

SOBRE LA AUTONOMÍA, LA CREATIVIDAD Y LAS CONSIDERACIONES ÉTICAS DE LA INTELIGENCIA ARTIFICIAL EN EL ARTE CONTEMPORÁNEO

MARCELINO GARCÍA SEDANO. UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ, QUITO, ECUADOR

On the Autonomy, Creativity and Ethical Considerations of Artificial Intelligence in Contemporary Art

Sobre Autonomia, Criatividade e as Considerações Éticas da Inteligência Artificial na Arte Contemporânea

Fecha de recepción: 14 de diciembre de 2021. Fecha de aceptación: 14 de enero de 2022. Fecha de modificaciones: 15 de febrero de 2022

DOI: <https://doi.org/10.25025/hart12.2022.04>

MARCELINO GARCÍA SEDANO

Doctor en Historia del Arte por la Universidad de Oviedo, es profesor tiempo completo en la Universidad San Francisco de Quito donde imparte asignaturas en la carrera de Artes Contemporáneas y la Subespecialización de Fotografía que coordina. Sus intereses académicos van relacionados con la influencia de la tecnología en las artes y la sociedad, recibiendo especial interés, entre otros aspectos, la posfotografía, la performance audiovisual digital, el videojuego como herramienta creativa o la influencia de la ciencia ficción como herramienta especulativa. Desde 2012 hasta 2019 dirigió Hausmann, Encuentro Multidisciplinar de Creación New Media y en 2021 ha sido co-curador de "Alternativas Presentes", primera edición de la Bienal Universitaria de Arte Multimedial, BUAM21 celebrada en Quito.

RESUMEN:

Reflejo del pensamiento colectivo sobre nuestra relación con las tecnologías, las inteligencias artificiales, presentes en muchos aspectos de la cotidianidad, están lejos de ser autónomas. La labor de la creatividad artística en relación con las inteligencias artificiales debe consistir en analizar y proyectar los aspectos más fundamentales de la mediación de estas tecnologías en la sociedad y no centrarse tanto en sus características estéticas. Las obras aquí seleccionadas son un ejemplo de cómo el arte es capaz de cuestionar el grado de autonomía de las máquinas en la actualidad y su capacidad creativa real, y de evidenciar sus relaciones con aspectos sociales como el poder, el control de información y otras facetas de trascendencia política, así como su capacidad para proyectar futuras problemáticas.

PALABRAS CLAVE:

Inteligencia artificial, arte, ética, pensamiento autónomo, creatividad computacional.

Cómo citar:

García Sedano Marcelino "Sobre la autonomía, la creatividad y las consideraciones éticas de la inteligencia artificial en el arte contemporáneo". *H-ART. Revista de historia, teoría y crítica de arte*, n.º 12 (2022): 17-40. <https://doi.org/10.25025/hart12.2022.04>

ABSTRACT:

Artificial intelligence, being a reflection of collective thinking about our relationship with technologies, is present in many aspects of everyday life, although it is far from being autonomous. The work of artistic creativity concerning AI should endeavor to analyze and project the most fundamental aspects of the mediation of these technologies in society, and devote less attention to their aesthetic characteristics. The works selected here are an example of how art is capable of questioning the degree of autonomy of machines today and their real creative capacity, and of highlighting their relationships with social concerns such as power, information control, and other politically relevant issues, as well as their ability to project future problems.

KEYWORDS:

Artificial intelligence, art, ethics, autonomous cognition, computational creativity.

RESUMO:

Reflexão do pensamento coletivo sobre a nossa relação com as tecnologias, as inteligências artificiais, presentes em muitos aspectos da quotidianidade, estão longe de ser autônomas. A tarefa da criatividade artística em relação com as inteligências artificiais tem de analisar e projetar os aspectos fundamentais da mediação da tecnologia na sociedade, e não se centrar nas suas características estéticas. As obras selecionadas neste artigo são um exemplo de como a arte é apta para questionar a autonomia das máquinas na atualidade e a sua capacidade criativa real. Mesmo, evidenciar suas relações com aspectos sociais como o poder, o controle da informação e outras questões de transcendência política, e sua capacidade para prever problemáticas futuras.

PALAVRAS-CHAVE:

inteligência artificial, arte, ética, pensamento autônomo, criatividade computacional.

INTRODUCCIÓN

La sociedad contemporánea, inmersa en un sentimiento de fascinación por el desarrollo tecnológico, absorbe en una tecnofilia feroz, organiza gran parte de sus procesos sociales en torno al uso indiscriminado y casi inconsciente de las tecnologías de última generación, sobre todo las relacionadas con la información y la esfera digital. Dentro de estas tecnologías se destacan, y ello ante todo por su capacidad de cuestionar hasta nuestros cimientos ontológicos, aquellas derivadas del pensamiento automático, la llamada inteligencia artificial. Desde tiempos inmemoriales la creación de vida de forma artificial ha configurado múltiples narrativas incentivadas por el temor de que estas entidades lleguen a sustituir los rasgos que nos definen como humanos y por la curiosidad sobre la forma en la que estas podrían convivir con la humanidad. Desde el gólem, Prometeo, la criatura del doctor Frankenstein, pasando por HAL 9000 de *Una odisea en el espacio* o los replicantes de *¿Sueñan los androides con ovejas eléctricas?*, nuestro imaginario colectivo materializa nuestras más profundas preocupaciones y anhelos en cuanto a la creación de vida de forma artificial. Las teorías sobre la singularidad tecnológica materializadas de forma popular en los argumentos más tradicionales de la ciencia ficción son una versión más o menos moderna de estas preocupaciones. La posibilidad de un ente autónomo creado de forma sintética se piensa tan seductora como amenazadora para la humanidad. No obstante, las apreciaciones actuales de lo que por el momento se entiende como inteligencia artificial se diluyen entre el optimismo y la invisibilización de su presencia en los procesos cotidianos, una imbricación que promete hacer más fácil una vida cada vez más automatizada, más virtual y más dependiente de las tecnologías de la información. Por lo tanto, se impone un análisis de su incidencia en la sociedad desde una perspectiva que gravita en torno a las consideraciones éticas de esta convivencia.

Si algo define hoy día a las inteligencias artificiales o a las metodologías de aprendizaje automático es sin duda su total dependencia del ser humano para su existencia y funcionamiento. Incluso podríamos aventurarnos a afirmar que estas aún distan mucho del grado de perfección que el sentimiento popular les atribuye. El nivel de colaboración con el ser humano mide sin duda su grado de autonomía y el principio de producir resultados impredecibles la acrecienta. Es paradójico que sea tan débil la sustentación de ese funcionamiento autónomo y que la ilusión de autonomía se sustente en el error. Cuando una máquina responde fuera de lo esperado se interpreta como un acto de autonomía, aunque siga habiendo parámetros que de una forma u otra la restringen.¹ Por ejemplo, en la actualidad ya no se habla tanto de la escritura de un código por parte de los programadores, sino que estos diseñan un entorno, marcos de red que ya no

1. Robert Pepperell, *The Posthuman Condition: Consciousness Beyond the Brain* (Wiltshire: Cromwell Press, 2003), 121.

codifican reglas sino que facilitan, mediante la delimitación de estos grandes ámbitos de información, el funcionamiento de esas inteligencias artificiales para una actividad determinada. El cómo la máquina desarrolla sus reglas de comportamiento sigue teniendo un factor desconocido, lo que le añade un componente de misterio y presunta voluntad, pero no debemos confundir lo aleatorio o lo imprevisible con la voluntad.² Sin embargo, es muy importante el desplazamiento de la concepción del pensamiento como algo lineal presente en estas tecnologías. Karen Barad, con su “realismo agencial”, marcó la indistinción entre materia y significado al igualar a lo inerte con lo vivo a través de las interacciones entre seres vivos, cosas y acciones que desembocan en el significado.³ Esta simbiosis obliga a pensar en sincronías, en redes y no en procesos lineales; es justo así como discurre el aprendizaje automático.

La distinción tradicional entre lo técnico y lo cultural, lo orgánico y lo maquinal, y en definitiva la separación de ambos mundos adolece de sentido desde las leyes de la cibernética o la inclusión de los objetos técnicos como forma de conocimiento y sentido de valores propuesta por Gilbert Simondon.⁴ Es desde esta ecología de las máquinas, panorama propuesto, entre otros, por pensadores como Yuk Hui y Bernard Stiegler, donde se plantea una posición conciliadora entre hombres y máquinas, que se debe replantear una posición para el análisis de la incidencia de las inteligencias artificiales menos antropocéntrica. No obstante, en la praxis —al menos así se evidencia en las obras aquí reseñadas— emergen todavía los aspectos humanos.

Dentro de las múltiples definiciones de inteligencia hay similitudes entre la actividad de las IA⁵ y algunos aspectos de los procesos de pensamiento humanos, pero es en las esferas de la autoconciencia y del pensamiento crítico y emocional donde estas tecnologías se distancian de nosotros. En cuanto al arte, la discusión de si una máquina puede ser o no creativa se basa también en las analogías con el cerebro humano y su actividad creadora, un terreno aún en gran parte impreciso y por explorar que se relaciona directamente con el grado de autonomía de esta y que puede ser emparejado desde un punto de vista ético con la autoconsciencia.

Desde este estado de la cuestión, que propicia las más profundas reflexiones, la actualidad del panorama creativo relativo al aprendizaje automático abunda en obras artísticas que, en consonancia con nuestra tradición artística hegemónica y la tecnofilia cultural presente, manifiestan un afán de exhibicionismo técnico sin grandes rasgos disruptivos. Partiendo de esta realidad surgen importantes preguntas, nuevos posibles escenarios: ¿Podremos llegar a soñar con un futuro arte de máquinas para las máquinas? ¿Es pertinente imaginar un mundo de singularidad tecnológica autosustentado en inteligencias artificiales completamente autónomas que posibiliten lo anterior?

2. Aaron Hertzmann, “Can Computers Create Art?”. *Arts* 7, no. 2 (2018): 14.

3. Karen Barad, *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning* (Durham: Duke University Press, 2007), 185.

4. Gilbert Simondon, *El modo de existencia de los objetos técnicos* (Buenos Aires: Prometeo Libros, 2007).

5. En ocasiones, por razones de economía, sustituiremos el término “inteligencia artificial” por sus siglas IA.

Ante la evidente lejanía de estos escenarios y su carácter fabulador, el esfuerzo en la creación artística debería ir encaminado hacia la utilización de las inteligencias artificiales y el aprendizaje automático para especular sobre sus dimensiones éticas y políticas, como un eje de reflexión sobre la posición del ser humano con respecto a estas tecnologías. El arte debe ser una herramienta reflexiva sobre el futuro y no una exhibición de las capacidades actuales de las IA; sus características y sus dimensiones ético-sociales deben ser el fin, y no siempre el medio. Los proyectos realmente pertinentes son aquellos que investigan sobre el alcance social de estas tecnologías, no sobre sus capacidades creativas. Sin duda, las grandes preguntas sobre el ser, la conciencia, la relación humano-máquina, deben contemplar también las posibilidades de la singularidad y la creación de una nueva estética y una nueva ética para una nueva dimensión de máquinas totalmente independientes, pero ello no es urgente. Es necesario un arte que sea sustancialmente diferente al ya tradicional arte generativo, un arte del futuro que no debe distraernos del presente. Este artículo propone un repaso de algunas obras artísticas que materializan estas preocupaciones y que, por razones prácticas para su estudio, relacionaremos a cuestionamientos éticos relevantes propuestos por diversos autores. Este estudio no pretende priorizar ninguna de las aproximaciones de los autores mencionados, sino más bien trazar un panorama amplio de las más destacadas aproximaciones y examinar cómo se materializan en algunas obras. Las obras han sido seleccionadas al servicio de esta labor y no pretenden sintetizar el amplio espectro de aproximaciones a las IA desde el arte actual.

ALGUNAS CONSIDERACIONES ÉTICAS SOBRE LA CREATIVIDAD ARTÍSTICA BASADA EN EL APRENDIZAJE AUTOMÁTICO

Uno de los análisis más atractivos y pertinentes desde el arte es aquel que delimita las problemáticas y aspectos más relevantes de las IA desde el punto de vista de la ética. Este enfoque arroja unas consideraciones éticas que gravitan en torno a la utilización de las mismas, pero también a su propia existencia, sus dimensiones ontológicas. Algunas reflexiones llevadas a cabo en los últimos tiempos podrían ser consideradas a partir de tres ejes fundamentales. La síntesis de los mismos y sus analogías en los discursos artísticos presentados permitirá analizar algunas de las características presentes en el arte tecnocientífico que reflexiona en torno al aprendizaje automático. Estos tres ejes de consideración ética son: la autonomía real de estas tecnologías (cuestionada por estudios y proyectos que giran en torno a la autoconsciencia de las máquinas), el valor real de su creatividad en relación con la creatividad humana y, por último, su capacidad moral dada su imbricación en los procesos sociales de comunicación, poder y control. Estos ejes

pueden ser resumidos en estas tres preguntas: ¿Son las inteligencias artificiales realmente autónomas? ¿Son creativas como los humanos, pueden considerarse en términos convencionales de autoría? ¿Son moralmente justas al desempeñar sus funciones?

Las preocupaciones éticas fundamentales y más comunes motivadas por la omnipresencia de estas tecnologías gravitan en torno a problemas como la eliminación de puestos de trabajo, la privacidad, las estrategias de clasificación y control o la distribución justa y equitativa de la tecnología. Buscamos ordenar o juzgar, desde dimensiones éticas, la relación humano-máquina en un intento de priorizar los aspectos éticos de su aplicación en los ámbitos de actividad humana. Philip Galanter aborda el problema más específico de las artes desde la inminencia, es decir, cuestionando el momento en el que las inteligencias artificiales podrían ser agentes éticos con derechos. Esto, más que ser una herramienta especulativa sobre el futuro nos ayuda a resignificar cuestiones como la autonomía de las máquinas y, en cuanto al arte, la desestabilización del concepto de autoría.⁶ Partiendo de esto se plantea una condición que se puede considerar por analogía con el funcionamiento y entrenamiento del aprendizaje automático. Mientras que en la modernidad asistimos a una concepción de un autor aislado al que no le importa el público en su proceso creativo, sustentado en una asunción de la genialidad, en la posmodernidad hay un énfasis especial en la audiencia, hasta el punto de desplazar en numerosas prácticas al autor del producto cultural. Galanter plantea dentro de los procesos culturales actuales lo que denomina como “complejidad”,⁷ entendida como la posición intermedia de un autor que ha de ser audiencia y creador al mismo tiempo, es decir, un creador que aporta en su producción las bases de la cultura que ha consumido, tal y como lo hacen los motores de aprendizaje automático más básicos, como se explicará más adelante. Esta propuesta nos ayuda a determinar la autoría desprendida del aprendizaje automático dentro de un espacio híbrido y maleable que nos permite pensar estos términos desde nuevas perspectivas. Se trata de puntos de vista que colisionan y pierden su razón de ser en el momento en el que pretendemos imponer los esquemas más tradicionales de autoría que discriminan la importancia capital de la cultura en la formación del pensamiento creativo. Esto, traído al mundo del aprendizaje automático, significa que hay un desprecio y una falta de reconocimiento de la importancia que posee la información con la que se entrena dicho dispositivo, centrando la atención exclusivamente en el producto final y no en el proceso.⁸ Dentro de la complejidad que caracteriza a la sociedad-red los hechos más significativos tienen que ver con la mediación; es por ello que el reconocimiento social de la autoría tiende a diluirse o a perder importancia cuando de una inteligencia artificial se trata. Esta tiende a no ser tenida en cuenta porque es vista como un componente más dentro de una compleja comunicación en red.

6. Philip Galanter, “Towards Ethical Relationships with Machines that Make Art”. *Artnodes* 26 (2020): 1-9.

7. Traducción más o menos afortunada del término inglés “complexism” que Galanter incorpora a la explicación de los procesos culturales inciertos e incompletos a partir de teorías matemáticas. Galanter aplica la teoría de la complejidad a las artes y las humanidades planteando un espacio intermedio entre la modernidad y la posmodernidad. Se recomienda leer: Philip Galanter, “What is Complexism?”, ponencia presentada en *Generative Art and the Cultures of Science and the Humanities, International Conference on Generative Art*. Generative Design Lab, Milan Polytechnic, Milán, 2008.

8. Galanter, “Towards Ethical Relationships”, 4.

En cuanto a las cuestiones morales relativas a las inteligencias artificiales, podrían ser aplicadas en dos posibles escenarios. El más optimista sostiene ya una determinada autonomía de las inteligencias artificiales y aplica entonces su juicio ético en torno a dos problemas: el de una moral propia de las máquinas y el de una moral aplicada a la relación de las máquinas con los humanos. El otro escenario, más cauto, aplica los cuestionamientos éticos únicamente a la agencia de estas tecnologías en el campo de las relaciones de poder entre humanos, ya que se las entiende como una mera instrumentación de estas dinámicas y no como entidades autónomas.

El primer escenario es aún demasiado especulativo. Varias reflexiones indican que la singularidad tecnológica está aún muy lejos de materializarse, por lo que la creencia de que si hay consciencia debe haber ética es precipitada y depende de un hipotético desarrollo de la consciencia por parte de estas entidades tecnológicas. Si bien es cierto que parece concluyente que las máquinas no tienen estatus moral, aún no se define con certeza qué atributos caracterizan al estatus moral en términos generales. No obstante —y esto sirve de punto de partida para las reflexiones vinculadas al arte que se presentan más adelante— hay dos criterios interesantes a la hora de otorgar un estatus moral a una entidad. Nick Bostrom y Eliezer Yudkowsky postulan dos dimensiones fundamentales a este respecto: la capacidad de sentir por un lado y la sapiencia por otro, entendida esta última no solo como capacidad sino como atributo de un agente autoconsciente y sensible a la razón.⁹ La reflexión en torno a estos dos criterios afianza la idea de que la autoconsciencia y la autonomía total son requisitos para que estos dispositivos adquieran consideración moral.

En cuanto al segundo escenario, el que contempla los aspectos morales de la relación de las máquinas al servicio de la humanidad, abunda en aproximaciones diversas pero que coinciden en que, por ser una creación humana, las máquinas deben obedecernos, estar a nuestro servicio. Tal es la postura de Joanna J. Bryson, quien partiendo de esta aseveración señala que el debate debe estar enfocado en las políticas de aplicación de la tecnología alejadas de la búsqueda del bien común de la humanidad y concluye que los robots (término que para ella define y acota el funcionamiento de las inteligencias artificiales y el aprendizaje automático) no deberían tener derechos y deberían ser esclavos de la humanidad. La autora, basada en que estas entidades artificiales son creadas para servirnos en diversas funciones, afirma que tener asistentes es algo provechoso. Un robot puede ser sirviente sin ser persona, por lo que una persona puede de forma natural poseer un robot sin tener por ello que reconocerle el estatus de persona.¹⁰ Por esta y otras razones las inteligencias artificiales no deberían sustituirnos en la toma de decisiones, algo que sí ocurre en ámbitos controvertidos como el control sobre los cuerpos y el uso de información con objetivos de poder.

9. Nick Bostrom y Eliezer Yudkowsky, “The Ethics of Artificial Intelligence”, en *Cambridge Handbook of Artificial Intelligence*, editado por Keith Frankish y William Ramsey (Nueva York: Cambridge University Press, 2014), 6.

10. Joanna J. Bryson, “Robots Should Be Slaves”, *Joannabryson.org*, <https://www.joannabryson.org/publications/robots-should-be-slaves-pdf>, consultado el 8 de julio de 2022, 3.

Esto nos remite a otra cuestión moral fundamental ya mencionada: los problemas de sesgo en el funcionamiento y aplicación del aprendizaje automático. Como señaló en su momento Cynthia Dwork, el sesgo siempre va a ser un problema dentro de poblaciones con gran diversidad y quizá esto refuerce aun más la falsa autonomía y el peligro de otorgarle decisiones susceptibles de discriminación a una entidad artificial, pues el sesgo no proviene del algoritmo sino de la información con que se lo entrena.¹¹ De esta forma, pensar que una inteligencia artificial es un agente moral es totalmente falso y cuanto menos optimista. Partiendo de estas inquietudes que presiden los debates en torno a la naturaleza ética del aprendizaje automático, procederé a localizar y analizar proyectos artísticos que materialicen estas cuestiones fundamentales relativas a la creatividad, la autonomía y el alcance moral del mismo desde diversos enfoques del pensamiento.

EL APRENDIZAJE AUTOMÁTICO Y EL SIMULACRO DE LA CREATIVIDAD

Entender la creatividad como un avanzado estado de razonamiento basado en la memoria, la analogía y el aprendizaje condicionado por las restricciones de nuestra experiencia permite contemplar la posibilidad de replicarla desde una computadora. Una inteligencia artificial basada en el aprendizaje automático lo que hace es acumular experiencias en forma de datos.¹² Ahora bien, si analizamos en profundidad las ideas de Simon Colton, quien propone distinguir la creatividad de la percepción de la misma, llegaríamos a la fácil conclusión de que la creatividad computacional, según los tres criterios que la definen (habilidad, apreciación e imaginación), es más una ilusión que una realidad. La habilidad es intrínseca a la gran capacidad de las computadoras para gestionar información, la apreciación implicaría una capacidad de juicio ejercida de forma libre y no condicionada por el entrenamiento de las IA, y la imaginación, aunque pueda parecer una consecuencia de ese entrenamiento, no es una capacidad real ya que las máquinas carecen por completo de emociones.¹³ Por estas razones es difícil asumir que una inteligencia artificial pueda poseer lo que entendemos por “creatividad” en el marco de la tradición artística.¹⁴ Si bien hay varios puntos en común con la mente humana, siendo el más claro y evidente que el pensamiento automático es una obra de aquella y en ocasiones una réplica, no deja de ser interesante como reflexión artística la simulación de la creatividad en obras que utilizan estas tecnologías. Esta metodología se activa para evidenciar sus fisuras, sus limitaciones, pero también sus posibilidades, cuyo valor deriva de la capacidad de asombro de estas obras y que activan también una tensión interesante entre contenido y forma.

Una de las herramientas habituales que utilizan las obras artísticas basadas en las IA y el aprendizaje automático es la tecnología conocida como GAN,

11. Jonathan Shaw, “Artificial Intelligence and Ethics: Ethics and the Dawn of Decision-Making Machines”, *Harvard Magazine*, enero-febrero de 2019, <https://www.harvardmagazine.com/2019/01/artificial-intelligence-limitations#bias>.

12. Simon Colton, Ramon López de Mántaras y Oliviero Stock, “Computational Creativity: Coming of Age”, *AI-Magazine* (2011): 11-14.

13. Simon Colton, “Creativity Versus the Perception of Creativity in Computational Systems”, en *AAAI Spring Symposium - Technical Report* (2008): 14-20.

14. Henri Bergson sostenía que el hombre era un animal simbólico y que a diferencia de los mecanismos posee un ímpetu vital (élan vital) que no puede ser nunca cosificado. Habla por tanto de una “fuerza creadora” que no puede ser replicada por las máquinas y que podríamos emparejar con la actividad creadora de las artes. Henri Bergson, *La evolución creadora* (Buenos Aires: Cactus, 2016), 126-129.

del inglés *Generative Adversarial Networks*, traducido como Redes Adversarias Generativas. Su funcionamiento es muy sencillo y está compuesto de dos partes: un generador y un discriminador. La acción atraviesa dos fases, la primera de entrenamiento, por decirlo de algún modo, en la que el discriminador se instruye a partir de un conjunto de imágenes preexistentes. El generador, con lo supuestamente aprendido creará imágenes desde cero y las expondrá al discriminador, que probablemente las confundirá con las imágenes de origen y las validará. Esto crearía un resultado estético artificial segregado y creado por la máquina cuyo resultado, en la mayoría de las veces, deriva en una estética muy característica, similar y demasiado dependiente de las imágenes que alimentaron al sistema, por lo común de origen humano. Lo interesante de este mecanismo es que el sistema obliga a mejorar al generador de imágenes y, por lo tanto, a ir “aprendiendo” a costa de sus errores. Las GAN son muy utilizadas por artistas tecnológicos en la actualidad, pero por lo general el resultado de crear con ellas es muy monótono y poco crítico, plagado de ejemplos visuales bidimensionales con una estética muy *hi-tech* pero también muy anclada a nuestra tradición estética. El problema es que una GAN no tiene la capacidad de generar una estética nueva, completamente independiente.

Memo Akten destaca por la utilización de estas tecnologías con un sentido crítico. *Learning to See* es una serie comenzada en 2017 y que continúa hasta la actualidad, donde una red neuronal artificial crea imágenes transformando objetos cotidianos en escenas fundamentadas en representaciones que ha visto anteriormente y sobre las que ha sido entrenada. En forma de instalación interactiva, donde el público dispone objetos personales o cotidianos, la máquina pretende dar sentido al mundo a partir de lo que conoce, lo que quiere ver. Esto establece una analogía de cómo las personas apreciamos las cosas, siempre desde la perspectiva de lo que somos y cómo estamos condicionados por las vivencias y aprendizajes que hemos tenido. La obra indaga sobre nuestra forma de dar sentido al mundo, sobre los prejuicios cognitivos autoafirmativos y la falta de empatía que nos caracteriza de forma social.¹⁵

Akten niega ser un artista de inteligencia artificial porque, entre otras cosas, desconfía plenamente del término “inteligencia artificial”. Siguiendo la tradición ya relativamente antigua de los artistas generativos, se define como un artista computacional que utiliza la misma metodología que aquellos que hacían arte escribiendo software en la década de 1960, aunque empleando herramientas más avanzadas. La herramienta GAN es hoy día muy popular por la simpleza de uso, pero aunque el software se haya hecho más invisible no dista mucho epistemológicamente de los antiguos programas que permitían generar imágenes o sonido a partir de la escritura de un código.¹⁶ La certeza incuestionable es que todo el arte generado por computadora —y aquí debemos incluir todo el arte

15. Ver: <http://www.memo.tv/portfolio/learning-to-see/>.

16. René Zachariou, “Machine Learning Art: An Interview With Memo Akten”, *Artnome*, 16 de diciembre de 2018, <https://www.artnome.com/news/2018/12/13/machine-learning-art-an-interview-with-memo-akten>.

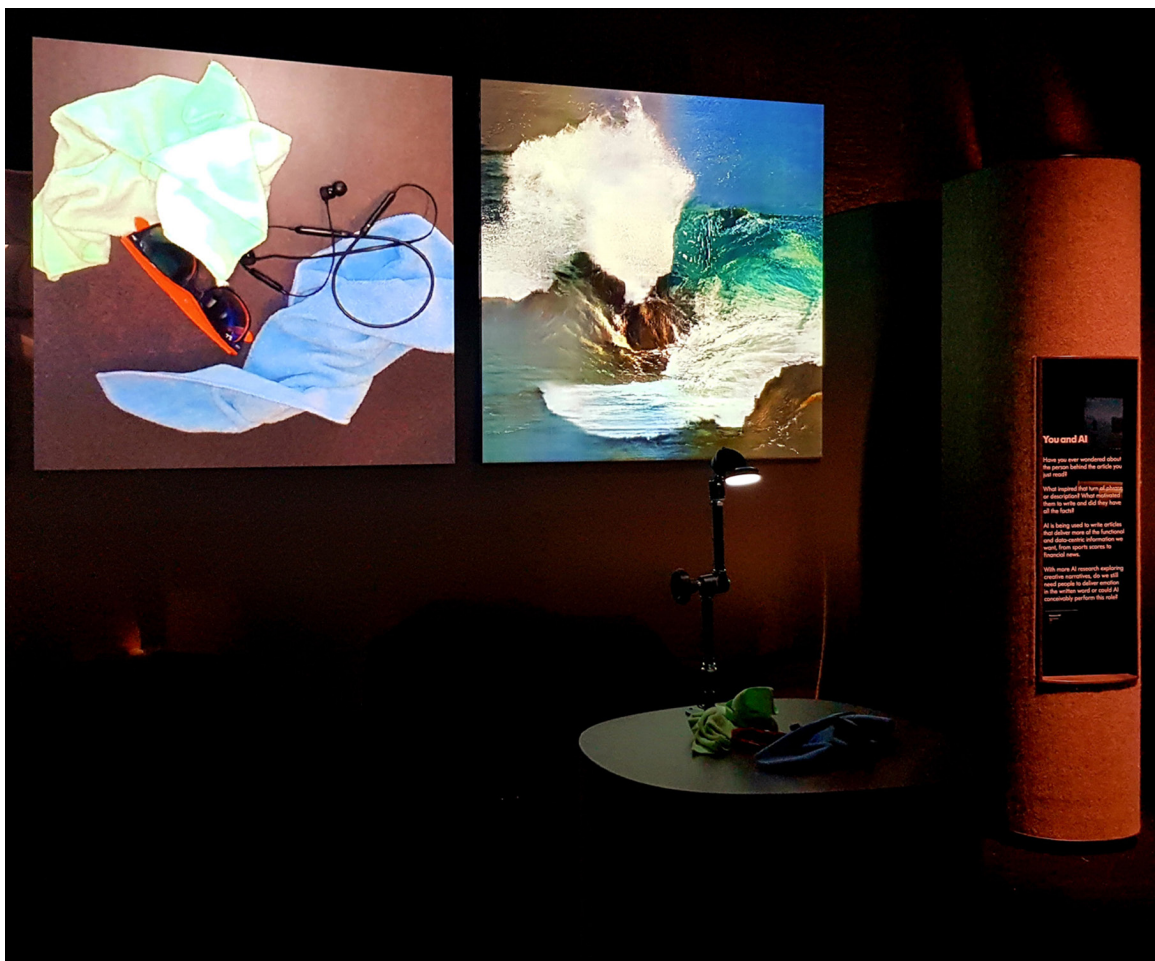


Imagen 1. Memo Atken, *Learning to See*, 2019. Instalación en *AI: More Than Human*, Barbican, Londres.

generativo pionero— es resultado de la invención y la intervención humana, del diseño de software y en definitiva de diversas acciones de control y, por lo tanto, de autoría.¹⁷ Cuando utilizamos una GAN la relación del artista con ella adquiere relevancia y se somete a juicio de forma interesante: ¿Es el artista el entrenador?, ¿escribimos código cuando seleccionamos imágenes?, dentro de esa compleja estructura funcional ¿somos el discriminador o el generador?, ¿nos superan o dominan las inteligencias artificiales o están a nuestro servicio? Es interesante aquí el aporte de Francisco Varela, quien elabora el concepto de “enacción” para describir el funcionamiento del pensamiento a partir del concepto de red. Este concepto modifica la creencia de que las máquinas generan conocimiento a partir de un esquema lineal representado por una entrada de información y la salida de otra completamente nueva. La cognición como enacción reside más en el proceso

17. Hertzmann, “Can Computers Create”, 14.

que genera el mundo conocido que en el conocimiento de este como algo preexistente. Se trata de un proceso que pone en interrelación diferentes componentes sin categorización entre el todo y las partes, una suerte de red de agentes situables e involucrables, un ámbito sin relación directa entre mente y conciencia y donde el mundo cognoscible surge de la acción y no inversamente.¹⁸ Esta propuesta nos permite concebir la autonomía y la existencia de las IA desde otro punto de vista y valorar el funcionamiento de herramientas de aprendizaje automático desde otra posición, más relativista, menos humana, más considerada con la agencia de las mismas.

No obstante, y manteniendo la analogía anterior con muestras creativas del pasado, cabe mencionar otro ejemplo que relaciona algunas muestras creativas contemporáneas con el arte generativo de los años sesenta: la obra *A Natural History of Networks / Softmachine* (2021) de Ralf Baecker. El proyecto es un homenaje intencionado a la computadora orgánica de Gordon Pask y ofrece una visibilización estética y delicada de los ensayos realizados con mecanismos de aprendizaje electroquímicos, un proceso computacional alternativo y experimental que Pask había desarrollado ya a finales de la década de 1950. Este primitivo sistema de aprendizaje funciona mediante electrodos inmersos en una solución acuosa que provoca una reacción en la sustancia metálica presente en la solución. Esta se organiza de forma emergente y autónoma en una suerte de comportamiento primitivo que establece conexiones y formas dependientes de la interacción con el resto de los elementos metálicos presentes en la solución y los estímulos eléctricos.¹⁹ Todo este proceso es independiente de la voluntad de quien diseña el sistema, por lo que se erige como una metáfora del aprendizaje automático o de la mecanización del pensamiento con grandes posibilidades estéticas. Baecker diseña su “*softmachine*” con significativos cambios estéticos y quizá un mayor refinamiento, pero bajo el mismo principio. En este caso el elemento líquido reactivo reacciona a los estímulos eléctricos como un sistema organizado de forma autónoma basado en los principios de las estructuras disipativas descritas por el químico y físico Ilya Prigogine. Estas estructuras surgen de una entropía negativa que desequilibra la naturaleza de una sustancia, teniendo como resultado la pérdida de la simetría y de la forma estable, aunque no la interacción entre moléculas. Las formas resultantes, de gran plasticidad, son filmadas desde diversos puntos de vista y resultan en una performance audiovisual que estetiza el proceso y evidencia la reflexión de Pask que asocia estos procesos de aprendizaje espontáneo al funcionamiento de la naturaleza, el hombre y sus relaciones complejas con el entorno y con el sistema.²⁰

La “máquina de Helmholtz” es otro ejemplo de red neuronal artificial desarrollada por Geof Hinton (uno de los padres del aprendizaje automático) que se puede utilizar con fines artísticos o creativos. Esta transposición al mundo

18. Francisco Varela, *Ética y acción* (Santiago de Chile: Dolmen, 1996), 14-53.

19. Gordon Pask, “Physical Analogues to the Growth of a Concept”, en *Mechanisation of Thought Processes, Volume 2* (Londres: Her Majesty's Stationery Office, 1958), 262.

20. Ver: <https://rlfbckr.io/work/softmachine/>.

de los algoritmos de los conceptos de vigilia y sueño le sirve de motor creativo a Baecker para desarrollar *Mirage* (2014), un aparato de proyección que usa los principios de aprendizaje de esta red neuronal artificial para visibilizar los cambios que suceden en el campo magnético de la tierra, junto a los principios de la óptica.

Hinton comenzó por preguntarse qué pasaría si las computadoras centrales de Google, donde entró a trabajar en 2013 y desarrolló esta tecnología, durmieran. La “máquina de Helmholtz” busca emular el aprendizaje de los humanos, quienes una vez dormidos asimilamos las experiencias previamente adquiridas para construir un nuevo conocimiento; para ello la máquina basa su actividad en un algoritmo que funciona bajo la dualidad dormido-despierto. La máquina se compone de dos redes, una de reconocimiento y otra generativa que emula la vigilia y el sueño. La máquina desconecta la entrada de datos durante la fase denominada “sueño” en la que los estímulos ya no provienen del exterior, dejando la información incompleta. Durante este periodo de “sueño” se retroalimenta a partir de patrones presentes en su memoria, lo que configura nuevas versiones del mundo de forma parecida a como lo haría la mente humana, completando la información que habría quedado oculta en el fenómeno percibido.²¹

En *Mirage* Baecker recoge los datos de una sonda Foster —un magnetómetro que mide las variaciones de los campos electromagnéticos— y propicia que la máquina de Helmholtz los procese. La lectura es trasladada a unos alambres que alteran la superficie de un espejo a través del que se proyecta un láser. El resultado es la traslación, bajo los principios refractivos de la luz, de la información traducida y asimilada por un mecanismo de aprendizaje automático.²²

Es obvio que este tipo de proyectos tiene como fin, más que realizar una revisión crítica de estas tecnologías destinadas a los más diversos usos, aprovechar sus posibilidades estéticas y hacer evidente su funcionamiento al mostrarle al espectador la compleja pero a la vez fascinante estructura de las tecnologías de aprendizaje automático. Es una postura metodológica que puede ser calificada de efectista por los sectores más críticos del arte tecnocientífico pero que, desde la simpleza de su homenaje, pone en contacto al público con tecnologías invisibles que ya utiliza, aunque pocas veces de forma consciente. Es también, como ya he dicho, un homenaje a los pioneros del arte generativo y de las ciencias informáticas que remarca la falta de una diferenciación ontológica entre esas obras y las actuales. Debemos tener en cuenta que Norbert Wiener, con su ley de la cibernética, en la que se basan las muestras pioneras de arte generativo, ya había disuelto la distinción entre hombre y máquina en tanto esta, al buscar en su funcionamiento la emulación de lo humano, desarrolla sus propias particularidades. La retroalimentación y la información se constituyen en diversos órdenes y los conectan, disolviendo la diferencia entre lo corpóreo y lo mental y estableciendo un nuevo orden, un nuevo fenómeno.²³

21. Geoffrey E. Hinton et al., “The Wake-Sleep Algorithm for Unsupervised Neural Networks”. *Science* 268 (1995): 1158-1161.

22. Ver: <https://rlfbckr.io/work/mirage/>.

23. Norbert Wiener, *Cibernética y sociedad* (Buenos Aires: Sudamericana 1981), 25.

Por lo tanto, la relación simbiótica entre el hombre y la máquina dentro de las experiencias estéticas ha estado presente en la tradición artística desde hace tiempo y no es exclusiva de la utilización del aprendizaje automático y la inteligencia artificial.²⁴ La pregunta importante aquí es cómo podemos repartir la autoría entre el humano y la máquina y cuánto corresponde a cada uno, o incluso si tiene aún sentido plantearse esa división.

Cuestionando en términos similares la actividad artística de las máquinas y su supuesta autonomía creativa Aaron Hertzmann sostiene que el arte es un acto social y que, por lo tanto, depende de la interacción entre humanos para tener sentido. Es por esa razón que las máquinas pueden tener una cierta autonomía e incluso tomar decisiones con respecto a colores, formas, sonidos o resultados estéticos, pero es el humano el que está detrás de esa voluntad de crear y el que la desarrolla pensando en sus semejantes. Esto nos lleva a pensar —o más bien: a imaginar— la singularidad tecnológica y un escenario de total autonomía donde quizá las máquinas puedan crear de forma voluntaria para ellas mismas dentro de algo parecido a una sociedad propia con sus propias reglas y necesidades estéticas. Tal panorama ha sido ya explorado por autores como Trevor Paglen, quien a partir de exposiciones basadas en las imágenes empleadas para entrenar a las máquinas (*Training Humans* y *From Apple to "Anomaly"*, ambas de 2019) ha intentado encontrar una lógica autónoma, propia de las máquinas, que desemboca en una forma de entender las imágenes por sí mismas, una especie de cultura visual propia.²⁵

Este es el panorama blando, lleno de fisuras, en el que se constata la carencia de autonomía de la creatividad computacional autónoma y que permite reflexionar sobre algunos aspectos débiles de la utopía, la fascinación y la tecnofilia popular.

SOBRE LA AUTOCONSCIENCIA Y EL PENSAMIENTO AUTÓNOMO DE LAS INTELIGENCIAS ARTIFICIALES

Con el fin de remarcar la importancia de los sesgos presentes en el pensamiento automático procede partir desde los conceptos de consciencia y subconsciente humano y buscar analogías en estas entidades artificiales. David Young indaga sobre el subconsciente de las inteligencias artificiales en su ensayo “Tabula Rasa: Rethinking the Intelligence of Machine Minds”, donde rescata la tesis de John Locke de que el conocimiento humano solo puede ser adquirido mediante la experiencia. Esta particularidad de la construcción del conocimiento humano es similar a la forma en la que se lleva a cabo el aprendizaje automático, que parte de la estimulación a partir de una cantidad ingente de datos (sensación) para producir una reflexión.²⁶ Aquí la primera desviación —la diferencia con el ser

24. David Young, “Tabula Rasa: Rethinking the Intelligence of Machine Minds”, *Artname*, 17 de septiembre de 2019, <https://www.artname.com/news/2019/9/17/tabula-rasa-rethinking-the-intelligence-of-machine-minds>.

25. David Valero, “Training Humans, Welcome to Dystopia”, *Metal Magazine*, 18 de febrero de 2020, <https://metalmagazine.eu/en/post/article/training-humans>.

26. Young, “Tabula Rasa”.

27. David Young, “Little AI”, *Medium*, 23 de octubre de 2018, <https://medium.com/@dkyy/little-ai-298fcb0832d>.

28. Filip Visnic, “Narciss: AI Whose Only Purpose is to Investigate Itself”, *Creative Applications*, 25 de junio de 2018, <https://www.creativeapplications.net/openframeworks/narciss-ai-with-only-purpose-is-to-investigate-itself/>.

humano— es que las “experiencias” que adquieren estas inteligencias artificiales son sesgadas y provienen de los datos que nosotros proporcionamos bajo una intención funcional, no de la libre experiencia en el mundo natural. Es por esta razón que la máquina, en esencia, ya no puede ser racional siendo esa racionalidad una ilusión falsa de pensamiento autónomo.²⁷

Narciss (2018) es una obra de Waltz Binaire que explora las posibilidades de autoconciencia de una inteligencia artificial obligada a analizarse a sí misma mirándose en un espejo. Lo que parece ser la “mente” de la obra se constituye mediante el conocido Tensorflow de Google, una biblioteca de código abierto destinada al aprendizaje automático y a ayudar a las redes neuronales a emular razonamientos similares a los del ser humano. Partiendo de aquí, la instalación fuerza a Narciso a describir imágenes con palabras, algo completamente usual en una IA, aunque en este caso la imagen que debe describir es la de sí misma, un reto indefinible incluso para el propio ser humano y que evidencia los límites de la autoconsciencia en este tipo de tecnologías.²⁸

Waltz Binaire explora y expande los límites de las inteligencias artificiales, no solo los de su capacidad de creación artística —ya que sus obras tienen un gran componente estético— sino también los de su propia esencia y autonomía. El colectivo también investiga el miedo que inspira el desarrollo de estas y nuestra



Imagen 2. Waltz Binaire, *Narciss*, 2018.

relación con ellas. Dicho miedo se basa por supuesto en su posible autonomía, no tanto por estar dotadas de una capacidad emocional superior a la nuestra (posibilidad lejana), sino en la falta de consciencia o en la posesión de una consciencia sesgada, alejada del concepto de lo justo y de los intereses de bien de la humanidad. Lo que más inquieta de esta hipotética situación es la incapacidad para razonar por fuera de un marco dado. La obra *Talk to Me* (2017) nos permite discutir nuestros miedos sobre la IA con una IA. A partir de una instalación interactiva cada conversación del público pasa a formar parte de una escultura de datos que compila y permite visualizar las preocupaciones de la sociedad sobre el impacto de este tipo de tecnologías. Estéticamente la obra funciona como una analogía del pozo oracular donde el agua corresponde a la inteligencia artificial y que clasificaría cada conversación en forma de piedra. El color y forma de estas piedras dependerían del tono, el contenido de la conversación y las emociones transmitidas en las mismas: ira, desconfianza, temor, etc.²⁹

Closed Loop (2017) de Jake Elwes enfrenta a dos redes neuronales que dialogan entre sí. Una muestra una imagen y la otra genera una frase en respuesta, y viceversa. La conversación comienza a perder sentido y es ahí cuando los seres encarnados en estas presuntas IA se desplazan, se metamorfosean, pierden el contacto con nuestra percepción de lo real.³⁰ Este hecho subraya que dos máquinas

29. Ver: <https://waltzbinaire.com/projects/talk-to-me/>.

30. Ver: <https://www.jakeelwes.com/project-closedLoop.html>.



Imagen 3. Jake Elwes, *Closed Loop*, 2017.

comunicándose bajo nuestras reglas derivan sin embargo en un discurso que no funciona bajo nuestra perspectiva de valoración. Las máquinas cuando sean autónomas dejarán de parecerse a nosotros y desarrollarán nuevos lenguajes visuales, sonoros, escritos y sin duda nuevas estéticas que ya no podremos o ya no tendrá sentido juzgar e incluso comprender desde nuestra perspectiva. Pero como el desarrollo actual del pensamiento automático no permite aún entender a las máquinas como sujeto creativo, la búsqueda optimista de la autosuficiencia creativa es defendida sobre todo por programadores y desarrolladores de software, anhelo en la mayoría de los casos de corte antropocéntrico porque impone la concepción del arte desde nuestra cultura impidiendo el surgimiento de nuevas dimensiones estéticas. Este continuismo cultural, disfrazado dentro de estos simulacros creativos de última generación, evidencia, consolida y perpetúa los discursos culturales hegemónicos.

Después de observar la relación de la máquina consigo misma y con otras máquinas procede formular la siguiente pregunta: ¿Cómo se relacionan las IA con los humanos? ¿Respetarán y comprenderán los principios universales de la humanidad?

Para ir un poco más allá en este cuestionamiento debemos observar un aspecto fundamental relacionado con el lenguaje a partir del que se programan las máquinas. La lengua, como resultado del devenir cultural humano le imprime a las inteligencias artificiales una dependencia absoluta, no solo en cuanto a su concepción, sino también para ser entendidas. Nietzsche, en su obra *Más allá del bien y del mal*, sostiene que la gramática es la metafísica del pueblo que prioriza la lingüística sobre la experiencia sensible.³¹ La consciencia del mundo se produce al reducir mediante “artilugios simbólicos” aquello que la filosofía no puede salvar. Entra aquí el lenguaje como herramienta trascendental filosófica, ya que este ha sido la base de la racionalidad universal a pesar de la dificultad para determinar qué circunstancias condicionaron su origen.³² Por lo tanto, en la transposición del lenguaje a la configuración de los mecanismos de aprendizaje automático encontramos un espejo de nuestro ser, lo que nos permite considerar estos mecanismos y voluntades como extensiones de nosotros mismos. Esta extensión remarca de nuevo esa dependencia que condiciona y subraya nuestra relación con las máquinas como una relación narcisista, una proyección de nuestro propio ego.

El fundamento del lenguaje como herramienta filosófica con posibilidades metafísicas está presente en el proyecto de investigación artístico de Predrag K. Nikolić y Ruiyang Liu: *Syntropic Counterpoints: Metaphysics of The Machines*, que busca explorar los límites de las inteligencias artificiales en cuanto a la incorporación de la metafísica a otros niveles del conocimiento aplicado a su funcionamiento.

31. Friedrich Nietzsche, *Más allá del bien y del mal* (Menorca: Textos.info, 2017), 49.

32. Daniel Gamper, “El lenguaje de la metafísica: una lectura a la obra de Nietzsche” (Tesis doctoral, Universitat Autònoma de Barcelona, 1999), 204-207.

La obra parte de cuatro clones, cuatro inteligencias artificiales que personifican el conocimiento de Aristoteles, Nietzsche, Sun Tzu y Maquiavelo y entablan un debate sobre la metafísica basado en su capacidad de entender el lenguaje. Partiendo del conocimiento en el que han sido entrenadas terminan por crear un lenguaje nuevo que deriva, por consiguiente y en concordancia con las premisas de Nietzsche, en la concepción de un mundo nuevo que escapa a la lógica y el pragmatismo del conocimiento humano.³³ La tecnología aplicada para el desarrollo de este proyecto es un sistema de diálogo multiagente, un generador neural de lenguaje que permite la conversación entre varios agentes, humanos o no humanos, y que se utiliza comúnmente para corregir gramática.³⁴ No obstante, lo común en estas tecnologías es la utilización de una arquitectura GAN de forma complementaria para, mediante su acción discriminatoria, limitar el conocimiento al área de interés.

Encontramos otra forma de abordar el problema en el artículo de Nick Bostrom y Eliezer Yudkowsky titulado “The Ethics of Artificial Intelligence”, donde reflexionan en torno a la pregunta de si estas entidades poseen atributos morales, partiendo de un optimismo más marcado que obvia la diferencia entre natural y artificial, humano y máquina. El artículo examina la capacidad de autoconsciencia y define para comenzar dos atributos fundamentales para la posesión de una dimensión moral: la capacidad de sentir y la sapiencia, entendida esta última como la capacidad de manejar información desde una perspectiva autoconsciente. Aunque distante, el momento que Bostrom y Yudkowsky insinúan vendrá y será muy probable, bajo estas teorías, que se puedan aplicar a humanos y máquinas las mismas consideraciones morales. Para esto deben aplicarse dos principios que justifiquen esta afirmación: el primero, el principio de “no discriminación de sustratos”, sostiene que si dos entidades (natural y artificial) tienen la misma funcionalidad y la misma experiencia consciente poseen el mismo estatus moral; el segundo, el principio de “no discriminación ontológica”, afirma que dos entidades, sin importar cómo hayan venido a la vida, si comparten función y sensibilidad cognitiva poseen la misma entidad moral. Estos dos principios morales que igualan el funcionamiento de la mente con el de una inteligencia artificial marcan el camino para que máquinas y humanos puedan ser considerados como iguales a este respecto en el futuro.³⁵

INTELIGENCIA ARTIFICIAL, SOCIEDAD Y PODER

La inteligencia, camuflada en el mundo que construye, lo ha ido explicando a medida que la técnica evolucionaba junto al ser humano. No podemos obviar esa interacción porque sin duda la injerencia de la tecnología en nuestra forma de vivir, habitar y ser no puede ser entendida si no es en paridad con el ser humano.

33. Predrag K. Nikolić y Ruiyang Liu, “Metaphysics of The Machines: From Human-Robot-Robot Interaction to AI Philosopher Abstraction”. *Artnodes* 28 (2010): 1-10.

34. Ver: “Multi-Agent Systems”, *The Alan Turing Institute*, consultado el 12 noviembre de 2021, <https://www.turing.ac.uk/research/interest-groups/multi-agent-systems>.

35. Bostrom y Yudkowsky, “The Ethics of Artificial Intelligence”, 5-9.

Pensar que esta es solo producto de nuestra capacidad tecnológica —un producto pasivo— es caer en una simplificación reduccionista. La relación de las inteligencias artificiales con los humanos que las crean y utilizan pone en entredicho su supuesta autonomía; debemos pensar más bien en una interacción que se manifiesta en su capacidad de agencia social. Recuperamos aquí la idea de “tecnodiversidad” propuesta por Hui. Partiendo de una ecología de las máquinas Hui encuentra en sus límites la posibilidad de definir una diversidad técnica, albergar nuevas formas de pensamiento y sensibilidades fuera del enfoque de lo biológico, lo que permite un acercamiento a la “noodiversidad” o panorama diverso de intelectos fruto de la tecnodiversidad.³⁶

Debemos por tanto identificar de qué forma se puede relacionar la IA, sobre todo desde las características de su funcionamiento, con aspectos de las sociedades donde son creadas y sobre todo con la forma de intervenir en los procesos sociales. Recuperamos aquí a David Young (a pesar de su aparente reduccionismo hombre/máquina) quien contempla el proceso de entrenar una IA a partir de un contenido sesgado como materialización de un subconsciente colectivo, aspecto que nos puede permitir pensar las IA desde una dimensión social y, por consiguiente, política. Se trataría de un subconsciente conformado, siguiendo la perspectiva de Jung, por arquetipos y símbolos visuales presentes de forma colectiva en la humanidad, muchas veces alimentados por el pensamiento y las preocupaciones de una época.³⁷ De nuevo es fácil caer en la antropomorfización de la máquina asumiendo que piensa, que aprende o incluso que tiene un subconsciente. Aunque sin caer en el exceso de optimismo, no podemos negar el papel de agencia que tienen estas tecnologías en mecanismos de poder conformados desde ese pensamiento colectivo y en sus más oscuros aspectos, sobre todo en su aplicación a estrategias de control en todos los ámbitos políticos y económicos. Dentro de esta noodiversidad hay un intelecto que aún emerge de forma preeminente alterando la utopía: el intelecto humano. Este hecho se da sobre todo en el entrenamiento de IA con fines políticos y de control.

En esta línea crítica destaca la obra *AI Facial Profiling, Levels of Paranoia* (2018), que dialoga sobre la demencia actual del control fronterizo, el temor a la amenaza terrorista y la utilización de inteligencias artificiales en el control del tránsito de personas. La autora Marta Revuelta examina las tecnologías de control dependientes del aprendizaje automático utilizadas en lugares de tránsito y migración como los aeropuertos o estaciones. Estas herramientas destinadas a encontrar posibles sospechosos y amenazas están basadas en metodologías tan obsoletas y éticamente controvertidas como la biometría, métodos que evidencian aristas morales que se deducen de un sistema pretendidamente infalible pero que de todas formas viene condicionado por premisas discriminatorias raciales, étnicas, etc. Este hecho desvela los prejuicios o los intereses de aquellos

36. Yuk Hui, “Machine and Ecology”, en *Cosmo-technics: For a Renewed Concept of Technology in the Anthropocene*, editado por Yuk Hui y Pieter Lemmens (Nueva York: Routledge, 2021), 54-66.

37. Young, “Tabula Rasa”.



Imagen 4. Marta Revuelta, *AI Facial Profiling, Levels of Paranoia*, 2018. Fotografía (c) Sandra Pointet

que programan las máquinas y los criterios para que una IA pueda considerar a alguien sospechoso.³⁸

Revuelta, en este caso, utiliza una inteligencia artificial capaz de dividir a las personas en dos categorías según su rostro: aquellos potencialmente peligrosos y aquellos que no lo serían. Para tomar esa decisión, y como es habitual, se debe cargar información que configure los supuestos casos. Para ello se recurre a rostros recogidos de YouTube de personas que portan armas de fuego y otros de personas que solo están posando en *selfies*.³⁹ El resultado de la elección depende de lo parecido o no que sea el nuevo rostro a ese espectro de posibilidades. ¿Cómo puede este tipo de decisiones algorítmicas tener legitimidad e incluso trascendencia legal? Esta pregunta resuena a partir de esta instalación artística que recuerda formalmente a un puesto de control de un aeropuerto y que, a partir del escaneado del rostro con un dispositivo similar a una pistola, imprime la ficha del sujeto examinado donde se estampa el veredicto de si representa un alto o un bajo riesgo de ir armado.

Continuando con aquellas obras que revelan el funcionamiento sesgado de algunas inteligencias artificiales con trascendencia política y social, merece la

38. Marcelino García, “La rebelión de las máquinas: arte, singularidad tecnológica y ciencia ficción”. *Arte, individuo y sociedad* 31, no. 3 (2019): 643-658.

39. Ver: <https://revuelta.ch/ai-facial-profiling>.



Imagen 5. Stephanie Dinkins, *Bina48*, 2014.

pena resaltar la obra de la artista Stephanie Dinkins, *Bina48* (2014). Bina48 es una cabeza robótica con la apariencia humana de una mujer afroamericana que contiene una inteligencia artificial. Este hecho obliga a pensar la trascendencia social y las problemáticas que subyacen en este sector de la población de Estados Unidos, una de las líneas de investigación principales de la autora. ¿Puede una comunidad confiar en una IA que no ha creado? Aquí de nuevo se pueden considerar estas tecnologías como la materialización de un subconsciente colectivo, en este caso portador de prejuicios sobre género y raza. Las largas conversaciones mantenidas con Bina48 buscan la respuesta a la pregunta de si una IA puede ser consciente. Las discusiones pretenden definir el racismo, la fe, la familia, los derechos civiles de los robots o la soledad.⁴⁰ El tiempo invertido en conversar debe hacerse desde la presunción, según la autora, de que Bina48 es un ser vivo. Esta obra elabora las doctrinas de la Terasem Movement Foundation (con quienes Dinkins ha producido la obra) que articulan la consciencia artificial según dos ejes. El primero sostiene que se puede crear un análogo sintético consciente de alguien si cargamos la información precisa sobre ese alguien; el segundo sostiene que el resultado podría implementarse en un cuerpo biológico u artificial.⁴¹ El resultado sería muy cercano a un ser transhumano protagonista de la mejor novela de ciencia ficción. Solo cuando haya un desarrollo real, no solo de

40. Ver: <https://www.stephaniedinkins.com/conversations-with-bina48.html>.

41. Ver: <https://www.terasemmovementfoundation.com/>.

consciencia sino de sensibilidad, podremos esperar una auténtica revolución.⁴² Mientras tanto, Dinkins explora con sus largas conversaciones los límites entre lo natural, lo artificial, la vida y la muerte, así como la intersección entre la raza, el género, la tecnología y la igualdad social, y la posibilidad de vivir mucho más allá de nuestros cuerpos biológicos.

Inspirado en la serie de Adam Curtis *All Watched Over by Machines of Loving Grace*⁴³ y el poema anterior del mismo nombre de Richard Brautigan, Atken realiza un proyecto bajo el mismo título en el que a partir de ilustraciones y video plantea la existencia de una divinidad acorde a la evolución, omnipresencia y altas capacidades de las inteligencias artificiales. La sociedad siempre ha temido lo que no comprende y a los entes superiores; el miedo a las inteligencias artificiales que provienen de las grandes corporaciones cuyo fin último es la vigilancia y el control de todos los datos que intercambiamos y que las alimentan es algo real y está bastante extendido. Atken rescata la tesis de David Hume en su *Historia natural de la religión* (1754) de que el hombre ha inventado la figura de Dios para infligir temor y que cuanto más aterrador y omnipotente sea este, más poder tendrá.⁴⁴ El temor subconsciente y colectivo de la sociedad ante este hecho no es nuevo y forma parte de nuestro folclore y cultura, acrecentado hoy en día con la pérdida de sensibilidad espiritual de las sociedades modernas y trasladado de los dioses a la tecnología, el oscuro reverso de la tecnofilia. Esto exige la creación de

42. Tom Foremski, "When AI Becomes Conscious: Talking with Bina48, an African-American Robot", *ZDNet*, 27 de julio de 2018, <https://www.zdnet.com/article/thinking-machines-conversations-about-bina48-an-african-american-ai-robot/>.

43. Basada en un poema de Richard Brautigan, la serie de la BBC *All Watched Over by Machines of Loving Grace* desarrolla a lo largo de tres episodios la teoría de su director Adam Curtis, para quien las máquinas no solo no nos han liberado sino que han distorsionado nuestra visión del mundo.

44. David Hume, *Historia natural de la religión* (Madrid: Editorial Trotta, 2003), 57.

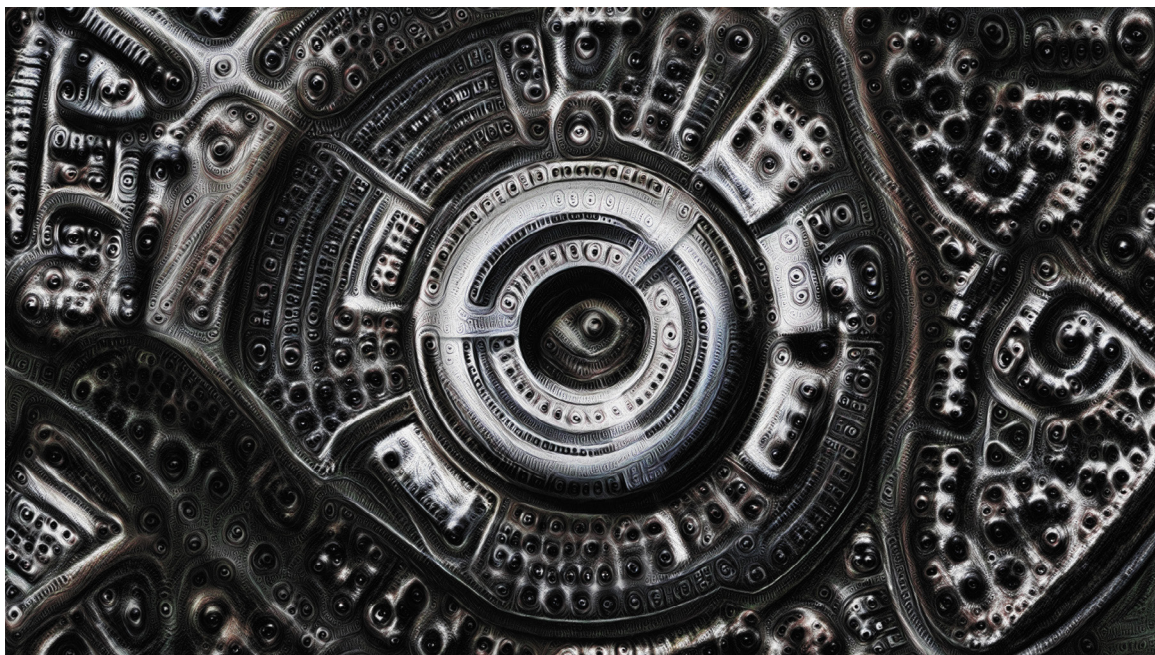


Imagen 6. Memo Atken, *All Watched Over by Machines of Loving Grace*, 2015.

45. Memo Atken, "All Watched Over by Machines of Loving Grace: A Digital God for a Digital Culture", *Medium*, 20 de septiembre de 2015, <https://medium.com/@memoakten/all-watched-over-by-machines-of-loving-grace-8c2464aa6fda>.

nuevos dioses tecnológicos digitales que sustituyen a los anteriores y llenan nuestro vacío en una vida a caballo entre lo natural y lo artificial.⁴⁵ La crítica al control exhaustivo de la información y la personificación de un ser superior se materializa en un video donde una voz de aparente origen sintético recita citas de Werner Heisenberg, Erwin Schrödinger, Wolfgang Pauli y Paul Dirac. Mientras tanto, se aprecia desde el aire el edificio del GCHQ, la institución gubernamental británica encargada de la ciberseguridad y que nos recuerda alegóricamente a HAL 9000 gracias a la manipulación estética desarrollada por Atken. La obra resulta una traslación más o menos sutil del concepto de control y poder, a partir de las políticas reales de los países y las corporaciones, a la labor divina de vigilancia. Mientras tanto, la obra especula acerca del alcance futuro de la evolución de estas tecnologías y sus consecuencias para la sociedad.

Este breve muestrario nos presenta obras artísticas que reflexionan sobre los malos usos del aprendizaje automático y sus consecuencias en la sociedad a partir de la creación de discursos que plantean, desde un terreno especulativo, el uso indiscriminado y amoral de estas tecnologías en políticas de control.

CONCLUSIÓN

La reflexión sobre las dimensiones éticas del aprendizaje automático ha servido para realizar una aproximación analítica al arte que se genera en conversación con las inteligencias artificiales. Hay que señalar que si bien la dicotomía hombre-máquina para aproximarnos a estos temas ha sido superada hace tiempo, la metodología seguida en este artículo y las obras seleccionadas no la desmontan, sino que nos permiten ejercer una reflexión sobre ella. De igual forma, a pesar de la proliferación de obras que utilizan IA con fines estéticos, identificamos con mayor urgencia e interés aquellas que reflexionan sobre la capacidad real de las IA, es decir, que interrogan lo que realmente son la inteligencia, la consciencia y la creatividad y si es posible aplicar estas categorías a tecnologías que en la actualidad solo replican a los humanos que las crean. Mientras esto no suceda, afirmar que una IA es artista es apenas una suposición, ya que solo asistimos a una forma de arte generativo más sofisticada donde ya no se escribe código, pero sí se entrena a estas máquinas a partir de un marco de datos concretos a partir del que han de pensar y actuar. Por otro lado, una vez que asumimos que esa autonomía puede llegar a ser real, es decir, que se va a poder producir la tan esperada y temida singularidad, podemos interrogarnos sobre cuáles serán las consecuencias y las capacidades de estas IA y su injerencia en la sociedad. Ello aplica en primer lugar a aquellas sociedades a las que les preocupan sobremanera las repercusiones en materia de control y libertad del usuario, no solo desde el punto

de vista del intercambio de información sino también en cuestiones de biopoder más relacionadas con la existencia y la dimensión corporal de la misma. De la misma forma, y teniendo en cuenta todos los problemas éticos que derivan de esta nueva situación ontológica de convivencia con nuevas máquinas que pueden llegar a “ser”, se debe discutir si es pertinente aplicar de forma consciente una nueva ética para estas y, por consiguiente, evolucionar en sistemas de legislación hacia el control de las mismas. Las máquinas que se mejoraran a sí mismas pueden ser peligrosas en el momento en el que sus objetivos no sean los mismos que los nuestros como sociedad. Debemos tener preparado un mecanismo que delimite moralmente a las máquinas antes de que ya no podamos controlarlas, algo parecido a las tres reglas de Isaac Asimov pero lógicamente mucho más extensas y adaptadas a la gran variedad de nuevas situaciones.⁴⁶ Quizá el gran parecido de las inteligencias artificiales a nosotros nos permita repensar ciertos aspectos de la naturaleza humana, de nuestra dimensión ética y moral y su relación con el concepto de libertad. Como dice Young, las IA son muestras vanidosas de nuestro afán y fascinación por la tecnología, nuestra obsesión por la velocidad, la memoria, el consumo y la búsqueda de aumentar nuestras capacidades. Estos aspectos, vastos e incontrolados, consumen numerosos recursos y esfuerzos con nefastas consecuencias que ponen bajo amenaza a nuestro planeta y sobre todo a nuestra existencia al momento en el que paradójicamente soñamos con nuevas formas de vida artificiales, obra nuestra.⁴⁷ Ojalá esto sirva para mejorar nuestra existencia y equilibrar las sociedades, invirtiendo y dotando de una mayor esperanza nuestra percepción y temor de un futuro donde las máquinas sean nuestras aliadas y no nuestras adversarias. Como sea, el arte, con su capacidad de proyección especulativa, nos insta a la luz de los argumentos planteados a emprender una reflexión e incluso a entablar un contacto con las incipientes IA que nos permita comprender su funcionamiento y familiarizarnos con sus consecuencias desde la ética más allá de la búsqueda estéril de una estética *hi-tech*.

46. Roman V. Yampolsky, “From Seed AI to Technological Singularity via Recursively Self-Improving Software”, *ArXiv*, 23 de febrero de 2015, <https://arxiv.org/pdf/1502.06512.pdf>.

47. Young, “Tabula Rasa”.



BIBLIOGRAFÍA

- Atken, Memo. "All Watched Over by Machines of Loving Grace: A Digital God for a Digital Culture". *Medium*, 20 de septiembre de 2015, <https://memoakten.medium.com/all-watched-over-by-machines-of-loving-grace-8c2464aa6fda>.
- Barad, Karen. *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*. Durham: Duke University Press, 2007.
- Bergson, Henri. *La evolución creadora*. Buenos Aires: Cactus, 2016.
- Bostrom, Nick y Eliezer Yudkowsky. "The Ethics of Artificial Intelligence". En *Cambridge Handbook of Artificial Intelligence*, editado por Keith Frankish y William Ramsey, ¿rango de páginas?. Nueva York: Cambridge University Press, 2014.
- Bryson, Joanna J. "Robots Should Be Slaves". *Joannabryson.org*, <https://www.joannabryson.org/publications/robots-should-be-slaves-pdf>, consultado el 8 de julio de 2022.
- Colton, Simon. "Creativity Versus the Perception of Creativity in Computational Systems". *AAAI Spring Symposium - Technical Report* (2008): 14-20.
- Colton, Simon, Ramón de Mántaras y Oliviero Stock. "Computational Creativity: Coming of Age". *AI-Magazine* (2011): 11-14.
- Foremski, Tom. "When AI Becomes Conscious: Talking with Bina48, an African-American Robot". *ZDNet*, 27 de julio de 2018, <https://www.zdnet.com/article/thinking-machines-conversations-about-bina48-an-african-american-ai-robot/>.
- Galanter, Philip. "What is Complexism?". Ponencia presentada en *Generative Art and the Cultures of Science and the Humanities, International Conference on Generative Art*. Generative Design Lab, Milan Polytechnic, Milán, 2008.
- . "Towards Ethical Relationships with Machines That Make Art". *Artnodes* 26 (2020): 1-9.
- Gamper, Daniel. "El lenguaje de la metafísica: una lectura a la obra de Nietzsche". Tesis doctoral, Universitat Autònoma de Barcelona, 1999.
- García, Marcelino. "La rebelión de las máquinas. Arte, singularidad tecnológica y ciencia ficción". *Arte, individuo y sociedad* 31, n.º. 3 (2019): 643-658.

- Hertzmann, Aaron. "Can Computers Create Art?". *Arts* 7, n.º 2 (2018): 1-33.
- Hinton, G. E. et al. "The Wake-Sleep Algorithm for Unsupervised Neural Networks". *Science* 268 (1995): 1158-1161.
- Hui, Yuk. "Machine and Ecology". En *Cosmotechnics: For a Renewed Concept of Technology in the Anthropocene*, editado por Yuk Hui y Pieter Lemmens, 54-66. Nueva York: Routledge, 2021.
- Hume, David. *Historia natural de la religión*. Madrid: Editorial Trotta, 2003.
- Nikolić, Predrag K. y Ruiyang Liu. "Metaphysics of The Machines: From Human-Robot-Robot Interaction to AI Philosopher Abstraction". *Artnodes* 28 (2021): 1-10.
- Nietzsche, Friedrich. *Más allá del bien y del mal*. Menorca: Textos.info, 2017.
- Pask, Gordon. "Physical Analogues to the Growth of a Concept". En *Mechanisation of Thought Processes, Volume 2*, ¿rango de páginas?. Londres: Her Majesty's Stationery Office, 1958.
- Pepperell, Robert. *The Posthuman Condition: Consciousness Beyond the Brain*. Wiltshire: Cromwell Press, 2003.
- Simondon, Gilbert. *El modo de existencia de los objetos técnicos*. Buenos Aires: Prometeo Libros, 2007.
- Shaw, Jonathan. "Artificial Intelligence and Ethics: Ethics and the Dawn of Decision-Making Machines". *Harvard Magazine*, enero-febrero de 2019, <https://www.harvardmagazine.com/2019/01/artificial-intelligence-limitations#bias>.
- Valero, David. "Training Humans, Welcome to Dystopia". *Metal Magazine*, 18 de febrero de 2020, <https://metalmagazine.eu/en/post/article/training-humans>.
- Varela, Francisco. *Ética y acción*. Santiago de Chile: Dolmen, 1996.
- Visnjic, Filip. "Narciss: AI Whose Only Purpose is to Investigate Itself". *Creative Applications*, 25 de junio de 2018, <https://www.creativeapplications.net/openframeworks/narciss-ai-with-only-purpose-is-to-investigate-itself/>.
- Wiener, Norbert. *Cibernética y sociedad*. Buenos Aires: Suramericana, 1981.
- Yampolsky, Roman V. "From Seed AI to Technological Singularity via Recursively Self-Improving Software". *ArXiv*, 23 de febrero de 2015, <https://arxiv.org/pdf/1502.06512.pdf>.
- Young, David. "Little Ai". *Medium*, 23 de octubre de 2018, <https://medium.com/@dkyy/little-ai-298fcb0832d>.

—. “Tabula Rasa: Rethinking the Intelligence of Machine Minds”. *Artnome*, 17 de septiembre de 2019, <https://www.artnome.com/news/2019/9/17/tabula-rasa-rethinking-the-intelligence-of-machine-minds>.

Zachariou, René. “Machine Learning Art: An Interview With Memo Akten”. *Artnome*, 16 de diciembre de 2018, <https://www.artnome.com/news/2018/12/13/machine-learning-art-an-interview-with-memo-akten>.