

COGNICIÓN ARTIFICIAL DELEGACIÓN DE INTELIGENCIA EN LA ERA DIGITAL

ARTIFICIAL COGNITION: INTELLIGENCE DELEGATION IN THE DIGITAL ERA

Andrés Pachón

Universidad de Coimbra

.....
Recibido: 26 05 2019

Aceptado: 14 07 2019

Publicado: 30 09 2019
.....

<http://doi.org/10.5281/zenodo.7656610>

Cómo citar este artículo

Pachón, A. (2019). Cognición artificial: delegación de inteligencia en la era digital. *ASRI. Arte y Sociedad. Revista de Investigación en Arte y Humanidades Digitales* (17), 93-107
Recuperado a partir de <http://www.revistaasri.com/article/view/5383>

Resumen

El fenómeno de la Inteligencia Artificial (IA) se presenta en la actualidad como un objeto de estudio relevante para pensar cómo se construye y se distribuye nuestra cognición entre los humanos y las tecnologías, una vez

que la IA modificará, de forma significativa, las formas en las que producimos y accedemos al conocimiento del mundo, al mismo tiempo que mantiene una posición controvertida sobre los límites entre lo humano y lo artificial.

Palabras clave

Inteligencia Artificial, ecologías cognitivas, caja negra, mito del progreso, neuromitología

Abstract

The Artificial Intelligence (AI) phenomenon is currently a relevant case study to understand how cognition is constructed and distributed between humans and technologies, considering that AI will

alter, significantly, the forms in which knowledge is produced and accessed while maintaining a controversial position on the limits between the human and the artificial.

Keywords

Artificial Intelligence, cognitive ecologies, black box, myth of progress, neuromythology.

1. Introducción

El fenómeno de la Inteligencia Artificial (IA) se presenta en la actualidad como un objeto de estudio relevante para pensar cómo se construye y se distribuye nuestra cognición entre los humanos y las tecnologías, una vez que la IA modificará, de forma significativa, las formas en las que producimos y accedemos al conocimiento del mundo, al mismo tiempo que mantiene una posición controvertida sobre los límites entre lo humano y lo artificial. Más allá de producir bienes que cubren la mera necesidad (como es definida la tecnología por la Modernidad (Pfaffenberger 1992)), la actividad sociotecnológica de la IA construirá símbolos materiales que enmarcan el comportamiento humano y que determinan nuestra *ecología cognitiva*¹. De esta forma, la IA se extenderá más allá del ámbito laboral (Sistemas Expertos), encontrándose presente en nuestra vida cotidiana a través de múltiples artefactos que ofrecen soluciones inteligentes y adaptables a cualquier necesidad: teléfonos inteligentes, redes sociales, motores de búsqueda como Google, coches inteligentes, diagnóstico de enfermedades, controles biométricos, asistencia por voz, etc.

El texto comenzará por abarcar la cuestión de la cognición humana como un proceso en construcción, ubicando las características específicas que determinan la ecología cognitiva de nuestra sociedad científica, dentro de la cual se enmarcará la era digital y en particular el nuevo paradigma cognitivo que supone la implementación masiva de la Inteligencia Artificial. A través del análisis del léxico empleado en la actividad sociotecnológica de la IA, centrándonos para ello en la fabricación de modelos de Redes Neuronales Multicapa, se identificarán las articulaciones mediante las cuales delegamos inteligencia en los artefactos tecnológicos donde se implementa la IA, lo cual modificará y determinará la forma en la que vemos el mundo, así como la forma en la que nos entendemos a nosotros mismos y a nuestras tecnologías. El objetivo de este texto es señalar la importancia de abrir las cajas negras² que se

¹ Aquí, la idea de *ecología cognitiva* se refiere al término empleado por Lévy: "Al desarrollar el concepto de ecología cognitiva, defenderé la idea de un colectivo pensante hombres-cosas, un colectivo dinámico poblado de singularidades actuantes y de subjetividades mutantes" (Lévy 1994, 13).

² Aquí, el término *caja negra* se refiere al proceso de *cajanegrizar*, concepto desarrollado por Latour: "Cuando una máquina funciona eficazmente, cuando se deja sentado un hecho cualquiera, basta con fijarse únicamente en los datos de entrada y los de salida, es decir, no hace falta fijarse en la complejidad interna del aparato o del hecho. [...] cuando más se agrandan y

producen con la operatividad (o buen funcionamiento) de las tecnologías de IA, para así conocer cómo se construye nuestra cognición, con el propósito de restar poder a aquellas personas e instituciones que gestionan los mediadores necesarios para que tengamos acceso al conocimiento del mundo. Por último, nos preguntaremos qué alternativas se nos ofrecen para futuras ecologías cognitivas, una vez que la solución no pasa por eliminar las tecnologías de IA, sino por multiplicar los mediadores que dan lugar a su desarrollo, evitando así perpetuar el mito del progreso (Latour 2001) propio de la Modernidad, que persigue el acceso a un conocimiento objetivo separado de la subjetividad de los sujetos.

2. Actividad sociotecnológica y cognición artificial

De acuerdo con Lévy (1994), nuevas temporalidades sociales y modos de conocimiento inéditos y específicos emergerán de la utilización de aquello que el autor denomina *tecnologías de la inteligencia*, como los computadores, la imprenta, la escritura o los métodos nemotécnicos de las sociedades orales. "Al deshacer y rehacer las ecologías cognitivas, las tecnologías intelectuales contribuyen para abalar los grandes pilares culturales que dirigen nuestra precepción de lo real" (Lévy 1994, 12), marcando así las categorías propias de la filosofía del conocimiento, como el mito, la ciencia, la teoría, la interpretación o la objetividad. Por ejemplo, el texto, producido por la tecnología de la escritura, se aísla de las condiciones particulares de su creación y de su recepción, disociando así el conocimiento de las personas que lo producen, separando el emisor del receptor, e imposibilitando la interacción entre ambos. Parafraseando a Lévy (1994), esta comunicación fuera de contexto provocaría que los autores crearan discursos que se bastasen por sí mismos, dando lugar a una ambición teórica y a una pretensión de universalidad. A diferencia del régimen de la oralidad primaria, donde no se disponía de ninguna técnica de registro exterior, más allá de la memoria por repetición que marcaba un tiempo cíclico (el mito), la escritura marcará un tiempo lineal, es decir, una sociedad *histórica* que se caracteriza por una pretensión de objetividad del recuerdo. Esto generó una preocupación por la verdad objetiva, elemento característico de las transmisiones escritas del saber y de la ciencia moderna, lo que modificará para siempre nuestra forma de acceder al conocimiento del mundo.

Según Latour (1998), será la especificidad de algunas aportaciones de la escritura y de las técnicas de visualización lo que determinó el origen y las características específicas de nuestra cultura científica, simples modificaciones que cambiarían los modos en que los grupos humanos discuten entre sí, utilizando para ello papel, signos, impresos, diagramas, etc. La clave de las transformaciones de nuestras sociedades científicas nos dirá Latour (1998), se encuentra en aquellos aspectos que contribuyeron a reunir, presentar, aumentar, favorecer y asegurar la fidelidad de nuevos aliados a la hora de explicar y convencer, para conseguir así el reconocimiento de la autoridad y originalidad del primer autor. Para convencer al oponente de un determinado hecho, el autor tendría que acompañar los textos de objetos que pudiera mostrar, y estos debían ser aliados fiables y estables que demostraran el enunciado. Para ello fue clave la creación de nuevas inscripciones visibles y combinables entre sí, que permitieran trasladar información en el espacio y en el tiempo sin que esta sufriera cambios en el camino. Las técnicas de visualización como la perspectiva, la cámara oscura, los grafismos (tablas,

difunden los sectores de la ciencia y de la tecnología que alcanzan el éxito, tanto más opacos y oscuros se vuelven" (Latour 2001, 362).

diagramas, etc.), los mapas o la fotografía serían estos móviles inmutables (Latour 1998), los cuales podían estar acompañados de etiquetas y leyendas que facilitarían su comparación, superposición y reproducción.

Las instituciones (economía, criminología, pedagogía, psiquiatría, medicina, etc.) que fueran capaces de acelerar la movilidad e intensificar la inmutabilidad de estas inscripciones, reuniendo el mayor número de ellas en categorías homogéneas, serían dotadas de una *consistencia óptica* (Latour 1998) que permitiría instaurar un poder a gran escala, pues son las técnicas de registro y de procesamiento de las representaciones las que posibilitan o condicionan determinadas evoluciones culturales. Podemos afirmar junto a Latour (1998) que la forma en la que contemplamos el mundo es lo que lo hace visible; es decir, una nueva cultura visual, como la instaurada por el nuevo paradigma cognitivo de la IA que será aquí analizado, redefinirá y determinará lo que significa ver y lo que hay que ver, al ofrecer nuevas formas de visualizar y clasificar la información.

Bajo estas premisas, propias de una antropología de la actividad sociotecnológica³, podemos decir que las capacidades cognitivas no se pueden explicar dividiendo mentes primitivas de mentes racionales, o separando culturas precientíficas y científicas, como si estas fronteras (categorías) fueran naturales en lugar de ser construidas. Para entender la evolución (o revolución(es)) científica y técnica, no podemos recurrir a respuestas *mentalistas*, pues el texto, la imprenta o el computador no aportan nada a la mente, sino que generan (y procederán de) una actividad sociotécnica que determinará una cierta ecología cognitiva, donde el estado de verdad solo puede ser visto como la producción contextualizada de actividades particulares. No hay nada especialmente objetivo en la ciencia, nos dirá Latour (2013), si no que se trata de una mediación que genera un tipo particular de transferencia de la información, donde algo es referenciando al mismo tiempo que se proyecta la idea de "una mente calculadora y un mundo calculable" (p. 427).

Esta característica de la mediación científica dará lugar a un *mito del progreso* que perdurará hasta nuestros días, extendiéndose "la certeza de que el pasado diferirá del futuro porque lo que antes se veía confuso aparecerá por fin clara y distintamente a nuestros ojos: la objetividad y la subjetividad dejarán de estar enmadradas" (Latour 2001, 238). Pero, lejos de este punto de vista estándar⁴ sobre el progreso científico y tecnológico, el paso del tiempo ha ido entrelazando cada vez más los sujetos con los objetos. Esto se debe a la multiplicación incesante de técnicas, las cuales no son un medio, sino mediadores (medios y fines al mismo tiempo) que nos hacen intercambiar propiedades con los objetos (Latour 2001). Esto producirá una *cadena de usos* (Lévy 1994) donde cada vez es más difícil encontrar el "primer motor" (Latour 2001, 217) o el "último utilizador" (Lévy 1994, 75), por lo que diferenciar objetos de sujetos en la actividad sociotecnológica se hace una tarea complicada. Los objetos, entre los que se encuentran las tecnologías, compuestas a su vez por otros tantos e innumerables objetos, ya no pueden ser vistos como algo externo a nuestra cognición, separado de nuestras acciones y de nuestra existencia, sino que deben ser entendidos como no humanos, en la medida en la que también actúan y comparten responsabilidades con nosotros. Esto dará lugar a aquello que Latour (2001) denominó

³ Para una guía de estudios antropológicos sobre la actividad tecnológica véase Pfaffenberger. (1992).

⁴ Véase la idea *punto de vista estándar de la Modernidad* desarrollada por Pfaffenberger (1992), la cual se relaciona con las ideas de *mito del progreso y solución moderna* (Latour 2001).

híbrido, un colectivo entre humanos y no humanos, donde la cognición será cada vez más distribuida⁵ entre los sujetos y aquello que ya no puede ser denominado simples objetos. De esta forma,

La acción es una propiedad de entidades asociadas. [...] El atribuir a un actor el papel de primer motor no debilita en modo alguno la necesidad de una composición de fuerzas para explicar la acción. [...] Sencillamente, la acción no es una propiedad atribuible a los humanos sino a *una asociación de actantes* [...]. (Latour 2001, 217 y 218)

Sostener esta actitud nos hará abandonar los binarios de objetividad/subjetividad, humano/no humano, visible/invisible propios de la Modernidad, y con ello nos distanciaremos de las promesas del *mito del progreso*, pues este nos impide ver el carácter mediador de la técnica y, como consecuencia, anula la composición que caracteriza nuestra ecología cognitiva. Por el contrario, debemos multiplicar los mediadores que las corporaciones e instituciones han absorbido y oscurecido en cajas negras. La importancia de abrir estas cajas negras radica en la necesidad de restar poder a aquellas instituciones, o dispositivos⁶, que regularán los artefactos que producen las nuevas formas de acceso al conocimiento. Al propagar la idea de progreso, estos dispositivos determinarán lo que es la realidad, normalizándose nuevas ecologías cognitivas que les permitirán ejercer un poder basado en el conocimiento.

3. La Inteligencia Artificial como paradigma cognitivo en la era digital

Una vez que aceptamos el carácter social del conocimiento podemos entender "las máquinas como un objeto legítimo para un estudio sociológico" (Woolgar, 1985, p. 558), y por tanto como un objeto de estudio cognitivo. Al afrontar con gran ambigüedad la frontera que separa al hombre de la tecnología, el fenómeno de la IA se ofrece como un objeto de estudio que puede ayudarnos a entender en qué medida la cognición y la inteligencia humana han mudado, o están mudando, con la continua expansión de las nuevas tecnologías de la informatización y la digitalización, donde la IA jugará un rol fundamental.

Desde que, en 1956, se acuñara el término por parte de un grupo de investigadores de matemáticas y ciencias de la Universidad de Dartmouth (Hanover, EEUU), la IA es entendida como el conjunto de técnicas que permiten a los computadores realizar una serie de tareas cognitivas, simulando así la inteligencia humana. Pero, al igual que sucede con toda actividad sociotecnológica, la IA sufrirá una serie de transformaciones a lo largo de su evolución, haciendo que su definición y sus objetivos no se mantengan estables. Si bien es cierto que el punto de partida de los actuales laboratorios de IA continúa siendo que los computadores aprendan de forma autónoma a programar tareas para cada uno de los potenciales escenarios de aplicación, ya no podemos decir que su principal objetivo sea el de simular las operaciones realizadas por la mente humana, y ni mucho menos emular una *cognición única y autónoma*. Es por ello, que términos como *Machine Learning* o *Deep Learning* han cobrado tanta fuerza en los últimos años, pues estos se refieren a subprogramas de la IA que permiten al software aprender y mejorar en el desarrollo de aquellas tareas que les son determinadas, gracias a la experiencia que consigue al *entrenarse* así mismo. Para comprender cómo estas transformaciones de la

⁵ Para una comprensión de *cognición distribuida* véase el estudio de antropología sobre marineros y navíos de guerra desarrollado por Hutchins (1995).

⁶ Aquí el concepto *dispositivo* se refiere al planteamiento teórico desarrollado por Foucault (1994).

IA determinan la ecología cognitiva actual, debemos analizar su evolución en el contexto de la era digital en la que esta tecnología se inscribe.

Los computadores están compuestos por capas de interfaces (que a su vez funcionarán como *cajas negras*) a través de los cuales accedemos a la información de forma simulada e interactiva. Tal y como ocurrió con otras tecnologías de visualización (como el microscopio o el telescopio; el diagrama o el mapa), la implementación del interfaz digital alargaría el campo de lo visible, provocando la movilización de muchos más aliados (set de datos). Según Lévy (1994), se instauraría así un conocimiento por simulación que será fundamental para la investigación científica, pues permitió ver y mostrar abstracciones de datos codificados digitalmente, estudiar fenómenos inaccesibles a la experiencia o predecir eventos que todavía no han tenido lugar. Entre estos interfaces podríamos ubicar los sistemas inteligentes de gestión de bases de datos, los cuales serían los primeros modelos de IA en ser implementados de forma funcional. Se trata de Sistemas Expertos que agilizan las tareas de profesionales en áreas específicas del conocimiento, como la medicina, la genética, la astronomía o la ingeniería. Estos sistemas parten de la información disponible y del núcleo de saberes establecidos previamente por los profesionales, para poder extraer nuevas conclusiones y así evolucionar con mayor rapidez en el conocimiento experto.

Con la expansión de la digitalización a todos los ámbitos de la vida, la producción y almacenamiento de datos crecería exponencialmente, lo que provocó un aceleramiento en el desarrollo de la IA. Al disponer de grandes y diversas bases de datos se pudieron entrenar y testar nuevos modelos de aprendizaje que, de forma simultánea, permitirían extraer la potencialidad inherente a todos estos datos, cuyo volumen, variabilidad y velocidad de crecimiento escapaban a las *ecologías cognitivas* predecesoras. Esto provocaría una modificación en las metas y objetivos de las investigaciones en IA: los intentos por emular computacionalmente las capacidades cognitivas humanas, cuya máxima pretensión fue generar una cognición única y autónoma, desembocaron en el desarrollo de un tipo de súper cognición (Manovich 2018a) que permitía analizar conjuntos de datos situados en todas partes. Actualmente podemos destacar tres grandes *subsets* de la IA que persiguen esta nueva meta, aquello que Manovich (2018b) definirá como tipos de cognición de la sociedad de la AI/data: la visualización de datos, el aprendizaje automático no supervisado y el aprendizaje automático supervisado. De esta forma, podríamos decir que los modelos de IA se han convertido en un nuevo paradigma cognitivo dentro del *conocimiento por simulación*, pues esta tecnología nos promete una rápida movilización, visualización e inscripción de información que hasta ahora escapaba a nuestro conocimiento.

Las técnicas de IA, como mediadores tecnológicos que son, vendrán a transformar y reforzar ciertas características de la *ecología cognitiva* de la era digital, al mismo tiempo que mantienen los *reflejos* de anteriores *tecnologías de la inteligencia*, como por ejemplo el caso del paradigma estadístico cuantitativo. Así, podríamos decir que dos de las características principales de la IA son el tiempo real y la delegación de inteligencia. El tiempo real, nos dirá Lévy (1994), se produce a partir del conocimiento por simulación a través de interfaces, un conocimiento operacional que es solicitado y reorganizado de forma puntual y en el momento (en tiempo real); se trata de obtener la información más fiable lo más deprisa posible mediante un acceso *online*, es decir, desde cualquier lugar. El tiempo de la informatización, por tanto, procura velocidad y operatividad, diferenciándose así del tiempo lineal de la historia y de la preocupación por la verdad crítica de las teorías, ambas características desencadenadas por la tecnología de la escritura (Lévy, 1994). Esto provocará que el conocimiento se encuentre en una

reorganización constante y en permanente metamorfosis, por lo que las teorías cederán terreno a los modelos, como es el caso de las propias técnicas de IA, las cuales son continuamente rectificadas y mejoradas.

Por tanto, la *ecología cognitiva* de la era digital, marcada actualmente por la escala de los datos digitales así como por el *tiempo real*, requiere una *inteligencia* que sea cualitativamente semejante a las capacidades cognitivas humanas (o mejor dicho, semejante a las ecologías cognitivas predecesoras, pues dichas capacidades pasan también por la movilidad y la fiabilidad obtenida de los móviles inmutables que ya configuraban nuestra cognición científica), "pero que opere en una escala cuantitativamente diferente" (Manovich 2018a). Es aquí donde el compendio de técnicas de la IA se ofrecerá como la respuesta para estas nuevas necesidades, una herramienta que nos promete un acceso al conocimiento libre de subjetivación, pues podremos delegar nuestra inteligencia en la máquina sin comprometer la objetividad mecanizada de esta.

Como vemos, estos argumentos nos devuelven a la idea estándar de la Modernidad, donde las tecnologías se ofrecen como la respuesta a necesidades específicas, y no como mediadoras que son. Así, nos encontramos nuevamente con el *mito del progreso*, donde los objetos (set de datos digitales en este caso) serán separados con mayor efectividad de los sujetos que operan con ellos, lo cual facilitará la constante redefinición de los correlatos operacionales de la IA⁷. Pero, tal y como se indicó anteriormente, esta separación entre sujetos y objetos no existe, y la tecnología no es simplemente un fin en sí misma, sino mediadora, por lo que si la IA se presenta como una solución que funciona cada vez mejor (con mayor operatividad) es precisamente, y paradójicamente, porque cada vez sabemos menos sobre su funcionamiento. Cuanto mejor funciona la IA más oscuras se vuelven las cajas negras que ocultan los *híbridos* y los mediadores necesarios para que se produzca esta delegación de inteligencia: programadores, hardware, diseñadores, interfaces, instituciones, algoritmos... El precio que pagar por la movilización de todos estos datos, en favor de un conocimiento operacional y en tiempo real, será una "opacidad algorítmica" (Burrell 2016. Citado en Elish y Boyd 2017, 14) que nos hará confiar "en un gran número de acciones delegadas que, por sí mismas, me impulsan a hacer cosas en favor de personas que no están aquí y de cuya existencia ni siquiera puede tener atisbo" (Latour 2001, 227).

Esto implicará un *poder-saber*⁸ que normalizará el poder de aquellos pocos sujetos e instituciones capaces de gestionar los algoritmos necesarios para que podamos ver y conocer en esta nueva *ecología cognitiva*, y en los cuales delegaremos decisiones y responsabilidades, lo cual provocará que estos algoritmos "*tengan éxito incluso cuando fallan*" (Magnet 2011, 3), pues, una vez implementados, permitirán ejercer un poder efectivo basado en la delegación de inteligencia y en la operatividad, independientemente de los resultados que ofrezcan. Por contra, y siguiendo la línea argumental planteada por Magnet (2011) en su estudio sobre los fallos de las tecnologías biométricas, podemos decir que al sustentarse en el *mito del progreso* las tecnologías de IA también "*fallan incluso cuando tienen éxito*" (Magnet 2011, 3).

Lo que a continuación se plantea es que podemos comenzar a abrir las *cajas negras* de la IA (siguiendo la cadena de delegaciones y mediaciones que en ellas se ocultan) a través del análisis del léxico

⁷ Para una relación de la *retórica del progreso* con el desarrollo de la IA y los Sistemas Expertos véase Woolgar (1985).

⁸ Aquí el concepto *poder-saber* se refiere al planteamiento teórico desarrollado por Foucault (1994).

empleado en la actividad sociotecnológica de la IA, el cual se extiende desde las prácticas de programación de los modelos hasta su implementación profesional y uso cotidiano. Se trata de un lenguaje humano/mecánico que, como veremos, se caracteriza por ser flexible y ambiguo (neuronas/artificiales, aprendizaje/automatizado, visión/computacional, inteligencia/artificial, *smart/phones...*). Su estudio puede ayudarnos a discernir los discursos en los que sustenta esta tecnología, y cómo estos modifican nuestra forma de entendernos a nosotros mismos y a nuestras tecnologías, al delegar en ellas funciones tan determinantes como la *inteligencia*, lo que sin duda afectará a nuestra *ecología cognitiva*.

4. Neuromitología y delegación de inteligencia: el léxico de la Inteligencia Artificial

Para comprender cómo actúa el lenguaje empleado en la actividad de la IA nos detendremos en las prácticas de programación de sus modelos⁹, particularmente en los subprogramas de la IA denominados *Deep Learning* o *Deep Neural Networks* (también llamados en español modelos de Red Neuronal Multicapa), una de las técnicas de aprendizaje automatizado más desarrollada e implementada en la actualidad. Entre estas redes neuronales artificiales encontramos el modelo *Convolutional Neural Network* (CNN), subprograma empleado en Visión Artificial para clasificación y segmentación de imágenes. A través de la CNN la máquina podrá reconocer y etiquetar imágenes que no ha visto con anterioridad, gracias a un entrenamiento (*Backpropagation* o algoritmos de aprendizaje supervisado) realizado a partir de una base de datos previamente adaptada y etiquetada. Las Redes Neuronales podrán incluso generar nuevas imágenes que no existían previamente, gracias a la combinación de las CNNs con otro tipo de redes neuronales, denominadas *Generative Adversarial Neural Network* (GANs).

Cuando se emplea el término Red Neuronal, los desarrolladores de IA se están refiriendo a la arquitectura de un software que es modelado a partir de la forma en la que se entiende que funcionan las redes neuronales del cerebro, en lugar de construirlo a través de instrucciones rígidas y predeterminadas por el programador, lo cual permitirá que no sea necesario programar todas las instrucciones necesarias para realizar una tarea. El programador implementará al modelo un algoritmo de aprendizaje adecuado a la tarea específica que se quiere desarrollar, "se diseñará la arquitectura de una Red Neuronal" (*creative coder*, comunicación personal, 8 de enero de 2019) y se entrenará con un set de cientos o miles de datos específicos, de tal forma que el modelo descubra, de manera autónoma, la forma óptima para *reconocer* nuevos inputs asociándolos a los outputs preestablecidos. Es decir, el algoritmo implementará automáticamente las variantes necesarias en el cálculo computacional para extraer patrones de los nuevos inputs, lo cual le permitirá generar un *output* correcto para cada input. Esto servirá, por ejemplo, para que el software pueda *reconocer* imágenes que no había visto antes, pero que comparten características con los datos con los que se entrenó el modelo.

⁹ Este análisis del lenguaje se ha realizado en base a un trabajo de campo colaborativo y experimental junto a un *creative coder*, cuyo resultado fue la programación de una serie de modelos de Machine Learning basados en Redes Neuronales Multicapa, específicamente dos modelos de CNN (*Convolutional Neural Network*) y una GAN (*Generative Adversarial Network*)

Estas Redes Neuronales nos interesan para el análisis del lenguaje sociotecnológico empleado en la actividad de la IA, pues en sus definiciones y prácticas encontramos un *léxico de la neuromitología* (Tallis, 2004) que nos permitirá comprender cómo se genera la delegación de *inteligencia* en esta tecnología. El término *neuromitología* será acuñado por Tallis (2004) para referirse a la manera en que la filosofía de la mente pasó a convertirse en una neurofilosofía, al asimilar, de forma pasiva y no crítica, los hábitos del lenguaje provenientes de la neurociencia. Tallis (2004) argumentará que fueron las ciencias cognitivas las que generaron el mito de que la ciencia neurológica moderna podía explicar lo que realmente es la consciencia, explotando el justificado prestigio de la neuroanatomía, la neurofísica, la neuroquímica y de otras neurociencias legítimas. En su intento por explicar la mente (consciencia), el discurso neurofilosófico dará un mal uso de los términos neurocientíficos, haciendo que el cerebro sea visto como mecánico y antropomórfico al mismo tiempo. Esto es, el cerebro es entendido como una maquinaria que ha sido diseñada para su utilidad funcional, y por tanto nuestras intenciones, proposiciones o pensamientos "son incorporados a las cadenas causales que simplemente pasan a través del tejido neuronal [...] y salen por el otro sitio" (Tallis 2004, 24), vinculando la entrada sensorial (input) con la salida motora (output). Tallis (2004) nos dirá que según este punto de vista **todo se reduce** a las conexiones neuronales que no están bajo la acción de un *yo como agente* (cerebro-mecánico), siendo el cerebro, sobre la base de su utilidad funcional y como resultado de una cadena causal, el que decide por mí (cerebro-antropomórfico).

La aceptación de estos modelos de consciencia por parte de las ciencias cognitivas, que unifican lo biológico (basado en las neurociencias) y lo computacional (basado en la Inteligencia Artificial), dependen del uso de ciertos términos ambiguos que tienen una multiplicidad de significados, un *léxico neuromitológico* que permite saltar de un área a otra con facilidad. Por ejemplo, la idea de que las máquinas tengan memoria, que almacenen información, que hagan cálculos o que detecten elementos, es aceptada de forma común por todos nosotros; "La característica más importante de esos términos es que tienen un pie en ambos campos: se pueden aplicar tanto a las máquinas como a los seres humanos, y su despliegue elude o erosiona las barreras entre el hombre y la máquina, entre la consciencia y lo mecánico" (Tallis 2004, 34).

Para ejemplificar el uso de este léxico neuromitológico en las prácticas de IA, usaremos un ejemplo clásico para iniciarse en la programación de Redes Neuronales Multicapa: la programación de una CNN que permite detectar números escritos a mano. Las imágenes de números contienen poca información, pocos píxeles en escala de grises, por lo que será más fácil seguir el funcionamiento de esta Red Neuronal que el de otras Redes que detectan imágenes más complejas, como imágenes fotográficas a color o imágenes tridimensionales. Además, este ejemplo servirá al programador para mostrar las grandes posibilidades que pueden ofrecer estos modelos, pues la máquina será capaz de reconocer algo tan manual, y que ofrece tantas variantes, como un número escrito a mano por una persona. La relación manualidad/automatismo que se destaca en este ejemplo de CNN convierte a este modelo en un objeto de estudio interesante para nuestro análisis del léxico y de la actividad sociotecnológica de la IA. A continuación, se transcribe un fragmento de un video explicativo muy popular entre aquellos que se están iniciando en la programación de los modelos de Red Neuronal¹⁰:

¹⁰ El análisis de este tipo de referencias puede servir para comprender la *economía explicativa* (Quintáis 2001) que realizan tanto los medios, como los expertos y divulgadores del conocimiento tecnológico, al presentar las técnicas y novedades de la Inteligencia Artificial.

Tu cerebro no tiene problemas en reconocer un 3 en una pequeña imagen de un 3 escrito a mano [...], es decir, hay muchas formas de escribir un 3 [...], que el cerebro pueda hacer esto sin esfuerzo es fascinante. [...] Las células fotosensibles que se disparan en el ojo cuando ves la imagen de un 3, son muy diferentes de las que se disparan cuando ves la imagen de otro 3, pero algo en la increíble e inteligente corteza visual que tienes dentro de ti, resuelve que las dos imágenes representan la misma idea, y, al mismo tiempo, reconoce otras imágenes con sus distintas ideas. [...] Realizar un programa que me diga que número piensa que es, nos muestra lo difícil que es esta tarea, lejos de lo cómicamente fácil que creíamos que era diferenciar un número de otro. [...] De esta forma, la Red Neuronal me permite automatizar la tarea al igual que lo hace el córtex cerebral (Sanderson 2017).

Esta explicación recurre a un lenguaje ambiguo que nos hace constantemente saltar y diluir la frontera que separa al hombre de la máquina. En primer lugar, se realiza un antropomorfismo del cerebro (el córtex cerebral es inteligente porque reconoce números, lo cual lo convierte en increíble), después se mentaliza la máquina (el programa me dice el número que piensa que es), y finalmente se acaba por mecanizar la mente (el córtex cerebral automatiza la tarea de reconocer números), lo cual ensalzará el valor de la Red Neuronal, pues esta consigue realizar una tarea compleja de forma automática, igual que lo haría el cerebro/mente. En este ejemplo se puede ver cómo hemos interiorizado un lenguaje que, a través de metáforas, facilita que los computadores sean empleados como modelos para la mente y viceversa.

Tallis (2004) nos dirá que la aceptación de este lenguaje neuromitológico nos hace simplificar los problemas filosóficos de la mente y la consciencia, lo cual nos hará asumir la existencia de cerebros antropomórficos y mecánicos que conocen la idea de 6, y que reconocen automáticamente todas las imágenes que se corresponden con esta idea. Por otro lado, nos encontramos con que este lenguaje propaga el mito de una inteligencia artificial con capacidades increíbles y una multitud de posibilidades (operatividad), pues esta tecnología es capaz de emular las complejas tareas que el cerebro realiza de forma automatizada. Pero si una máquina conoce la idea de 6 (para reconocer posteriormente todas las posibles imágenes del 6) es porque nosotros se lo hemos determinado; somos nosotros los humanos quienes hemos construido esta idea de 6, dentro de nuestra cultura particular y gracias al desarrollo de otras tecnologías que preceden a esta máquina, como es el caso de la propia tecnología de la escritura. Puede que el uso metafórico de este lenguaje comience de forma consciente por parte de los especialistas ("una forma de hablar, porque necesitamos de metáforas para explicar lo que hacemos" (*creative coder*, comunicación personal, 8 de enero de 2019)), pero la base que sustenta este lenguaje antropomórfico se puede olvidar, haciendo que "las ropas metafóricas en las que se envuelve el pensamiento se conviertan en su piel" (Tallis 2004, 36).

Tallis (2004) nos dirá que podemos entender estas atribuciones metafóricas de intencionalidad como el resultado de una relación cultural que establecemos con los artefactos tecnológicos que desarrollamos y a los cuales extendemos nuestros propósitos, un animismo que puede llegar a hacernos olvidar que las máquinas no son más que autómatas inconscientes. Pero "un programa antifetichista" (Latour, 2001, 227) basado en este presupuesto solo reivindicaría la mano humana que se encuentra detrás de estas técnicas, lo cual sería un planteamiento asimétrico e incompleto pues, como vimos, los humanos ya no

están solos, delegamos continuamente acciones en los demás actantes¹¹ (no humanos), los cuales comparten nuestra existencia y la modifican, y sin cuya mediación estaríamos perdidos y discapacitados cognitivamente. En cualquier caso, podemos decir que el lenguaje neuromitológico no es el resultado del empleo de simples metáforas, ni la consecuencia de un animismo:

Las técnicas modifican el fondo de lo que expresamos, no sólo su forma. Las técnicas tienen significado, pero lo generan gracias a un tipo especial de articulación [...] que sobrepasa los límites que el sentido común traza entre los signos y las cosas. [...] Abandonamos la relativa comodidad de las metáforas lingüísticas y nos adentramos en un territorio desconocido, [...] permanecemos en la esfera del significado, pero hemos abandonado la del discurso (Latour 2001, 222-224).

Latour (2001) nos dirá que es a través de las modificaciones producidas por las técnicas que los no humanos son socializados, de modo que puedan entablar relación con los humanos. De esta forma, las traducciones y modificaciones que se producen con la evolución de las técnicas de la IA, no solo establecen nuevas metas y funciones para esta tecnología (de una cognición única a un Sistema Experto, y de este a una supercognición), sino que genera un cambio en el fondo mismo de aquello que se expresa, lo que convierte al léxico neuromitológico en un tipo de disyunción más que en simples metáforas (Latour 2001). Las técnicas que componen la IA actuarán como modificadoras de forma y fondo, pues en el momento en el que el significado de esta tecnología cambia (cuando los cálculos computacionales pasan a significar *inteligencia artificial*, o cuando la Inteligente Artificial pasa a significar supercognición), se produce una traducción que modificará la naturaleza misma del significado (lo que significa inteligencia, o lo que significa cognición): "el asunto que constituía el objeto de la expresión ha quedado modificado" (Latour 2001, 224). Por tanto, podemos entender este léxico como el resultado de un proceso de socialización que consiste en el intercambio de propiedades entre los humanos y los no humanos, donde algunas características de la inteligencia humana se convierten en una máquina de cálculo computacional, y, como consecuencia, algunas características de la máquina se convierten en humanas.

De esta forma, "la distinción entre apariencia superficial y esencia subyacente, en este caso la distinción entre 'lo que realmente es la inteligencia' y 'meros signos de inteligencia'" (Woolgar 1985, 564) se diluye. Puede que para un usuario común el impacto que genera esta tecnología allí donde se implemente sea el mismo, se presente como una máquina con meros signos de inteligencia o como una máquina que es inteligente, pero lo cierto es que esta diferencia genera un cambio en las expectativas y las metas de los usuarios. La primera forma (la IA presenta meros signos de inteligencia) apela a un usuario activo, que puede sacar beneficios del automatismo de esta tecnología allí donde la cognición humana (o la anterior ecología cognitiva) no podía llegar; mientras que la segunda (la IA se presenta como *lo que es ser inteligente*) apela a una entidad autónoma, que opera bajo la objetividad del cálculo algorítmico y con la adaptabilidad del raciocinio humano, modificando así el significado mismo de inteligencia. Esta cualidad, que hasta ahora era exclusivamente humana, se caracterizará por ser automática, objetiva y mecanizada. Al modificar el fondo de la expresión y presentarse como *lo que*

¹¹ Aquí el término *actante* se refiere al binomio de actuación *actor-actante* (humano-no humano), concepto desarrollado por Latour: "Dado que en inglés el uso de la palabra 'actor' se suele restringir a los humanos, la palabra 'actante', tomada en préstamo de la semiótica, se utiliza a veces como el elemento lingüístico que permite incluir a los no humanos en la definición" (Latour 2001, 361).

es ser inteligente, estas tecnologías nos prometen el acceso a un conocimiento libre de subjetivación, tanto del mundo exterior (las cosas, convertidas en sets de datos) como del funcionamiento de nuestro mundo interior (los sujetos, convertidos en una cadena de estímulos neuronales automatizados).

5. Conclusiones preliminares: por una futura reterritorialización de nuestra máquina abstracta cognitiva

El léxico neuromitológico empleado durante el trabajo de programación se propagará en nuestra vida cotidiana a través de nuevos artefactos en los cuales también delegaremos inteligencia y responsabilidades¹² (teléfonos inteligentes, cámaras fotográficas, coches inteligentes...). Pero en muchas ocasiones no seremos conscientes de estar utilizando tecnología de IA, pues esta se encuentra cajaneada en los distintos mediadores tecnológicos que empleamos a diario. Además, estas cajas negras se multiplicarán en la medida en que deleguemos más funciones en la tecnología y según se establezcan nuevas metas para la IA, lo cual creará nuevos artefactos que se separarán de los objetivos predecesores, provocando que la IA acabe por operar de forma silenciosa e invisible. Siguiendo el planteamiento de Manovich (2018a) podemos decir que las redes sociales, los motores de búsqueda o las *apps* se sirven del análisis cuantitativo de los datos extraídos de Internet para predecir nuestros gustos e intereses, y así recomendar y/o sugerir contactos, fotografías, videos, música, noticias o *fake news*, lo cual influirá en nuestras decisiones e imaginaciones. Por tanto, como se ha pretendido demostrar en este texto, podemos afirmar que el paradigma cognitivo de la IA nos ofrece un conocimiento operativo en la medida en que fomentamos, de forma más o menos consciente, la dataficación de nuestra realidad (*Big Data*), lo que provocará un mayor poder entre aquellas instituciones que gestionan los mediadores necesarios para su funcionamiento, encerrándolos en cajas negras que propagarán el mito del progreso.

Es por ello por lo que se hace necesario multiplicar los mediadores que dan lugar a esta tecnología, con el objetivo de tener un mayor (y mejor) conocimiento de cómo se configura nuestra ecología cognitiva actual. Es especialmente interesante y significativo el estudio de las *cajas negras* presentes en la actividad sociotecnológica de la IA, una vez que este término es empleado en la propia práctica de programación de Redes Neuronales: caja negra, capas ocultas, "donde ocurre la magia" o allí donde "el software aprende a comportarse inteligentemente" (*creative coder*, comunicación personal, 1 de noviembre de 2018) serán algunas de las expresiones usadas por los programadores para referirse a los cálculos algorítmicos en los que delegan inteligencia. Estos cálculos o capas ocultas de la Red Neuronal establecerán patrones que relacionan nuevos *inputs* con los *outputs* preestablecidos (categorías y/o etiquetas), lo que hará que el software aprenda a reconocer y clasificar nuevas imágenes, sonidos, textos, etc. Al tratarse de cálculos imposibles de revisar por un humano, se generará un desconocimiento, incluso entre los expertos, de aquello que sucede en el interior de la *caja negra* de los modelos. Es decir, aún siendo funcional (o precisamente debido a su funcionalidad) existirá un desconocimiento de los patrones que hacen que la IA tome ciertas decisiones.

¹² Véase el *spot* publicitario del smartphone Huawei P20 PRO: "Inteligencia que desafía la fotografía *smartphone*. El primer *smartphone* del mundo con triple cámara asistida por Inteligencia Artificial" (Huawei Mobile España, 2018).

De esta forma, y aunque el automatismo de la IA se haya convertido en un pilar fundamental para el progreso de las nuevas tecnologías de la informatización, así como para el devenir de nuestra era digital, no podemos limitar u obviar nuestra responsabilidad ante sus resultados, pues esto propagaría el paradigma cuantitativo de las ciencias sociales del siglo XIX y XX, donde el ideal de las matemáticas sociales (Gould 2017), basado en la clasificación de la realidad en sistemas cuantificables, reductores y manejables, prometía un conocimiento que ocultaba la ideología y el poder subyacente a estas prácticas. Pero al igual que sucede hoy con la IA, era y es la mano humana, junto a las técnicas y las cosas (no humanos), la que establece, fija y normaliza las categorías con las que pensamos el mundo, mediante un proceso de visualización y cognición (Latour 1998) que tuvo lugar gracias a las inscripciones-mediadores que Latour (1998) describe como móviles inmutables. Por ejemplo, las categorías supuestamente neutrales y evidentes, aquello que denominamos *de sentido común*, como la raza o los tipos de enfermedad, surgen de contextos políticos y sesgos culturales que tienen necesidad de movilizar ciertas inscripciones. Por tanto, podemos afirmar que la programación de la IA no puede separarse de la actividad sociotécnica en la que se desarrolla y a la cual modifica.

En este sentido, sigue siendo un reto para los expertos en Visión Artificial (subprograma de la IA donde se inscriben las CNNs y las GANs anteriormente señaladas) entender con precisión qué es lo que provoca que la máquina clasifique correctamente ciertas imágenes pero se equivoque con otras, lo cual conlleva una serie de consecuencias graves (de carácter ético, como el racismo implícito en los controles de seguridad biométricos y de reconocimiento facial (Magnet 2011), o fatales, como las muertes en accidentes que implican vehículos inteligentes). Por ello, los investigadores en IA están desarrollando de forma urgente mapeos de los datos visuales que utilizan estos sistemas, para comprender los patrones que estos softwares generan durante el proceso de predicción. El objetivo de estas investigaciones es "ver a través de los ojos" de la red" (Carter, Armstrong y Schubert 2019), hacer comprensibles las capas ocultas de las Redes Neuronales Multicapa (abrir la caja negra, desvelar la magia) para intentar responder a la siguiente pregunta: "¿qué han aprendido estas redes que les permite clasificar tan bien las imágenes?" (Carter et al. 2019).

Más allá de encontrar respuestas para fallos concretos, o de separar el buen del mal funcionamiento de los modelos de Visión Artificial (lo cual solo suscribiría *el punto de vista estándar* de la Modernidad), estos softwares de descajaneigrización pueden ser particularmente interesantes para especular sobre futuras ecologías cognitivas. Lo que proponemos aquí es una investigación que se sirva del potencial del paradigma cognitivo de la IA a través de la experimentación con estos softwares de mapeo¹³, con el objetivo de acceder a nuevas articulaciones entre las categorías y sus significados, en lugar de fijarlas y automatizarlas, escapando así a la herencia estadística¹⁴. En resumen, al concienciarnos de la artificiosidad y la distribución de nuestra cognición, al comprender cómo se genera la delegación de *inteligencia* que realizamos en las tecnologías de nuestra era digital, al abrir el mayor número posible de *cajas negras* que conforman la IA, y al visualizar las articulaciones algorítmicas que se dan en el interior

¹³ Actualmente me encuentro desarrollando una investigación teórico-práctica basada en la experimentación con softwares de visualización y análisis de características de Redes Neuronales.

¹⁴ Para una relación entre la herencia estadística y las técnicas de IA véase Manovich (2018a, 2018b)

de las Redes Neuronales que configuran la Visión Artificial, podremos experimentar con los límites y las posibilidades de la IA, con el objetivo de reterritorializar nuestra propia máquina abstracta¹⁵ cognitiva.

Referencias Bibliográficas

- Carter, S., Armstrong, Z. y Schubert, L. (2019) *Activation Atlases* [Atlas de activación]. Distill.
DOI: 10.23915/distill.00015
- Deleuze, G. y Guattari, F. (2015) *Mil mesetas* (José Vázquez Pérez, trad.). Pre-textos. (Obra original publicada en 1980)
- Elish, M.C. y Boyd, D. (2017) "Situating methods in the magic of Big Data and AI", *Communication Monographs*. DOI: 10.1080/03637751.2017.1375130
- Foucault, M. (1994) *L'Histoire de la sexualité 1. La volonté de savoir*. Gallimard
- Gould, S. (2017) *La falsa medida del Hombre* (Ricardo Pochtar y Antonio Desmots, trad.) Editorial Planeta. (Obra original publicada en 1996)
- Huawei Mobile España. (27 de marzo de 2018) *Huawei P20 / Huawei P20 PRO: la fotografía renace hoy* [archivo de video].
Recuperado de https://www.youtube.com/watch?time_continue=28&v=XqNSdjs3THg (Fecha de consulta: 15 de abril de 2019).
- Hutchins, E. (1995) *Cognition in the Wild* [Cognición en la naturaleza]. The MIT Press.
- Latour, B. (1998) "Visualización y cognición: pensando con los ojos y con las manos", *La balsa de la Medusa*. nº 45-46.
- Latour, B. (2001) *La esperanza de Pandora* (Tomás Fernández Aúz, trad.). Gedisa. (Obra original publicada en 1999)
- Latour, B. (2013) "How to be Iconophilic in Art, Science, and Religion?", en: Jones, C.A. y Galison, P. (ed.). *Picturing Science Producing Art*. Routledge.
- Lévy, P. (1994) *As Tecnologias da Inteligência. O Futuro do Pensamento na Era da Informática* (Fernanda Barão, trad.) [Las tecnologías de la inteligencia. El futuro del pensamiento en la era de la informática]. Lisboa, Portugal: Piaget. (Obra original publicada en 1990).
- Magnet, S. A. (2011) *When Biometrics Fail. Gender, Race, and the Technology of Identity* [Cuando la biometría falla. Género, raza, y la tecnología de la identidad]. Durham y Londres, Reino Unido: Duke University Press.
- Manovich, L. (2018a) *AI Aesthetics* [Estética de la IA]. Moscú, Rusia: Strelka Press. Recuperado de <https://www.amazon.com>
- Manovich, L. (2018b) "Can We Think Without Categories?" [¿Podemos pensar sin categorías?] en: Reichert, R. y Fuchs, M. (ed.). "Rethinking AI: Neural Networks, Biometrics and the New Artificial Intelligence". *Digital Culture & Society (DCS)*. Vol. 4, (1).
- Pfaffenberger, B. (1992) "Social Anthropology of Technology" [Antropología social de la tecnología]. *Annu. Rev. Anthropology*. 21: 491-516.

¹⁵ Aquí el concepto *máquina abstracta* se refiere al planteamiento teórico desarrollado por Deleuze y Guattari (2015) en referencia al funcionamiento de nuestras experiencias perceptivas y cognitivas cotidianas, al enfrentarnos al reconocimiento y categorización del rostro del otro, marcado por el poder implícito que conlleva nuestra necesidad de imponer normas de significancia y de subjetivación.

- Quintáis, L. (2001) "Modernidade e pesadelo: um ensaio sobre ciência, localização cerebral e economia explicativa" [Modernidad y pesadilla: un ensayo sobre ciencia, localización cerebral y economía explicativa]. *Antropologia portuguesa*. Vol. 18.
- Sanderson, G. (5 de octubre de 2017) *But what is a Neural Network?* [Pero ¿qué es una Red Neuronal?] [archivo de video].
Recuperado de <https://www.3blue1brown.com/videos-blog/2017/10/9/neural-network> (Fecha de consulta: 20 de febrero de 2019).
- Tallis, R. (2004) *Why the Mind is not a computer. A pocket lexicon of neuromythology* [Por qué la mente no es un computador. Un léxico de bolsillo de la neuromitología]. Imprint Academic.
- Woolgar, S. (1985) "Why not a Sociology of Machines? The Case of Sociology and Artificial Intelligence" [¿Por qué no una sociología de las máquinas? El caso de la sociología y la Inteligencia Artificial]. *Sociology*. Vol. 19, (4).

BIO



Andrés Pachón (Madrid, 1985) es formado en Antropología (máster en Antropología Social y Cultural, Universidad de Coimbra (Portugal), 2019) y en Arte Contemporáneo (licenciatura en Bellas Artes, UCM - CES Felipe II, 2008; magíster en Teoría y Práctica del Arte Contemporáneo, UCM, 2009). Durante los últimos años ha desarrollado una práctica visual que reflexiona sobre la construcción del imaginario colonial a través del uso de archivos fotográficos en las prácticas antropológicas y etnográficas, colaborando para ello con instituciones como el Musée du Quai Branly de París, el Museo Nacional de Antropología de Madrid o el Archivo Fotográfico del Museo de Arte de Lima (MALI) en Perú. Sus últimos trabajos se centran en el papel que desempeña la mediación tecnológica en la construcción del conocimiento, estableciendo relaciones entre las prácticas del s. XIX y XX y los usos tecnológicos del presente, como es el caso de su actual investigación sobre la actividad sociotécnica de los sistemas de Visión Computacional por Inteligencia Artificial.