



DESAFÍOS DE LOS SISTEMAS Y APLICACIONES DE LA IA PARA LA EDUCACIÓN Y LA FORMACIÓN DE CIUDADANÍA

CHALLENGES OF AI SYSTEMS AND APPLICATIONS FOR EDUCATION AND CITIZENSHIP FORMATION

LUZ MARINA BARRETO 

UNIVERSIDAD CENTRAL DE VENEZUELA, CARACAS, VENEZUELA

Fecha de recepción: 27 marzo 2024

Fecha de aceptación: 24 mayo 2024

RESUMEN

Este artículo refiere los desafíos que los nuevos sistemas de inteligencia artificial generativa plantean a nuestros conceptos de conocimiento, realidad y verdad y, por extensión, a aquellos lugares que tradicionalmente han tenido la tarea de definirlos, a saber, las universidades modernas. Presenta la idea de que la información no puede ser homologada al conocimiento sin despojarlo de su significado filosófico genuino, en tanto que una definición de la realidad entendida como portadora de referentes verdaderos. Como resultado de esta confusión entre información y conocimiento verdadero, los sistemas informáticos que pueden gestionar exitosamente sentido semántico, tal y como los sistemas de aprendizaje automático han logrado hacerlo de manera espectacular en las últimas décadas, fallarán de todos modos en la tarea básica de promover consensos racionales entre personas. Esta insuficiencia es debida a los límites de nuestra capacidad de modelar la realidad basándonos solamente en funciones lógico-matemáticas, de modo que el papel que ha de desempeñar una universidad que goce de suficiente financiamiento y apoyo para la investigación es ineludible para comprender y regular todos los sistemas de inteligencia artificial.

PALABRAS CLAVE: Inteligencia Artificial; Universidad; Educación; Ciudadanía.

ABSTRACT

This paper addresses the challenges posed to our concepts of knowledge, reality and truth (and by extension to the places traditionally endowed with the task of defining them, i.e., modern universities) by the new systems of generative artificial intelligence. It conveys the idea that information cannot be understood as knowledge without emptying the latter from its genuine philosophical meaning as a conceptualization of reality that carries true meaning. As a result of the confusion between information and true knowledge, systems that can successfully manage semantic meaning, as all digital and machine learning systems have spectacularly learned to do in the last decades, will nevertheless fail in their capacity to bring about rational consensus among persons. This happens because our capacity to model reality solely with logic and mathematical functions is limited and therefore my paper stresses the role a well-funded and unabashed university has to play in understanding and regulating all systems of artificial intelligence.

KEY WORDS: Artificial Intelligence; University; Education; Citizenship.

1. INTRODUCCIÓN

Los problemas teóricos y conceptuales de los desarrollos de la inteligencia artificial y a la posibilidad de ampliar los sistemas actualmente existentes en dirección de una inteligencia artificial *general*, es decir, una que permita a los modelos de aprendizaje automático pasar de un dominio de objeto a otro, son problemas que atañen a la reflexión filosófica, en la medida en que ella es una disciplina de fundamentación. De hecho, la filosofía del siglo XX se perfila como una reflexión sobre las crisis en la fundamentación del conocimiento científico que impulsaron muchas de las innovaciones que se produjeron durante las primeras décadas del siglo (Barreto, 2023a).

La filosofía, como disciplina de fundamentación, se ocupa de los conceptos básicos que definen los dominios de objeto de las distintas ciencias y los analiza para ampliar el posible campo de aplicación de una teoría vigente que comienza a perder flexibilidad o es puesta en cuestión por nueva evidencia empírica que no puede explicar (Heidegger, M., 1927; Popper, 2023; Gilson, 2004; Barreto, 2023a). De este modo, la filosofía del siglo XX es básicamente una reflexión sobre los fundamentos del conocimiento y, por ello, el problema de la racionalidad del conocimiento teórico no ha dejado de encarar los desarrollos de las ciencias de la computación e información en su relación con la teoría del conocimiento y su capacidad de alcanzar la verdad y definir la naturaleza de la realidad (Lonergan, 2005; Lucas, 1961).

El problema de fondo, y el tema del presente artículo, es si la ganancia de información puede definirse conceptualmente como ganancia del conocimiento. Mi hipótesis es que la respuesta que podemos darle a este problema tiene consecuencias en dos ámbitos: en el de la posibilidad de desarrollar una inteligencia artificial general y en el de la posibilidad de definir el papel y los desafíos a los cuales se enfrenta el modelo tradicional de universidad como generadora de conocimiento innovador y, por ello, a su capacidad de contribuir al desarrollo de una ciudadanía responsable y reflexiva.

2. Desafíos de la IA

La metodología utilizada para la redacción del presente artículo es el análisis conceptual filosófico de fuentes primarias y secundarias (Wittgenstein, 1939; Strawson, 1992).

El primer desafío que la IA plantea a la universidad y a otras comunidades académicas de argumentación como generadoras de conocimiento, así como también a los sistemas democráticos que delimitan un espacio público institucional para la discusión política y la creación de ciudadanía, puede definirse como sigue: los sistemas de IA preparan los conjuntos de datos fragmentando como valores y parámetros discretos el conocimiento previamente estructurado por el sistema formal de educación universitaria, entre otros substratos culturales que sustentan lo que consideramos conocimiento o saber compartido (Barreto, 2018; 2022). En este sentido, lo que hace la IA es aprovechar las capacidades que tienen muchas personas, en particular jóvenes, para anotar o etiquetar conjuntos de datos, competencia que han adquirido porque han pasado por un sistema de educación formal en

sus países de origen o tienen una capacidad especial para expresar de modo racional sus experiencias y saberes. Pero no sólo esto. Es también muy importante notar que el conocimiento, como sabemos los filósofos y lo analizamos desde hace milenios y de forma más sistemática desde tiempos más recientes es, en realidad, nuestra respuesta racional a los problemas que nos presenta nuestro entorno, tanto natural como social. Nuestra razón resuelve problemas y esas soluciones son justamente racionales en la medida en que son *razonables*, en que se apoyan en aquellos que comparten con nosotros un mundo común y pueden articular comunicativamente con nosotros las distintas vertientes del problema y sus soluciones y, además, ven los problemas dentro de un contexto más amplio que se entiende como “realidad” (Spitzer, 2015; Barreto, 2023 b). Interpretar un solo dato aislado de la realidad requiere que quien interpreta esté familiarizado con el mundo contextual que le da *sentido*, un sentido del que Claude Shannon (Shannon, 1948) pensaba que se podía prescindir en el análisis de la comunicación. Por esta razón, todos los sistemas de IA necesitan anotadores o etiquetadores.

Estos anotadores por lo general se encuentran en la sombra o trabajan en países como Colombia o la India, entre otros, incluyendo, por supuesto, nuestro país Venezuela. Su contribución se oculta, disimula, o no se comprenden bien sus reales alcances, por distintas razones. En primer lugar, porque su presencia, por ejemplo en aplicaciones de la IA semántica como ChatGPT que aprovecha no sólo la competencia de etiquetadores humanos, sino también documentos académicos y de otro tipo dispersos por la red, disminuye el valor de las acciones de las distintas empresas que impulsan estos desarrollos, como Google, OpenIA, Meta, etc., en la medida en que este reconocimiento pudiera volverlas vulnerables a alteraciones de tipo político, como huelgas o estados de descontento de sus trabajadores.

Muchas de las aplicaciones de IA están interesadas en hacer creer a sus inversores que son, de algún modo, total y mágicamente automatizadas y, por lo tanto, inmunes a ese tipo de problemas humanos, de modo que el retorno de la inversión estaría asegurado y 100% garantizado. El hecho de que estas aplicaciones aspiren o ya coticen en la bolsa de valores de Nueva York, en caso de que no hayan sido intervenidas ya por un fondo de capital privado como en el caso de OpenAI, es un incentivo enorme para prolongar la ilusión de que ellas prescindan de los mecanismos habituales de generación de conocimiento verdadero y que rendirán dividendos a sus inversores *per secula seculorum*, aprovechando que existen trabajadores desempleados que son graduados universitarios en muchos países que están dispuestos a generar bases de datos por un salario de subsistencia. Un resultado de este tipo de reconocimiento o insight es que la capacidad de estas aplicaciones para generar conocimiento original es necesariamente limitada.

Es sólo cuestión de tiempo para que muchas de estas aplicaciones terminen por exprimir la originalidad de las fuentes de información que alimentan sus sistemas y se vuelva necesario, por lo tanto, regresar a los estilos de investigación científica rigurosa y profunda que sólo puede ofrecer un entorno reflexivo, plenamente comunicativo y genuinamente autónomo como el que promueven las universidades y no *cualquier* tipo de universidad, es oportuno recordar.

El segundo gran desafío que plantean estos sistemas tiene que ver con la estructura misma de la sociedad democrática y las fuentes que animan la responsabilidad ciudadana.

Las técnicas de tratamiento analítico y estadístico que se desarrollaron a inicios del siglo XX para encontrar sentido en grandes cantidades de datos discretos o conglomerados de elementos se usan ahora para analizar el comportamiento de conjuntos de usuarios en redes sociales, entendidos como nodos en una gran red. La enorme confluencia de millones de usuarios sobre una red social produce un fenómeno conocido como redes libres de escala, es decir, conjuntos de datos complejos que confluyen en una enorme red que carece de una medida media y por lo tanto producen distribuciones de comunidades de usuarios muy fragmentadas o con distintas distribuciones de probabilidad posibles para contenidos semánticos y otros atributos (Caldarelli, 2007).

Y aquí sucede lo siguiente: debido a la complejización de las sociedades y al aumento de ciudadanos educados por un sistema formal en sociedades contemporáneas, el ágora pública que define una voluntad política común no se constituye ya en espacios institucionales en los que todos pueden participar en igualdad de condiciones. Las redes sociales online se convierten entonces en el espacio público al que todos intentan contribuir en pos de la creación de una voluntad política común. Pero, como a partir de un cierto número crítico de usuarios o nodos o vértices, la red pierde la media y se convierte en una red libre de escala, el conjunto total de la red se fragmenta en grupos discretos.

Este es un resultado muy extraño que contradice un poco, o hasta dónde alcanzo a ver, la intuición de Francis Galton en relación con la “sabiduría de las multitudes”, una manera intuitiva de interpretar el teorema del límite central, de acuerdo con el cual valores en grandes números confluyen siempre a un valor medio. Distribuciones de Gauss o normales tienden a aparecer por doquier en el mundo natural, de modo que números grandes tienden a distribuirse alrededor de un valor promedio. Pero esto no sucede así en grandes redes sociales, en donde, como se ha dicho, la red se fragmenta en grupos discretos. Cuando el espacio de opiniones en las redes sociales se modela conforme a una función, como el espacio total produce una red libre de escala, el gráfico de la función no es una curva normal o una curva de campana, sino un gráfico de leyes de potencia. En este caso, la desviación de la función tiene lo que se llama una “cola pesada”, por contraste con la desviación que decae exponencialmente en las funciones que dibujan una curva normal o de campana. En este último caso, la varianza se acota de modo más o menos preciso para los valores medios, mientras que, en los gráficos de leyes de potencia, por contraste, la ausencia de escala hace que se puedan encontrar valores variables a todo lo largo del espacio geométrico, sin que la desviación caiga a cero. En un ejemplo muy elocuente dado por el matemático húngaro-americano Albert Laszlo Barabási, si la estatura humana se modelara no con una curva normal, como suele hacerse en la vida real, sino con un gráfico de leyes de potencia, todo el mundo sería más o menos tan chiquito como un elfo de cuento de hadas y de vez en cuando encontraríamos en el espacio gente que mide, no sé, cien metros, los verdaderos valores extremos de la función (Barabási, 2018).

Como sugieren Nicola Perra y Luis Rocha en un importante artículo publicado en el 2019, esta fragmentación de la red social en grupúsculos discretos puede suceder incluso sin que sea intencional, simplemente por el efecto de los tipos de algoritmos utilizados, y ello encierra inevitablemente a grupos enteros de usuarios en cámaras de eco o de resonancia en donde, por ejemplo, la tierra es plana, mi candidato favorito ganó indudablemente las

elecciones, o las vacunas son letales, todas opiniones que se salen de una media razonable. Como sugieren algunos estudios, al parecer mientras más encerrado está un grupo en un determinado contexto, más divergente respecto de la red llega a ser, dado que esta cerrazón es un producto de la focalización del grupo en un determinado entorno de sentido (Perra y Rocha, 2019).

Por ejemplo, un banco de peces que descubre una fuente de comida se aglutinará alrededor de esta en el inmenso océano, sin que otros bancos de peces de la misma especie lleguen a enterarse. Podemos suponer que un banco de peces específico se aglomera alrededor de la fuente de comida no sólo por efecto de los mensajes sensoriales que los peces reconocen, sino por los mensajes que emiten los peces vecinos. Lo mismo sucede en una red social: los mensajes angustiados de un grupo humano que experimenta un terremoto en un determinado lugar del mundo se reproducirán hasta un determinado número crítico, luego del cual el mensaje pudiera llegar a carecer totalmente de interés para un grupo de usuarios muy alejado del epicentro de la catástrofe. De este modo, podemos suponer que la fragmentación es un efecto natural de la complejidad de la red que se relaciona con distintos contextos y responde racionalmente a ellos.

Si la fragmentación es un resultado natural de la complejidad de las relaciones humanas con el entorno entonces es probable que, como sugiere la investigadora de la Universidad de California Davis Raissa De Souza, no tenga remedio y sea irreversible, tal vez porque es un aumento neto de la entropía de un sistema de información (D'Souza, 2015). Por esta razón, las mismas empresas de redes sociales la explotan y venden su capacidad de agrupar usuarios en distintos grupos a terceros con diversos intereses políticos o de mercadeo. Buena parte de la alteración y crispación del clima político venezolano y latinoamericano, en donde prevalece una voluntad de no ponerse de acuerdo con el adversario sino de suprimirlo, es un producto directo de la manipulación intencional de las conversaciones que tienen lugar en las redes sociales y que son dominadas por subcontratistas de estas empresas o por grupos de activistas virtuales o “cybertroops” que son contratados por distintas agencias internacionales para servir a uno u otro grupo político por relativamente poco dinero. Por supuesto, estas empresas también venden como publicidad su capacidad de conectar usuarios en campañas de mercadeo.

Este tipo de tareas de manipulación de la opinión pública en torno a un tema de interés político general se hace explotando distintas medidas de centralidad de usuarios o nodos en una red. En particular, se pueden cortar los ejes que conectan un usuario que exhiba una medida elevada de *centralidad intermedia* con otros grupos, con lo que se consigue que el mensaje no se transmita de un grupo a otro que pudiera estar interesado (Granovetter, 1973). De este modo, un mensaje importante pudiera morir allí mismo y quedar sepultado debajo de todos aquellos que subcontratistas interesados sí quieren que circulen por la red.

Esto, que todo el mundo ha visto que sucede, es muy fácil de hacer si uno es dueño de la red y responde, por un monto económico dado, a la petición de un tercero interesado. El poder definir usuarios con una medida elevada de centralidad intermedia *es* sin duda alguna el fundamento real del poder de las redes sociales online.

3. DISCUSIÓN

3.1. Ganancia de información vs. Ganancia de conocimiento.

Para responder a la pregunta de cómo los desarrollos recientes y no tan recientes de la inteligencia artificial pudieran afectar al sistema de educación y al desarrollo pleno de una ciudadanía responsable, tal vez es útil dar un paso atrás y preguntarnos en qué consiste en realidad un modelo de inteligencia artificial.

La inteligencia artificial es una evolución subsidiaria de la capacidad técnica que surgió a finales del siglo XIX e inicios del siglo XX para digitalizar información y enviarla a través de canales eléctricos o codificándola en transistores ordenados con operadores lógicos. El problema de cómo comprimir la información de manera eficiente, eliminando el ruido, la probabilidad de que *bits* de información pudieran perderse o confundirse entre los otros posibles mensajes en un sistema, y otros tipos de obstáculos de carácter físico, ocuparon muchos de los esfuerzos de tipo técnico que se encuentran en el origen de los actuales sistemas de IA.

Un momento central en este itinerario es la idea, que fue tomando forma antes de la primera mitad del siglo XX gracias al ingenio de Ralph Hartley y, crucialmente poco tiempo después, Claude Shannon, conforme a la cual la entropía de un mensaje, es decir, el desorden físico que dificulta su reconocimiento entre otros mensajes posibles o a causa del ruido, se puede medir matemáticamente como la sumatoria de su probabilidad por el producto de su logaritmo. La idea subyacente es que se puede medir la información como una cantidad física, con independencia de cualquier significado o inteligibilidad de la misma.

Precisamente por ello, Claude Shannon, el padre de la teoría de la información, llega a pensar al inicio de su trabajo como matemático en los laboratorios Bell, que un manejo eficiente de grandes cantidades de información tiene que prescindir de su condición de portadoras de verdadero conocimiento: “El problema fundamental de la comunicación – escribe en su artículo seminal de 1948– es el de reproducir en un punto, bien sea exactamente o aproximadamente, el mensaje seleccionado en otro punto”. Y agrega: “Frecuentemente los mensajes tienen *sentido*” (Shannon, 1948).

Aquí parecía más intuitivo creer que mientras más símbolos se usan, más informativo es el mensaje. En realidad, lo que sucede es lo contrario: mientras más grande sea el pool de símbolos de dónde se debe elegir, más informativo son los poquitos elegidos, más informativa o inesperada es cada selección y es esa sorpresa la que hay que intentar medir de alguna manera. Información se opone así pues conceptualmente a la entropía, al desorden de los símbolos.

La ecuación que define Hartley, y que Shannon precisará después, para medir la cantidad de información que transmite un mensaje es, entonces:

$$H = n \log s$$

En donde H es la cantidad de información, n es el número de símbolos en el mensaje y s el número de símbolos en el lenguaje disponible que debemos elegir para expresar el mensaje. Es decir, H es la cantidad de información, n el número de símbolos transmitidos y

s es el tamaño o “size” del alfabeto. Mientras menos símbolos hay en un lenguaje, más símbolos deben ser transmitidos para que el mensaje sea informativo, es decir, más redundante debe ser. Pero, por contraste, esto implica también que mientras más podemos elegir, mayor es la incertidumbre y menor es la información (Gleick, 2011).

Como ingeniero, el problema para Hartley era cómo optimizar la transmisión de sonidos con ondas electromagnéticas a lo largo de un cable o a través del aire. Y aquí se trataba no de la transmisión de ruido, sino de información, sonidos cuantificados como información.

Por ejemplo, supongamos que queremos definir la cantidad de información que tiene 4 letras elegidas de un alfabeto de 28 (“Hola”):

$$\log 28=4.80 \text{ (O sea: } 2^{4.8}=27.85)$$

Según la fórmula de Hartley este resultado se multiplica $\times 4= 19.2$

El logaritmo hace más manejable la cifra real de combinaciones posibles, que es un exponente:

$$28^4=614.656 \text{ (O sea, } 2^{19.23} \text{ aprox.)}$$

Mientras más bajo es ese número en la fórmula, más informativo es el mensaje, según Hartley. Su fórmula significa que la cantidad de información no es proporcional al tamaño del alfabeto. La relación es logarítmica: para doblar la cantidad de información, es necesario elevar al cuadrado el tamaño del alfabeto (Gleick, 2011).

La cantidad que describe la información no es proporcional al tamaño bruto del alfabeto, sino a las combinaciones y elecciones posibles de la palabra. Se trata de hacer abstracción del sentido o semántica tradicional del mensaje para comprender el fenómeno de la comunicación como el resultado de simples combinatorias de elementos que se deben atinar cada vez. La comunicación ya no se ve entonces como la esforzada o laboriosa construcción de un sentido dentro de una comunidad de hablantes, sino como la búsqueda de la combinación más eficiente o económica de símbolos para transmitir lo que uno quiere decir del modo más rápido posible.

Por su parte, afirma igualmente Gleick, Claude Shannon reflexiona también desde muy temprano sobre la paradoja de que un mensaje, para ser comprendido, conlleva una cierta cantidad de redundancia. Desde el inicio, está interesado en los aspectos estadísticos y probabilísticos que acompañan la transmisión de mensajes y cómo estos aspectos contribuyen a su inteligibilidad.

Shannon concibe una solución más elegante que la de Hartley, y mide la información como una suma de probabilidades para todos los estados posibles ($p_1+p_2+\dots+p_n$), con una función logarítmica base 2 que mide, como en el primer caso, la improbabilidad del mensaje y ofrece, entonces, también una idea de su carácter de no ser previsible, de ser inesperado:

$$H = -\sum p_i \log p_i$$

En donde p_i es la probabilidad de cada mensaje i . La medición de la información transmitida mide elección e incertidumbre. Es, por lo tanto, una medida de entropía, en el

sentido de que, a mayor entropía, menor información, dado que todos los estados saldrían con igual probabilidad (Shannon, 1948).

Poder medir de modo objetivo la información que se transmite a través de un medio físico vuelven muy importantes los desarrollos técnicos de hardware y los desarrollos matemáticos y analíticos de software para codificar, comprimir y transmitir adecuadamente inmensos caudales de información.

Paralelamente a estos esfuerzos que se hacen en la teoría de la computación y en la física, esfuerzos que están haciendo matemáticos e ingenieros electricistas, por el lado de la lógica y la filosofía del conocimiento y de la ciencia se está tratando, desde inicios del siglo XX, de discernir modelos lógicos y ontológicos que nos ayuden a definir parcelas de la realidad de manera definitiva, no por cierto por afán totalitario, sino para facilitar, como en el primer caso, los entendimientos y la comunicación racionales entre personas que quieren ponerse de acuerdo sobre que algo es el caso, o sencillamente sobre que algo *es* algo. Tanto aquí como allá se trata del problema de cómo facilitar la transmisión de “inteligencia” e información de modo que sean más fáciles la comunicación y los acuerdos racionales.

No obstante, en 1931, con su *teorema de la incompletitud*, el lógico y matemático de origen checo Kurt Gödel descubre que no es posible *probar* que un enunciado o proposición que sabemos verdadera dentro de un sistema de pruebas de tipo lógico-matemático dado es verdadero o falso. Es decir, logra demostrar que existen enunciados dentro de un sistema lógico que el sistema nos dice que son verdaderos pero que no se pueden probar dentro de ese mismo sistema lógico o que son indecibles en ese sistema lógico-axiomático, con lo que surge inmediatamente la pregunta ¿cómo sabemos que es verdadero lo que no se puede probar dentro de ese sistema? Este resultado es muy extraño, dado que Gödel de algún modo demuestra que cuando mostramos lógicamente que la verdad de un enunciado no se puede probar dentro de un sistema lógico y nos damos cuenta de que esa incapacidad de hacerlo es *verdad*, necesariamente trascendemos el sistema en cuestión, lo vemos desde una perspectiva más elevada, como desde arriba. O dicho de otro modo, vemos que una demostración lógica no se puede hacer pero no sabemos cómo es que lo sabemos, porque no es el sistema lógico quien nos lo dice, dado que no puede probarlo, sino una perspectiva más elevada que no sabemos muy bien de dónde sale, una perspectiva que la *trasciende*. Por eso su teorema se denomina “teorema de incompletitud”, porque ningún sistema lógico consistente puede ser suficientemente expresivo o completo, es decir, demostrar o probar todos los enunciados o axiomas del sistema.

Una manera de interpretar este resultado tan curioso, una en la que Gödel mismo estuvo trabajando hacia el final de su vida, es que los *insights* o entendimientos que resultan de aplicar pruebas lógicas a enunciados deben ganarse o formularse de otro modo, a saber, a través de un procedimiento discursivo y reflexivo de mayor envergadura o amplitud teórica y filosófica, intentando así enunciar de manera más amplia aquello que entendemos pero que no sabemos expresar con las únicas herramientas que ofrece la lógica y la matemática. Volveremos sobre esta intuición, que tiene importantes consecuencias para la aspiración actual de construir sistemas de inteligencia artificial general, que son sistemas que definen un conjunto de funciones matemáticas para describir un espacio de análisis dado, así como

las tendrá también para la ontología y el realismo metafísico como fundamentos de nuestros conceptos de realidad, racionalidad y verdad (Spitzer, 2015; Barreto, 2023b).

De este modo, y de forma inexorable tal vez por la divergencia de ambos enfoques, durante todo el siglo XX la distinción entre, por una parte, el manejo eficiente de la información y, por la otra, el análisis de cómo puede comunicarse conocimiento, no hará sino ampliarse. Pese a que los computistas y matemáticos como Claude Shannon, Ralph Hartley y John von Neumann coinciden en Congresos internacionales de física con sus pares lógicos, filósofos de la ciencia y físicos teóricos, los esfuerzos en dirección de una cada vez mayor optimización de los algoritmos y técnicas para manejar información y extraer algún *insight* de ella sigue su camino sin volverse a topar demasiado con la discusión sobre la racionalidad y las teorías de la verdad que domina la reflexión de la lógica y la filosofía de finales del siglo XX y entrado este siglo XXI.

El resultado más importante de estos esfuerzos técnicos es, sin duda alguna, la digitalización de toda la información disponible y su transmisión a través de la Internet, que nos hunde bajo una montaña enorme de informaciones, sin que nos hayamos vuelto más eficaces a la hora de entenderlas o usarlas para promover entendimientos recíprocos. Entonces, cabe preguntarnos: ¿aquello que se transmite de modo tan eficiente y rápido, sin atender su sentido, su significado semántico, es *conocimiento*, contribuye al aumento del caudal de nuestros conocimientos sobre la realidad?

3.2. Algunos ejemplos de aplicaciones de la IA

Una manera obvia de responder esta pregunta es simplemente constatar que ahora es posible comprimir los datos discretos de información de la clase de un profesor y enviarla a un lugar remoto para que sus estudiantes puedan escucharla allí o más tarde. Pero, en otro sentido, ya no tan obvio o trivial, es posible por otro lado fragmentar como elementos discretos la información contenida en esa clase y someterla a distintas técnicas analíticas para ganar entendimientos en relación con aquello que es realmente importante, original o relevante dentro de la clase y, al compararlos con otros elementos discretos emanados de las alocuciones de otros usuarios de una herramienta online de educación a distancia, discernir el ámbito temático del que se ocupa la clase transmitida o sus características más importantes.

Es justamente este otro tipo de aplicación el que revela el potencial de la IA para obtener conocimiento novedoso. Un buen ejemplo es el descubrimiento en el 2020, por investigadores del MIT, de un antibiótico nuevo destinado a enfrentar la infección causada por bacterias resistentes. Para conseguirlo, los investigadores agruparon en un sistema de inteligencia artificial un conjunto de datos de entrenamiento de dos mil moléculas conocidas. Tales conjuntos se llaman “datos de entrenamiento” porque se trata de que el sistema “aprenda” las características que los responsables del sistema automatizado consideran relevantes para resolver un problema. Por eso, estos sistemas se denominan sistemas de aprendizaje automático o de *machine learning*. En este ejemplo, los datos de entrenamiento codificaron información de cada una de ellas, como por ejemplo el peso atómico de esas moléculas o los tipos de enlaces que definen propiedades que inhiben el crecimiento de bacterias. Lo interesante es que la implementación de distintos algoritmos sobre este conjunto de datos fue capaz de identificar atributos o propiedades del conjunto que hasta entonces

habían eludido la capacidad de categorización humana que se realizaba siguiendo métodos estándar o tradicionales de clasificación del conocimiento, tales como el laborioso análisis guiado por ensayo y error sobre cientos de moléculas conocidas, el cual se suele hacer, por decirlo así, con lupa.

Una vez analizado este conjunto de datos, los investigadores del MIT lo aplicaron sobre una biblioteca de 61.000 moléculas contenidas en distintos componentes conocidos, como medicamentos o productos naturales. Entre esas 61.000 moléculas se encontró *una* que cumplía con criterios de no-toxicidad y otros. Los investigadores la bautizaron “Halicina” en homenaje al robot de la película clásica *2001 Odisea en el espacio*, de Stanley Kubrick, que se llamaba Hal. Se trata, pues, de la obtención de nuevo y efectivo conocimiento hecho por un sistema de IA (Kissinger et al., 2022).

Los tres tipos de métodos de aprendizaje automático usados por la inteligencia artificial trabajan sobre datos preparados por seres humanos que anotan, etiquetan y deciden cuáles son los parámetros más importantes del modelo. Estos métodos suelen dividirse en tres: métodos de aprendizaje automático supervisado (que definen tareas de clasificación), no-supervisados (que definen tareas de agrupación con etiquetado impreciso) y por refuerzo (que son básicamente modelos de optimización). También las redes neuronales y sus derivados (entre los que se encuentra ChatGPT) pueden entenderse como un método de clasificación basado en datos previamente etiquetados. El objetivo de todos ellos es que el modelo, como hemos visto con el ejemplo de la Halicina, pueda predecir etiquetas a partir de los datos previamente anotados como relevantes en las tareas de reconocimiento o clasificación.

De este modo, podemos ver aquí que la digitalización de información en datos discretos, un enfoque sugerido primero por Ralph Hartley que se impuso rápidamente dentro de la comunidad que quería poder representar matemáticamente los flujos de comunicación, ofreció una manera directa de manejarla que es crucial para el desarrollo de los sistemas de inteligencia artificial. Permite ver la *inteligencia* (otra palabra para información) como un producto que es susceptible de ser comparado estadísticamente con conocimiento emanado de otras fuentes posibles, presuponiendo que aquellas informaciones más frecuentes, o más frecuentemente ligadas a otras que aparecerían en el espacio público de opiniones en conjunción con ellas, expresarían una forma de “consenso racional”.

Este tipo de estrategia distribuye atributos o propiedades discretos de un conjunto de datos en una matriz $m \times n$, por ejemplo, en donde las columnas codifican n -distintos valores en los parámetros del modelo, seguida por la implementación de diferentes algoritmos basados en técnicas inspiradas por la aplicación de la estadística mecánica, el álgebra lineal y el cálculo de probabilidades, que se habían desarrollado para encarar los desafíos planteados por la física de partículas al inicio del siglo 20.

Este tipo de tratamiento de datos discretos que analiza sus frecuencias estadísticas, en efecto, se vuelve necesario desde 1905, fecha en la que Einstein escribió un artículo sobre el movimiento Browniano, el movimiento aleatorio de pequeñas partículas en un fluido. Mostró que ese movimiento era el resultado de la energía calórica de las moléculas, cuya existencia probó (Einstein, 1905). Las partículas, que eran visibles en un microscopio, son de hecho

bombardeadas por colisiones moleculares y por eso es que se mueven del modo como lo hacen. El desafío de tipo técnico se convierte entonces en cómo modelar estas relaciones y colisiones de carácter dinámico. Como resultado de los estudios sobre el movimiento browniano, se definió y comenzó a utilizar una nueva herramienta técnica de tipo matemático, la estadística mecánica, diseñada para paliar las insuficiencias de la mecánica clásica para explicar el movimiento kinético de objetos en sistemas dinámicos complejos. Todos los pioneros de la termodinámica, empezando por Rudolph Clausius, Lord Kelvin, Maxwell, Boltzmann o Gibbs, se dieron cuenta de que el calor es el producto de miles de moléculas o átomos en movimiento. Y que, aunque el movimiento de una molécula tiene que seguir las leyes de Newton, cuando hay demasiadas interacciones tenemos que utilizar una disciplina distinta a la mecánica clásica. Esa es la tarea de la estadística mecánica, que traza un puente entre los detalles microscópicos y macroscópicos de una sustancia y que ha permitido el desarrollo de un nuevo paradigma ontológico en el que grandes conglomerados de elementos pueden estudiarse de modo no simplemente analítico (Anderson, 1972).

Técnicas de estadística mecánica se usan ahora para descubrir nuevos atributos en la parcela de la realidad que los parámetros del modelo mapean o codifican (Castellano, Fortunato, & Loreto, 2009). Desde el punto de vista de los métodos usados por las distintas ciencias, que la filosofía estudia para comprender sus alcances y su posible universalización a otros ámbitos, el método de extracción del conocimiento de la IA realmente constituye un despeje de los puntos de vista tradicionales que la filosofía del conocimiento y de la ciencia han estudiado a lo largo de sus milenios de historia para comprender nuestros conceptos de verdad, racionalidad y realidad. La pregunta importante para nosotros los filósofos profesionales, sin embargo, es si el uso de técnicas de estadística mecánica para encarar los consensos producidos por el manejo computarizado de la información puede reducir a la teoría de la computación y, de nuevo, a inicios de este siglo XXI, al nominalismo lógico y a la matemática, nuestros conceptos de verdad, racionalidad y realidad. Pienso que esto no es posible; pero el problema es que “allá afuera” hay mucha gente que cree realmente que este tipo de conceptos puede reducirse a los tipos de fenómenos físicos que encara la estadística mecánica. Tómese por ejemplo el caso de Stephen Wolfram, el creador de la popular aplicación de matemática Wolfram Alpha, quien recientemente expuso en una conferencia de TED una visión metafísica de que es posible reducir la estructura subyacente del universo a términos matemáticos (Wolfram, 2023).

Cuando uno entiende esto, comprende también inmediatamente que la capacidad de los sistemas de IA para generar conocimiento novedoso u original, y no meramente basado en la frecuencia estadística de conocimiento producido de manera usual en universidades y otras comunidades de argumentación, es ciertamente relativa.

4. CONCLUSIONES

Como podemos ver, las distintas aplicaciones de inteligencia artificial tienen una gran capacidad para aprovechar el conocimiento tanto teórico como práctico que produce una sociedad compleja y, en particular, las universidades y otros sistemas de educación formal o de cultura política. No obstante, por su propia naturaleza, estos sistemas poseen claras limitaciones. Entre aquellas que ya he mencionado: la IA, que está basada en técnicas y

procedimientos de índole lógico-matemático, no puede modelar por sí sola, y sobre todo sin la ayuda de etiquetadores humanos, la realidad. Los seres humanos tenemos una capacidad que todavía no se comprende muy bien de crear modelos de realidad que relacionan entre sí conceptos que se aplican a ámbitos distintos y que crean un modelo heurístico del conocimiento que es universal, verdadero y definido causalmente, es decir, no simplemente se limita a establecer correlaciones (Lonergan, 2005, Spitzer, 2015).

La capacidad de tomar decisiones racionales individuales y comunicarlas a un grupo que puede hacerlas suyas, lo que llamaba el filósofo alemán Martin Heidegger (1927) nuestra condición de ser entes que tienen una ontología, una capacidad que desde la antigüedad Aristóteles entiende como el objeto propio de la metafísica que estudia nuestros modos de decir el ser, de definir o clasificar algo como algo, *es* lo que los sistemas de IA no pueden o no saben modelar.

Ese tipo especial de conocimiento general, que intenta implementarse en redes de inteligencia artificial *general* o GAI, no se puede mapear con un sistema de IA, aunque es verdad que se están haciendo ahora enormes esfuerzos en esta dirección. Por esta razón, en primer lugar, toda universidad que se respete y, en especial, nuestra Universidad Central de Venezuela, debería hacer esfuerzos para proteger la docencia y la investigación que se aboca a definir, producir, preservar y transmitir verdadero conocimiento, conocimiento de frontera y conocimiento con valor emancipatorio.

En segundo lugar, nuestra universidad debe atender los desafíos y peligros que acechan a la sociedad cuyos espacios y discusiones públicos han sido, pudiéramos decir, confiscados por redes sociales online que han sustituido la aspiración de alcanzar acuerdos en pos de la constitución de una voluntad política común, que dota de legitimidad a las decisiones que se toman en los poderes públicos, por cámaras de resonancia de grupos fragmentados que suelen terminar cocinándose en su propia salsa. Para ello, la universidad debe ofrecer una oportuna y completa asesoría en relación con el manejo y comprensión de los modos de operación de las redes sociales y proponer a los poderes legislativos principios jurídicos que protejan los datos de usuarios y su uso indebido por parte de los que gestionan una red, teniendo en vistas la promoción de una educación ciudadana que procure acuerdos mínimos de carácter razonable (es decir, no motivados por el extremismo provocado por las redes) alrededor de temas de interés común.

En tercer lugar, debería también ayudar a los cuerpos legislativos a aprobar leyes que impidan que empresas públicas y privadas que crean distintas aplicaciones violen los derechos humanos y las protecciones constitucionales de sus usuarios finales, como por ejemplo la prohibición de discriminación por etnicidad, credo o posición política.

En cuarto lugar, una sociedad responsable debe también controlar la creación de armas de guerra informadas por sistemas de IA y, más importante aún, la promoción de conflictos o confrontaciones cuya única motivación está basada en las conclusiones a las que ha llegado un modelo de IA. Esto último viene sucediendo mucho desde hace tiempo, a saber, el uso que terceros interesados hacen de la información que ofrecen las redes para pagar a “influencers” que explotan una diferencia entre grupos humanos con fines bélicos o de confrontación política, ahondándolas aún más e impidiendo o volviendo imposible un

encuentro cordial de opiniones encontradas. Todos los que en redes sociales o chats de whatsapp nos vemos bombardeados una y otra vez por los mismos mensajes machacones por parte de un mismo individuo o grupos de individuos, que la literatura científica sobre el tema llama “zelotes” (Mobilia, 2003), tenemos que aprender a no asumir tan de prisa que este tipo de paranoia en la red no se hace por una suma de dinero o a cuenta de intereses de terceros.

En definitiva, se trata de adaptarnos al aumento de nuestro poder computacional y de aprender a manejar conjuntos de datos con miles de parámetros, sin olvidar que la verdad es algo muy distinto que sólo puede cultivarse en un entorno comunicativo verdaderamente autónomo, distendido, abierto a la creatividad y a la realidad entendida como un todo, libre de manipulaciones y deformaciones opacas o que no podamos comprender bien, y que la ciencia y la filosofía desde hace tiempo han mostrado que converge (ese entorno comunicativo) hacia la razonabilidad, la cordialidad y la concordia en las opiniones.

REFERENCIAS

- Anderson, P. (1972). More is different. *Revista Science*, Vol. 177. 4047.
- Barabási, A. L. (2018). *The Formula*. Hachette Book Group.
- Barreto, L.M. (2018). *Estudio de dos paradigmas de modelado de tópicos aplicados sobre un conjunto de datos tomados de una red social*. Tesis de Maestría en Ciencias de la Computación. Universidad Central de Venezuela. <http://saber.ucv.ve/handle/10872/20186>
- Barreto, L.M. (2022). Rational consensus and deliberative democracy in complex societies dominated by online social media interactions. *Revista Episteme* NS, 42. http://saber.ucv.ve/ojs/index.php/rev_ens/article/view/26591
- Barreto, L.M. (2023 a). Una escuela alemana. Los orígenes de la Escuela de Filosofía de la UCV. *Revista Trópico Absoluto. Revista de Crítica Pensamiento e Ideas*. <https://n9.cl/4whs1>
- Barreto, L.M. (2023b). Sobre la tal vez imposible universalidad del conocimiento. *Revista Trópico Absoluto. Revista de Crítica Pensamiento e Ideas*. <https://n9.cl/4jlr8>
- Barreto, L.M. y Núñez, H. (2020). Una exploración de la semántica del concepto de ‘Derechos’ en Twitter. *Revista Apuntes Filosóficos*, 30(48). http://saber.ucv.ve/ojs/index.php/rev_af/article/view/23753/144814489981
- Caldarelli, G. (2007). *Scale-free Networks. Complex Webs in Nature and Technology*. Oxford University Press.
- Castellano, F. y Loreto (2009). Statistical physics of social dynamics. *Rev. Mod. Phys.* The Americal Physical Society, No. 81.
- D’Souza, R.M. (2015). *The Mathematics of Influence and Opinion Analysis*. *You Tube*, 2015, Canal @UCDavisSocialSciences.
- Einstein, A. (1905). Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen Der Physik*, 322:8, 549-560. <https://onlinelibrary.wiley.com/doi/10.1002/andp.19053220806>
- Gilson, É. (2004, orig. 1948). *La unidad de la experiencia filosófica*. Rialp.

- Gleick, J. (2011). *The information: a history, a theory, a flood*. Pantheon Books, 2011.
- Granovetter, M.S. (1973). The Strength of Weak Ties. *American Journal of Sociology*, 78:6, 1360-1380.
- Heidegger, M. (orig. 1927). *Ser y tiempo*. Fondo de Cultura Económica.
- Kissinger, H., Schmidt E. y Huttenlocher, D. (2022). *The Age of AI*. John Murray Publishers.
- Lonergan, B. (2005). *Insight, a study on human understanding*. University of Toronto Press.
- Lucas, J. R. (1961). Minds, Machines and Gödel. *Philosophy*, XXXVI.
- Mobilia, M. (2003). Does a Single Zealot Affect an Infinite Group of Voters? *Phys. Rev. Lett.* 91, 028701. <https://doi.org/10.1103/PhysRevLett.91.028701>
- Perra, N., Rocha, L.E.C. (2019). Modelling opinion dynamics in the age of algorithmic personalisation. *Scientific Reports* 9, 7261 <https://doi.org/10.1038/s41598-019-43830-2>
- Popper, K., (1963). *Conjeturas y refutaciones. El desarrollo del conocimiento científico*. Paidós Básica.
- Shannon, C. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal*, 27(3).
- Spitzer S.J., Robert J. (2015). *The Soul's Upward Yearning*, Ignatius Press. San Francisco.
- Strawson, P.F. (1992). *Analysis and Metaphysics*. Oxford University Press.
- Wittgenstein, Ludwig, (2021, orig. 1939). *Investigaciones filosóficas*. Editorial Trotta.
- Wolfram, S. (2023). *How to Think Computationally About AI, the Universe and Everything*. TED, Youtube: <https://www.youtube.com/watch?v=fLMZAHyrpyo&t=423s>.

Luz Marina Barreto. Profesora Titular y Jefa del Departamento de Filosofía Teórica de la Escuela de Filosofía de la UCV. Doctora en Filosofía de la Universidad Libre de Berlín, Magister Scientiarum en Ciencias de la Computación de la Universidad Central de Venezuela, Diplomada en Teología de la UCAB y Licenciada en Filosofía de la Universidad Central de Venezuela. Directora del Instituto de Filología “Andrés Bello” de la Universidad Central de Venezuela y Directora del Doctorado de Filosofía de la misma universidad.



Todos los contenidos de esta revista se distribuyen bajo una licencia de uso y distribución “**Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional**”. Puede consultar desde aquí la [versión informativa](#) y el [texto legal](#) de la licencia. Esta circunstancia ha de hacerse constar expresamente de esta forma cuando sea necesario.