



Volumen 6
año 2024

“Semiótica, vida digital e Inteligencia Artificial”

Artículos Temáticos

DOI 10.30972/nea.617581

Artículo recibido 10 de enero del 2024
Aprobado 10 de abril de 2024

Cómo citar este artículo: Gastaldello, D. (2024). Semiótica e I+D de la IA. El salto hacia la CA (conciencia artificial). *Ñeatá. Revista digital del Grupo de Estudios Semio-discursivos (GESEM, SGCyT-UNNE)*, 6, pp. 1-11. <https://doi.org/10.30972/nea.617581>

Semiótica e I+D de la IA. El salto hacia la CA (conciencia artificial)

Semiotics and R&D of AI. The leap towards AC (artificial consciousness)

Gastaldello, Daniel Clemente
clemente.gastaldello@gmail.com
Universidad Nacional del Litoral

Es Doctor en Semiótica por la Universidad Nacional de Córdoba, titular de las cátedras de Semiótica General y de los Seminarios de Semióticas Específicas en la Universidad Nacional del Litoral, y se desempeña como docente de cursos de posgrado del Doctorado en Artes (Universidad Nacional de Córdoba). Asimismo, es director del Grupo de Investigación en Semiótica (Universidad Nacional del Litoral), director de la colección editorial Estudios Semióticos (Ediciones UNL), miembro de la Comisión directiva de la Asociación Argentina de Semiótica y, también, Investigador Categorizado III Especializado en Semiótica e Inteligencia Artificial.

Resumen

A pesar de los avances que nos circundan en materia de IA, desde diagnósticos médicos hasta la vigilancia detallada, los algoritmos plantean limitaciones de la IA para pensar de manera independiente (conciencia artificial), lo que subraya la necesidad de comprender sus futuros procesos de abducción y enfocarse en los datos, modelos y personas para cultivar una inteligencia artificial más creativa y autónoma, materia de la que nos ocuparemos en las notas que siguen.

Palabras clave

Inteligencia artificial - conciencia artificial - semiosis - abducción

Abstract

Despite the surrounding advancements in AI, from medical diagnoses to detailed surveillance, algorithms pose limitations for AI to think independently (artificial consciousness), highlighting the necessity to comprehend its future processes of abduction and focus on data, models, and individuals to foster a more creative and autonomous artificial intelligence, a subject we will delve into in the following notes.

Keywords

Artificial Intelligence - Artificial Consciousness - semiosis - abduction

Y debido a que en toda la Galaxia no habían encontrado nada más precioso que la mente, alentaron por doquier su amanecer. Se convirtieron en granjeros en los campos de las estrellas; sembraron, y a veces cosecharon.

2001: A Space Odyssey

Stanley Kubrick, 1968

En este artículo se examinan las limitaciones que encuentra la IA desde el punto de vista del análisis semiótico, en el marco de corrientes filosóficas contemporáneas que presuponen que la tecnología mejorará la condición humana y marcará el progreso futuro de la especie. La semiótica enfrenta el reto de la IA al centrarse en cómo el manejo masivo de datos y las respuestas abductivas al mundo generan significado en el contexto humano, no en la mera interpretación a través de software o algoritmos. Aunque el enfoque de la IA es esclarecedor, la semiótica resulta mucho más pertinente para comprender la cognición humana, ya que explora los signos como constitutivos de la inteligencia, no como representaciones artificiales de la misma.

1. Antecedentes y estado actual de la I+D en IA

Entre las décadas de 1950 y 1960 se dio forma al campo de la inteligencia artificial como una investigación independiente. Alan Turing estableció los fundamentos para evaluar la inteligencia de las máquinas con su icónico Test de Turing, cuestionando si una máquina podría demostrar un comportamiento inteligente similar al humano. En 1956, John McCarthy acuñó el término “inteligencia artificial” y lideró la Conferencia de Dartmouth, donde se reunieron expertos de diversas áreas del conocimiento para explorar la posibilidad de creación de máquinas inteligentes. Creemos que, durante este período, se desarrollaron los primeros programas de inteligencia artificial enfocados en resolver problemas a través del razonamiento lógico-deductivo con resultados promisorios. Pero este mismo razonamiento lógico y sus resultados arribaron, paradójicamente, a una de las primeras crisis de la investigación en inteligencia artificial. Tal como lo explica Erik Larson (2021):

After the failure of what critics dubbed ‘good old-fashioned artificial intelligence’, which dominated AI before the modern era (up through the 1990s), AI scientists abandoned deductive approaches to inference en masse. Indeed, many younger readers might find it strange that something like ‘rules’ and deductive approaches to AI were ever taken seriously by practitioners in the field. (p. 115)¹

Frente a estas restricciones, los expertos en ciencia comenzaron a investigar distintas vías, desde el empleo de redes neuronales hasta el aprovechamiento extensivo de datos para enseñar a

1 “Después del fracaso de lo que los críticos denominaron ‘la inteligencia artificial buena y antigua’, que dominó la IA antes de la era moderna (hasta los años 90), los científicos en IA abandonaron en masa los enfoques deductivos para la inferencia. De hecho, muchos lectores más jóvenes podrían encontrar extraño que algo como ‘reglas’ y enfoques deductivos para la IA alguna vez fueran tomados en serio por los profesionales en este campo.”

las máquinas a desempeñar tareas específicas con mayor exactitud, campo también reconocido como “aprendizaje automático” o *machine learning*. Esta fusión entre el análisis de datos y el razonamiento lógico posibilita resolver desafíos que demandan tanto conocimiento explícito como la habilidad de identificar patrones a partir de ejemplos. A partir de la década de 2010, el “aprendizaje profundo” (*deep learning*) tomó una forma de redes neuronales profundas, transformando el desarrollo de la IA al lograr avances considerables en el procesamiento e interpretación de imágenes y el procesamiento del lenguaje natural. Pero Larson volvió sobre la importancia de la abducción en este proceso:

Since the three known types of inference are not reducible to each other but are distinct, and abductive inference is required for general intelligence, purely inductively inspired techniques like machine learning remain inadequate, no matter how fast computers get (...) In open-ended scenarios requiring knowledge about the world like language understanding, abduction is central and irreplaceable. (Larson, 2021, p. 231)²

Porque para pensar de manera inteligente es necesario *Actuar-humanamente*, y esto implica tener en cuenta la complejidad semiótica de los datos como así también la creatividad como un subproceso mental que opera con esos datos. En las líneas que siguen abordaremos estos dos temas desde los aportes de la semiótica.

| 2. Datos, dispositivos y semiosis

Debemos distinguir a la IA tanto como un producto como un proceso. En medio de dicha definición, podríamos acordar que se trata de un *dispositivo* siguiendo la definición clásica de Traversa:

(...) si hay discurso, se pone en juego una materialidad sígnica y una técnica para producirla, un espacio asimétrico, que solo el circunstancial actor puede suponer homogéneo. El analista, para aprehenderlo en su movimiento, debe apelar a una distancia, que se puede patentizar en la descripción de entidades que podemos designar como dispositivos. (Traversa, 2001, p. 47)

Esto merece un extenso capítulo sobre la naturaleza material de los datos (que resumiremos más adelante), pero en esta instancia sólo queremos dejar asentado que no abordaremos en detalle pese a que no desconocemos la problemática en la que se inscribiría la discusión sobre la ontología de los datos que la IA pone a circular. Por el momento nos basta con consignar que reconocemos la doble naturaleza de producto y proceso de la IA, como así también la compleja trama de los procesos que implica.

Como podemos notar, el campo de investigación y de acción de la IA es amplio, y abarca desde la definición conflictiva de su ontología hasta el desarrollo de algoritmos específicos para tareas puntuales. Todas ellas están abocadas a comprender a la vez que a poner en funcionamiento entidades inteligentes. Estos subcampos de desarrollo se basan en datos y generan datos. Esto es, son dispositivos

2 “Dado que los tres tipos conocidos de inferencia no son reducibles entre sí sino distintos, y la inferencia abductiva es necesaria para la inteligencia general, técnicas puramente inductivas como el aprendizaje automático siguen siendo insuficientes, sin importar cuán rápidas sean las computadoras (...) En escenarios de naturaleza abierta que requieren conocimiento sobre el mundo, como la comprensión del lenguaje, la abducción es central e irremplazable.” *Ob. cit.*

que, en términos semióticos, generan y hacen circular sentidos. Lo que comúnmente se conoce como *machine learning*, sería un subcampo de las ciencias de la IA cuyo objetivo es desarrollar técnicas que permitan que los sistemas computacionales aprendan. Se entiende que un “agente” aprende cuando su desempeño mejora con la experiencia y mediante el uso de datos, es decir, cuando la habilidad no estaba presente en su genotipo o rasgos de nacimiento.

En el aprendizaje de máquinas un computador observa datos, construye un modelo basado en esos datos y utiliza ese modelo a la vez como una hipótesis acerca del mundo y una pieza de software que puede resolver problemas. (Russell y Norvig, 1995, p. 651)

Estos datos son, en términos semióticos, una materialidad significativa: pueden ser imágenes, sonidos, texto lingüístico, redes, tablas o cualquier representación que se nos ocurra. Son lo que Verón llamó, a fines del siglo XX, “paquetes de sentido”:

Toda producción de sentido, en efecto, tiene una manifestación material. Esta materialidad del sentido define una condición esencial, el punto de partida necesario de todo estudio empírico de la producción de sentido. Siempre partimos de ‘paquetes’ de material sensible investidos de sentido que son productos; con otras palabras, partimos siempre de configuraciones de sentido identificadas sobre un soporte material (texto lingüístico, imagen, sistema de acción cuyo soporte es el cuerpo, etcétera) que son fragmentos de semiosis. (Verón, 1987, p. 126)

Por ende, estos datos que se tramitan en el *machine learning* son parte, fragmentos de una semiosis, y están necesaria y previamente investidos de sentido.

Cuando hablamos de producción de sentido hablamos de cosas que remiten a materialidades, en el modo más raso en que se puede entender ese término, es decir, cosas que están en el mundo al alcance de nuestros analizadores biológicos. Cada uno de ellos, a partir de los estímulos que les sirven de fuente, nos habilita para detectar múltiples diferencias, distinguiéndose rangos de procesos cognitivos en los que intervienen, los cuales no se sitúan en espacios sociales ni son portadores de jerarquías homogéneas. (...) Esta materialidad se extiende hasta lo recóndito de la producción onírica, la que se libra al mundo en la palabra del soñante –unos sonidos o unos trazados gráficos– al alcance de los oídos o los ojos de alguien, que los retoma luego en otra palabra. (...) Si algo hace sentido, entonces, lo hace en tanto se integra en una sucesión de un fragmento previo de semiosis, suerte de estaciones, donde una producción implica un reconocimiento a través de un paquete de materias sensibles (...) los que se derivan de configuraciones distintas de aplicación de la misma técnica, que potencialmente se abre sobre campos de producción de sentidos diversos. (Traversa, 2001, pp. 234-237)

La pregunta que debemos hacernos entonces en esta instancia es si el dato/materialidad significativa que aprende la IA contiene este fragmento de sentido o si está desnaturalizado, desencajado, alienado de esa semiosis inicial. En la educación humana es sencillo verlo (cfr. Ana Camblong y Froilán Fernández cuando estudiaron la alfabetización semiótica en las fronteras³). Pero los algoritmos de aprendizaje

3 “(...) nuestro discurso no está exento de controversias sociales e históricas; no habla desde arriba ni desde afuera, nuestro pensamiento asume su inscripción temporal y su modelización situada.” (Camblong, 2014, p. 1) y “el proceso

automático están, paradójicamente, diseñados en una contradicción básica: por un lado, están escritos por especialistas en la materia (sospechamos, suponemos, que hay expertos en educación detrás de los algoritmos que ayudan a hacer eficientes los programas, llevando saberes de un área de conocimiento como es el aprendizaje humano al universo de las máquinas, a los fines de agilizar procesos e incluso abaratar costos) y por otro lado, se aspira a un modelo de la enunciación antropomórfico (se tiene por objetivo lograr emular las formas de enunciar del ser humano, como si la verdad se lograra en y desde la voluntad enunciativa del ser humano y no hubiera otras formas de inteligir y de confrontarse con el mundo, como si los datos y técnicas que pasan por el lenguaje fueran lo único que existiera e importara). Esta contradicción, o mejor dicho, esta paradoja de sujetos que enseñan a máquinas a replicar el mundo tal como lo modela su lenguaje ya ha dado sus advertencias en los llamados “sesgos” de programación, donde la ideología inadvertida y naturalizada se muestra en resultados inesperados. Más adelante veremos que los sesgos se inscriben en una lógica inductiva bajo la cual son creados los algoritmos de programación. Una vez más, la verdad se percibe arraigada en el objeto, más que en la semiosis que le confiere sentido, ya sea este posible, contingente o transitorio, entre otros significados que un dato procesado mediante alguna técnica puede adoptar. Tal como lo sostuvo Keith Darlington:

esta comprensión es necesaria porque aunque un programa de IA sea capaz de resolver un problema con un nivel de competencia equiparable al de un humano, es posible que carezca de un contexto más amplio para juzgar el valor de su solución. Dicho de otro modo, en la actualidad estos sistemas son incapaces de entender el significado... (Darlington, 2021, p. 1)

En síntesis, podríamos decir que el estado actual de avance de la IA que éste intenta emular al lenguaje, no al dato, a la cosa misma que es sometida a un proceso de transformación. Y no al lenguaje en su completud acaso enciclopédica, o en lo paragramático (Kristeva, 1969a) sino apenas en lo sintagmático: en la definición única, esto es, en el significado o en el sentido denotado (ni siquiera en el sentido connotado, que implicaría una semiosis mayor, un fragmento de la cultura en un tiempo y un espacio dados). Los datos son fundamentales en el entrenamiento de sistemas de aprendizaje automático, pues son la base para determinar la precisión de las conclusiones. Es crucial comprender que estos sistemas no están diseñados únicamente para hacer predicciones con los datos de entrenamiento, sino que se espera que puedan generalizar y ofrecer conclusiones precisas incluso en situaciones no vistas previamente. La capacidad de generalización es esencial para que los modelos sean útiles más allá de contextos conocidos. En general, la premisa clave de estos sistemas es que los datos de prueba serán similares pero no idénticos a los datos de entrenamiento. No es suficiente crear bases de datos representativas o modelos justos desde un punto de vista específico. Los sistemas de IA son desarrollados por personas con sus propias perspectivas, prejuicios y sesgos que pueden influir en la creación y evaluación de estos modelos. Si los equipos de trabajo carecen de diversidad para reflejar diversas perspectivas, es probable que ni siquiera identifiquen la presencia de sesgos, por lo que corregirlos podría ser difícil de lograr.

El objetivo no es detener el progreso de la inteligencia artificial como campo de I+D, sino que tanto los usuarios como los desarrolladores sean conscientes de sus limitaciones. Creemos fundamental tener en cuenta estas limitaciones al concebir y usar estas tecnologías.

alfabetizador puesto (...) en su propia constitución abarca, involucra y compromete todos los signos, no sólo el lenguaje; a la inversa, el lenguaje permanece tramado en los flujos semióticos continuos que lo sustentan y le dan sentido.” (Camblong y Fernández, 2012, p. 10)

3. La IA y el método abductivo

La IA puede instituirse, con todo derecho, como un método nuevo de investigación impactante y actualizado sobre la cognición humana, con implicaciones evidentes para las teorías semióticas, las prácticas y la representación mental. No obstante, hasta el momento, parece que ha recibido escasa atención analítica por parte de los semiólogos. Parecería ser que el propósito fundamental que nos ofrece la IA es representar y comprender en profundidad diversas facetas de la mente humana (como la percepción, la cognición y las emociones) e incluso crear mentes artificiales que puedan superar la inteligencia humana en un futuro no tan lejano. Una perspectiva semiótica podría enriquecer la I+D de la IA al ofrecer una comprensión relevante de cómo piensa semióticamente el ser humano, cuestionando posibles intentos de replicarla.

Deberíamos desentrañar algunas nociones en tanto semiólogos que nos permitan, a partir de nuestro campo de conocimiento específico, hacer algunos aportes a este problema de los alcances de la IA en lo que respecta al pensamiento creativo, trayendo al debate a otras disciplinas que se han abocado al mismo problema, como las ciencias de la comunicación, la filosofía y los estudios culturales. Las nociones de inteligencia y de IA ya fueron suficientemente elaboradas y discutidas por autores como Norvig y Russell (1995):

The field of artificial intelligence, or AI, goes further still: it attempts not just to understand but also to build intelligent entities (...) a huge variety of subfields, ranging from the general (learning and perception) to the specific, such as playing chess, proving mathematical theorems, writing poetry, driving a car on a crowded street, and diagnosing diseases. AI is relevant to any intellectual task; it is truly a universal field. (Norvig y Russell, 1995, p. 1)⁴

Cuando los autores ingresan en el detalle de la noción de IA realizan una segmentación de acciones y aspectos: por un lado, encontramos una diferencia entre los verbos pensar/actuar como humanos (lo que señala una diferencia notoria entre ser/parecer) y, por otro lado, la precisión de esas acciones en adverbios tales como humanamente/racionalmente (lo que nos indica que lo humano y lo racional no tienen por qué ser sinónimos, sino por el contrario, pueden oponerse y marcar un rumbo diferente de interpretación de acciones):

- *Pensar-humanamente* se refiere al esfuerzo de crear computadoras que puedan pensar como los humanos, desarrollando máquinas con mentes completas y literales. Este enfoque abarca la automatización de actividades asociadas con el pensamiento humano, como la toma de decisiones, la resolución de problemas y el aprendizaje. “El emocionante esfuerzo por hacer que las computadoras piensen... máquinas con mentes, en el sentido completo y literal” (Haugeland, 1985); “La automatización de actividades que asociamos con el pensamiento humano, actividades como la toma de decisiones, resolución de problemas, aprendizaje...”

4 “El campo de la inteligencia artificial, o IA, va más allá: no solo intenta comprender, sino también construir entidades inteligentes. (...) Aborda una amplia variedad de subcampos, desde aspectos generales como el aprendizaje y la percepción, hasta actividades específicas como jugar ajedrez, demostrar teoremas matemáticos, escribir poesía, conducir en calles congestionadas y diagnosticar enfermedades. La IA es relevante para cualquier tarea intelectual; es verdaderamente un campo universal”. *Ob. Cit.*

(Bellman, 1978).

- *Pensar-racionalmente* implica el estudio de las facultades mentales utilizando modelos computacionales. Este enfoque se centra en analizar los procesos mentales para comprender cómo se percibe, razona y actúa el ser humano, utilizando herramientas matemáticas y de ingeniería para comprender estos procesos cognitivos. “El estudio de las facultades mentales a través del uso de modelos computacionales” (Charniak y McDermott, 1985); “El estudio de los cálculos que hacen posible percibir, razonar y actuar” (Winston, 1992).
- *Actuar-humanamente* se refiere al arte de crear máquinas capaces de realizar funciones que normalmente requieren inteligencia humana. Este enfoque se concentra en diseñar sistemas artificiales que puedan llevar a cabo tareas que los humanos realizan, centrándose en la replicación de la inteligencia humana con tecnologías. “El arte de crear máquinas que realizan funciones que requieren inteligencia cuando son realizadas por personas” (Kurzweil, 1990); “El estudio de cómo hacer que las computadoras realicen tareas en las que, en este momento, las personas son mejores” (Rich y Knight, 1991).
- *Actuar-racionalmente* implica el estudio del diseño de agentes inteligentes o artefactos que pueden tomar decisiones y realizar acciones de manera racional, utilizando la inteligencia computacional para lograr comportamientos considerados racionales en función de los objetivos y conocimientos disponibles. “La Inteligencia Computacional es el estudio del diseño de agentes inteligentes” (Poole et al., 1998); “La IA... se ocupa del comportamiento inteligente en artefactos” (Nilsson, 1998).

Ahora bien, ¿dónde creemos que radican las principales limitaciones de la IA una vez que entendemos que los sesgos son superados y los datos son incluidos en semiosis donde adquieren sentidos mayores?

Para responder a esta pregunta debemos situarnos en otro contexto teórico primero, y entender cómo funciona la mente cuando razona. Siguiendo la tipología de Charles S. Peirce (1903) sobre los tipos de razonamiento, identificamos que se encuentra en la abducción. Peirce habla de tres tipos de razonamiento: la deducción (silogismo), la inducción y abducción:

- La *deducción* es un proceso de razonamiento en el que llegamos a una conclusión específica a partir de premisas generales o conocimientos previos.

En la deducción, o razonamiento necesario, partimos de un estado hipotético de cosas que definimos en ciertos aspectos abstractos. Entre las características a las que no prestamos atención alguna en este modo de argumento, está la de si la hipótesis de nuestras premisas se conforma, o no, más o menos al estado de cosas del mundo exterior. (Peirce, 1903, párr.11)

En este tipo de razonamiento se utiliza un patrón lógico en el que si las premisas son verdaderas, la conclusión necesariamente también lo es.

- La *inducción*, en cambio, es un proceso de razonamiento en el que se infiere una generalización o patrón a partir de observaciones o datos concretos.

La inducción consiste en partir de una teoría, deduciendo de la misma predicciones de los fenómenos, y observando estos fenómenos con vistas a comprobar lo aproximadamente que concuerdan con la teoría. La justificación para creer que una teoría experiencial que ha sido sometida a un determinado número de test experimentales quedará también corroborada en el futuro cercano como lo ha sido hasta ahora por otros test como éstos, es la de que mediante la prosecución constante de este método a la larga tenemos que averiguar el estado real de la cuestión. (Peirce, 1903, párr. 20)

A diferencia de la deducción, la conclusión en la inducción no es necesariamente verdadera, sino que se basa en una probabilidad y puede cambiar cuando se encuentra un caso o dato que no concuerda con la teoría. Una parte importante de los sesgos que sostienen las inteligencias artificiales provienen de esta lógica inductiva de funcionamiento.

- La *abducción*, finalmente, es un proceso de razonamiento en el que el investigador sugiere una explicación posible para un fenómeno, arriesgando una hipótesis a partir de datos:

La abducción es el proceso de formar una hipótesis explicativa. Es la única operación lógica que introduce alguna idea nueva; pues la inducción no hace más que determinar un valor, y la deducción desarrolla meramente las consecuencias necesarias de una pura hipótesis. La deducción prueba que algo tiene que ser; la inducción muestra que algo es actualmente operativo; la abducción sugiere meramente que algo puede ser. Su única justificación es la de que a partir de su sugerencia la deducción puede extraer una predicción que puede comprobarse mediante inducción, y que, si podemos llegar a aprender algo o a entender en absoluto los fenómenos, esto tiene que conseguirse mediante la abducción. Por lo que puedo saber, no puede darse ninguna razón cualquiera que sea a favor de la misma; y no necesita razón alguna, ya que sólo ofrece sugerencias. (Peirce, 1903, párr. 21)

Este procedimiento implica un proceso creativo de deducción que busca formular nuevas explicaciones hipotéticas: a partir de la inferencia de ciertos datos, leyes y suposiciones, se derivan proposiciones creíbles. Esta acción puede considerarse como el resultado obtenido (lo inferido) o como un procedimiento que construye o elige hipótesis respaldado por la inducción.

Si seguimos el razonamiento de las últimas palabras del autor, veremos que se trata de un tipo de razonamiento inferencial que se emplea cuando se enfrenta a circunstancias desconocidas o ambiguas. En ellas la deducción y la inferencia son imposibles de realizar no por ser inexactas, sino porque no aportan ningún dato nuevo o información que no se tuviera *a priori* del experimento.

Recordemos que un avance significativo en la I+D de la IA consistiría en crear una inteligencia autónoma, “El arte de crear máquinas que realizan funciones que requieren inteligencia cuando son realizadas por personas” (Kurzweil, 1990). El arte o creatividad han sido asumidos adoptando una postura completamente antropocéntrica como uno de los rasgos característicos de la IA. Esta noción la

concibe como una forma avanzada de resolver problemas de manera veloz o en cantidades extensas de datos, mientras que, en realidad, la creatividad debe generar relaciones novedosas entre distintos tipos de conocimiento, dado que la IA es tanto un proceso como un producto. Debido a que comúnmente se le ha vinculado a procesos artísticos se le ha atribuido una cualidad que tiende hacia lo subjetivo, lo que la ha calificado como no reproducible y difícil de formalizar.

Según Larson “la IA carece de una teoría fundamental: una teoría de la inferencia abductiva” (Larson, 2021, p. 189). Y según lo expresado no hace mucho tiempo por los semiólogos canadienses Stéphanie Walsh Matthews y Marcel Danesi: “es poco probable que la IA pase alguna vez una prueba de abducción, ya que las inferencias son impredecibles y se basan en experiencias corporales” (Walsh y Danesi, 2019, p. 210).

Estas opiniones subrayan la necesidad de un enfoque más holístico para el desarrollo de IA, reconociendo la importancia de comprender las inferencias abductivas y las experiencias corporales en el proceso de toma de decisiones de la misma.

4. Apertura a nuevas ideas y debates

Retomando la reflexión de Carlos Scolari (2023), las limitaciones de los sistemas actuales de IA tienen su raíz en un problema lógico-semiótico, dificultando la reproducción de las complejas operaciones del cerebro humano. Actualmente, nos apoyamos en máquinas inductivas poderosas, impulsadas por grandes conjuntos de datos y el creciente poder de procesamiento de computadoras, mientras que la IA busca extraer conclusiones. Una comprensión más profunda de la semiótica humana, aportada por la agenda semiótica, podría desafiar los intentos de modelar completamente la IA, no sólo trayendo la abducción a la escena de la programación de los algoritmos, incluso como una forma de razonamiento práctico que puede apoyar inclusive la implementación de la cognición distribuida. También la noción de semiosis propuesta por Peirce podría colaborar en un salto significativo en la I+D de la IA. Tal vez la gran masividad de datos pueda así elaborar sentidos en su diseño de una IA más eficiente y éticamente orientada. Asumir esta perspectiva se vuelve pertinente para comprender las limitaciones de enfoques reduccionistas que se centran únicamente en la psicología o en la función de la inteligencia como entidad en sí misma. Creemos que, desde esta posición, podríamos expandir la noción de “inteligencia” hacia otras prácticas y operaciones y evaluar hasta qué punto podemos considerar a la “inteligencia artificial” como una forma de inteligencia genuina. O tal vez deberíamos habilitar una nueva categoría, la de *conciencia artificial*, un estado de la intelección nuevo que pueda asignar a la IA la experiencia subjetiva. Si bien para los neurocientíficos esta experiencia es imposible de ser asumida por una IA, debido a que no guardan una conexión profunda y sensorial con el contexto, ni tienen el sistema tálamo-cortical (un componente vital para la conciencia en mamíferos), ni cuentan con la compleja arquitectura y funcionamiento de las neuronas y sus interacciones físicas, químicas y biológicas... estamos lejos aún de comprender la mente en su plenitud. Seguramente serán varias disciplinas (filosofía, ciencias de la información, semiótica, medicina, robótica cognitiva, ética... entre muchas otras) las que se aunarán para trabajar juntas en este futuro paso que debe dar la IA para volverse autónoma e independiente de los protocolos de programación, hacia una CA. Dejamos anotado finalmente que, para ello deberá dar un giro en la dirección inductiva en que viene pensando la I+D en la actualidad y pasarse a la abducción tal como la planteó Peirce:

La abducción es el proceso de formar una hipótesis explicativa. Es la única operación lógica que introduce alguna idea nueva; pues la inducción no hace más que determinar

un valor, y la deducción desarrolla meramente las consecuencias necesarias de una pura hipótesis. La deducción prueba que algo tiene que ser; la inducción muestra que algo es actualmente operativo; la abducción sugiere meramente que algo puede ser. Su única justificación es la de que a partir de su sugerencia la deducción puede extraer una predicción que puede comprobarse mediante inducción, y que, si podemos llegar a aprender algo o a entender en absoluto los fenómenos, esto tiene que conseguirse mediante la abducción. (Peirce, 1903, párr. 21)

Sin duda, se espera la colaboración de diversas disciplinas, desde la filosofía hasta la ética, en el camino hacia la autonomía de la IA. Este futuro nos impulsa a cambiar nuestra perspectiva inductiva por una “sugerente” abducción, un proceso que abre el camino para vislumbrar nuevas ideas y nos permite comprender, ya sea superficialmente o en profundidad, nuevos procesos cognitivos que hasta ahora desconocemos. La IA nos recuerda que cada sugerencia tiene el potencial de revelar los secretos que ansiamos comprender del mundo que nos rodea y también de nuestra propia cognición y de nuestra creatividad.

Bibliografía

Bellman, R. (1978). *An Introduction to Artificial Intelligence: Can Computers Think?* Boyd & Fraser Publishing Company.

Camblong, A. y Fernández, F. (2012). *Alfabetización semiótica en las fronteras*. Misiones, UNAM.

Camblong, A. (2014). *Semiótica de fronteras: dimensiones y pasiones territoriales* en www.artes.unne.edu.ar Misiones, UNAM.

Charniak, E. y McDermott, D. (1985). *Introduction to Artificial Intelligence*. Addison-Wesley.

Darlington, K. (2021). “El papel de la consciencia artificial en los sistemas de IA” en *OpenMind BBVA / Apuntes científicos. Top 10: Así fue la ciencia en 2023* URL <https://www.bbvaopenmind.com/tecnologia/inteligencia-artificial/consciencia-artificial-y-sistemas-ia/>

Haugeland, J. (1985). *Artificial Intelligence: The Very Idea*. Boyd & Fraser Publishing Company.

Kristeva, J. (1969a). *Semiótica I. Recherches pour une sémanalyse*. Editions du Seuil Madrid, Editorial Fundamentos.

Kurzweil, R. (1990). *The Age of Intelligent Machines*. Massachusetts, MIT Press.

Larson, E. J. (2021). *The Myth of Artificial Intelligence. Why computers can't think the way we do*. New York, Shackleton Books.

Matthews, S. W. y Danesi, M. (2019). “AI: A Semiotic Perspective” en *Chinese Semiotic Studies*.

Vol 15. Nro 2. Walter de Gruyter GmbH, Genthiner Straße 13. Berlin. <https://doi.org/10.1515/css-2019-0013>

Nilsson, N. J. (1998). *Artificial Intelligence: A New Synthesis*. Morgan Kaufmann Publishers Inc. 340 Pine Street, San Francisco, Sixth Floor.

Norvig, P. y Russell, S. (1995). *Artificial Intelligence: A Modern Approach*. Londres, Pearson. Ed 2020.

Peirce, Ch. S. (1903) “Tres tipos de razonamiento” (Lecciones de Harvard sobre el pragmatismo, Lección VI), C. S. Peirce (1903) Traducción castellana y notas de José Vericat. En: Charles S. Peirce. *El hombre, un signo (El pragmatismo de Peirce)*, J. Vericat (tr., intr. y notas), Crítica, Barcelona, 1988, pp. 123-141. “On Three Types of Reasoning” corresponde a CP 5. 151-179.

Poole, D., Mackworth, A. y Goebel, R. (1998). *Computational intelligence: A logical approach*. Oxford, UK, Oxford University Press.

Rich, E. y Knight, K. (1991). *Artificial Intelligence* (2da ed.). New York, McGraw-Hill.

Scolari, C. (2023). “El mal que aqueja a las IA es la abducción” en *Hipermediaciones*. Agosto 21, 2023. <https://hipermediaciones.com/2023/08/21/el-mal-que-aqueja-a-las-ia-es-la-abduccion/>

Traversa, O. (2001). “Aproximaciones a la noción de dispositivo” en *Signo y Señal*, (12), 231-247. <https://doi.org/10.34096/sys.n12.5612>

Verón, E. (1987). *La semiosis social*. Barcelona, Gedisa.

Winston, P. H. (1992). *Artificial Intelligence* (3era ed.). Londres, Addison-Wesley.