

<https://artnodes.uoc.edu>

ARTÍCULO

**NODO «IA, ARTE Y DISEÑO: CUESTIONANDO EL APRENDIZAJE AUTOMÁTICO»**

# Identificación, clasificación y control: estrechos vínculos analizados desde las prácticas artísticas en el corazón de la inteligencia artificial

**Hugo Felipe Idárraga**

Universidad de Los Andes

Fecha de presentación: marzo de 2020

Fecha de aceptación: junio de 2020

Fecha de publicación: julio de 2020

## Cita recomendada

Idárraga, Hugo F. 2020. «Identificación, clasificación y control: estrechos vínculos analizados desde las prácticas artísticas en el corazón de la inteligencia artificial». In: Andrés Burbano; Ruth West (coord.) « IA, arte y diseño: Cuestionando el aprendizaje automático ». *Artnodes*, núm. 26: xx-xx. UOC. [Fecha de consulta: dd/mm/aa]. <http://doi.org/10.7238/a.v0i26.3361>



Los textos publicados en esta revista están sujetos –si no se indica lo contrario– a una licencia de Reconocimiento 4.0 Internacional de Creative Commons. La licencia completa se puede consultar en [https://creativecommons.org/licenses/by/4.0/deed.es\\_ES](https://creativecommons.org/licenses/by/4.0/deed.es_ES).

## Resumen

La inteligencia artificial (IA) está antecedida de una historia que se remonta a antiguos esfuerzos por crear seres con movimiento e inteligencias artificiales. El objetivo de este artículo es destacar cómo en esta historia, que incluye los desarrollos actuales en el campo del machine learning (ML) y de las redes neuronales profundas (RNP), las tareas de vigilancia y control por medio de la identificación y clasificación de personas, cosas o eventos del mundo, han sido medulares tanto en los mitos o las teorías sobre autómatas, homúnculos, androides, robots o cibernéticos, así como en los diferentes intentos por materializarlos. Se afirmará, además, que los deseos y esfuerzos por imaginar y crear estos seres no son ni inocentes ni circunstanciales; más bien, provienen en gran medida de una visión patriarcal del mundo en la que todo lo existente debe ser sometido a una vigilancia que asegure un control sobre lo real. Para ello se recurrirá,

por un lado, a historias que evidencien el predominio de estas tareas y la visión del mundo que detrás de ellas se esconde, y, por el otro, a prácticas artísticas y representaciones estéticas que han cuestionado el funcionamiento identificatorio y clasificatorio de la IA como medio de vigilancia y control.

### Palabras clave

Inteligencia artificial, arte, identificación, clasificación, vigilancia, control, adversarial examples.

### *Identification, classification and control: close ties analysed in reference to artistic practices in the heart of artificial intelligence*

#### Abstract

*Artificial intelligence (AI) is preceded by a history going all the way back to former efforts in creating beings with artificial movement and intelligence. The objective of this article is to highlight how, in this history, which includes the current developments in the fields of Machine Learning (ML) and Deep Neural Networks (DNNs), the tasks of surveillance and control through the identification and classification of people, things or events in the world, have been so central in the myths or theories on automatons, homunculi, androids, robots or cyborgs, as well as in the various attempts to make them a reality. It will be asserted, furthermore, that the wishes and efforts to imagine and create these beings are neither innocent nor circumstantial; rather, that to a large extent they come from a patriarchal vision of the world in which everything existing must be subjected to a surveillance that ensures control over that which is real. To achieve this, reference will be made on the one hand to histories which show the predominance of these tasks and the vision of the world hidden behind them, and, on the other, to artistic practices and aesthetic representations which have questioned the identifying and classifying operations of AI as a means of surveillance and control.*

#### Keywords

*artificial intelligence, art, identification, classification, surveillance, control, adversarial examples*

## Introducción

El arte, desde sus más remotas manifestaciones, ha señalado una infinidad de lugares desde donde despellejar la realidad de su ilegítima normalidad. La crítica artística de los fenómenos técnicos hace parte de la construcción de estos lugares. Actualmente, el fenómeno de la IA, globalizado y cotidiano, debe enfrentarse a las manifestaciones estéticas que cuestionan críticamente su funcionamiento, apuntando a sus fundamentos históricos y teóricos, con la intención de desarraigar su normalización elogiosa por parte de corporaciones y gobiernos que impulsan su desarrollo y aprovechan sus beneficios. Mientras que una revisión histórica de estos desarrollos permite identificar algunas ideas recurrentes, fijas en la mirada de una sociedad patriarcal que necesita de unos y otros como propiedades que deben ser organizadas y controladas, las prácticas artísticas desencajan la normalidad con la que estas ideas han sido abrigadas «permitiéndonos imaginar

diferentes posibilidades y diferentes futuros de cómo podemos ser empleando estas tecnologías»<sup>1</sup> (Núñez, 2018).

## Identificación y clasificación: relaciones con la vigilancia y el control

Los tempranos intentos en la creación de seres artificiales se concentraron en autómatas con movimientos mecánicamente autopropulsados. La *Iliada* presenta a Hefesto como constructor de una diversidad de autómatas (Homero, 2010). Entre ellos, Hefesto construyó a Talos, «el primer robot», hecho de bronce y «cuyo deber es patrullar tres veces al día las playas de Creta» (McCorduck, 2004: 5). En el año de 1966, el robot *Shakey*, desarrollado por la empresa Panoramic, una de las principales contratistas de la CIA y activa participante en el Project MK-Ultra, fue considerado «el primer robot del mundo con intelligen-

1. Todas las traducciones de las citas del inglés al español son del autor.

cia artificial» (Raviv, 2020). Allí, el profesor Woodrow Wilson Bledsoe desarrolló el primer sistema de reconocimiento facial, partiendo de las teorías de Alphonse Bertillon. Su objetivo: «ayudar a las agencias de seguridad a examinar rápidamente las bases de datos de fotografías y retratos, buscando coincidencias» (Raviv, 2020). Igual que Talos, el reconocimiento facial es empleado con propósitos de vigilancia, identificando y clasificando para permitir o denegar el acceso a lugares físicos o virtuales. El sistema de Bledsoe fue entrenado exclusivamente con fotos de personas caucásicas, excluyendo las de otras razas. El artista Trevor Paglen y la investigadora Kate Crawford, en su exposición *Training Humans* de 2019, desentrañan los prejuicios raciales y morales insertos en los set de datos, analizando y denunciando las imágenes empleadas para entrenar los sistemas de identificación y clasificación, así como las teorías supuestamente científicas que los legitiman (Paglen y Crawford, 2019).

«Las computadoras de HP son racistas» (Bunz, 2009). Joy Buolamwini fue una de las primeras artistas en denunciar estos prejuicios raciales en los sistemas de IA. Sus proyectos *Gender Shades* y *AI, Ain't I a Woman?*, ambos del año 2018, exploran la parcialidad de los sistemas de reconocimiento facial que, por ejemplo, causan una mayor persecución policial sobre la población afro de EE.UU. o discriminan negativamente a las mujeres. Mimi Onuoha ha criticado los conjuntos de datos empleados en estos sistemas. En su obra escultórica *Classification.01*, de 2017, dos paréntesis de neón se iluminan cuando el espectador cumple con las medidas de identificación algorítmica determinadas de antemano por la artista, en una especie de emparejamiento maquínico coherente con los prejuicios de estos sistemas.

Durante la Alta Edad Media, la metáfora representada en la *natura artifex* evidencia la importancia del neoplatonismo como fundamento para comprender la creación de autómatas. Los autómatas son pensados así como una creación que implicaba tanto la habilidad humana como el conocimiento de las formas ideales por medio de la acción de la naturaleza. Estas creaciones son misteriosas y exigían no solo un conocimiento artesanal, sino también el entendimiento al que únicamente filósofos o hechiceros podían acceder, como aquel que aparece en los *romans antiques* *Le Roman d'Éneas* (ca. 1160) y *Le Roman de Troie* (ca. 1165).

En *Le Roman de Troie*, Benoit de Sainte-Maure reconstruye la Guerra de Troya. Allí se describe la alcoba de convalecencia de Héctor. De entre sus diferentes maravillas destacan dos autómatas. El primero interpreta instrumentos musicales y esparce flores por la habitación, determinando lo estéticamente correcto. El segundo carga un incensario que cura la enfermedad y el dolor a quien aspira su aroma, y su tarea consiste en identificar y corregir todo comportamiento descortés, decidiendo sobre lo moralmente adecuado e inadecuado. En *Le Roman d'Alexandre* (ca. 1180), dos hombres de oro vigilan un puente amenazado por el ejército enemigo. Asimismo, en la continuación de Perceval, dos figuras de oro y plata resguardan la tienda de Alardin, caballero venido de

tierras distantes. De estas figuras, una está encargada de identificar y evitar el ingreso de villanos, mientras que la segunda posee un arpa que tañe de manera discordante cuando una doncella que ha perdido su virginidad desea acceder a la tienda. Los autómatas, así, «realizan una función disciplinaria. Su capacidad para vigilar, proteger y castigar a las personas es tan asombrosa y maravillosa como su valor estético o el misterio de su creación» (Truitt, 2015: 60).

No obstante, la vigilancia y el control llevados a cabo por estos seres y sistemas inteligentes no han sido desplegados sin resistencia. El artista Adam Harvey, en su proyecto *CV Dazzle*, del año 2010, intenta evadir el control de la visión algorítmica por medio de audaces maquillajes no convencionales. Asimismo, el proyecto de Sterling Crispin, *Data-Mask*, de 2013, aprovecha el funcionamiento de los algoritmos de reconocimiento facial en una estrategia de ingeniería inversa para engañar a la máquina con caras irreconocibles para el ser humano y, sin embargo, legibles por el algoritmo, impidiendo cualquier tipo de identificación efectiva.

Por otro lado, en general la creación de autómatas buscó delegar ciertas tareas humanas que realizaban comúnmente siervos y esclavos. Para el estagirita, «(...) si los peines por sí mismos tejiesen, y la pluma por sí misma tocara la cítara, ni los oficiales tendrían necesidad de ministros, ni los señores de siervos» (Aristóteles, 2004: 20). De esta forma, los autómatas se vinculan con la servidumbre, vigilada por el señor de la casa, o del gobierno, pero también con la vigilancia y el control social, en esa patria potestad recordada por Platón (2011) que solo podía ser ejercida por los hombres-ciudadanos. Por eso, para Truitt, los autómatas «no solo imponen límites físicos, sino también sociales» (2015: 62).

## Homúnculos, robots, magia y poder

Pigmalión, rey de Chipre, surge como otro vector de fantasías relacionadas con un mundo controlado y controlable desde la hegemónica visión patriarcal del universo. Desesperado por no encontrar a una mujer perfecta, decide diseñar una que le agrade plenamente. Galatea, el amor de Pigmalión, fue humanizada por la diosa Afrodita. De allí se originan diferentes mitologías sobre el homúnculo femenino, que es una creación principalmente masculina.

Teofrasto Paracelso fue uno de los primeros alquimistas en reivindicar la creación de homúnculos, sugiriendo su crecimiento por fuera del cuerpo femenino, despreciando así el papel de la mujer en la creación de vida y uniéndolo, de esta forma, la magia y la mecanización del ser humano.

El libro de Hubert Dreyfus *What Computers Can't Do: A Critique of Artificial Reason* (1972), primero publicado bajo el nombre de *Alchemy and Artificial Intelligence* (1965), aborda el misterio persistente que se esconde detrás de toda creación de vida artificial. Su libro fue escrito como crítica a aquella declaración de John

McCarthy según la cual las diferencias entre el ser humano y la máquina son «ilusorias». La comparación entre IA y magia, que fue la base de Dreyfus para criticar este excesivo optimismo, le sirvió también de excusa para revelar cómo, a través de la eficacia de la IA —al igual que la alquimia en sus inicios—, se esconde una confianza que socava los esfuerzos por analizar científicamente los más crípticos y oscuros procedimientos que ocurren en el interior de estos sistemas, legitimando el poder de quienes los desarrollan e implementan. Asimismo Campolo y Crawford (2020) afirman que vivimos un resurgimiento del *techno*-optimismo mágico, parecido al optimismo alquímico en sus inicios. Por ejemplo, según Stuart J. Russell, «recién estamos comenzando a obtener una comprensión teórica de cuándo y por qué la hipótesis de aprendizaje profundo es correcta, pero en gran medida sigue siendo una especie de magia» (Ford, 2018: 42, citado en Campolo y Crawford, 2020: 2).

¿Cómo se vinculan la clasificación, la magia y el poder político que de allí se deriva? Según Campolo y Crawford, la IA, que nace en un campo asociado al desencantamiento del mundo por medio de una explicación racional de este, se ha refugiado, sin embargo, en oscuras explicaciones por medio de un lenguaje reencantado. Este reencantamiento refuerza el poder de aquellos que diseñan, comercializan y despliegan estos sistemas, al tiempo que disminuye el de la vasta mayoría de personas que son identificadas, clasificadas y previsualizadas. «Es una forma de poder sin conocimiento», afirman Campolo y Crawford (2020: 5).

Trevor Paglen intenta develar los misterios de los proyectos del complejo industrial-militar de EE.UU en eso que él mismo ha llamado como el *Black World*. En su libro *Blank Spots on the map*, de 2009, o en su proyecto *Unmarked Planes and Hidden Geographies*, de 2006, Paglen trata de arrojar luz sobre aquellos misterios que, con mucha eficacia, han sido escondidos a los ojos del público en proyectos destinados a la vigilancia y el control social. Sin embargo, es el artista Mario Klingemann quien, sin recurrir a explicaciones mistificadas, experimenta con las RNP y los resultados inesperados que ellas producen, intentando entrever estéticamente lo que ocurre en el interior de ellas, en una técnica que él mismo ha denominado como *Neural Glitch*. Esta serie de obras de 2018 son el resultado de un intento por emancipar los límites de la percepción tanto humana como algorítmica, indagando en aquellas propiedades intrigantes de la IA, provocando que los modelos «malinterpreten los datos de entrada en formas interesantes, algunas de las cuales pueden ser explicadas como atisbos de creatividad autónoma» (Klingemann, 2018).

## El control, fundamento de la visión tecnológica patriarcal

Bajo la figura del homúnculo también se han pensado seres artificiales con vida y pensamiento propio. Cuarenta años después de la

muerte de Paracelso surge la figura de Joseph Golem, una creación del rabino Judah ben Loew por medio de artificios mágicos. Además de prestar servicios de vigilancia alrededor del templo, el Golem fue empleado como esclavo. Estas tareas de carácter servil sembraron en él la idea de superar a su creador, rebelándose frente a su poder, en una temática que se repetirá no solo en la modernidad literaria europea, sino también en múltiples representaciones contemporáneas, como en *Odisea en el espacio* (1968), *Blade Runner* (1982), *Terminator* (1984), *Matrix* (1999) o *Yo, Robot* (2004), en las que la IA se subleva frente a su amo. En el fondo, como veremos más adelante, este miedo de la máquina rebelada es también el miedo al trabajador rebelado, a la mujer rebelada, a la sociedad entera rebelada contra sus amos.

De entre los personajes modernos que han encarnado el espíritu de este mito está el mismo Frankenstein. Como subraya Ruha Benjamin respecto al carácter servil de Frankenstein en el contexto de la etimología de los robots, estos «transmiten una agitación continua sobre la dominación humana sobre otros humanos. (...) La dominación social caracterizó el laboratorio cultural en el que los robots fueron originalmente imaginados. Y, técnicamente, las personas fueron los primeros robots» (Benjamin, 2019: 55). Y también los últimos. Hoy en día, una de las consignas de los trabajadores en empresas ultratecnificadas es «no somos robots. Para satisfacer a la máquina, los trabajadores se sienten forzados a convertirse ellos mismos en máquinas» (Dzieza, 2020).

El pensamiento del amo y el esclavo, la idea de sirvientes y vigilantes como formas que asegurasen el orden por medio de la identificación y clasificación, está encarnado en estos y los posteriores seres artificiales, así como en los proyectos subsiguientes dedicados a la IA. Es precisamente a ello, como parte de esta herencia histórica masculina del amo-esclavo, a lo que apunta Anna Everett en su incómoda relación computacional: «Al iniciar mi computadora me confronto con el texto basado en DOS que me da una pausa... “Pri. Master Disk, Pri. Slave Disk, Sec. Master, Sec. Slave”». (Everett, 2002, citada en Benjamin, 2019: 55).

El mito de Pigmalión se conecta con el argumento de la famosa serie *Ghost in the Shell* (1995) a través de las muñecas de Hans Bellmer, empleadas como modelos visuales de las androides de Hanka Robotica. Bellmer creó desde 1933 mujeres artificiales «(...) con posibilidades anatómicas que son capaces de recrear las elevaciones de la pasión o aún de inventar nuevos deseos» (Brown, 2008: 235). Publicadas en el sexto número de la famosa revista surrealista *Minotauro*, las fotos de las muñecas de Bellmer querían servir como una crítica política de los cuerpos; sin embargo, las motivaciones de Bellmer terminaron siendo objeto de una crítica sobre la dominación del cuerpo femenino: «vinculando sus fantasías de chicas adolescentes y poderosas seductoras con los temas de la nostalgia y el erotismo, y relacionando su trabajo con el ambivalente deseo y repulsión de los surrealistas franceses

por el cuerpo femenino» (Brown, 2008: 235). Como en el mito de Pigmalión, Bellmer proyecta los deseos masculinos sobre un cuerpo del que no se tiene posesión. En estos seres, según las historiadoras del arte Therese Lichtenstein y Sidra Stich, «el cuerpo entero puede ser montado y desmontado como una máquina» (Brown, 2008: 236), de la forma misma en que son tratadas las androides en *Ghost in the Shell*.

## Androides y ciborgs en la antesala del control social

La traducción al latín de los escritos árabes de la Escuela de Alejandría introdujo en el occidente cristianizado el pensamiento mecanicista (Truitt, 2015). El famoso ajedrecista de Kempelen, El Turco, se convirtió en el vínculo entre la creación de autómatas mecánicos y las primigenias máquinas de cálculo como parte de los primeros intentos por imitar la inteligencia humana.

Desde el ajedrecista de Kempelen hasta el jugador de damas de Arthur Samuel y el primer programa verdaderamente funcional de un jugador de ajedrez de Alex Bernstein, el juego ha sido un tema esencial en la historia de la creación de seres artificiales inteligentes. El juego ha sido un medio privilegiado para escenificar una dinámica social basada en la identificación y clasificación de los deseos y movimientos de un enemigo, tanto en el campo de batalla como en las relaciones sociales. John von Neumann y Oscar Morgenster, en su teoría de juegos, quisieron desarrollar formas de procesamiento de información que pudieran ser aplicadas a la lucha entre enemigos —no necesariamente limitada al campo militar—, caracterizados como «un operador perfectamente inteligente y despiadado» (Galison, 1994: 159).

Lo que hay detrás del jugador artificial es una idea que ya desde Ramon Lull se revelaría posible: recrear la inteligencia por medios mecánicos, pues los humanos, como diría Descartes, son máquinas maravillosas, pero máquinas, a pesar de todo. Si los seres humanos pueden ser pensados e interpretados como mecanismos, el ser humano puede ser conocido y controlado, de la misma forma en la que se conoce y se entiende una máquina. Esta idea calará profundo en la forma de pensar de muchos de los impulsores de la cibernética y la IA. Norbert Wiener y Arturo Rosenbluth, por ejemplo, afirmarán en su *Purposeful and Non-purposeful Behavior* que, «como objetos de indagación científica, los humanos no difieren de las máquinas» (Galison, 1994: 250). Pero más que igualar la máquina y el humano, para el artista Wayne McGregor, entre otros, debemos pensar la IA como una forma de expandir nuestras capacidades. En su obra *Living Archive*, de 2018, McGregor crea una coreografía a partir de un diálogo entre su trabajo, los bailarines y las herramientas de IA, generando movimientos inéditos a partir de la identificación de patrones en sus obras.

Aquella separación entre mentecuerpo, sin embargo, ha sobrevivido en los desarrollos actuales de la IA, con sus evidentes consecuencias políticas. Esta separación, como fundamento mecánico-filosófico del entendimiento del ser humano, es portadora de un poder patriarcal que se ha impuesto en el campo de la ciencia y la técnica, pero también en otros campos estratégicos de la vida: «La separación mente-cuerpo es, por supuesto, fundamental para la tradición patriarcal occidental, esencial no solo para la teología, sino también para la política, la economía y todas las demás áreas donde las mujeres y los trabajadores han sido sistemáticamente subordinados» (McCorduck, 2004: 40). Al simplificar la complejidad biológica humana para igualar a las máquinas y el humano, esta separación entre cuerpo y mente ha legitimado la superioridad intelectual de un grupo social privilegiado por el género, la raza y el poder.

David Hume, en su *Tratado de la naturaleza humana*, argumenta que, en el mismo sentido que lo había hecho John Locke un siglo antes, la mente puede ser explicada por medio de leyes universalmente válidas, en una dinámica patriarcal de la ciencia que Donna Haraway denominará como la cultura de la no cultura, destacada en el trabajo de Robert Boyle, donde se evidencia el valor masculino y caballeresco de la autoinvisibilidad como «la forma científica específicamente moderna, europea y masculina de la virtud de la modestia». De esta manera, destaca Haraway, el hombre de ciencia de la temprana modernidad europea «está dotado con el importante poder de establecer los hechos», pues su «subjetividad es su objetividad». (Haraway, 2004: 35). Abrogándose el conocimiento de la verdad a partir de su testimonio modesto, el hombre no solo será el juez legítimo de lo verdadero, también será el padre genuino de un mundo tecnificado.

Todos estos complejos intentos por crear seres artificiales autónomos e inteligentes encuentran sus más fructíferos desarrollos durante y después de la Segunda Guerra Mundial. Peter Galison (1994) analiza cómo la cibernética, entendida como una de las ramas que contribuyó decisivamente en el desarrollo del campo de la IA, estuvo atravesada por la visión del mundo de Norbert Wiener.

Desde ya antes del inicio de esta guerra, Wiener estaba interesado en ofrecer sus conocimientos para el desarrollo de estrategias defensivas militares, especialmente en el campo de las defensas antiaéreas. Al iniciar la guerra y «durante los años siguientes, la atención de Wiener se interesó principalmente en el problema de cómo destruir el aeroplano enemigo», en lo que se denominó como el *Antiaircraft Predictor* (AA), un problema que necesitaba de la visión de la máquina para identificar, clasificar y atacar al adversario. Es precisamente sobre la base de esta lógica del enemigo, bajo la cual se imaginaba Wiener el funcionamiento del mundo en la posguerra, que nace la cibernética: «Finalmente, el predictor AA, junto con sus nociones de ingeniería asociadas a sistemas de retroalimentación y cajas negras, se convirtió, para Wiener, en el modelo para una comprensión cibernética del universo» (Galison, 1994: 229). Además de este enemigo, emerge un tercero, que Galison llama Enemy

Other<sup>2</sup>, calculado en los laboratorios del MIT a partir de tres disciplinas articuladas: «investigación de operaciones, teoría de juegos (...) y cibernética».

Los servomecanismos, identificados por Le Roy Archibald MacColl con «sistemas esclavos que emplean retroalimentación» (MacColl, 1945: 8), son esenciales en la teoría cibernética. Para Wiener, «el enemigo servomecánico se convirtió (...) en el prototipo de la psicología humana y, en últimas, de toda la naturaleza humana» (Galison, 1994: 233). El ML, «el mayor experimento de clasificación en la historia humana» (Crawford 2017), es heredero directo de una idea de control sobre el mundo, postulada por Wiener: «la idea de Wiener fue utilizar redes eléctricas para determinar, varios segundos por adelantado, dónde podría ocurrir un ataque aéreo y emplear ese conocimiento para dirigir el fuego de la artillería» (Galison, 1994: 234). La identificación y la clasificación fueron campos cruciales en la aplicación de esta visión de la guerra y de la sociedad. «De particular importancia, sostuvieron que su clasificación rehabilitaba el “propósito” y la “teleología” al ponerlos bajo los auspicios de un “análisis conductista uniforme” que era igualmente aplicable a los organismos vivos y las máquinas» (Galison, 1994: 245).

La cibernética, como una rama fundamental para el desarrollo contemporáneo de la IA, es postulada por Wiener no solo como una forma efectiva de controlar las acciones del enemigo militar, sino también de ese *Enemy Other* aplicado a la sociedad misma: «los mensajes pueden ser enviados con el propósito de explorar el universo, pero pueden también ser enviados con la intención de controlarlo» (Galison, 1994: 256).

Paul Edwards (1997) demuestra, además, cómo estas ideas fueron fundamentales en la psicología cognitiva y en el campo de la IA: «Las teorías científicas de la inteligencia artificial y de la psicología cognitiva también forman núcleos de ideologías sobre la mente humana como máquinas manipulables» (Edwards, 1997: 21). Así, el computador se convirtió en una «metáfora culturalmente central para el control, para el análisis científico y para la mente» (1997:28).

En este contexto de un mundo analizado y controlado por medio de servomecanismos de retroalimentación computacionales, la película *Colossus: The Forbin Project*, lanzada públicamente en el año de 1970, recrea la construcción de una superinteligencia computacional por parte del científico Charles Forbin, que tiene la capacidad de monitorear todas las formas de comunicación en plena Guerra Fría, en un presentimiento tecnológico materializado por medio de la IA en el programa PRISM, de la NSA. Una escena de la película de *Batman, The Dark Knight*, de 2008, recrea visualmente esta vigilancia y control panóptico con un Morgan Freeman frente a una pared abarrotada de pantallas que presentan modelos 3D en tiempo real, basados en la

geolocalización de los datos generados por los teléfonos móviles de todos los ciudadanos de Gotham City.

Para Steve Anderson, el asunto de la vigilancia y del control no es en el fondo un simple problema tecnológico; más bien, se trata de una visión del mundo materializada, en nuestro caso, en los actuales sistemas de IA y, en particular, en las RNP: «Aún la vigilancia computacional no es primariamente un asunto tecnológico; está más propiamente enmarcada como el ejercicio de poder por parte de estructuras institucionales sobre grupos e individuos» (Anderson, 2017: 133).

Las RNP, sin embargo, han creado su propio accidente. Los *Adversarial Examples*, en el campo de la visión algorítmica, son instancias que impiden la correcta identificación y clasificación de personas, cosas y acciones (Szegedy *et al.*, 2014). En mi proyecto de investigación visual intitulado *Inspección Entheo-Rítmica*, de 2020<sup>33</sup>, un vídeo construido a partir de un set de imágenes diseñadas por Nguyen, Yosinski y Clune (2015), me he permitido reflexionar sobre los límites de la vigilancia y el control, al experimentar con la identificación y la clasificación de los objetos que allí supuestamente surgen a la luz de la percepción del modelo algorítmico (ver fig. 1).

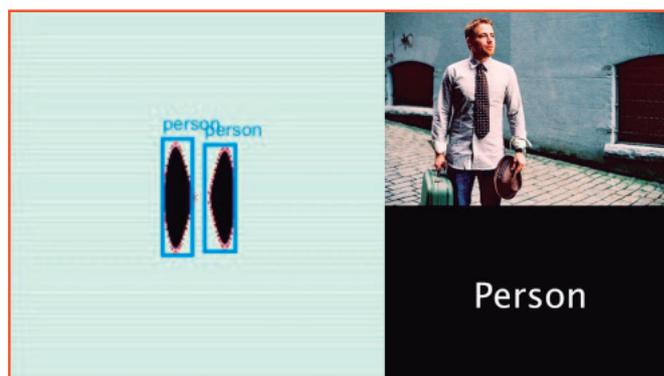


Fig. 1. Fotograma del vídeo resultado del proyecto de investigación *Inspección Entheo-Rítmica*.

Las más de cinco mil imágenes que componen el set de datos de Nguyen, Yosinski y Clune fueron analizadas por el modelo de identificación y clasificación COCO-SSD —que es un conjunto de datos de detección, segmentación y subtítulos de objetos a gran escala—, por medio de la aplicación Run-way, ejecutada en un servidor local. Las imágenes que efectivamente fueron etiquetadas por el modelo se aislaron y agruparon con el objetivo de contrastarlas visualmente con fotos de objetos de la misma categoría, reconocibles por la visión humana y alojadas en el set de datos de entrenamiento de COCO-SSD. Así, el espectador puede experimentar el abismo que aquí separa

2. «Enemigo otro».

3. <http://www.hugoidarraga.info> <https://www.youtube.com/watch?v=qGKAq03T6Zs>

la percepción visual de las RNP, de la percepción visual humana, contrastando la realidad creada por cada uno de estos modelos perceptivos, ahondando en los límites y errores de la visión algorítmica, e imaginando narrativas alternativas a las tareas de identificación y clasificación que estos sistemas cumplen con el objetivo de vigilar y controlar.

En un sentido parecido a los trabajos de Adam Harvey, Sterling Crispin o de Joselyn McDonald y su *Mother Protect Me*, de 2019, *Inspección Entheo-Rítmica* reflexiona sobre la realidad algorítmica y las maneras en las que la máquina podría alucinar un mundo irreconocible para el ser humano. *Inspección Entheo-Rítmica*, además de generar tensiones críticas entre la percepción humana y la percepción maquínica, pone presente las diversas formas en las que las tareas de identificación y clasificación pueden ser *hackeadas*, aprovechando el error inherente a las RNP, abordando críticamente el alcance de la vigilancia y el control de los sistemas algorítmicos.

## Conclusiones

Este artículo ha intentado demostrar que, junto a las tareas serviles y de entretenimiento, las de vigilancia y control social han sido históricamente predominantes en lo que atañe a la consecución de seres con inteligencia autónoma; que, para realizarlas, la identificación y la clasificación son esenciales para comprender su funcionamiento. Además, que estas tareas emergen de una visión patriarcal del mundo, en el que personas, animales y cosas deben estar bajo la disposición de su poder. Con ello se ha pretendido proponer una discusión *acerca de* los proyectos de vigilancia y control que hoy se encarnan en los sistemas de ML y las RNP, reclamando un interés especial por las consecuencias políticas y sociales que se derivan de su implementación cada vez más cotidiana y globalizada, a la vez que los intereses que se esconden detrás de sus publicitados beneficios. Las explicaciones mágicas que han empezado a emerger para tratar de aclarar el misterio detrás de las RNP terminan, como se ha afirmado, excusando la responsabilidad que sujetos e instituciones deberían asumir respecto a las consecuencias de la implementación masiva de estos sistemas.

Las prácticas artísticas aquí vinculadas proponen no solo una mirada crítica sobre el funcionamiento de la IA, también ofrecen alternativas, por un lado, sobre las maneras diferentes en las que podemos relacionarnos con estas tecnologías y, por el otro, sobre las posibilidades estéticas y políticas que la interacción entre humanos e IA puede albergar. Desnudando la neutralidad de los set de datos, deslegitimando el automatismo de los algoritmos como excusa de la responsabilidad de quienes diseñan y desarrollan, pero también criticando los supuestos que se encarnan en sus funcionamientos, el arte en el campo de la IA construye lugares desde donde podemos pensar y crear otras realidades.

Los límites de este artículo impiden abordar en profundidad todas las problemáticas relacionadas con la identificación y la clasificación, con la vigilancia y el control, por ejemplo, sobre los actores involucrados, sobre las muy diversas tecnologías que encarnan actualmente esta visión del mundo o las más recientes teorías sobre el aprendizaje profundo, supervisado o no. Sin embargo, espero haber abordado los puntos neurálgicos de los que en un futuro inmediato dependerá nuestra relación con la tecnología, así como el tipo de sociedades que junto a ella queremos promover.

## Referencias bibliográficas

- Anderson, Steve F. 2017. *Technologies of Vision: The War Between Data and Images*. Cambridge, Massachusetts: The MIT Press.
- Aristóteles. 2004. *Política*. Madrid: Gredos.
- Benjamin, Ruha. 2019. *Race After Technology: Abolitionist Tools for the New Jim Code*. 1.ª ed. Nueva York: Polity.
- Bunz, Mercedes. 2009. «Are Hewlett-Packard computers really racist?». *The Guardian*, December 23, <https://www.theguardian.com/media/pda/2009/dec/23/hewlett-packard>.
- Brown, Steven T. 2008. «Machinic Desires: Hans Bellmer's Dolls and the Technological Uncanny in *Ghost in the Shell 2: Innocence*». *Mechademia* 3(1): 222-53. <https://doi.org/10.1353/mec.0.0088>.
- Campolo, Alexander y Kate Crawford. 2020. «Enchanted Determinism: Power without Responsibility in Artificial Intelligence». *Engaging Science, Technology, and Society* 6(0): 1-19. <https://doi.org/10.17351/ests2020.277>.
- Crawford, Kate. 2017. «The Trouble with Bias». *Keynote lecture held at NIPS*, Conference at Long Beach, California, December 4-9. [https://www.youtube.com/watch?v=fMym\\_BKWQzk](https://www.youtube.com/watch?v=fMym_BKWQzk).
- Dreyfus, Hubert L. 1965. *Alchemy and Artificial Intelligence*. Santa Monica, CA: RAND Corporation. <https://www.rand.org/pubs/papers/P3244.html>.
- Dreyfus, Hubert L. 1972. *What computers can't do: a critique of artificial reason*. New York: Harper & Row.
- Dzieza, Josh. 2020. «How hard will the robots make us work?». *The Verge*, February 27. <https://www.theverge.com/2020/2/27/21155254/automation-robots-unemployment-jobs-vs-human-google-amazon>.
- Edwards, Paul N. 1997. *The Closed World: Computers and the Politics of Discourse in Cold War America*. Cambridge, Mass. London: MIT Press Ltd.
- Galison, Peter. 1994. «The Ontology of the Enemy: Norbert Wiener and the Cybernetic Vision». *Critical Inquiry* 21(1): 228-266.
- Homero. 2010. *Iliada*. Madrid: MESTAS Ediciones.
- Haraway, Donna Jeane. 2004. *Testigo\_Modesto@Segundo\_Milenio. HombreHembra@\_Conoce\_Incoración@*. Barcelona: Editorial UOC, S.L.

- Klingemann, Mario. 2018. «Neural Glitch». *Quasimondo*. *Accesado*, March 2, 2020. <http://underdestruction.com/2018/10/28/neural-glitch/>.
- Nguyen, Anh, Jason Yosinski y Jeff Clune. 2015. «Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images». *arXiv:1412.1897 [cs]*, April. <http://arxiv.org/abs/1412.1897>.
- MacColl, Le Roy Archibald. 1945. *Fundamental Theory of Servomechanisms*. 1.a ed. First Printing. D. Van Nostrand Company, inc.
- McCorduck, Pamela. 2004. *Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence*. 2.a ed. Natick, Mass: A K Peters/CRC Press.
- na, Mimi. 2018. «How the Arts Help Us Understand Our Relationship With Technology», video. [https://www.youtube.com/watch?v=\\_JJ60j\\_u-Qw](https://www.youtube.com/watch?v=_JJ60j_u-Qw).
- Paglen, Trevor y Kate Crawford. 2019. «Excavating AI». <https://www.excavating.ai>.
- Platón. 2011. *La República o El Estado*. Traducido por Patricio de Azcárate Corral. Madrid: Austral.
- Raviv, Shaun. 2020. «The Secret History of Facial Recognition». *Wired Magazine*, January 21. <https://www.wired.com/story/secret-history-facial-recognition/>.
- Truitt, Elly R. 2015. *Medieval Robots: Mechanism, Magic, Nature, and Art*. Pennsylvania:University of Pennsylvania Press. <http://muse.jhu.edu/book/39875/>.
- Szegedy, Christian, Wojciech Zaremba, Ilya Sutskever, Joan Bruna, Dumitru Erhan, Ian Goodfellow y
- Rob Fergus. 2014. «Intriguing properties of neural networks». *arXiv:1312.6199 [cs]*, February. <http://arxiv.org/abs/1312.6199>.

## CV

---



### Hugo Felipe Idárraga

Universidad de Los Andes

[h.idarraga@uniandes.edu.co](mailto:h.idarraga@uniandes.edu.co)

Filósofo de la Universidad Nacional de Colombia y Magister en Comunicación de la Pontificia Universidad Javeriana. Interesado en temas de teoría de la imagen, visión por computador e inteligencia artificial. Actualmente estudiante de la Maestría en Humanidades Digitales de la Universidad de Los Andes.