



Reflexiones sobre la democracia y la verdad en la era de la IA (La posverdad y la normalización de las mayorías)

*Reflections on democracy and truth in the era of AI
(Post-Truth and The normalization of the majorities)*

Guillermo José Mañón Garibay*

Recibido: 25 de abril, 2024. Aceptado: 6 de septiembre, 2024.

Resumen: Desde una cierta perspectiva, la IA y la democracia están en tensión, sobre todo cuando se considera que los algoritmos utilizados están diseñados con el fin de ganar y/o socavar la confianza en proyectos comunitarios. Ciertamente, desde tiempos antiguos, el pensamiento colectivo y la opinión pública han estado modelados por discursos persuasivos que aspiran a cambiar la opinión o a ponerla a la deriva. Pero ahora, en la era de la posverdad e IA, es cuando el sector público y privado (i. e., el poder institucional y factico) echan mano de herramientas como el Chatbot o el Chat GPT con el fin de validar sus intereses sesgados.

Palabras clave: Filosofía, democracia, verdad, Inteligencia Artificial, opinión pública.

Abstract: From a certain perspective, AI and democracy are in tension, especially when considering that the algorithms used are designed with the aim of gaining and/or undermining trust in community projects. Certainly, since ancient times, collective thought and public opinion have been shaped by persuasive speeches that aspire to change opinion or set it adrift. But now,

* Doctor en Filosofía por la Universidad de Rostock, Alemania. Maestro en Filosofía por la Universidad Alexander Von Humboldt de Berlín, Alemania. Investigador Titular de Tiempo Completo en el Instituto de Investigaciones Jurídicas, UNAM. Miembro del Sistema Nacional de Investigadores: Nivel I. Líneas de investigación: Bioética, filosofía y objeción de conciencia.

in the era of post-truth and AI, is when the public and private sector (i.e., institutional and factual power) use tools such as the Chatbot or the GPT Chat in order to validate their biased interests.

Keywords: Philosophy, Democracy, truth, artificial intelligence, public opinion.

INTRODUCCIÓN

La era de la posverdad surge de la mano de la IA-generativa; un maridaje que engendra a la “megamáquina” como sociedad totémica, jerárquica e inflexible, de la que hablaba desde 1967 Lewis Mumford¹.

Si con la IA se puede hacer decir a alguien cualquier cosa, entonces, es necesario o bien contar con mejores fuentes y herramientas de información o bien considerar el advenimiento inevitable de la era de la posverdad y sus peligros concomitantes. La era de la posverdad no es el momento en que la mentira parece tan verdadera como la verdad misma, sino el tiempo en que la verdad ya no importa porque ésta no es una y única, sino múltiple y acomodaticia.

La simulación puede ser muy poderosa si se analiza la convergencia entre el poder de esta tecnología y el poder político y económico. Entonces, la simulación inteligente tiene una efectividad política y económica sin igual.

¿Qué pasa cuando en la política se emplean modelos digitales con la capacidad de ejercer influencia en el gran público? ¿Basta una serie de ChatsBots para modificar opiniones e incluso valores? El caso ejemplar tuvo lugar en el año 2000 con los videojuegos los Sims², cuyos creadores experimentaron exitosamente con avatares y decisiones inteligentes simuladas, a la manera en que, en 1938, Orson Welles logró también un engaño perfecto con su *Guerra de los mundos*.

LA DEMOCRACIA Y LA IA

Desde una cierta perspectiva, la IA y la democracia están en tensión, sobre todo cuando se considera que los algoritmos utilizados están diseñados con el fin de ganar y/o socavar la confianza en proyectos comunitarios. Ciertamente, desde tiempos antiguos, el pensamiento colectivo y la opinión pública han estado modelados por discursos persuasivos que aspiran a cambiar la opinión o a ponerla a la deriva³. Pero ahora, en la era de la posverdad e IA, es cuando el sector público y

¹ Mumford plantea en sus libros el conocido dilema entre civilización y cultura, entre desarrollo técnico-científico y humanismo, y sus repercusiones en la biopolítica y biotecnología. Un problema desde el siglo XIX, porque, si en el pasado, la tecnología había estado subordinada a la inteligencia humana, ahora, sin embargo, parecería lo contrario: la inteligencia humana subordinada a la tecnología. Para Mumford, esto tiene como resultado la *Megamáquina*, la convergencia de tecnología, economía y política o la instrumentalización de la dos últimas por la primera.

² Ver: <https://www.ea.com/es-mx/games/the-sims/the-sims-4>

³ Henning Ottmann (2001), en su libro sobre el pensamiento político de los sofistas en la antigüedad griega, hace ver claramente cómo ellos utilizaban la retórica y argumentación para alcanzar sus metas políticas.

privado (i. e., el poder institucional y factico) echan mano de herramientas como el Chatbot o el Chat GPT con el fin de validar sus intereses sesgados.

POPULISMO DIGITAL

El concepto *deepfake* (gran impostura⁴) revela la capacidad de mentir y el ocaso de la rectitud política. La propagación acelerada de contenidos falsos y de alta verosimilitud contrasta con la lentitud en el desarrollo de herramientas para detectarlos (sumado a la ausencia de derechos sobre imágenes). El peligro es encontrarse una y otra vez con la prueba de Alan Turing, o sea, con el desafío de la máquina embaucadora que crea o genera imágenes y voces indistinguibles, infunde motivaciones ocultas, esconde su propósito original y apela a emociones recelosas, como la vergüenza, lástima o el miedo.

Porque más allá de la máquina de Turing, existe la inclinación a creer en algo superior, en someterse dócilmente a una autoridad, a obedecer sin cuestionar, siempre y cuando las circunstancias emocionales sean las indicadas⁵. Esto altera el sentido y/o socaba la confianza a través de figuras o expresiones ficticias, pero dotadas de poder emocional. Lo que desafía el fenómeno social de la comunicación y la capacidad de discernimiento, ya que, aunque sea falso el video, las emociones causadas son verdaderas.

La sociedad se acostumbra a ciertas conductas, sobre todo cuando vienen de la mano de figuras populares. De eso se aprovecha el *populismo digital*, que promueve la admiración por figuras creadas en las redes sociales o las interacciones masivas falsas que favorecen intereses políticos difíciles de verificar. Así es como la biopolítica conquista la gobernanza de naciones y se transforma en antropotécnica⁶.

Hoy día, la posverdad cohesiona a la población mediante la verdad y el lenguaje de la *deepfake*. Se multiplican las pruebas de Turing, se multiplican los desafíos a la sensibilidad y pensamiento crítico, desaparecen los límites entre ficción y realidad. ¿Dónde queda la responsabilidad por apelar al narcisismo o a las tendencias más radicales? La convergencia de la biopolítica con la tecnología de la IA engendra la herramienta perfecta de desinformación, la era de la posverdad, sin criterios definidos, falsos ídolos y procesos democráticos simulados.

Esto arroja el siguiente problema: el de la verdad. ¿Qué ha pasado con ella en la era de la posverdad?

⁴ Lo relevante de *Deepfake* en este trabajo no es tanto la “capacidad de mentir” como la intención de hacerlo (Tampoco cómo se define o concibe la verdad). Claro que cuando se afirma que hay una mentira se asume que ésta es falsa, que hay una intención de hacer pasar la falsedad por verdad y que hay algo que es verdadero. Pero, de todo esto, lo más importante en este trabajo es la intención de mentir. Por tanto, la teoría de verdad que se defiende es, hasta cierto punto, irrelevante.

⁵ La denuncia del hombre y su miedo a la libertad (o su deseo a ser engañado) está presente en el arte de Rene Magritte (1898-1967). Sobre la serie de pinturas tituladas “la condición humana” de Rene Magritte, ver: <https://historia-arte.com/obras/la-condicion-humana-de-magritte>.

⁶ El sociólogo alemán Thomas Lemke ha llevado a cabo con sus estudios (2013) la relación entre la biopolítica y la biotecnología y la antropotécnica.

LA VERDAD FILOSÓFICA Y LA VERDAD COMO FUNCIÓN

La realidad virtual y la IA-generativa ponen en la palestra el problema de la verdad. Y la alegoría de la caverna, la manera en que este problema debe ser abordado, a saber: como un problema entre realidad y apariencia, y no tanto entre verdad y falsedad⁷. Dentro de las teorías de la verdad existen perogrulladas como el que la verdad es la totalidad o la verdad es la existencia o, la más común, que la verdad es correspondencia⁸. De todas ellas, ¿cuál se ajusta mejor al problema que suscita la realidad virtual y la IA? Porque con la realidad virtual y la IA el problema no es el de distinguir entre verdad y falsedad, sino entre verdad y apariencia engañosa, entre verdad y disimulo.

Es sabido que la verdad es un atributo de las sentencias y no de las cosas: “el pasto es verde”, puede ser verdadero o falso; pero “la yerba verde”, llamada “pasto”, es, sin predicado (lógico) alguno. En el caso de la caverna de Platón, lo que está en juego es la susceptibilidad al engaño y/o a tomar la apariencia como la realidad misma. Platón propone más que el problema de la verdad en sí, el temple para encarar la verdad o seguirse engañando. Entonces, verdad/falsedad, más allá de ser atributos de las sentencias (i. e., oraciones declarativas), son “atributos” de las personas que desean o no ser veraces o falsarios. El desenlace es muy claro en el libro VII de la República: aquel que logra salir de la caverna y ver las cosas como son, vuelve al interior con dificultad y contra su voluntad, con la tarea de enfrentar el rechazo y oposición de sus congéneres a ser despabilados y convertidos a la rectitud, al mundo fuera de la caverna.

¿Cuál es específicamente el problema de la verdad en el caso de la IA? Por un lado, el de la actitud adecuada; porque la verdad tiene una función orientadora que permite saber cómo actuar adecuadamente en determinadas situaciones y con determinadas expectativas. Quién cree en A⁹, confía y otorga credibilidad a actuar según sus creencias, o sea, confiando en que así se cumplirán ciertas expectativas. Por eso, es importante que una foto o video sean reales o verdaderos, porque, además de su aquiescencia, comprometen a las personas a determinadas acciones. Sin embargo, es sabido que, en la guerra y en la política, la primera víctima es la verdad, ya que aquello que importa es la animosidad y no la evidencia.

Cuando en el contexto de la realidad virtual y de la IA se pone en la palestra el problema de la verdad, resulta baladí insistir en que ésta se predica sólo de sentencias declarativas y no de las cosas, porque el talante de la cuestión es de carácter ontológico (y no lógico o epistemológico): ¿son reales o irreales, dadas o imaginadas las cosas?, pero no: ¿son verdaderas o falsas? Ahora bien; cuando lo imaginado se presenta como lo real y lo real como imaginado, con el fin de engañar, entonces, no están en juego las palabras y las cosas, sino las intenciones y los emprendimientos. Por tanto, el debate no alude a si una sentencia como “el pasto es verde” es verdadera y otra como “el pasto no es verde” es falsa; porque no está en juego la relación semántica entre palabras y cosas, sino entre las palabras y la impostura.

⁷ Antes bien: verdad v.s. verisimilitud o verdad y apariencia engañosa de verdad.

⁸ La clásica concepción de la verdad como correspondencia la hizo santo Tomás de Aquino: *adaequatio intellectus ad rem*. O: *Veritas consistit in adaequatione intellectus et rei, sicut supra dictum est*. Ver: S.Th. I, q. 16, a. 1.

⁹ Platón analiza el problema de la verdad y el conocimiento por lo menos en 4 diálogos, a saber: Fedón, Teeteto, Sofista y Parménides. En resumen, puede decirse que conocer A, es creer que A, tener razones para creer que A, y que sea el caso que A.

En el criterio de verdad por correspondencia se dice que una afirmación verdadera es aquella que corresponde uno a uno con la realidad¹⁰; pero, en el caso de la falsedad de la virtualidad (o los engendros de la IA), no hay siquiera esta posibilidad, porque la virtualidad se presenta como la realidad sin más. La realidad virtual y la IA no son meros ardides, e insistir en el desafío al criterio de verdad sería como preguntarse si existe una verdad más allá del lenguaje o más allá de la realidad, lo que es absurdo. Entonces, dos cosas: o bien la verdad como adecuación es insuficiente (exactamente porque lo que se cuestiona es lo real y no la adecuación), o bien los criterios de verdad/falsedad no aplican al caso de la virtualidad y de la IA frente a otra distinta, como la realidad del sueño y la vigilia. En este caso, verdad/falsedad se cuecen aparte, a saber: en una nueva realidad. Si no hay acceso a lo real fuera del lenguaje (sobre esta o aquella realidad), ¿cómo entonces calificar una afirmación sobre la IA como verdadera o falsa respecto a otra distinta? Sí se le puede calificar de engañosa, pero sólo cuando suplanta a otra, porque, entonces sí, la virtual es apariencia o disimulo de lo real-real.

SOCIOLOGÍA FUNCIONAL

Si los criterios filosóficos de verdad no sirven; entonces, en lugar de la filosofía habría mejor que incursionar en la sociología funcional; de este modo, la pregunta es, sociológicamente hablando, “¿qué función cumple la verdad?”, y no “¿qué es la verdad?”, del mismo modo que se pregunta por la función social que cumplen fenómenos como el del arte, dinero o del amor¹¹. ¿Resuelve la verdad algún problema social?

Según Luhmann, socialmente hablando, la verdad de una sentencia desempeña un papel en cuanto logra consensos. En términos de teoría de sistemas, como la de Niklas Luhmann¹², la verdad existe cuando dos interlocutores están de acuerdo en que una comunicación seleccionada debe ser considerada como una experiencia del mundo. Ciertamente, la comunicación de suyo concita el acuerdo, y aún más cuando hay acuerdo sobre lo que es verdadero (o falso), porque, entonces, ambos interlocutores funcionan socialmente. El consenso sobre lo verdadero da lugar a la realidad común; y no sólo como acto comunicativo, sino como acto social de un conjunto de personas, del entorno y de las consecuencias futuras de la comunidad entera¹³.

La verdad por consenso resuelve los problemas de la complejidad (por reducción de perspectivas) y allana la coordinación de la negociación y el desempeño social. El acuerdo sobre qué es verdad y qué no facilita cualquier negociación. El mundo de lo social es más simple cuando hay consenso.

Entonces, más allá de la semántica está la pragmática y lo convencional. Cuando millones de personas creen en el *deepfake*, y esto facilita su interacción, hay verdad (verdad por consenso).

¹⁰ Como en el caso extremo del atomismo lógico (del primer Wittgenstein, i. e., del *Tractatus logico-philosophicus*), según el cual la verdad de una declaración reside en el análisis de sus elementos componentes y de su correspondencia uno a uno con la realidad.

¹¹ Niklas Luhmann analiza el fenómeno del amor en dos obras, la más relevante aquí es *Liebe als Passion*.

¹² En biología (ecología) y sociología el enfoque sistémico es ineludible, como muestra Luhmann en *Systemtheorie der Gesellschaft*.

¹³ En este sentido, hay una “religión verdadera”, unos “ritos mágicos” curativos verdaderos (o auténticos), verdaderos hábitos y costumbres (i. e., nacionales v. s. importados).

No hay que temer a las verdades-falsas, sino al absurdo de la teoría de la verdad entre sistemas inconmensurables entre sí¹⁴.

Para Luhmann, la verdad por consenso asegura el éxito en la medida en que opera (en términos sociales). La distinción verdadero/falso existe funcionalmente, para adaptarse a los engranajes de la sociedad y, por el contrario, la mentira dificulta el engranaje (aunque únicamente sea en uno de sus segmentos¹⁵).

La reticencia para aceptar la verdad como función reside en que, en la sociedad actual, tan grande y compleja, sólo se admita un principio o verdad que la abarque toda por completo, ya que esto daría lugar a una especie de dictadura ilustrada o del pensamiento único (incluso cuando se trata del científico). Pero la verdad con minúsculas, con un alcance limitado a uno sólo de sus subsistemas sociales¹⁶, permite el juego de interacciones sociales a la vez que la pluralidad y tolerancia de enfoques; porque a nadie extraña que la verdad sea indispensable en la ciencia, pero irrelevante en el arte, en las actividades de esparcimiento y la política¹⁷.

DE LA DEMOCRACIA DIGITAL

En democracia, los contenidos están en competencia, lo que significa que el éxito de unos sobre otros determina la verdad. En las dictaduras y autocracias es la fuerza y no el consenso lo determinante. La realidad virtual y la IA-generativa no empeoran las cosas al repetir “mentiras”, anular o secuestrar los medios de comunicación, porque esto siempre ha sucedido y porque la relación entre IA y verdad, más allá de la realidad virtual, es irrelevante.

Si en la política es irrelevante la verdad, entonces no se concita acuerdo a través de la razón sino de la pasión y el engaño. ¿Qué tan inteligente es la IA para crear el engaño?

LA INTELIGENCIA DE LA IA PARA CREAR NORMALIDAD Y MAYORÍAS

En la antigüedad helena, los oráculos conectaban al hombre con los dioses y con los misterios del incierto futuro. Hoy día, el lugar de la pitonisa lo ocupa la IA, porque se cree que posee la capacidad de predecir el clima, realizar un pronóstico médico a partir del historial clínico y seleccionar música

¹⁴ Sobre el tema de la inconmensurabilidad entre paradigmas es necesario atender a Th. Kuhn (1922-1996).

¹⁵ La importancia y concepción de la verdad en la sociedad de físicos es muy otra a la verdad religiosa entre feligreses. Ver: Luhmann, Niklas, *Die Gesellschaft der Gesellschaft*.

¹⁶ Contra la sociedad vista monóticamente está la concepción de la misma como un conjunto con muchos subconjuntos o sub-sistemas, tema de Luhmann en *Die Gesellschaft der Gesellschaft*.

¹⁷ Tanto el deporte como las actividades de tiempo libre (ocio) habían sido descuidadas hasta que llegaron los estudios de Norbert Elías y Eric Dunning e hicieron ver su papel en el proceso de civilización.

o videos anticipando los antojos personales. Todo esto hace pensar en un adivino moderno con la capacidad de anticipar la toma de decisiones, desde las más triviales hasta las medulares¹⁸.

Hasta ahora parecía que todas estas eran tareas propias y exclusivas de la inteligencia humana, porque demandaban raciocinio y creatividad (incluso las artísticas). Sin embargo, parece que la IA ha tomado el relevo y que el algoritmo computacional tiene la habilidad de crear contenido lo suficientemente creativo e inteligente como para deslumbrar incluso a los más expertos. ¿Cómo lo hace? ¿Es verdaderamente inteligente? ¿Qué significa ser inteligente?¹⁹

LA INTELIGENCIA Y CREATIVIDAD HUMANA

Una de las características de la inteligencia humana es la creatividad. Y desde Descartes, la posibilidad de máquinas inteligentes (pese a ser por definición predecibles, inflexibles y limitadas) está en la palestra de la reflexión filosófica. Hoy día, ese atributo se le adjudica a la Inteligencia Artificial (IA), por eso no es baladí preguntarse qué tan inteligente o creativa es la IA.

Inteligencia, creatividad y/o conciencia, son conceptos ampliamente discutidos por la filosofía, siempre teniendo presente la advertencia agustina que dice: “si no me lo preguntas lo sé; si me lo preguntas, no lo sé”²⁰. Lo que significa que existe una indudable dificultad en definir estos conceptos, pese a que todo mundo cree saber lo que significan cuando los usa. Más allá de esto, existe un temor a que la IA sea verdaderamente inteligente y creativa y arrebatase al hombre su lugar privilegiado dentro del universo.

Si se indaga la inteligencia por el lado de su creatividad, entonces, *prima facie*, creatividad (de la IA) significa la capacidad de producir algo nuevo a partir de ciertos elementos dados. Alguien más puntilloso podría considerar la novedad no sólo respecto a los elementos componentes, sino frente a todo lo habido hasta entonces²¹. En este caso, escasamente existiría la creatividad en el género humano y sólo en el divino (i. e., crear de la nada). Si se hace a un lado este requerimiento desmedido, ciertamente faltaría por averiguar si la novedad es respecto al uso o a la concepción (i. e., fuera de su aplicabilidad inmediata).

¹⁸ Empleadores japoneses, preocupados por la rotación de personal en sus empresas, podrán con una app de la IA anticipar quiénes serán sus próximos empleados, esto revela el estudio de: https://www.dw.com/es/herramienta-de-ia-logra-predecir-cuándo-renunciarás-a-tu-empleo/a-68871440?fbclid=IwZXh0bgNhZW0CMTEAA-R1AEBYtfJv7V66s9BmFOqW-9wPoG7rM_5xMT2QCbgNMTXORXb8IVMAaOMc_aem_AdlfPzmVMnUJFSe1pV-VIsXHRkVeeYXpnpADaba2tm8SVC80E3uxAl1cxgWIN4CWvIkmdXTVatzXXQMg07aAV_Yi0

¹⁹ Las llamadas redes neuronales de los ordenadores asemejan la compleja conexión de las neuronas cerebrales y los procesos mentales que hacen posible. Dentro del *Machine Learning*, el lenguaje Python, unos de los más eficientes.

²⁰ Es muy conocida la preocupación de Aurelio Agustín por el lenguaje en su obra *Confesiones*, libro XI.

²¹ La novedad debería ser considerada también respecto a los diferentes tipos de inteligencia (y, por qué no, las teorías didácticas). Las inteligencias múltiples han sido analizadas, en general, por Howard Gardner; mientras que Robert Sternberg ha estudiado la diversificación de la inteligencia por su contexto cultural; y, finalmente, a Daniel Goleman y Antonio Damasio se deben los trabajos sobre la diversificación de la inteligencia por su interacción con las emociones.

¿Qué tan inteligente y creativa es la IA?

Vencer en un juego de ajedrez a partir del aprendizaje de reglas y estrategias dadas, pintar un cuadro a partir de indicaciones verbales (y sus asociaciones simbólicas previas), o componer música a partir del acervo digital en la red ¿es o no inteligencia creativa?

Los mitos sobre la creatividad como una caja oscura (i. e., inspiración, revelación o unión mística) provienen de una antropología filosófica que considera al hombre como algo tan extraordinario como la imagen misma de Dios (Dios el absoluto creador y el hombre su remedo).

Hoy día, las concepciones del hombre predominantes no son religiosas o filosóficas, sino biológicas y computacionales. Por ello, en la indagación sobre el cerebro y sus funciones superiores afloran conceptos como *pensamiento de diseño*, *redes neuronales*, *procesamiento analítico de datos*, *sistemas cognitivos computacionales*, *sistemas expertos o dinámicos*, etc., obtenidos principalmente de la teoría de sistemas computacionales.

Lo que tiene una historia que se remonta a 1943, cuando Alan Turing (1912-1954), al analizar la máquina nazi *Enigma*, diseñada para enviar mensajes encriptados durante la Gran Guerra, se percató de su similitud con la capacidad humana de pensar y crear un discurso coherente. Así nació su interés por averiguar hasta qué punto las máquinas pueden pensar o imitar procesos mentales inteligentes. Turing, en 1950, realizó una prueba para evaluar la capacidad de la máquina para conversar con personas, sin que estas percibieran el fraude²². Si las personas no notaban la diferencia, entonces la máquina pasaba la prueba (bot conversacional) y podía decirse que pensaba (o imitaba procesos mentales inteligentes). El experimento tuvo éxito y Turing celebró el fin de la era mentalista y el advenimiento de la fisicalista o de las explicaciones racionales sobre los procesos mentales.

No mucho tiempo después vino el mentís a los logros de Turing con *The Chinese Room Argument* del filósofo John Searle (1980²³) y el *Stochastic Parrots* de la lingüista Emily Bender (2021²⁴), que pusieron de manifiesto el fiasco de la máquina de Turing, capaz de armar discursos sin comprenderlos. Claro que, aun admitiendo que una máquina no piensa, sí engaña al interlocutor simulando que lo hace.

LA CREATIVIDAD DE LA IA

La pregunta es si la IA puede liberarse de su circunscripción a una programación dada y saltar a la dimensión de lo original, entendido como inesperado, no antes calculado, ni previsto o predeterminado. Lo que implica, además del acopio y la transformación de contenidos, el juicio o evaluación subjetiva. Porque todo juicio de creatividad está sujeto a una buena dosis de subjetividad, ya que supone una mayor o menor red de conexiones entre el producto artístico presente y todos aquellos del pasado e, incluso, su influencia en los posibles venideros.

²² Esto es: simular inteligencia en la medida en que sostiene una conversación.

²³ El clásico argumento del cuarto chino se debe al americano John Searle, y, aunque sea de los años 80s del s. XX, sigue siendo muy interesante.

²⁴ El *mono estocástico* de Emily Bender tiene en mi opinión mayor interés que el texto de Searle, porque lo hace desde la lingüística.

En el caso de la IA-generativa que, considerada por definición como incapaz de subjetividad y reducida a su condición de objeto, se le considera imposibilitada de inteligencia y creatividad efectivas o sólo de manera metafórica o figurada. Sin embargo, si se recapacita sobre el producto de una IA-generativa, se concederá, sin incurrir en una petición de principio, que, como cualquier otro agente creativo, la IA agrega una novedad, una porción de algo hasta ese momento desconocido, aunque conservando algo del pasado. Porque toda concepción artística sirve de punto de partida a otra (s) nueva (s), por lo que la novedad siempre aludirá a sus predecesores.

De esta forma, la inteligencia creativa deberá descansar en cuatro aspectos: materia prima, instructivo, procesamiento y evaluación. Cuanto más abundante y variada sean las materias primas y más pormenorizadas las instrucciones, mayor será la probabilidad de generar un resultado creativo²⁵. Y como crear de la nada es imposible (crear algo absolutamente nuevo), entonces hay que considerar la creatividad (y la novedad concomitante) también como una cuestión subjetiva de enfoque y/o evaluación. La evaluación informada sobre el producto generado (p. ej., una pintura, sus materiales, su técnica, historia y contexto), dispondrán mejor al crítico de arte a dictaminar si algo es más o menos novedoso.

¿Es la creatividad humana reemplazable por la IA-generativa?

En un modesto sentido, la IA sí es creativa, porque libera al artista de las restricciones materiales y le permiten cristalizar diferentes ocurrencias de su imaginación. Nadie niega que los principales materiales de cualquier artista son su experiencia, conocimiento y habilidad; pero nunca le han faltado las herramientas. Por ello, con la invención de la IA, la creatividad convierte al artista en una especie de director de orquesta, situado a un nivel superior al resto de los músicos, pero precisando de cada uno de ellos para lograr la interpretación deseada²⁶.

Si alguien considera que con la IA se renuncia a parte de la capacidad creativa del artista (en beneficio de la IA), se puede argumentar en contra que con la IA se logra un mejor efecto creativo, como en el caso del director de orquesta, con la habilidad de tocar todos los instrumentos, que cede su lugar a cada músico para alcanzar un mejor efecto interpretativo.

No hay que echar en el olvido que, en todo momento, el hombre ha sido capaz de combinar, experimentar y transformar con ayuda de las máquinas/herramientas la naturaleza, antes que con sus manos desnudas. Con la IA, la creatividad será una distribución de tareas entre el hombre y la máquina inteligente, lo que repercutirá en el aumento de la capacidad exploratoria, combinatoria y transformatoria de ambos.

Además de la creatividad artística, la capacidad de la IA se cuestiona frente al engaño o simulación de actos mentales superiores.

²⁵ Inversamente proporcional a esto sería el efecto creativo divino: crear de la nada (Génesis, 1, 27).

²⁶ O sea: los datos más la combinación algorítmica. El artista es quien decide (aún) cuando la obra final está lista.

LA IA Y EL ENTRENAMIENTO ALGORÍTMICO

Para responder a la pregunta(s) sobre la inteligencia de la IA, hay que descomponerla en tres secciones: primera, la de los datos previamente elegidos y recopilados; segunda, la del “entrenamiento” o formación de patrones de análisis y procesamiento de datos (algoritmos)²⁷; y, por último, la interacción con el usuario (humano o máquina) en forma de una aplicación.

Entre los datos, el entrenamiento algorítmico y la interacción, el misterio de la creatividad e inteligencia de la IA reside principalmente en el diseño de los algoritmos cada vez más sofisticados, como el chatGPT (de OpenAI) o Gemini Pro, Quillbot, Claude, Chatsonic, etc. Todos provistos de la capacidad de analizar y generar imágenes, componer música, responder preguntas y hasta escribir textos completos y especializados.

La apariencia de disponer de competencias humanas reside en su capacidad de crear contenido coherente y reflexionar vis-à-vis con especialistas sobre problemas actuales. Lo que ha llevado a preguntar qué tan creativa e inteligente es la IA. Durante mucho tiempo, la creatividad e inteligencia humana han sido una especie de caja negra que encierran un misterio difícil o imposible de explicar. Y se ha pensado que los dispositivos inteligentes creados por mano humana desentrañan este misterio (i. e., A. Turing, et al.). Un texto redactado por la IA hace pensar en la creatividad de un especialista por el vocabulario técnico, la exactitud gramatical y la coherencia y pertinencia de las ideas.

IA VIS-À VIS ESPECIALISTAS

Se puede plantear como caso de estudio el de la criminalidad y narcotráfico: un experto consultado pediría detalles sobre el problema para enfocar mejor su análisis; la IA asociará automáticamente criminalidad con todo tipo de delincuencia, violencia, tráfico indiscriminado de drogas, sin considerar otras posibilidades distintas a las de su programación. Lo que remite de la *caja oscura* del pasado a la *caja cerrada* o el mencionado *loro estocástico* del presente. El sesgo se debe a que, probablemente, en los datos que alimentaron la IA estaban estos conceptos juntos, sin considerar enfoques diversos (como pobreza, enfermedad, contextos recreativos, etc.).

El que el ChatGPT, por ejemplo, genere una respuesta sin demandar más información adicional o sin tomar en cuenta otros contextos y perspectivas, además del criminal (y no el económico o medicinal), ha tenido como consecuencia el que se califique a ChatGPT como el ya mencionado *loro estocástico*²⁸ en la *habitación China*²⁹, por su capacidad de generar contenidos sin comprenderlos. Ciertamente, la capacidad de generar una repuesta a cualquier tema da la impresión de inteligencia o creatividad e, incluso, de profetizar el futuro. Aunque esa impresión de *inteligencia* hace pensar en todos los problemas que conllevan las decisiones tomadas con base en algoritmos, como en el pasado en los oráculos y profecías, donde la Sibila de Cumas ofrecía respuestas poco claras o enigmáticas a sus desapacibles devotos. Por ello, estas decisiones tomadas con la herramienta

²⁷ Con ello se está en condiciones de anticipar una respuesta frente a una nueva información equiparable a la dada.

²⁸ Emily Menon (2021) hizo ver con su experimento que puede existir la capacidad de generar textos sin comprender el contenido.

²⁹ John Searle (1980) expondría la forma en que es posible conversar sin entender el sentido de la conversación.

ChatGPT no exime a nadie de una dosis de precaución y de tener presente dos cosas: el tipo de datos que nutren a estos sistemas y los algoritmos que la organizan. No hacerlo llevaría a sesgos en la información y en la toma errónea de decisiones. Entonces, no es de sorprender que un texto sobre delincuencia y narcotráfico ignore perspectivas y enfoques importantes, no contemplados en los datos y algoritmos.

La capacidad de la IA para generar “nueva” información depende de los datos previamente suministrados y del procesamiento algorítmico. Por ello, la IA no representa una bola de cristal para la adivinación, sino una herramienta con las restricciones propias del programador, de la información que la alimenta y del algoritmo. Como cualquier otra herramienta, la IA es eficiente en la ayuda para resolver tareas, pero sin la comprensión de la materia (o de sus matices más finos). Su uso es de indudable utilidad, pero de la mano de la razón crítica y consciente de sus restricciones y limitaciones.

IA GENERATIVA Y DERECHO PENAL

Pese a esta última advertencia, sigue presente el hecho de que, en el mundo de hoy, todo está sometido a la revolución digital. La forma como nos comunicamos, desplazamos y consumimos está controlada por un código, cada día más y más inteligente, sofisticado y omnipresente. Y a esto que se le llama Inteligencia Artificial (IA) que, pese a su instrumentalidad, está lejos de ser algo objetivo y neutral.

El uso de la tecnología en la vida cotidiana plantea la siguiente pregunta: ¿quién toma las decisiones correctas (las mejores decisiones posibles)? Si un auto inteligente está equipado con IA y de repente se haya en el peor escenario posible (p. ej.: fallan los frenos), ¿se dirigirá la IA del auto a tomar la “mejor decisión” y cambiar la dirección para estrellarse en un poste (a costa de la vida del conductor), o seguirá de frente a través el cruce en luz roja, plétórico de transeúntes, matando al mayor número de ellos, pero sin dañar al dueño y conductor? Estos problemas de ética práctica, delicia de los profesores utilitaristas norteamericanos, constituyen parte de la problemática incomoda por el avance de la IA.

SESGOS ALGORÍTMICOS DE LA IA

Nadie puede negar que los ingenieros-diseñadores de la IA incorporan en la programación sus propias ideas (tipo educadores), creencias y sesgos culturales, respecto a cómo es y debería de funcionar el mundo. Estos sesgos algorítmicos siempre han formado parte de la informática, y desde 1948, el matemático Joseph Weizenbaum (1977)³⁰ advirtió sobre las inesperadas (o indeseadas) consecuencias del código algorítmico.

³⁰ Para advertencias de este tipo está el titán Prometeo, quien representa desde el origen de la cultura helena la luz, fuego o inteligencia, que corona la creación del hombre, junto con sus peligros. Esta misma idea de *Prometeo* continúa a lo largo de los siglos: primero, en Hesíodo, después en Esquilo, Platón, Esopo, hasta el s. XIX con Goethe, Nietzsche y Mary Shelley o André Gide en el s. XX.

Porque, cuenta aparte de los programas ya existentes de espionaje o distorsión premeditada de la información (de probada inmoralidad), la IA tiene un sinfín de aplicación polémicas, como los referentes a los sistemas de evaluación automatizados que seleccionan por etnia o sexo con el fin de prevenir un delito, obtener un puesto laboral o recibir un programa social o una membresía universitaria, ¿y quién puede prever los futuros alcances de la IA y las revoluciones tecnológico-sociales a que dará lugar?

IA GENERATIVA

Primeramente, el aprendizaje automatizado de la IA generativa plantea un cambio de paradigma a partir de los nuevos códigos y aplicaciones. De una secuencia de instrucciones a seguir (tipo receta de cocina), el aprendizaje automatizado de la computadora (IA generativa) puede generar nuevos códigos partiendo de la información acumulada, a la manera de una cocinera experimentada que sin riesgo alguno se aparta de las instrucciones e introduce innovaciones al cocinar.

Se han creado aplicaciones antes inimaginables, como imitar voz e imagen para después presentarlas en actitudes o contextos diferentes o, el ya mencionado, conducir un auto en condiciones extremas de tránsito citadino. Entonces, el precio a pagar no son sólo los sesgos del programador, sino también los sesgos adicionales aprendidos por estos desarrollos tecnológicos (i. e., entrenamiento posterior); porque 1) la IA tiene como base modelos de “caja negra”³¹, 2) porque no se entiende bien a bien su alcance y funcionamiento y 3) porque los programas de IA (como cualquier otro) son susceptibles de ataques maliciosos.

Por todo esto, la IA y sus perfeccionamientos son dignos de una buena novela de terror al permitir al sistema tomar la iniciativa en caso, por ejemplo, de un fallo en los frenos, contar transeúntes, dirigirse a izquierda o derecha y calcular, por anticipado, los daños y así hacer un bien o un menor mal³². Además, hay una alta probabilidad de que no exista una sola decisión 100% correcta dentro de una situación catalogada como “extrema” y que, por eso, haya siempre un margen de error considerable y no sancionable (por no haber responsables). Obviamente, las máquinas o sistemas inteligentes están impedidos de reparar o reponer el daño, y aunque la responsabilidad esté en el programador, cuando el sistema IA-generativo es capaz de aprender³³ y tomar otras decisiones no contempladas por el programador, ¿quién es el culpable?

Los sistemas IA carecen de personalidad jurídica susceptible a castigo, aprendizaje moral y reposición del daño, sin embargo, su uso conlleva consecuencias judiciales, no deseadas u ocultas; porque cuando una persona toma la decisión A, puede tener en mente las consecuencias B y C, pero no el resto de las otras posibles como D y E, inevitables o indeseadas (y hasta cierto punto

³¹ Por modelos tipo “caja negra” se entienden sistemas o mecanismos donde se conoce la entrada (input) y la salida (output), pero se desconoce el procedimiento intermedio.

³² Ver: máquinas morales del MIT, iniciativa incoada por el Instituto Tecnológico de Massachusetts con la finalidad de reunir opiniones sobre la capacidad de tomar decisiones morales por los dispositivos o sistemas inteligentes. Ver: <https://www.moralmachine.net/hl/es>

³³ La diferencia entre Machine Learning (ML) y Deep Learning (DL) es interesante y pertinente en este trabajo: en el caso de ML, los programas informáticos utilizan algoritmos dados con el fin de crear u organizar patrones de datos. En el caso de DL, si bien forma parte de ML, constituye su evolución ya que genera (nuevos) algoritmos que crean nuevos patrones de datos automáticamente. Por ello, es lo más parecido a la forma de aprender del ser humano.

ignoradas). Esto no tiene por qué ser distinto en el caso en la IA-generativa. Las instituciones equipadas con programas IA de reconocimiento facial para identificar delincuentes no están libres de prejuicios, porque el peligro se encuentra en su concepción inicial. ¿O es posible una IA libre de prejuicios? ¿Cómo sería eso? ¿Qué puede significar? ¿Puede la IA eludir los prejuicios presentes en todo ser humano? ¿Por qué sería la IA mejor que el hombre? La experiencia es que la IA amplifica los sesgos existentes en lugar de evitarlos. Para entender esto es necesario reflexionar sobre los “accesorios” de la IA, a saber: los datos históricos, la fuente de su operación y aprendizaje.

La creación, recopilación y selección de datos es llevada a cabo por personas con poder y saber, disponiendo de datos personales en abundancia, como los registros históricos de contratación, los antecedentes penales propios y familiares, el origen étnico, el ingreso mensual, el grado de estudios, los hábitos de consumo, los diagramas de movilidad, más el acervo de juicios y prejuicios del programador.

Ejemplo de esto es Tay³⁴, un ChatBot de Microsoft, que aprendía de las entradas de Twitter y tuvo que ser eliminado por sus fabricantes porque aprendió a mandar mensajes antisemitas. O el programa de reconocimiento facial de Google Nest Aware³⁵, que comenzó a clasificar rostros de negros como si fueran gorilas. O el mismo programa de Amazon³⁶, que clasificó a mujeres de piel oscura como hombres.

PROGRAMAS DE IA PARA PREVENIR EL DELITO

PredPol³⁷ es un programa de IA diseñado para la policía de los Ángeles con el fin de identificar los posibles focos de delincuencia, pero también para perseguir y prevenir el delito. El programa alerta sobre los barrios que deben ser patrullados y a qué personas prestar mayor atención. La justificación del programa IA para prevenir el delito es de sobra conocido: los humanos son imperfectos, cometen errores y, por eso, es mejor delegar las tareas de vigilancia en robots de IA, y atenerse a sus predicciones.

El sesgo racial y clasista de PredPol residió en su concepción, a saber: los lugares y las personas que en pasado habían causado conflictos sociales. No es de sorprender, entonces, que PredPo haya tenido como resultado el patrullaje excesivo de barrios con población negra o hispana.

Esta IA se basó en los datos históricos de detenciones pasadas, sin cuestionar el desenlace de los mismos ni tampoco las causas sociales. Por ello, el programa IA PredPol reflejó el racismo y clasismo de sus programadores, manifestados a gran escala e inevitablemente amplificando sus defectos.

³⁴ Ver: https://www.bbc.com/mundo/noticias/2016/03/160325_tecnologia_microsoft_tay_bot_adolescente_inteligencia_artificial_racista_xenofoba_lb

³⁵ Ver: <https://news.un.org/es/story/2020/07/1477531>

³⁶ Ver: <https://www.sandiegouniontribune.com/en-espanol/noticias/story/2020-06-10/amazon-prohibe-uso-de-su-programa-de-reconocimiento-facial>

³⁷ Ver: <https://www.xataka.com/aplicaciones/como-funciona-predpol-el-software-que-dice-predecir-donde-van-a-sucedcrimenes>

PREDPOL Y NORMALIZI.NG.ES, UNA HISTORIA DE IMÁGENES Y PREJUICIOS

PredPol tiene un derivado o excrecencia, el sitio *Normalizi.ng.es*³⁸, creado para responder a la pregunta sobre las características del rostro normal.

Este proyecto experimental online analiza la forma cómo el cerebro elabora y determina el criterio de normalidad a través del aprendizaje reflejo o automatizado. Todas las personas analizan y clasifican cada rostro que ven de manera constante y mecánica, y *Normalizi.ng.es* cree que este mismo proceso lo puede llevar a cabo la tecnología computacional gracias a los estudios de procesos automatizados y a su capacidad de formar patrones algorítmicos de “normalidad”. A través del aprendizaje automatizado, el reconocimiento facial puede utilizarse para clasificar y predecir la conducta humana con base en lo que es normal o esperado, algo muy útil para los aparatos de control social.

HISTORIA DE LA NORMALIDAD

El proyecto de investigación de *Normalizi.ng.es* no es nuevo, el intento de contar con criterios de normalidad comenzó por lo menos en el siglo XIX, con el aumento poblacional de las urbes capitales y los problemas sociales concomitantes.

Por encargo de Napoleón III, Georges-Eugène Haussman cambió drásticamente el aspecto de París, debido a la inmigración, la revolución industrial y la urbanización. La demanda de mano de obra reasentó a los campesinos en la gran ciudad y exigió la construcción de viviendas y barrios completos. Estos rápidos cambios demográficos y urbanísticos de la metrópoli tuvieron como consecuencia una creciente división entre ricos burgueses e indigentes trabajadores, la inevitable quiebra de la unidad social y la intranquilidad en las calles; porque los parisinos de abolengo, obligados a convivir día a día con los advenedizos provincianos, sentían una amenaza constante a su patrimonio.

La simple amenaza reformó la manera de operar de la policía, quien inventó la foto policial para identificar más fácilmente a los (posibles) criminales. Antes del crecimiento de las urbes capitales, la policía sabía a quién estaba arrestando, conocía el barrio donde vivía, a quienes frecuentaban y cuál era su círculo familiar más cercano. Con el vertiginoso aumento de nuevas caras migrantes, la situación de la policía cambió y comenzó a usar las nuevas tecnologías, como la fotografía para documentar arrestos. Lo que derivó en un cúmulo de datos inutilizables, hasta que el oficial Alphonse Bertillon decidió ordenar estos datos mediante la *normalización* de las fotos policiales y creó la ficha de arresto estándar, frontal y lateral. Foto icónica de la criminalidad.

El siguiente paso fue medir, cuantificar y clasificar el cuerpo y añadir a la ficha de arresto los datos de la identificación. Alphonse Bertillon introdujo el modelo piloto, llamado *Le Potrait Parlé* (retrato hablado), que revelaría la identidad de una persona en caso de que el presunto criminal se negara a colaborar. Este modelo de archivo, con decenas de fichas de identificación, presentaba distintos tipos de narices, bocas, orejas y demás rasgos faciales, clasificados según los tipos más comunes de personas arrestadas o de presuntos criminales.

³⁸ Ver: <https://normalizi.ng>

Ciertamente, para Alphonse Bertillon simplemente se trataba de un sistema taxonómico para la identificación: el *Bertillonage*, el cual se difundió rápidamente en Europa y los Estados Unidos de Norteamérica como práctica científica de vanguardia que clasificaba criminales por medio de fichas modelo de identificación. Y aunque más tarde el *Bertillonage* fue relegado a un segundo plano por las huellas digitales (una tecnología de identificación más simple y precisa), las fotografías criminales continuaron esculpiendo la imagen del criminal común (i. e., de los ojos, nariz, boca, estatura, color de piel, etc.). A estos avances criminalísticos de Alphonse Bertillon, y de las huellas digitales, se sumaron los estudios e inventos del inglés Francis Galton (1822-1911).

¿QUÉ ES NORMAL?

Al prurito analítico y taxonómico del *Bertillonage* se unió el interés eugenésico de Francis Galton (2004). Este fue un multitallento británico³⁹ que despilfarró su tiempo y creatividad en desarrollar la pseudociencia eugenésica y fisiognómica, a través de sumar a las fotos criminales y las huellas digitales, la estadística. Con ello, creó un patrón de normalidad estadística, o distribución estandarizada de la criminalidad, según sexo, color de piel, nivel económico, residencia, procedencia, etc.

Con Galton, eventos delictivos, aparentemente fortuitos e inconexos, revelaban patrones de conducta representables gráficamente en forma de curva o campana (curva-de-campana), donde la cima de la curva se interpretaba como la coincidencia normativa y los extremos como su desviación. Esta gráfica de campana encarnó físicamente en un mecanismo ingenioso y arbitrario diseñado por Galton: una caja, con dos compartimentos; de un lado, un sinnúmero de abalorios; en medio, una única salida para todos ellos, y, en el otro extremo, una serie de canaletas dispuestas paralelamente. Cuando las cuentas se hacían pasar por la abertura, éstas se introducían en las canaletas al otro lado de la entrada. Los abalorios caían “azarosamente” dentro de las distintas canaletas, y, aunque era difícil predecir donde caerían exactamente, la forma de la caja y el lugar de la abertura propiciaban una distribución que repetía la curva de campana (i. e., más abalorios en las canaletas centrales y menos en las laterales). A esta caja o representación física de la curva se le bautizó con el nombre de su creador: caja de Galton. La distribución de los abalorios revelaba -según Galton- la coincidencia normativa, así como su desviación: a una menor dispersión de los abalorios, mayor coincidencia y normalidad; a mayor dispersión, mayor alejamiento y anormalidad.

Con la curva-campana-caja de Galton se fijó el criterio de normalidad estadística, lo que con el tiempo se convirtió en algo más que una medida estándar para criminalistas, porque se le utilizó para todos los aspectos de la vida, y abarcó a las ciencias naturales y sociales y a todas las disciplinas donde hubiera requerimientos taxonómicos de normalidad e irregularidad.

El éxito de los trabajos de Galton, a finales del siglo XIX, permitió hacer de uso corriente la palabra *normal*. A ello colaboraron sociólogos como Emilio Durkheim, quien aplicó la palabra normal-normalidad a la sociedad humana y a los fenómenos culturales en general.

³⁹ Galton Estudió medicina antropología, genética, psicología, filosofía, en el contexto de La frenología o craneología, creada por Franz Joseph Gall (1758-1828) (y J. Caspar Spurzheim y George Combe). La frenología tuvo como fin estudiar la relación entre inteligencia y cerebro, lo que permitió la investigación de la inteligencia con fundamento neurofisiológico.

ROSTRO NORMAL Y LA CRIANZA PRESCRIPTIVA

La normalidad unió los destinos de Galton y Bertillon. En 1893, el inglés visitó al francés en su laboratorio, y aunque el primero no estaba tan interesado en identificar a criminales como en predecir la criminalidad, el *Bertillonage* le pareció un método útil para predecir quién podía llegar a ser un criminal si se le perfeccionaba con la estadística. Naturalmente, había primero que analizar fotos, identificar y clasificar criminales, para luego predecir la criminalidad estadísticamente. Por ello, Galton inventó una técnica más: el retrato compuesto. Fotografiando a distintos sujetos con la misma película a baja exposición, obtuvo una imagen mixta, de rostros sobrepuestos, y que revelaban un arquetipo de criminal. Galton fotografió a oficiales del ejército, ingenieros, padres de familia y criminales (también a enfermos mentales), todo con el fin de saber el aspecto de la gente normal y anormal, y así predecir desviaciones sociales. Criminales y enfermos mentales probados (sentenciados o confesos) ofrecerían el criterio de normalidad criminal o patológica.

El paso siguiente de Galton fue rebasar el ámbito experimental y concebir una sociedad mejor con ayuda de la eugenesia y la fisiognómica; porque la documentación, clasificación y la estadística, permitirían la crianza darwiniana, esto es: seleccionar y eliminar a los elementos perniciosos de la sociedad y propiciar la proliferación del resto. De esta forma, por un lado, la criminología no quedaría en una simple elucubración y, por otro, la utopía social dejaría de ser una lejana quimera.

Armado de esto, Galton fundó el movimiento eugenésico, que pretendía dirigir la evolución a través de la competencia y la selección, pero en términos prescriptivos (¿qué tipo de hombre es deseable y qué tipo no). Los estudios de Darwin ofrecieron fundamento científico sobre la variabilidad en la naturaleza y los de Galton fundamento para la selección-prescriptiva y una sociedad mejor.

Además de la eugenesia, Galton y sus discípulos practicaron también la fisiognómica, otra pseudociencia útil para evaluar los rasgos de la personalidad a partir de la simple apariencia física⁴⁰. Ambas, en mancomunidad, hacían finalmente factible “predecir” el potencial criminal de una persona, ya que contaban con el método de Bertillon (medir y registrar), las estadísticas y evaluaciones sobre la apariencia física. Además, el noble fin de contribuir al mejoramiento social justificaba la eliminación de locos, criminales y enfermos, con base en la ciencia darwiniana⁴¹; porque los eugenistas eran una nueva especie de sociodarwinistas⁴², solamente que confiaban más en la crianza que en la selección natural para asegurar el progreso humano.

La idea del progreso⁴³, típica del siglo XVIII y XIX, y la esperanza en una sociedad mejor (típica del pensamiento Occidental), contaba ahora con bases científicas y con una meta clara: la pureza de la herencia, evitando que locos, criminales y enfermos de toda laya trajeran niños al mundo.

Las teorías eugenistas se probaron en toda Europa Occidental, pero no fue sino hasta que arribaron a los Estados Unidos de América donde se implementaron políticas y leyes de esterilización

⁴⁰ No sólo en Inglaterra o Francia se realizaron estos tipos de estudio, también el médico italiano Cesare Lombroso (1835-1909), fundador de la escuela de criminología positivista en Italia, investigó el crimen con base en la biología y sus caracteres físicos normales.

⁴¹ Además de la ciencia taxonómica y estadística, la genética evolutiva.

⁴² Al origen de las especies siguió el origen del hombre, donde Darwin expuso sus ideas de la evolución de los valores y culturas.

⁴³ La idea de progreso, idea fundamental del siglo XVIII y XIX, solamente es concebible dentro de una concepción del tiempo lineal o vectorial. Por eso, ésta no existió en la Antigüedad grecolatina.

para que los *Freaks*⁴⁴ no tuvieran descendencia. Este “racismo científico”, con base eugenista y fisiognómica, alcanzó su punto crítico en la primera mitad del siglo XX, cuando Hitler aprovechó a los eugenistas norteamericanos como fuente de inspiración para las políticas nazis encargada de salvaguardar la pureza de la raza aria alemana a través del genocidio de judíos, gitanos, homosexuales y discapacitados.

Baste apostillar que solamente pasada la Segunda Guerra Mundial fue cuando las ideas eugenistas y fisiognómicas sufrieron condena internacional; si bien no han desaparecido completamente, porque están ancladas en las teorías de la herencia y normalidad estadística.

LA IA Y LA NORMALIDAD

El análisis de rostros ha tenido su vuelta gracias a las mejoras en las estadísticas y en el análisis de datos (rostros) vía la IA. Los científicos de datos son los herederos de los Galton y Bertillon, mediante la codificación de datos y clasificación estadística. Algunos se centran en la clasificación e identificación, tipo Bertillon (identificación, archivo y clasificación forense). Otros, discípulos de Galton, desean estadísticamente hacer proyecciones futuras, también con base en patrones conductuales del pasado.

Llevando la tecnología al límite, la identificación, clasificación y proyección tiene nuevos alcances, por ejemplo, en 2021, debido a las presiones de los reguladores de Facebook, se abandonó el reconocimiento facial para identificar usuarios en fotos de todo tipo y, en sustitución, se implementó el servicio *Automatic Alt Text*⁴⁵, que analiza imágenes y generar automáticamente textos con la descripción de sus rasgos faciales más relevantes.

Los centros de datos acumulan más y más información, con la posibilidad de crear un retrato cada vez más detallado de nuestras caras y conductas pasadas, que a su vez alimentan las predicciones algorítmicas de nuestra apariencia y conducta futura. Al igual que los retratos superpuestos de Galton, la imagen de conducta normal se construye a partir de múltiples ejemplos superpuestos de los diferentes rostros del pasado. Cada “abalorio” de información se mete a una *caja negra algorítmica* para incoar su camino y distribución hacia la normalidad. El resultado es una curva de campana, pero de la conducta futura normalizada.

El criterio sobre lo normal, creado por los datos superpuestos de cientos de personas, se basta a sí mismo, porque no sólo predice el futuro, sino que también lo dicta o prescribe. En todos los ámbitos en los que la predicción es la apuesta más segura, apostar a una desviación representa un enorme riesgo (ya sea en los juegos de azar, en las finanzas o en la política y gobernanza). De esta forma, la predicción basada en datos y algoritmos normaliza el presente y, sobre todo, el futuro, al impedir el cambio.

Por eso, los algoritmos de aprendizaje automatizado son conservadores desde su concepción, ya que solamente predicen patrones realizados en el pasado. Además, la disputa sobre la privacidad en línea no influye en los derechos sobre la privacidad de actividades llevadas a cabo en el pasado. Los aparatos de vigilancia algorítmica no están tan interesados en las actividades

⁴⁴ La película norteamericana de 1932, *The Freaks*, dirigida por T. Browning, presenta un circo itinerante compuesto de *Freaks* o seres deformes. Por eso, no tienen lugar de residencia fijo y solamente sirven para el escarnio.

⁴⁵ Ver: <https://alttext.ai/>

realizadas en el pasado como en las futuras. A esto se debe su interés en las imágenes o “retratos compuestos” (á la Galton), que proyectan una imagen capaz de definir el camino a la normalidad o a la sospechosa desviación.

ONLINE NORMALIZI.NG

Como ya se mencionó, en la vida cotidiana, los cerebros clasifican todo tipo de apariencias humanas, y la memoria captura, analiza y clasifica cada fisonomía (clasifica y es clasificada). Pero no es consciente de la manera cómo lo hace; y exactamente de esto se trata el proyecto *Normalizi.ng.es*, a saber: entender cómo se establece quién es normal y quién no (o no tanto).

Los participantes en el proyecto *Normalizi.ng.es* pasan por tres etapas comunes a la estadística temprana. La primera, poner el rostro dentro de un marco: tomarse una foto. Segundo paso, clasificar: se les muestra un registro de participantes anteriores y se enlistan los caracteres (narices, bocas, ojos) para decidir cuáles son más parecidos (normales) y cuáles no. Tercer paso, análisis: se analizan rostros normales o más comunes. Luego, todo se registra o guarda en una ficha tipo *Bertillonage* (tipo policial), con puntuaciones dinámicas de normalización. Después se agregan los rostros de todos los participantes con un mapa algorítmico de normalidad, inspirado en un retrato hablado (*Portrait parlé*). Y así el resultado es un mapa de la normalidad (*the Normalizi.ng Map*⁴⁶), que utiliza algoritmo de análisis facial y agrupa cuadricularmente todos los rostros en dos dimensiones y con las características más familiares.

Cuanto más alto es el puntaje, más ancho es el marco en blanco de la foto en el *Normalizi.ng Map*. A medida en que se suman participantes, el algoritmo se actualiza y se vuelve a dibujar el mapa. El mapa presenta dos conglomerados, identificados binariamente, como masculino/femenino. En el centro y en los márgenes del cuadro se forman distintas constelaciones de imágenes, por ejemplo: a) conglomerado de piel negra (unisex), b) conglomerado asiático (que puede ir del chino al latinoamericano), etc. Así, por cada característica, resultan distintos conglomerados: del tipo estandarizado mujer, hombre, raza negra o caucásica, niño o infantil, anciano barbudo, etc.

Cada vez que se repite la construcción del mapa de imágenes se repiten y amplifican las divisiones y la clasificación, y se demuestra que toda clasificación algorítmica es racista, sexista, clasista, etc., porque corresponde a la discriminación social presente en todo el mundo (sea por género, color de piel, sexo, etc.)

El algoritmo es el vigilante policíaco que nos observa a través de nuestros estereotipos y nunca habla con iniciativa propia. Pero, cabe preguntar: ¿es la normalidad un bien a custodiar? ¿En eso reside la utopía, en sistematizar, automatizar y amplificar los prejuicios? Si es así, entonces el Internet nunca dejará de ser una aldea, aunque los rostros poco familiares sean interesantes y plenos de viabilidad.

⁴⁶ Ver: <https://normalizi.ng>

CONTROL DE RIESGOS

Ciertamente se trabaja para reducir los riesgos y errores de la IA-generativa (como el de prejuicio racial), por ejemplo, a través de la creación de una lista como la *Awful Artificial Intelligence*⁴⁷, creada por David Dao, que tiene como objetivo hacer conciencia de los problemas de discriminación, desinformación y vigilancia omnipresente. O los proyectos tecnológicos descentralizados, tipo *Blockchain*⁴⁸, que exploran la forma de limitar ciertos desarrollos de la IA⁴⁹ que pudieran apoderarse de la vida del hombre, controlando el uso de los datos no consentidos y/o poniéndolos a disposición de todo el mundo. O programas como *Stop Killer Robots*⁵⁰, que también desean prohibir desarrollos de la IA perjudiciales.

CONCLUSIÓN

Parece claro que el control social automatizado beneficia al Estado, y es de preguntarse si la discriminación también. ¿Qué sucedería si no se atajan estos sesgos y problemas de la IA? ¿Qué tipo de Estado y sociedad se tendría?

Estados autoritarios, como el chino o Norcoreano, han llevado al límite el uso sin cortapisas de la IA, como la vigilancia omnímoda e intemporal, el acopio de todo tipo de huellas digitales o físicas, los diagramas de movilidad, las historial judicial, médico y laboral, etc. Sus programas no solo reconocen la identidad particular de las personas, además, el tipo de sexualidad, la proclividad a ciertas enfermedades y adicciones, y las posibles actividades delictivas (aunadas a ciertos rasgos faciales).

Ciertamente, en los Estados totalitarios, no sólo se desea saber, sino además, predecir quienes de sus ciudadanos no seguirán las normas y prescripciones, para alimentar sus sistemas de crédito social, asignar puntos automatizados y otorgan o negar derechos según la puntuación. A partir de la IA pueden crear programas sociales de incentivos y castigos para ciertos ciudadanos y minorías étnicas (p. ej.: el caso de los Rohinyá de Myanmar⁵¹ o los Uigures de la República Popular China⁵²).

Por lo anterior, es una obligación comprender y limitar los avances tecnológicos actuales de la IA-generativa. Crear reglas para su desarrollo seguro, porque la IA es una herramienta que soluciona problemas, pero también que sirve para reforzar o ampliar los ya existentes. Por eso, es loable el desarrollo constitucional de los neuroderechos en la Carta Magna de Chile⁵³ (ejemplo para Latinoamérica).

⁴⁷ Ver: <https://medium.com/@daviddao/awful-artificial-intelligence-99b8220746c>

⁴⁸ Ver: <https://blockchain-revolution.com/>

⁴⁹ Ver: <https://redbioetica.com.ar/limitar-la-inteligencia-artificial/>

⁵⁰ Ver: <https://www.stopkillerrobots.org/es/>

⁵¹ Ver: <https://www.acnur.org/mx/emergencias/emergencia-rohingyas>

⁵² Ver: <https://www.humanium.org/es/una-eliminacion-etnica-en-china-el-abominable-destino-de-los-ninos-uigures/>

⁵³ La Constitución chilena dice a la letra: “Capítulo III. Artículo 19, numeral: 1º.- (...) El desarrollo científico y tecnológico estará al servicio de las personas y se llevará a cabo con respeto a la vida y a la integridad física y psíquica. La ley regulará los requisitos, condiciones y restricciones para su utilización en las personas, debiendo resguardar especialmente la actividad cerebral, así como la información proveniente de ella (...)”.

BIBLIOGRAFÍA

- Bischof, Andreas (2017). *Soziale Maschinen bauen*. Bielefeld
- Boden, Margaret (2006). *Mind as Machine. A History of Cognitive Science*. New York
- Galton, Francis (2004). *Hereditary Genius: An Inquiry Into Its Laws and Consequences*. Whitefish Montana, Kessinger Publishing.
- Gillham, Nicholas (2001). *Wright, A Life of Sir Francis Galton*. Oxford, Oxford University Press.
- Lemke, Thomas (2013). *Biopolitik zur Einführung*. Junius Verlag. Hamburg.
- Lenzen, Manuela (2023). *Künstliche Intelligenz. Fakten, Chancen, Risiken*. München, C.H. BECK Verlag.
- Luhmann, Niklas (1994). *Liebe als Passion*. Frankfurt am Main, Suhrkamp Verlag.
- Luhmann, Niklas (1987). *Soziale Systeme*. Frankfurt am Main, Suhrkamp Verlag.
- Luhmann, Niklas (1996). *Die Gesellschaft der Gesellschaft*. Frankfurt am Main, Suhrkamp Verlag.
- Luhmann, Niklas (1975). *Systemtheorie der Gesellschaft*. Frankfurt am Main, Suhrkamp Verlag.
- Luhmann, Niklas (1983). *Legitimation durch Verfahren*. Frankfurt am Main, Suhrkamp Verlag.
- Luhmann, Niklas (1999). *Gesellschafts-Struktur und Semantik*. Frankfurt am Main, Suhrkamp Verlag.
- Mau, Steffen, (2017). *Das Metrische Wir. Über Quantifizierung des Sozialen*. Berlin.
- Menon, Emily, (2021). "On the Dangers of Stochastic Parrots: Can Language Models be too Big?" (En los peligros de los loros estocásticos: ¿son los modelos de idiomas demasiado grandes?), <https://dl.acm.org/doi/10.1145/3442188.3445922>.
- Mumford, Lewis (2010). *El mito de la máquina. Técnica y evolución humana*. Pepitas de calabaza, España.
- O'Neill, Cathy, (2016). *Weapons of Math Destruction. How Big Data Increases Inequality and Threatens Democracy*. Crown Publisher, New York.
- Ottmann, Henning, (2001). *Geschichte des politischen Denkens. Band 1/1, Die Griechen von Homer bis Sokrates*. J.B. Metzler Verlag. Stuttgart.
- Rid, Thomas, (2016). *Maschinendämmerung. Eine Kurze Geschichte der Kybernetik*. Berlin.
- Searle, John, (1980). "The Chinese Room Argument" (El argumento del cuarto chino), <https://plato.stanford.edu/entries/chinese-room/>.
- Turing, Alan, (1967). *Kann eine Maschine denken? Kursbuch 8*, Ditzinger, Reclam Verlag.
- Weizenbaum, Joseph, (1977). *Die Macht der Computer und die Ohnmacht der Vernunft*. Frankfurt am Main, Suhrkamp Verlag
- Zuboff, Shoshana, (2018). *Das Zeitalter des Überwachungskapitalismus*. Frankfurt am Main.
- Zweig, Katharina, (2019). *Ein Algorithmus hat kein Taktgefühl. Wo künstliche Intelligenz sich irrt, warum uns das betrifft und was wir dagegen tun können*. München.