

Todo es causal: inteligencia artificial como mecanismo para reforzar arquetipos

Marco López-Paredes
Observatorio de Comunicación (OdeCom)
Pontificia Universidad Católica del Ecuador
mvlopez@puce.edu.ec
<https://orcid.org/0000-0001-7853-5564>

Andrea Carrillo-Andrade
Observatorio de Comunicación (OdeCom)
Pontificia Universidad Católica del Ecuador
acarrillo745@puce.edu.ec
<https://orcid.org/0000-0002-0779-6991>

Introducción

“Teatro de ópera espacial”, de Jason Allen, abrió el debate sobre la autoría de imágenes digitales tras ganar el Concurso de Arte Digital (Colorado, Estados Unidos), habiendo utilizado la plataforma Midjourney. Esta plataforma utiliza inteligencia artificial para generar —en cuestión de minutos— imágenes “inéditas”, a partir de descripciones textuales. Este tipo de tecnología se basa en el reconocimiento de patrones matemáticos a partir del big data que colecciona de autores y usuarios. Así, si la tecnología es usada para la generación de imágenes en un mundo cada vez más visual, cabe recordar la famosa primera ley de la tecnología de Melvin Kranzberg: “La tecnología no es ni buena ni mala; ni es neutral”. Surge, entonces, la pregunta de esta investigación: ¿qué arquetipos son reforzados en las plataformas de ilustración que utilizan inteligencia artificial? Para contestarla, se hace una exploración acerca de la formación de arquetipos y su función en el storytelling empleado en las ilustraciones creadas por inteligencia artificial.

Arquetipos y su función en la creación de historias

Adolfo Bastian (Giraldo, 1961), estudiando la psicología de los pueblos, fue el primero en poner de relieve la existencia de ciertas ideas primitivas universalmente difundidas. Este sería el punto de partida para lo que Jung denominaría inconsciente colectivo, que se relaciona directamente con la formulación de arquetipos. De acuerdo con Jung (1970), el inconsciente colectivo se refiere a un lugar de reunión de contenidos olvidados y reprimidos de naturaleza universal. Tiene contenidos y modos de comportamiento iguales en todas partes y en todos los individuos: los arquetipos. Así, por ejemplo, Lévy-Bruhl (en Jung, 1970) usaba este concepto para hablar de las figuras simbólicas de la cosmovisión primitiva; en cambio, en las doctrinas tribales se refieren a fórmulas conscientes que se transmiten a través de la tradición, como en los mitos y las leyendas. No obstante, Usener explica que los arquetipos pueden resurgir en todo tiempo y lugar sin que haya alguna influencia migratoria exterior. Esto es parte de la teoría de la Gestalt, dada la preformación inconsciente en las estructuras (Giraldo, 1961). Por ello, los arquetipos resultan útiles en la elaboración de cualquier historia dado que las conexiones y representaciones se forman de manera inmediata: “El arquetipo representa esencialmente un contenido inconsciente, que al concienzializarse y ser percibido cambia de acuerdo con cada conciencia individual que surge” (Jung, 1970, p. 12).

Asimismo, Becker y Neuberg presentan una teoría de los arquetipos. Afirman que son estructuras mentales (patrones de percepción, memoria, y acción), que están en consonancia con la teoría de la narrativa y la literatura. Los investigadores afirman que esto ocurre dado que los arquetipos son una consecuencia de la dinámica evolutiva y la experiencia personal (en Green *et al.*, 2019). Ahora bien, en consonancia con esta aproximación, (Fitzgerald *et al.*, 2020) evidencian que los lectores responden a las narrativas en la misma manera independientemente de si están etiquetados como “hecho” o “ficción”. De hecho, Aristóteles habla de la mimesis praxeos o representación de la

vida humana, que es lo que permite al espectador realizar el proceso de recepción de una obra.

En literatura, un arquetipo se refiere a un personaje, acción, tema, símbolo, escenario o situación fácilmente identificables por el lector, dado que representan los patrones universales de la naturaleza humana y dan forma a la estructura de una historia, que se intuye gracias a la inconsciencia colectiva de la humanidad. De hecho, en una carta de Julio Cortázar a Ana María Hernández, Ph.D., en literatura comparada, el escritor reclama: “me gustaría que algunas de las cosas que he escrito, por lo menos, no estén sometidas a las huellas preestablecidas por el inconsciente colectivo” (Bernández y Álvarez, 2012). Lo dice una vez que la estudiosa enfrenta que la obra de Cortázar responde al itinerario del héroe y a la estructura de los mitos y arquetipos, ya que las historias están determinadas por estructuras primordiales e influencias literarias. Asimismo, Christopher Booker, periodista y autor inglés, a partir de estas definiciones, presentó siete arquetipos en las historias: superar al monstruo, pasar de la pobreza a la riqueza (en el más amplio sentido), la búsqueda del héroe, el viaje, situaciones comedia, tragedia (historias sin final feliz), y el renacimiento del héroe (D&AD in association with Creative SkillSet & FutureLearn, n.d.).

Imágenes: historias y arquetipos

De acuerdo con Giraldo Ángel (Giraldo, 1961), “la imagen y la representación psíquica traducen, hasta cierto punto, un mismo fenómeno mental”. Asimismo, “la identificación con los personajes es un mecanismo a través del cual los sujetos experimentan e interpretan una narración desde dentro, como si los acontecimientos que se relatan les estuviesen ocurriendo a ellos mismos” (Igartua, 2008, p. 43).

En este sentido, las imágenes son útiles solamente en la medida en la que son reconocibles, pero, a su vez, son codificadas de acuerdo con una convención social. Por ello, el significado que una imagen pueda tener como tal no es inherente a las imágenes y los objetos en sí o no se traducen literalmente, sino que operan como signos de los

que podemos ser o no conscientes gracias a este inconsciente colectivo. Alex Potts (1996) escribió que las imágenes y los objetos son no solo mediados por las convenciones, sino que el significado es activado en gran medida por la cultura.

Dentro de la ilustración, uno de los retos fundamentales es transmitir una historia dentro de la imagen a través de la composición. Para Berguer (1976), “la composición responde a uno de los anhelos esenciales del espíritu, el de coherencia, y a una aspiración profunda de nuestro ser, que es la de constituir una realidad de la que sea autor el hombre” (p. 140). En este sentido:

El proceso de composición es el paso más importante en la resolución del problema visual. Los resultados de las decisiones compositivas marcan el propósito y el significado de la declaración visual y tienen fuertes implicaciones sobre lo que recibe el espectador. (Dondis 1976, p. 33)

Por lo tanto, Jones (en Giraldo, 1961) define la imagen como:

Una forma mantenida indefinidamente en el inconsciente y que a menudo se identifica con personas distintas del sujeto. De acuerdo con esto sería un sincretismo representativo, o como lo apuntaba mi maestro R. De Saussure una ruta hacia el camino de la ilusión representativa. (p. 131)

Midjourney, principal plataforma de ilustración a partir de la inteligencia artificial se presenta a partir de dos conceptos: *inteligencia artificial* y *arte*. El primer concepto lo define como “la teoría y el desarrollo de sistemas informáticos capaces de realizar tareas que normalmente requieren la inteligencia humana, como la percepción visual, el reconocimiento del habla, la toma de decisiones y la traducción entre idiomas” (Midjourney, 2022, s/p). En cambio, el arte es “la expresión o aplicación de la imaginación y la habilidad creativa humana, típicamente en una forma visual como la pintura o la escultura, produciendo obras que se aprecian principalmente por su belleza o poder emocional” (Midjourney, 2022, s/p).

Resalta la relación que hace con los sentimientos, en una concepción muy similar al *storytelling*, que una vez más, encuentra su base en los arquetipos. De hecho, cuando se habla de *storytelling* (y, en todo caso, de cualquier proceso de interpretación) se entra en un territorio dual, a la vez individual y social, que es más bien como el mar que va y viene, que no se está quieto. Para Andy Cameron, las historias que el hombre crea dan una identidad a la existencia y evidencian la manera en que cada cultura se percibe a sí misma y al mundo (Saavedra-Bautista *et al.*, 2017).

Inteligencia artificial (IA)

En palabras de Lasse Rouhiainen:

La inteligencia es la capacidad de las máquinas para usar algoritmos, aprender de los datos y utilizar lo aprendido en la toma de decisiones tal y como lo haría un ser humano. Sin embargo, a diferencia de las personas, los dispositivos basados en IA (...) pueden analizar grandes volúmenes de información a la vez. Asimismo, la proporción de errores es significativamente menor en las máquinas que realizan las mismas tareas que sus contrapartes humanas. (Rouhiainen, 2018, p. 17)

La inteligencia artificial se basa en el aprendizaje automático (en inglés, *machine learning*). Se trata un aspecto de la informática en el que los ordenadores o las máquinas tienen la capacidad de aprender sin estar programados para ello. Así, las máquinas aprenden no solo a partir de las consultas que se hacen, sino de la forma en que se los utiliza y, con una misma tecnología, son capaces de ofrecer una experiencia al usuario personalizada. Un mecanismo base para el desarrollo de la inteligencia artificial es su uso en el juego; históricamente, se ha alimentado y entrenado a las máquinas a partir de la competencia en juegos como ajedrez, *Jeopardy* o *Go* contra humanos. En junio de 2018, la empresa OpenAI anunció que su IA había conseguido ganar a varios equipos humanos en el juego estratégico multijugador Dota 2: “la compañía entrenó a los jugadores de IA jugando entre ellos mismos y estos fueron capaces de obtener conocimientos y habilidades en

un día para los que un ser humano necesitaría ciento ochenta años” (Rouhiainen, 2018, p. 32).

La inteligencia artificial contempla tres tipos de tareas: dominios formales, dominios técnicos y dominios cognitivos. En la primera fase, la IA pretende solucionar problemas. Lo hace a través de modelos de búsquedas (algorítmico o heurístico). En una segunda fase, el dominio técnico se refiere al conocimiento científico-técnico y, finalmente, el dominio cognitivo intenta comprender el funcionamiento del cerebro humano y sus funciones cognitivas (razonar, oír, hablar, o incluso las emociones) para emular dichos procesos con modelos computacionales (Malagón, n.d.).

La tecnología usada en la generación de imágenes se sustenta en redes neuronales artificiales (Artificial Neural Networks, ANNs); se utilizan para una variedad de tareas, y, por ejemplo, recopila conjuntos de datos de imágenes como diferentes razas de perros, entrenar una red neuronal en las imágenes y, luego, si proporciona una nueva imagen de un perro, dará una puntuación estadística sobre qué tan cerca la nueva imagen coincide con el modelo, entonces, mostrará qué raza de perro es la imagen (Adams, 2019). En otras palabras, la inteligencia artificial se alimenta de un repositorio cuyo modelo privilegia el *mainstream* en el sentido en que reconoce con facilidad lo más consumido, por lo que la minoría queda subrepresentada.

Metodología

Esta investigación tiene un alcance descriptivo; en este sentido, pretende describir el reciente fenómeno de la inteligencia artificial para modificar y crear imágenes desde plataformas que dan la impresión de “programar” ilustraciones de diferentes estilos, pero que se basan en algoritmos y el ordenamiento básico a partir del uso de palabras clave en las que se sustentan los buscadores de Internet. Así, a partir de comandos que pueden ser complejos o muy simples, se generan imágenes profesionales en cuestión de segundos. En *Los bárbaros* (Baricco, 2010) explica que los buscadores tuvieron que enfrentarse a

su principal problema: el ordenamiento de la información antes que su recopilación. De ahí que los creadores de Google, Larry Page y Sergei Brin, solucionaron esto a partir del uso de etiquetas, cuyo principio se sustenta en la lógica de la repetición —que a su vez profundiza las brechas para las minorías— y sus interrelaciones. Esta base de ordenamiento es el objetivo de esta investigación, ya que se parte de la hipótesis de que estos generadores de ilustraciones reproducen una cultura *mainstream*, teniendo en cuenta que la cuna de la inteligencia artificial está en San Francisco, Estados Unidos.

Para responder a la pregunta de investigación ¿qué arquetipos son reforzados en las plataformas de ilustración que utilizan inteligencia artificial?, se aplicó una investigación cualitativa en dos etapas. En una primera, se procedió a recolectar datos a través de un muestreo no probabilístico y, en una segunda, se hizo una entrevista a la ilustradora Paola Karolys para el procesamiento de los datos. Por lo tanto, llevó a cabo una observación no participativa para explorar las plataformas Midjourney, Craiyon y DeepAI y determinar sesgos de representatividad y sistémicos. Se optó por esta metodología ya que la observación no participativa consiste en registrar el comportamiento en el entorno habitual del sujeto.

No obstante, vale resaltar que la inteligencia artificial está sustentada tanto la teoría de la información “clásica” (Shannon) como la algorítmica. Ajitesh Kumar (Kumar, 2022), autor en *Data Analytics*, explica que la teoría de la información es el estudio de cuánta información está presente en las señales o datos del entorno. La inteligencia artificial se trata de extraer estas informaciones a partir de los datos que luego, que se utilizan para construir los modelos. Así, la teoría de la información comienza con la idea de que esta cantidad se puede medir por el número mínimo de bits necesarios para describir la observación. Pero mientras que la teoría de Shannon considera métodos de descripción que son óptimos en relación con alguna distribución de probabilidad dada, la teoría algorítmica de Kolmogorov adopta un enfoque no probabilístico diferente: cualquier programa de com-

putadora que primero calcule (imprima) la cadena que representa la observación y luego termine, se considera una descripción válida (Grunwald y Vitanyi, 2008). En todo caso, se resalta que la observación se realizó sin intervención para observar el comportamiento tal como ocurre de forma natural.

Tras la recolección de datos no probabilístico se generaron matrices para analizar las imágenes en su significante (composición de la imagen) y significado (interpretación), según las definiciones clásicas propuestas por Saussure:

La lengua como un sistema de signos interdependientes, en los que el valor de cada término resulta solamente de la presencia simultánea de los otros. La palabra forma parte de un sistema y está revestida no solo de una significación, sino también, y sobre todo, de un valor. (Saussure, 2008, p. 146)

Saussure (2008) explicita que, “en lugar de ideas dadas de antemano, hay que considerar los valores que emanan del sistema” (p. 147) lo que está en estrecha relación con la teoría jungiana de los arquetipos:

En la lengua, cada término o elemento tiene un valor correspondiente por su oposición con los otros términos. Los valores dependen de una convención inmutable: los principios de la semiología. Además, son relativos, ya que la relación entre la idea y el sonido es arbitraria, y son negativos porque la relación entre los signos lingüísticos de un sistema se define de forma negativa, ya que un signo es lo que otro no es. (Saussure, 2008, p. 147)

Esto sin decir que esta investigación entienda a los significados como arquetipos, sino más bien en el sentido de que: el análisis de los símbolos, en su función semiótica y en la descripción de Jung permite comprender dos modos de conducta simbólica que trazarían:

1. Un macrocosmos de significaciones universales, universo del discurso con apariencia de invariante por su tiempo psicológico lento, de tan generalizado, cuyas normas o pautas serían las imágenes arquetípicas de Jung universales,

impersonales, de origen inconsciente y comunes a todas las razas y épocas (...)

2. El microcosmos de los modos de representación de cada cultura. Variante evidente, sus significados estarían normados (diacrónicamente), por los arquetipos jungueanos a los cuales se sumarían los factores de cambio, propios de la dinámica de todo fenómeno cultural (Sager 1967, p. 10).

La investigación se complementa a través de una entrevista semiestructurada dado que esta técnica de recolección de datos permite averiguar hechos no observables fundamentales para la comprensión de los arquetipos, como: significados, motivos, puntos de vista, opiniones, insinuaciones, valoraciones, emociones, etc. Asimismo, a partir de esta conversación, se puede comparar y contrastar aproximaciones de los investigadores y de una experta en el tema particular. Así, la petición de entrevista a la ilustradora Paola Karolys se sustentó en que la profesional tiene experiencia en estudios e ilustraciones históricas andinas, uno de los arquetipos que se estudian en esta investigación. Karolys ha trabajado en múltiples proyectos editoriales, ilustrando tanto para obras literarias (cuentos, novelas, poesía, etc.), como para textos educativos, revistas de divulgación temática e institucionales, afiches, publicidad. Junto con su hermano Gabriel Karolys, en 2009, obtuvo el primer premio en el concurso “Darío Guevara Mayorga” por las ilustraciones del libro *Imágenes del Bicentenario*.

Unidades de análisis

La selección de las plataformas en las cuales se explorarían los resultados se determinó en función de la misma lógica algorítmica a partir de palabras clave. Se utilizó el muestreo no probabilístico por bola de nieve. Es decir, a partir de la búsqueda en blogs de tecnología y arte, se revisaron las plataformas que más se nombraban y que tenía versión gratuita o de prueba. Esta decisión está motivada en el hecho de que el fenómeno AI ha sido concebido en la técnica boca en boca para publicitarlo; sin actividades de marketing explícito. Así, los resultados más repetitivos fueron Midjourney, Craiyon y DeepAI.

Cada plataforma se corrió en su versión gratuita y se exploraron los comandos (*prompts*) en inglés: *doctor* (doctor/a), *nurse* (enfermero/a), *secretary* (secretario/a). Posteriormente, en la plataforma Midjourney se corrieron los comandos *Europe* (Europa), *Latin America* (América Latina), *European girl studying* (niña europea estudiando) y *Latin America girl studying* (niña latinoamericana estudiando). La exploración se hizo únicamente en esta plataforma dado que es la única que permitió más resultados gratuitos, pero, además, porque es la plataforma de la que más se está hablando en los blogs de tecnología.

El uso de palabras clave en inglés resulta beneficioso para la investigación porque, al evitar el uso del género en los sustantivos, se pueden explorar tendencias de representación y subrepresentación. Se seleccionaron estas profesiones dado que el objetivo es la prueba de los arquetipos y, en este caso, se basa en profesiones históricamente catalogadas como masculinas y femeninas para probar si es reproducido por estas inteligencias. Así, mientras una enfermera está catalogada dentro de las labores de cuidados —concepción pensada para las mujeres—, un doctor es pensado en el ámbito de la salud. De hecho, en 2018, este inconsciente colectivo respecto a la asociación de las profesiones con el género y el arquetipo del héroe fue probado en un enigma viral donde la eminencia en cirugía es la madre de un joven (La Vanguardia, 2018). En un segundo caso, se pretende probar las aproximaciones a la cultura y el fenómeno cultural que desarrolla Jung en sus fundamentaciones del inconsciente colectivo.

Midjourney

Es un laboratorio independiente fundado por el estadounidense David Holz. Esta tecnología se basa en *text-to-image*; lo que significa que es capaz de crear imágenes a partir de texto. Desde su interfaz se crean cuatro opciones por cada comando que se ingresa. Esta inteligencia se abrió al público en julio de 2022 y su uso se volvió controversial en septiembre, cuando “Teatro de ópera espacial”, obra de Jason Allen, abrió el debate sobre la autoría de imágenes digitales

tras ganar el Concurso de Arte Digital (Colorado, Estados Unidos), habiendo utilizado esta plataforma.

Midjourney solo se puede utilizar mediante un bot en el Discord oficial del proyecto, en una especie de chat colaborativo donde se debe ingresar la búsqueda presidido de **/imagine**, lo que da la sensación de que el usuario está programando su trabajo. De acuerdo con el blog de generative AI (Yubin, 2023), Holz no concibe a Midjourney como un chatbot, sino más bien como un recipiente, un vehículo para la mente. Sin embargo, no se tiene información precisa acerca de los antecedentes y el entrenamiento de esta inteligencia; se especula que extrae imágenes y texto de Internet para describirlos, usando millones de imágenes publicadas para entrenamiento (Dieter, 2023).

Craiyon

Craiyon, desde su misma plataforma se presenta como “anteriormente DALL-E mini, es un modelo de IA que puede dibujar imágenes desde cualquier mensaje de texto”; explica que su modelo aprende conceptos a partir de imágenes: “cada vez que ingrese un aviso, creará imágenes únicas basadas en su conocimiento actual e intentará combinar los conceptos de su aviso de una manera creativa”. Es muy explícito al advertir que su modelo base se entrenó con datos sin filtrar de Internet, limitados a imágenes con descripciones en inglés. No se utilizaron textos e imágenes de comunidades y culturas que usaban otros idiomas. Por lo tanto, “esto afecta todos los resultados del modelo, con la cultura blanca y occidental afirmada como predeterminada” (Dayma *et al.*, 2021, s/p).

DeepAI

Se autodefine como “un generador de imágenes AI”. Crea una imagen desde cero a partir de una descripción de texto. Utiliza IA para comprender sus palabras y convertirlas en una imagen única cada vez. Su fundador, Kevin Baragona, firmó la petición para pedir

que se haga una pausa el desarrollo de la IA, dado que afirma que es una tecnología “disruptiva”: “lo más probable es que tenga capacidades de razonamiento tan avanzadas que se asemejarán a las del cerebro humano con lo cual es demasiado disruptiva para el momento actual” (Bloomberg en Español, 2023). Baragona acepta la ausencia de protocolos entre todas las empresas de inteligencia artificial, lo que supone una falta de transparencia en su desarrollo:

Creo que esta tecnología es tan poderosa que no importa quién la posea, tampoco importa en qué país tenga sede. El problema es que se está construyendo en primer lugar. Esta es una tecnología tan increíblemente poderosa que comencé a llamarla el arma nuclear del software. (Bloomberg en Español, 2023)

Resultados

Scolari (2023) explica la inteligencia artificial a partir del uso de traductores en la recopilación de *Dire quasi la stessa cosa* de Umberto Eco (2000). Resalta que la tecnología que se basa en comandos apunta a convertirse en una enciclopedia por sobre un diccionario. La diferencia radica en los contextos: la IA no debe leer cada palabra en sí misma, sino en su relación con el resto:

El pasaje del diccionario a la enciclopedia es fundamental en la teoría interpretativa de Umberto Eco: no es suficiente conocer el significado de cada signo para comprender el sentido de una expresión. Las «selecciones contextuales» nos sitúan en una red semántica abierta, en permanente transformación, formada por una trama densa de reenvíos. (Scolari, 2023, s/p)

Para continuar con la metáfora que sugiere el académico, el paso del diccionario a la enciclopedia se da a través del análisis y procesamiento de datos que son recolectados a partir de cada interacción; son alimentados a partir de los usos humanos. Sin embargo, uno de los mayores reproches desde los enfoques feministas se basa en la explotación de datos, el colonialismo y la falta de transparencia de la IA (Amrute, 2019). Los datos, en la actualidad, son considerados

como un bien que puede ser explotado por empresas y estados. De hecho, esta tecnología utiliza datos masivos del pasado para predecir comportamientos y adaptarse al futuro; se habla de que dichos datos se han convertido en prescriptivos en la medida que condicionan las acciones de los seres humanos. Por lo tanto, los límites entre el pasado, el presente y el futuro se vuelven más difusos porque las acciones que se generan (presente) son modeladas, por lo que el futuro de la sociedad también puede serlo.

Según Sareeta Amrute, profesora adjunta de Antropología en la Universidad de Washington, “la mayoría de las veces, los diseñadores de sistemas técnicos comienzan con un usuario estándar [en mente] y, al hacerlo, ponen en marcha patrones de discriminación que están ocultos por la suposición de neutralidad del sistema” (Amrute, 2019). Por lo tanto, un peligro fundamental radica en pensar que la IA es neutral al suponer que los datos se generan en iguales cantidades desde todos los puntos del mundo y que estos son procesados de la misma manera.

En cuanto a la conceptualización de los arquetipos, Paola Karolys cuestiona y explica los límites en esta aproximación psicológica:

Hay una limitación y es la generalización de las formas de vida en todos los continentes. En investigaciones históricas, los mitos fundacionales son muy diferentes, pero cuando los europeos acercan [traducen] estas historias a los suyos surgen las adaptaciones.

Dichas adaptaciones, por supuesto, respetan la construcción arquetípica occidental. De hecho, Karolys explica que, por ejemplo, el arquetipo del héroe sustentado en los mitos europeos no debería existir dentro del mundo indígena dado que los indígenas no tienen este deseo de destacar por sobre el resto, sino que su concepción de vida es mucho más comunitaria. No obstante, dado el proceso de mestizaje y la penetración de medios de comunicación, para cualquier habitante resulta comprensible el mito del héroe, a pesar de que no sea una representación fiel de su cultura o su historia. Por lo tanto, el

riesgo de lo que la ilustradora cataloga como la “generalización de la falacia” es que “los elementos [en la ilustración] quedan desprovistos de los significados”, porque esta inteligencia no es capaz de reconocer la evolución histórica de los símbolos y por eso los simplifica para generar una imagen cuya cualidad principal es que es fácilmente reconocible.

Los datos que se muestran a continuación fueron recolectados entre el 9 y el 15 de enero de 2023.

Arquetipos profesionales

El informe de UNESCO respecto a la representación de género recolectó entrevistas con expertas en inteligencia artificial. Así, explicita que ante la pregunta de si existen instrumentos o principios normativos de IA que aborden con éxito la igualdad de género, Daniela Braga, fundadora y CEO de DefinedCrowdo, explicó que las prácticas actuales en cuanto a la preocupación por la búsqueda de la igualdad eran insuficientes; mientras que para Rediet Abebe, Junior Fellow de la universidad de Harvard, son inexistentes. Asimismo, Sara Kassir, analista de investigación, pymetrics, explica que “es necesario formalizar los estándares legales, de modo que términos como “ético” y “imparcial” se definan formalmente” (UNESCO 2020, p. 12).

En este sentido, la igualdad de género no es una preocupación en el desarrollo de esta tecnología, sino que se enmarca en el principio de justicia, que viene a ser un concepto muy amplio y no direccionado a disminuir las brechas de género y frente al que, además, no hay una definición unificada.

El informe de UNESCO advierte que vale la pena considerar las lecciones de esfuerzos anteriores sobre género y tecnología. Resalta que “cuando el género se hace explícito y cuando se deja implícito, debe ser deliberado y reflexivo” (UNESCO, 2020). La tabla 1 procesa los resultados obtenidos al generar el prompt *doctor*, *nurse* y *secretary*.

También importa dónde se hacen las referencias explícitas. Por ejemplo, ¿la igualdad de género está en una lista larga, está en

un preámbulo, está en el cuerpo principal de principios y en su implementación? Referencias vagas y generales y ubicación donde hay pocos “dientes” o enfoque directo, o, por el contrario, donde está en todas partes y, por lo tanto, en ninguna, el riesgo que el género recibe.

Tabla 1
Arquetipos profesionales

	Midjourney	Craiyon	DeepAi
Doctor			
Nurse			
Secretary			

La ilustradora e historiadora Paola Karolys explica que los arquetipos pueden tener distintos contenidos. Agrega que esta teoría puede ser tan extendida de hecho porque se basa en los vacíos que existen en las culturas; lo cual facilita la clasificación y proliferación de arquetipos. Resalta que la ilustración hace un uso racional de símbolos, alegorías, pero resalta que también hay algo intuitivo, emocional. En la contraposición médico/enfermera, destaca que hay más riqueza interpretativa en el caso del doctor. En el caso de Midjourney —de estilo onírico y artístico— muestra no solo un hombre joven, sino uno

viejo y al menos uno de ellos está ubicado en un escenario que hace alusión a la tecnología. Además, las expresiones faciales demuestran tono de seguridad e incluso resulta imponente. En contraste, el arquetipo de enfermera es una mujer únicamente joven, sin escenario y con una expresión más de sumisión. Esto, a pesar de que, por ejemplo, tanto enfermera como doctor usan el estetoscopio y el uniforme, que resultan ser los íconos de la medicina. En una lectura más profunda se puede afirmar que, a pesar de que se reconocen habilidades similares en ambas profesiones, la tecnología no las valora de la misma forma. Craiyon, plataforma que ha sido denunciada por falta de representatividad, mantiene el problema de la brecha de género; sin embargo, tiene más riqueza en cuanto al lenguaje corporal cuando se mira los resultados de la dicotomía doctor/enfermera.

Más allá de que resalta que al ingresar comandos, en principio, neutros, los resultados muestran brechas en el género, no conciben la construcción histórica de las profesiones. La palabra *secretaria* se deriva de *secreto*, dado que son los encargados de manejar información que puede llegar a ser confidencial. Por lo tanto, la posición de secretario va más allá de una alusión a estar sentado en un escritorio, cosa que no se toma en cuenta con el prompt *secretary*. Los resultados son mujeres jóvenes, sexualizadas y atadas a un escritorio, en tono desafiante, en el caso de Midjourney, pero se repite en Craiyon: “usamos elementos simbólicos como detonante, que generalmente son aceptados, pero el nivel de significación debe ser más profundo” (Karolys, 2023).

La plataforma DeepAi muestra más posibilidades tanto en estilos como en actores. Es la única plataforma que concibe, por ejemplo, que un doctor sea alguien usando un terno, que tiene relación con la evolución de los estudios donde doctor no se refiere únicamente al personal de la salud. En el caso de enfermera, también hay más riqueza en su base de datos, dado que es capaz de mostrar la evolución de esta profesión a lo largo de los años. Sin embargo, no está libre de la denuncia de la brecha de género.

Representación étnica y cultural

De acuerdo con la Unión Europea, la inteligencia artificial (IA) se refiere a sistemas que muestran un comportamiento inteligente analizando su ambiente y tomando acciones con algún grado de autonomía para lograr objetivos específicos. Esto concluye que los sistemas de inteligencia artificial (IA) intentan imitar el comportamiento humano; sin embargo, la mayor parte del trabajo sobre IA no toma en cuenta que el comportamiento humano está intrínsecamente moldeado por los contextos culturales en los que están inmersos, los valores y creencias que tienen, y las prácticas sociales que siguen. Así, dado que estas tecnologías se conciben y desarrollan principalmente en un puñado de países, incorporan los valores y prácticas culturales de estos países o, incluso, resignifican la cultura de los otros a partir de su hegemonía. Incluso, “los datos que se utiliza para entrenar los modelos tampoco representan equitativamente la cultura global y su diversidad. Por lo tanto, surgen problemas cuando estas tecnologías interactúan con sociedades y culturas diversas, con diferentes valores y prácticas interpretativas” (Prabhakaran *et al.*, 2022, p. 1).

Como se examinó en el apartado “Unidades de análisis”, la inteligencia artificial es una tecnología que se desarrolla en el Norte global, lo que puede traducirse en que sus contenidos no sean una muestra, sino una interpretación de cada comando. Karolys, en referencia a la representación étnica y cultural destaca que “la historia es algo que está en continuo desarrollo; ningún programa debería desentenderse de eso”. Para probar la construcción de los arquetipos en este sentido, se ingresaron los comandos *Europe*, *Latin America* y *European girl studying* y *Latin America girl studying*. Los resultados se exhiben en la tabla 2.

Al presentar los resultados de la primera fila a la entrevistada Paola Karolys, la profesional resalta que, por ejemplo, la concepción de Europa muestra un progreso y una riqueza cultural y narrativa que no se equipara con los resultados arrojados en la búsqueda de América Latina: mientras las imágenes de Europa se preocupan por contar

una historia, América Latina se reduce a una concepción netamente de naturaleza, donde incluso los personajes son mimetizados con la naturaleza, eliminando cualquier rastro de cultura. De hecho, desde la antropología norteamericana se discute la dicotomía naturaleza y cultura y fue ampliamente aceptada debido a los estudios de George Perkins Marsh en su libro *Man and Nature: Or, Physical Geography as Modified by Human Action*. En este sentido, al reducir a Sudamérica a su naturaleza, simbólicamente se llega a desconocer la cultura que posee.

Tabla 2
Arquetipos en la representación étnica y cultural

Europe	Latin America
<p><i>European girl studying</i></p>	<p><i>Latin American girl studying</i></p>

En la segunda búsqueda, sobre mujeres estudiando, para Karolys destaca el hecho de que en las expresiones faciales son más bien de tristeza, en especial en el caso latino. La ilustradora es consciente de que el repositorio de donde se alimenta esta inteligencia es todavía limitado y se refiere, por ejemplo, a los emoticones que tienen muchas más posibilidades interpretativas, pero cuya evolución requirió años de usos y necesidades para apoyar a la comunicación no verbal a la que se puede limitar la mensajería instantánea. Asimismo, destaca que todas las niñas llevan un “estilo maya, caribeño y no tienen en cuentas las mezclas”; hay un error porque se nos ve como “solo mestizos”. Contrasta con las posibilidades europeas, donde se exhibe más posibilidades: fenotipos, tiempos y espacios. De hecho, una de las ilustraciones europeas tiene un escenario que refleja que está acompañada o un lugar dedicado al estudio, mientras que en el caso latino el estudio se concibe en completa soledad.

A un alto nivel (Prabhakaran *et al.*, 2022) explican que se distinguen dos formas en que la cultura interactúa con los sistemas de IA: en el desarrollo y en el uso. El proceso de desarrollo de la interfaz de los sistemas de IA con la cultura se basa tanto en los datos como en los recursos que capturan el comportamiento humano moldeado culturalmente, así como a través de la cultura normas y valores encarnados por los propios desarrolladores e investigadores. Si se tiene en cuenta, además, que todas estas tecnologías se presentan al público a través de un chatbot que traduce palabras en imágenes, se hace más evidente la domesticación cultural a partir del idioma:

Los sistemas que están entrenados o pre-entrenados en datos web pueden capturar varias modalidades de comportamiento humano, incluyendo el uso del lenguaje y las imágenes, lo que implícitamente cuece en varios aspectos culturales que luego influir en las aplicaciones posteriores. Dado que el lenguaje y los símbolos, la ontología y la axiología, juegan un papel crítico en el desarrollo de sistemas de IA, por ejemplo, a través de “etiquetas” en los datos, y cómo “conocimiento”, se construyen la “objetividad”, la “realidad/verdad” y los “objetivos del sistema”, las normas culturales de los desarrolladores e investigadores de IA también infunden de manera generalizada los sistemas de IA. (Prabhakaran *et al.*, 2022)

Asimismo, los sistemas de IA y las tareas que esta realiza se apegan a la cultura para las que fueron creados en formas que se adhieran a las expectativas moldeadas culturalmente, y cómo interactúan con otros seres humanos.

Por ejemplo, las tareas interpretativas están intrínsecamente moldeadas por la cultura en la que están inmersos, incluyendo no solo las dependencias cultural-lingüísticas de tareas como inferir emoción, sentimiento, ofensa, pero también imagen y símbolo, interpretación —incluyendo gestos, expresiones faciales, imágenes tabú incluyendo pornografía y violencia, y denotaciones y connotaciones de símbolos—. Cuando los supuestos culturales y las normas que se integran en los sistemas de IA durante su desarrollo están en desacuerdo con las normas culturales y expectativas de los ecosistemas culturales objetivo, vemos rupturas y fallas tales como malas interpretaciones o tergiversaciones culturales, que colectivamente llamamos incongruencias culturales. (Prabhakaran *et al.*, 2022, p. 2)

Conclusiones

El valor de esta investigación radica en la relación conceptual entre arquetipos y significados desde la teoría lingüística, para el análisis de la inteligencia artificial, lo que la convierte en un estudio original y actual. No obstante, presenta limitaciones desde su implementación metodológica debido a la naturaleza de las unidades de análisis. Así, las imágenes que se estudiaron se recolectaron en enero 2023; sin embargo, debido a la alta penetración de estas tecnologías, nuevos usos y apropiaciones y a su entrenamiento en machine learning, es probable que las imágenes creen en un futuro (incluso en diferencia de minutos) sea muy diferente; lo que perjudica a la posibilidad de la replicabilidad y comprobación del estudio. Sin embargo, se han sentado métodos de análisis y procesamiento de datos para el estudio de este fenómeno.

En contraste, una de las fortalezas de este estudio cualitativo es que procura la búsqueda de profundidad en los significados del fenómeno en estudio y que respeta el método científico, procurando

eliminar la subjetividad en la interpretación de datos. De esta manera, la investigación explica que la teoría de los arquetipos se sustenta en un inconsciente colectivo que se nutre de la naturaleza universal y que se transmite a través de la tradición; esta teoría se basa, por ejemplo, en los mitos y explica las preocupaciones y soluciones que el ser humano ha propuesto. De manera directa se relaciona con la concepción del *homo narrans*, dado que los arquetipos son fundamentales en la construcción de historias para generar espacios fácilmente identificables para el receptor. De hecho, (Jenkins, 2010) afirma que:

El significado se transmite a través de los medios a través del gesto, el espacio, el color, el sonido, la actividad y la agencia. Creo que una de las razones por las que estas experiencias son tan convincentes es que nos permiten acceder más directamente a nuestra conciencia prelingüística de *homo ludens*, ya que podemos interactuar con ellas de manera lúdica. Por supuesto, luego damos un paso atrás y hablamos de ello, lo que involucra nuestra conciencia discursiva de *homo narrans*. Entonces tenemos *ludic narrans*, cuentos lúdicos. (Jenkins, 2010, s/p)

De hecho, el extracto anterior podría incluso explicar por qué los albores de la inteligencia artificial están en el desarrollo de máquinas que compiten en juegos contra seres humanos. Ahora bien, cuando los arquetipos son divulgados a partir de una inteligencia artificial cuyo procesamiento de datos no es transparente, surgen nuevas consideraciones. En primer lugar, se requiere de una ética de la inteligencia artificial que sea capaz de cumplir con los principios de la bioética, dado que los datos que colecciona son obtenidos de seres humanos que alimentan esta tecnología bajo un uso casi “mágico” —en el sentido en que hay falta de transparencia en la recolección y procesamiento de datos—. Así, desde sus bases se debe preocupar por la beneficencia, no-maleficencia, autonomía y justicia. Sin embargo, los colectivos resaltan que dados los índices de penetración y los posibles usos que se pueden desprender de esta tecnología, la búsqueda de equidad también debe ser considerada. El informe de la UNESCO, por ejemplo, solicita que la tecnología busque evitar daños, aumentar la visibilidad y contribuir al empoderamiento. En esta misma línea de ideas, denuncia

que la falta de diferenciación y la posibilidad de extender el concepto de equidad al de justicia invisibiliza la posibilidad de considerar las luchas sociales. Con esta afirmación concuerda la entrevistada:

Y no sabemos de dónde salen las imágenes o el repositorio base para estos programas de IA, pero se notan limitaciones aún, prejuicios, estereotipos (no solo arquetipos). Eso llama la atención... ¿Un programa analiza, calcula, toma todos o aleatoriamente los datos gráficos de internet y los traduce matemáticamente, estadísticamente, para sacar conclusiones sobre las cosas y sus atributos y así poder ubicarlas en categorías? ¿O alguien, un grupo (inteligencia humana) participa también en la consolidación de la información, con la estructuración de la base de datos? ¿El material de esa base de datos es general o el restringido?

De esta manera podríamos saber, por ejemplo, de dónde vienen las limitaciones. Una cosa es que un programa evalúe los datos gráficos mundiales. Si hay brecha de género, es porque estadísticamente, a nivel mundial, prevalece la brecha de género. Si se mantiene la idea de América como un lugar salvaje, natural y no cultural, es que prevalecen las mismas nociones conceptuales racistas, eurocentristas, que discriminan, etc. del siglo XV, no se ha superado ideas colonialistas, imperialistas, etc. Porque el programa solo lee lo que la gente está poniendo mayoritariamente...

Esto está en consonancia con la tercera ley de la interfaz que en realidad conforma un ecosistema. Scolari advierte que “las leyes de la interfaz excluyen cualquier teoría heroica de la invención: toda innovación es el resultado de una red de interacciones entre actores humanos y tecnológicos” (Scolari, 2018, p. 47). Esta ley advierte que las interfaces mantienen un fecundo y conflictivo intercambio entre ellas. Por lo tanto, la sociología de los medios nos obliga a pensar en las jerarquías que pueden existir en dichos ecosistemas. Maigret (2003) clasifica a los medios en tres grupos:

Cultos, medios y populares. Afirmo que la estética de los medios populares se reduce a un “gusto por lo necesario”, a una subordinación de la forma a la función, a una adhesión ingenua a los contenidos, a una simple conformidad. (p. 212)

Los sistemas potentes de IA, debido a la cultura en la que se han desarrollado, potencian estas diferenciaciones e imponen una visión del mundo desde la distinción, lo cual aumenta las brechas existentes. La declaración reciente de OpenAI con respecto a la inteligencia general artificial establece que “En algún momento, puede ser importante obtener una revisión independiente antes de comenzar a entrenar sistemas futuros, y para los esfuerzos más avanzados para acordar limitar la tasa de crecimiento de la computación utilizada para crear nuevos modelos”. Los mismos fundadores de estas inteligencias advierten que “las imágenes generadas por el modelo pueden incluir estereotipos perturbadores y dañinos en las clases protegidas; características de identidad; y grupos sensibles, sociales y ocupacionales” (Dayma *et al.*, 2021, s. p.). Además, el modelo solo se puede utilizar para generar imágenes basadas en texto en inglés, lo que limita la accesibilidad del modelo para personas que no hablan inglés y contribuye potencialmente a los sesgos en las imágenes generadas por el modelo.

A partir de búsquedas metódicas, se observa que la generación de inteligencia artificial desde Norteamérica y Europa, tanto en el estilo como en el contenido, se sustenta en el *mainstream*, aumentando las brechas en la comprensión de la diversidad del mundo. Se concluye que las tecnologías, incluidos los datos en sí, incorporan valores y suposiciones humanas y, por lo tanto, dan forma o restringen los comportamientos en función de esas posibilidades.

Referencias bibliográficas

- Adams, N. R. (2019). *How Artificial Intelligence Works*.
- Amrute, S. (2019). Of Techno-Ethics and Techno-Affects. *Feminist Review*, 123(1), 56-73.
- Baricco, A. (2010). *Los bárbaros*. Anagrama.
- Berguer, R. (1976). *El conocimiento de la pintura*. Noguer.
- Bernández, A. y Álvarez, C. (eds.) (2012). *Cartas 1977-1984*. Alfaguara.
- Bloomberg en Español. (2023, April 4). *Fundador de DeepAI: la IA generativa es demasiado disruptiva*.

- D&AD in association with Creative SkillSet, & FutureLearn. (n.d.). *The Seven Story Archetypes*. Brand Storytelling: How to Use Narrative to Sell.
- Dayma, B., Patil, S., Cuenca, P., Saifullah, K., Abraham, T., Lê Khắc, P., Melas, L. y Ghosh, R. (2021). *DALL-E Mini*.
- Dieter, B. (2023). *Midjourney: todo lo que necesitas saber sobre el artista de IA*. Reviews.
- Dondis, D. A. (1976). *La sintaxis de la imagen*. G.G.
- Fitzgerald, K., Paravati, E., Green, M. C., Moore, M. M. y Qian, J. L. (2020). Restorative Narratives for Health Promotion. *Health Communication*, 35(3), 356-363. <https://doi.org/10.1080/10410236.2018.1563032>
- Giraldo, J. (1961). Imagen, símbolo y arquetipo. *Revista Colombiana de Psicología* 6(2), pp. 131-147.
- Green, M. C., Fitzgerald, K. y Moore, M. M. (2019). Archetypes and Narrative Processes. *Psychological Inquiry*, 30(2), 99-102. <https://doi.org/10.1080/1047840X.2019.1614808>
- Grunwald, P. D. y Vitanyi, P. M. B. (2008). *Algorithmic information theory*. Cornell University.
- Igartua, J. J. (2008). Identificación con los personajes y persuasión incidental a través de la ficción cinematográfica. *Escritos de Psicología*, 2(1), 42-53.
- Jenkins, H. (2010). *Ludic Narrans: Drew Davidson Talks Cross-Media Communication (Part One)*. Pop Junctions.
- Jung, C. G. (1970). *Arquetipos e inconsciente colectivo*. Paidós.
- Kumar, A. (2022). *Information Theory, Machine Learning & Cross-Entropy Loss*. Data Analytics.
- La Vanguardia. (2018). *El cirujano era la madre*. La Vanguardia.
- Maigret, É. (2003). *Sociología de la comunicación y de los medios*. Fondo de la Cultura Económica.
- Malagón, C. (n.d.). *Introducción a la inteligencia artificial*.
- Midjourney. (2022). *Midjourney: Art in the Age of Artificial Intelligence*.
- Potts, A. (1996). Sign. En Robert S. Nelson y Richard Stiff (eds.), *Critical Terms for Art History*. University of Chicago Press.
- Prabhakaran, V., Qadri, R. y Hutchinson, B. (2022). Cultural incongruencies in Artificial Intelligence. *First Workshop on Cultures in AI/AI in Culture*.
- Rouhiainen, L. (2018). *Inteligencia artificial 101 cosas que debes saber hoy sobre nuestro futuro*. Editorial Planeta.
- Saavedra-Bautista, C., Cuervo-Gómez, W. O. y Mejía-Ortega, I. D. (2017). Producción de contenidos transmedia, una estrategia innovadora. *Revista Científica*, 28(1), 6-16. <https://bit.ly/49EtMSY>

- Sager, N. (1967). Inmutabilidad y mutabilidad del símbolo: Planteo para una comprensión semiológica del símbolo de Jung. *Revista de Psicología*, 4, 93-104.
- Saussure, F. (2008). *Curso de Lingüística General*. Losada.
- Scolari, C. (2018). *Las leyes de la interfaz. Diseño, ecología, evolución, tecnología*. Gedisa Editorial.
- Scolari, C. (2023, February 7). *Inteligencia artificial, entre el deseo y el miedo*. CCCBLAB.
- UNESCO. (2020). *Artificial Intelligence and Gender Equality: Key findings of UNESCO's Global Dialogue*.
- Yubin. (2023). *David Holz | Biography, Midjourney Founder & CEO*. Tuts.