022 33-40) que tal categoría se desvanece en el “idealismo dickeano”, ya que

En última instancia, para conocer las leyes a las cuales obedece un mundo en Dick, no hay que buscar establecer relaciones constantes entre los fenómenos que lo componen, sino más bien sondear en las profundidades del psiquismo que controla sus apariencias. En tal caso, durante una misión interestelar, se necesitará más un psicoanalista que un físico (2022 35).

⁴ Indicamos sus nombres en cursivas para no repetir la fórmula de “capítulo y nombre del capítulo”.

⁵ Porque sería extraño comprender a Dick, de buenas a primeras, como un filósofo al estilo de un moderno. Ahora bien, con la exposición de Lapoujade se podría ver en el autor de *Ubik* la actitud de un verdadero filósofo en tanto que él se posiciona como un creador de mundos o, más bien, como un *bricoleur*, alguien cuyo poder está en sus manos y es capaz de ensamblar fragmentos de mundos (2022, 140-143). Más allá del mundo dado y posible de interpretar encontramos la acción de Dick, que construye mundos sin la seriedad de una necesidad dada y cuyo colapso muchas veces depende de la irresponsabilidad desobediente de sus personajes. Un delirio psicótico o un exceso de drogas bien podrían ser la razón de que el mundo de alguien se creara, pero el mundo de otro colapsara mientras ese efecto está activo. Y, por supuesto, no hablamos de una irresponsabilidad moral, sino existencial.

⁶ Y esto será algo constante a lo largo de las referencias a las cuestiones que se presentan a través de la historia de la filosofía.

Aquí la causalidad no tiene una potencia necesaria en tanto que los personajes principales no encuentran ningún fundamento último. He ahí la razón para que en *La cosa pensante* (2022 41-51) el autor separe a PKD de la certeza cartesiana: “Si el mundo desaparece, eso quiere decir que el yo también ha perdido toda sustancialidad, que está disuelto, dando prueba de la estrecha correlación entre yo y el mundo” (2022 42). Mas, si el mundo está en estrecha relación con el sujeto, entonces todos nuestros trastornos psíquicos son igual de capaces de modificar la existencia. El fondo pulsional colectivo es capaz de ir más allá de la certeza y el delirio posibilita constantes curas⁷ respecto a una sociedad capitalista que produce paranoicos. Y si dudamos de la realidad como producto de nosotros, ¿por qué no dudar de nosotros mismos también?

Es esa pregunta la que le permite a Lapoujade advertir que la CF de PKD no se agota en la creación de mundos científicos, sino que también es una narrativa de *Lo fantástico* (pp. 53-63), de una confrontación permanente con lo irracional que incluso puede perturbar la distinción filosófica entre lo real y lo irreal. Tal es la razón de que Jason Taverner⁸ pueda ser percibido en un mundo donde él es totalmente inexistente y aun así ni siquiera un heideggeriano *in-der-welt-sein* podría funcionar como un criterio primordial para distinguir entre mundos. No hay espacios posibles en los que la duda no afecte la capacidad de juzgar respecto de algo. *Ubik* (1969) refleja con excelencia la incerteza irracional de la imposibilidad del juicio certero: ¿quién está realmente vivo?

Pensar la condición de estar vivos o muertos recuerda a la clásica cuestión filosófica sobre el ser o el devenir. Se supone que la entidad que es, nosotros si se quiere, está sometida al devenir. Esa sería la razón de que cada día podamos morir. No obstante, en *Ubik* se presenta una “Regresión acelerada y envejecimiento precipitado” que “destruyen la sustancia misma del mundo” (2022 66)⁹. Lo interesante es que *Entropía y regresión* (65-74) presenta un gesto bastante alejado de la CF convencional: el avance tecnocientífico y una existencia más allá de la humana no asegura que el progreso nos ofrezca un desenlace feliz. Cuando hablamos de imponer el ser sobre el devenir también hablamos de un solipsismo delirante, advirtiendo que en este

Había un peligro relativo al mundo del juicio, el de vivir solo *en un único mundo*, impedir toda comunicación entre los mundos, lo que es para Dick, la negación de todo pluriverso. Pero este peligro es inseparable del otro: *el que solo vive en un mundo*

⁷ En un sentido freudiano.

⁸ El personaje principal de *Flow My Tears, the Policeman Said*.

⁹ La capacidad de Dick por hacer desmoronar los mundos se ilustra con cosas que retroceden a formas tecnológicas anteriores pero lo que progresa solamente acelera el dejar de existir.

quiere la destrucción de todos los mundos, incluido el suyo. Tal vez es el sentido último de la fórmula “no hay alternativa” (TINA), cuando se trata de destruir todos los mundos en provecho de uno solo (2022 71)

El citar a Margaret Thatcher muestra un gesto clave en la lectura que Lapoujade realiza de Dick. No hay que olvidar que el ciudadano de Disneylandia era un teórico profundamente crítico con el capitalismo, y muchas veces sus novelas muestran mundos en donde, aunque no haya alternativa, todavía es posible enfrentarse al realismo capitalista al estilo del filósofo Mark Fisher. Existimos, por lo tanto, en una constante lucha por el poder, y como bien sabemos desde Foucault, el poder se ejerce. Esa es la razón por la que en *Los poseedores de mundos* (2022 75-85) se establece el siguiente “teorema dickiano: *todo mundo pertenece a un psiquismo*, y su variante: *todo mundo pertenece a varios psiquismos, los cuales componen entonces un mundo colectivo*” cuyo corolario sería: “*un mundo pertenece a quien produce o controla sus apariencias*” (2022 75). Habitamos en un mundo de apariencias, pero ya no en una apariencia al estilo de Platón, no es una apariencia ontológica que quiere mostrar un mundo más allá del que nos permite existir. Repitiendo la fórmula heideggeriana: “No hay ser-en-el-mundo porque estamos siempre ya en un mundo de otro” (2022 78). La religión y el capitalismo, con ayuda de los dispositivos de poder en una era postdisciplinar controlan la información que nos posibilita entendernos en los mundos como sujetos. La insistencia de Lapoujade en la figura de los *dealers*¹⁰ es clave para explicar los modos en que se poseen los mundos: “Gobernar significa entonces imponer la autoridad de una forma exclusiva de realidad” (2022 84). Y los *dealers* son los representantes tanto del capitalismo como de la religión, pues en su ejercicio de repartir instancias de delirios son capaces de eliminar a la realidad en tanto que preexistente como un presupuesto. El control tiene como objeto al psiquismo por medio de la información.

De ahí que en *Los mundos artificiales* (2022 87-98), el autor apueste —en contra de Frederic Jameson (2008)— que en realidad PKD no está intentando restituir una actualidad posthumana a su presente, sino que más bien quiere “revelar el aspecto falso o ficticio del presente del que Dick es contemporáneo” (2022 93). Incluso, la lectura de Lapoujade sobre Dick podría considerarse anticolonialista si consideramos que

El colono no es solamente aquel que se apropia de la tierra y de sus ocupantes; es también el que les impone una nueva realidad, como los misioneros imponen su dios

¹⁰ Sí, de drogas, porque en la obra de PKD, las drogas son instrumentos claves para la colisión de mundos.

a las poblaciones infieles o heréticas, con la diferencia de que, en Dick, se utilizan medios tecnológicos o farmacológicos (2022 96).

La más clara confirmación de lo anterior aparece cuando PKD afirma que “en nuestra sociedad hay que pagar para ser reducido a la esclavitud, lo que es el mayor de los insultos” (2022 94). Una aseveración tal permite entender por qué cuando nos encontramos con humanos en las novelas de Dick, en realidad, nos encontramos con androides. Las nuevas formas de sujeción llevan al extremo las formas clásicas de humanismo, la racionalidad deviene incansable y el androide le permite al humano ya no luchar por la dominación de los mundos. Más esclavos que los robots, la inhumanidad se revela como un elemento determinante de los psiquismos que ya no luchan contra el realismo capitalista. En esta línea de pensamiento es que en *El hombre digital (o qué es un androide)* [2022 99-115] el autor del ensayo realiza lectura casi nietzscheana de la comprensión que Dick tiene de la religión. Al comprender las religiones como “fuerzas políticas del porvenir” (2022 112) se está refiriendo a ellas como “la Mentira fundamental”¹¹ de la que Dick habla en sus *Exegesis* (2011), en tanto que funcionan como negadoras de este mundo para promover otros.

Pero frente a los que no luchan, suelen aparecer los paranoicos capaces de construirse a sí mismos en un contexto literario que refleja la sociedad en la que Dick se movía, entre la guerra fría, el macartismo, el caso Watergate o, en breve, una *American way of Life*. De esta manera, los que intentan escapar de la ideología de turno participan de la *Cacería y paranoia* (2022 117-124) en la que todo esfuerzo por ir más allá de lo que nos esclaviza nos convierte inmediatamente en enemigos. Rick Deckard, de *Do Androids Dream of Electric Sheep?* (1968) es un ejemplo paradigmático de lo que intenta explicar Lapoujade. El que retira androides, en el aparecer de su paranoia, no sabe si él mismo es también “algo” que debe ser retirado. Y peor aún, como bien refleja la película de Ridley Scott basada en la mencionada novela, ¿los creadores de mundos artificiales no son incluso capaces de acecharnos en nuestros sueños? Nadie está a salvo.

Aunque también tiene sentido preguntarse: ¿de qué deberíamos salvarnos si tampoco es que seamos seres vivos que disputamos la comprensión del mundo? Las obras de Dick describen innumerables personajes que existen entre la vida y la muerte. No hay una diferencia categorial entre ellas, pues se puede estar en un estado de semivida, a veces muerto, otras veces ser llamado a la vida o incluso consumir la vida de otros o de uno mismo en diferentes líneas temporales. Así, *Entre vida y muerte* (2022 125-133) se encarga de detallar la lucha que

¹¹ La influencia de la *Genealogía de la Moral* (1887) en este capítulo es evidente en tanto que el ascetismo se entiende como una forma de cuidado de sí, o bien, en el contexto de Dick, como una forma de devenir inhumano.

hemos venido describiendo —por la realidad— ya no tanto en las antípodas de la vida y la muerte, sino que dominando la comprensión de la existencia *entre* las mencionadas categorías. La sobrevivencia es política y se trata, sobre todo, de imponerse sobre las subjetividades paranoides que podrían intentar entrar en guerra con los que imponen el control, aunque eso implique extender el dominio más allá de la vida.

No deja de ser interesante que entre tanto pesimismo de ciencia ficción frente al realismo capitalista todavía existan alternativas para PKD. En *Hacer bricolaje (o la variable aleatoria)* [2022 135-149] Lapoujade realiza una descripción clave tanto para el tipo de CF que escribe Dick como también sobre el tipo de personaje que cargan con el desarrollo de la novela. Sus relatos suelen ocurrir luego de que la catástrofe ya tuvo lugar (2022 136), de manera que los personajes de ese mundo se enfrentan a un mundo desconocido, jamás ocupando una posición especial en la jerarquía social. Los personajes dickianos aparecen ahí donde la preexistencia de mundos no está dada y así tampoco lo están las posibilidades de sus protagonistas. En este sentido es que la meta de los personajes en estas obras

[...] no es apropiarse de un mundo fijando las condiciones de su expansión, sino ensamblar fragmentos de mundos heterogéneos para circular entre ellos; el héroe dickiano es siempre un individuo modesto, pero cuyo carácter imprevisible de “partícula aleatoria” lo vuelve superior a las grandes figuras de dominación política y social (2022 143).

Así, incluso un ciudadano común, pero delirante, puede reparar o mejorar su existencia situándose por sobre las ruinas de otros mundos que ya colapsaron. Y si el autor de las obras de CF que comentamos es considerado él mismo como un personaje que también intentó levantar mundos cuando el colapso capitalista ya ocurrió, entonces no debería ser extraño también estimarlo como un filósofo. En su jovial esfuerzo por destruir las categorías que sostienen los relatos modernos encontramos también una cierta remembranza a la última figura del espíritu de Nietzsche. Es cierto que los personajes de PKD carecen de seriedad (2022 146) y es que el delirio no puede controlarse desde criterios racionales. La fuerza de la irresponsabilidad con la que él, en tanto que escritor, pero también en tanto que personaje de sus propias novelas, deja entrever una alternativa incluso cuando el mundo hizo todo lo posible para creer que no existían más alternativas.

Por eso, y sin vacilar, *La alteración de los mundos* debería ser tenido en cuenta como una lectura mínima para todos los que estamos interesados en intentar leer literatura desde la filosofía sin caer en los lugares comunes a los que tan acostumbrados nos tiene la producción ensayística de corte derrideano. El delirio

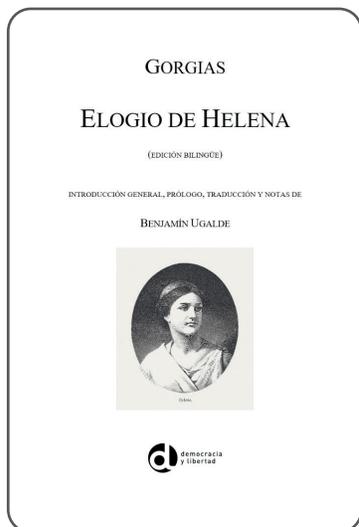
mesiánico encuentra en la pluma del Lapoujade un rendimiento filosófico que logra estabilizar uno de los primeros y más serios aportes académicos en torno a los presupuestos intelectuales de PKD.



Bibliografía

- Dick, P. K. *Eye in the Sky*. Ace Books, 1957.
———. *Flow My Tears, the Policeman Said*. Doubleday, 1974.
———. *Do Androids Dream of Electric Sheep?* Doubleday, 1968.
———. *Ubik*. Doubleday, 1969.
Nietzsche, F. *Genealogía de la Moral*. Tecnos, 2016.





Gorgias. Trad. Ugalde Benjamín.
Elogio de Helena.

Santiago: Democracia y libertad. 2023.

HUGO ALARCÓN-ACUÑA

Universidad de Chile, Chile
felalvarez@ug.uchile.cl

ORCID: <https://orcid.org/0000-0003-1870-6477>

La traducción de textos filosóficos constituye un requisito indispensable para la filosofía. En efecto, no resulta evidente bajo qué criterios un discurso en sede filosófica puede ser traspasado de una lengua a otra sin perder la expresividad, el estilo y el contenido significativo que tiene el original. Signo de esto parece ser el hecho de que no tenemos muy claro cómo ha de proceder quien se pretende ocupar como traductor de un texto filosófico; qué reglas orientativas debe emplear con el propósito de obtener criterios de corrección que le permitan evaluar su traducción como ‘buena’ o ‘mala’. Ahora bien, aún cuando la posibilidad de una teoría universal de la traducción permanece como una incógnita bastante poco plausible, la tradición nos ofrece una indicación en sede negativa, que al decir de Cicerón se expresa en el conocidísimo apotegma “*non verbum pro verbo*”. Vale decir, no se trata de traducir palabra por palabra, coma por coma y, en fin, símbolo por símbolo. Antes bien, de lo que se trata es de tomarse en serio el texto original bajo la perspectiva de totalidad en la que este está enmarcado, y que siempre va más allá del conjunto de símbolos expresos en el texto, y a partir de ahí obtener una orientación para llevar a cabo la ocupación

CÓMO CITAR ESTE ARTÍCULO:

En APA: Alarcón-Acuña, H. (2024). Gorgias. Trad. Ugalde Benjamín. Elogio de Helena. Santiago: Democracia y libertad. 2023. *Resonancias*, (17), 167-170. DOI: 10.5354/0719-790X.2024.74181

En MLA: Alarcón-Acuña, H. “Gorgias. Trad. Ugalde Benjamín. *Elogio de Helena*. Santiago: Democracia y libertad, 2023.” *Resonancias*, n.º 17, julio de 2024, pp. 167-170. DOI: 10.5354/0719-790X.2024.74181

de traducirlo. Nuevamente, en palabras de Cicerón “*Non enim ea me adnumerare lectori putavi oportere, sed tamquam appendere*”¹.

Hago esta indicación porque me parece que la traducción que Ugalde (2023) hace del *Elogio de Helena* se enmarca, a sabiendas de ello o no, en esta tradición de traductores que remonta hasta los rétores antiguos. Esta observación se vuelve más solícita cuando constatamos que el texto en cuestión es uno pocos de los sofistas que se mantuvo íntegro hasta nuestra actualidad. La traducción en comentario debe arrogarse el mérito de conservar un “estricto apego al sentido del texto griego” (p. 49), pero también cuidando la expresividad y el estilo que deben tener un discurso epidíctico que toma por propósito la exoneración de responsabilidad de Helena frente a las causas que conllevaron a la guerra de Troya.

Con todo, Ugalde entiende que una lectura de un texto de Gorgias aparece inmediatamente circunscrita bajo prejuicios hermenéuticos que tenemos sobre los denominados ‘sofistas’. Ciertamente, la inserción de la introducción general y el prólogo permiten explicitar y, por cierto, problematizar tales prejuicios que obliteran la comprensión que solemos tener de estos autores. Este es uno de los logros más acabados de la traducción, además de ella misma.

Ugalde es lúcido al constatar que bajo el rótulo ‘sofistas’ no estamos ante una unidad doctrinaria común propia de una escuela filosófica (p. 9). Más bien, lo que nos ha sido legado es la constatación de los miembros de este movimiento bajo el discurso filosófico de Platón, pero a partir de una consideración eminentemente desfavorable. Es justamente en este punto donde recae la procedencia del prejuicio que desfigura la evaluación filosófica que pudiésemos hacer sobre los sofistas, puesto que o bien son vistos como autores carentes de cualquier valor filosófico, o bien son considerados como individuos cuyo discurso no tiene pretensión de verdad (p. 10.). En tal sentido, la traducción que lleva a cabo Ugalde se asigna la difícilísima labor de mostrar el potencial filosófico que tiene el *Elogio*. En otras palabras, la introducción y el prólogo pretenden dejar de manifiesto que esta es una traducción de un texto *filosófico*, al consignar, por una parte, la vigencia que tiene para ciertas vertientes de filosofía política (pp. 27-33) y, por otro, mostrar que los sofistas se enmarcan en un diálogo directo con los primeros filósofos griegos con relación a cuestiones de recorte epistemológico, ontológico, ético y político (pp. 22-25).

Se trata, en el fondo, de hacerse cargo de la comprensión histórica que anticipa *ex ante* el sentido que le pudiésemos otorgar a un texto de un movimiento que la tradición a relegado como no-filosófico, y mostrar contra esto que los sofistas merecen un tipo de atención e interés que no sea anticuario, sino que

¹ “Puesto que no he creído oportuno que el lector las cuente [las palabras], sino que las pese”.

sean sopesados en virtud de su propio mérito filosófico, como puede ser la teoría del *lógos* gorgiana que está incoada a lo largo del *Elogio* o la distinción entre *nómos* y *physis*, que también recorre la totalidad del texto. Así vistas las cosas, la traducción de Ugalde demuestra una sensibilidad hermenéutica al tomar fina nota de que el ejercicio de traducir un texto debe hacerse cargo, a su vez, de elementos que trascienden la esfera de símbolos significativos que contiene este, ya que la posibilidad de que un texto sea comprensible es siempre tributaria de una interpretación históricamente legada. A la práctica de la traducción le subyace siempre una labor interpretativa, y Ugalde es sumamente conciente de esto.

Después de estas breves indicaciones, paso a comentar brevemente el modo en el que Ugalde pensó la construcción del texto en cuestión, ya que la elección de una introducción general y un prólogo que antecedan a la traducción misma cumple una función que va en consonancia con lo anteriormente mencionado. Ciertamente, el libro presenta una direccionalidad que va desde lo más general hacia lo más particular, ya que la introducción tiene por propósito explicitar la concepción filosófica propia de los sofistas, así como sus discusiones con otros exponentes como los físicos o el propio Sócrates. Al respecto, Ugalde ofrece observaciones interesantes que pretenden concebir a este como un sofista (pp. 28-29). La tesis gruesa que recorre la totalidad de la introducción general es que los sofistas fueron pensadores cuya riqueza filosófica no se agota en la inculcación de técnicas destinadas a la persuasión, dado que de un modo igual de enérgico se abocaron a la ontología y la epistemología (pp. 22-24), como también a la política en el sentido de pensadores de la *polis* (pp. 25-258).

El prólogo se ciñe exclusivamente a la figura de Gorgias de Leontinos; este se enmarca en el *background* hermenéutico establecido en la introducción, en la medida que Ugalde ha sido enfático en llamar la atención acerca del potencial filosófico que tiene el discurso filosófico de los sofistas, lo cual incide directamente en la disposición de quien lee el prólogo y el objeto de este mismo: Gorgias y su *Elogio*. Lo que nos encontramos acá es una contextualización de la función epidíctica y apologética del *Elogio*, así como también la teoría del arte retórico inscrita en el mismo (pp. 38-39), del mismo modo se ofrece una vista panorámica del mito de Helena y la situación en la que se encuentra su figura en el momento en el que Gorgias decide escribir su discurso sobre ella (pp. 40-41).

Hay un segundo aspecto del prólogo que merece atención, que alude directamente a la teoría gorgiana de la retórica y del lenguaje (pp. 43-49). La premisa de la que parte concierne a la fuerza creativa y generadora de experiencia que tiene la palabra (*lógos*). Según las palabras de Ugalde “Así, la palabra adquiere un rol protagónico en la filosofía del lenguaje de Gorgias. Al no haber un sustrato o esencia ontológica —tal como lo plantea el sofista en su tratado *Sobre lo que no es*— la palabra, el *lógos*, es la que viene a ocupar este lugar” (p. 46). La función

que Gorgias le asigna al lenguaje, en el fondo, consiste en la producción significativa de las cosas para el ser humano (p. 47). Esto se vuelve más interesante aún si uno toma en consideración, tal como lo hace Ugalde, de que uno de los candidatos que Gorgias concibe como responsables de los actos de Helena no es sino la palabra que con su fuerza nos imprime emociones que nos lleva a ciertos campos de acciones, querámoslo o no.

Por último, la traducción toma como referencia las ediciones de Diels y Kranz (1951-1952), y la de Untersteiner (1954-1962) para establecer el texto crítico. Del mismo modo, Ugalde se basa en las versiones de MacDowell (1982) y Donaldi (1982). Con una versión bilingüe que permite cotejar la traducción con su original, Ugalde emplea un lenguaje comprensible para el lector actual sin perder de vista el sentido y, muchísimo más importante aún, el estilo del discurso de Gorgias. Además, el conjunto de notas incluidas en la traducción permite al lector comprender las razones y motivos que llevaron al traductor por las palabras empleadas, así como también permiten ir siguiendo el argumento esbozado por el propio Gorgias.

El texto que nos presenta Ugalde constituye un genuino aporte para quienes se quieran adentrar en el estudio del pensamiento de los sofistas, en general, y de Gorgias, en particular. A riesgo de caer en la majadería, Ugalde nos muestra un excelente ejemplo de que traducir un texto filosófico no debe consistir en la reproducción de la misma cantidad de símbolos en el mismo orden en el que estaban en el original. Antes bien, este debe ser ‘sopesado’ en favor del sentido y del estilo que el discurso en cuestión pretende expresar. Podemos decir con bastante seguridad que la traducción de Ugalde satisface con creces este criterio de corrección propio de la práctica del traductor.



Bibliografía

- Diels, H. y Kranz, W. *Die Fragmente der Vorsokratiker*. vol. 3. Weidmann. 1951-1952.
 Gorgias. Trad. Donadi, F. *Encomio di Elena*. L'Erma di Bretschneider. 1982.
 Gorgias. Trad. MacDowell, D.M. *Encomium of Helen*. Bristol Classical Press. 1982.
 Gorgias. Trad. Ugalde, B. *Elogio de Helena*. Democracia y libertad. 2023.
 Untersteiner, M. *Sofisti. Testimonianze e frammenti*. Vol. 4. La Nuova Italia. 1954-1962.

