

IA, música, estilo: ¿puede haber demasiada cultura?¹

Sunil Manghani

Universidad de Southampton
S.Manghani@soton.ac.uk

Recibido: 29/10/2024

Aceptado: 10/11/2024

Resumen

En el centro de este artículo hay una pregunta sobre la aplicación artificial del *estilo* y su importancia para un modo *avanzado* de producción cultural. Vale la pena señalar que el estilo no es algo que esté fácilmente sujeto a la ley de derechos de autor. A medida que los modelos de IA se vuelven más sofisticados, un nuevo enfoque generativo del estilo plantea el potencial de una (sobre)producción masiva de cultura. En el momento de escribir este artículo, los modelos que convierten de texto a video, aunque no son de uso frecuente, están empezando sin duda a ganar en funcionalidad y sofisticación, acercándonos a un futuro potencialmente «sin memoria», como lo predijeron los teóricos posmodernos.

Palabras clave: inteligencia artificial; música; estilo; creación.

AI, Music, Style: Can there be too much culture?

Abstract

At the heart of this article is a question about the artificial replication of *style* and its significance for an *advanced* mode of cultural production. It is worth noting, style is not something that is readily subject to copyright law. As AI models become more sophisticated, a new generative approach to style poses the potential for a massive (over-)production of culture. At the time of writing, text-to-video models, while not prevalent in use, are certainly beginning to gain in function and sophistication; offering a glimpse into a potentially ‘memory-less’ future as foretold by postmodern theorists.

Keywords: artificial intelligence; music; style; creation.

IA, música, estilo: pode haver cultura em excesso?

Resumo

No centro deste artigo existe uma questão sobre a aplicação artificial do *estilo* e sua importância para um modo *avançado* de produção cultural. É importante notar que estilo não é algo que esteja facilmente sujeito às leis de direitos autorais. À medida que os modelos de IA se tornam mais sofisticados, uma nova abordagem generativa do estilo propõe a possibilidade de uma (super)produção massiva de cultura. No momento da redação deste artigo, os modelos que convertem texto em vídeo, embora não sejam amplamente utilizados, estão certamente começando a ganhar funcionalidade e sofisticação, aproximando-nos de um futuro potencialmente «sem memória», como previram os teóricos pós-modernos.

Palavras-chave: inteligência artificial; música; estilo; criação

¹ Traducción del inglés: Álvaro Lema Mosca.

El legado de Warhol

I've got a Brillo box and I say it's art
It's the same one you can buy at any supermarket
'Cause I've got the style it takes
Style It Takes (1990), LOU REED Y JOHN CAGE

A mediados del siglo XX, Andy Warhol hizo pedazos para siempre nuestra comprensión del «carácter artístico» del estilo. Siguiendo las anteriores intervenciones de Duchamp, que reparaba en el amplio *contexto* de una obra de arte, Warhol adoptó un modo «industrioso», célebre por su exposición de cajas Brillo de supermercado en 1964. Al convertir el objeto producido en masa en una obra del arte, desafió las nociones preconcebidas sobre los límites entre lo artístico y lo cotidiano. Si hubiera sido un problema limitado solo al mundo del arte, tal vez no hubiera sido tan significativo.

Sin embargo, el legado de Warhol atraviesa más ampliamente la cultura. Llegó a personificar la capacidad de *afectar* el estilo, es decir, de engendrar una nueva capacidad para influir, cambiar o alterar un estilo *existente*. Eso se acerca más hacia una lectura posmoderna. Después de Warhol ya no ha sido necesario crear nuevas expresiones, sino que ha sido posible (o *únicamente* posible) replantear, reposicionar y reenvasar lo que ya existía. Llevado al extremo, el argumento es que ya no hay nada nuevo, solo nuevas formas de contextualizar y situar. Hoy en día, a través de la avalancha de tableros de Pinterest, feeds de Instagram y videos de TikTok y YouTube vemos la repetición masiva de poses, filtros y memes. Cada uno de ellos son indudablemente actos individuales de creatividad, pero podría decirse que hay escasez de originalidad. El «arte» está menos en la generación de nuevos significados y más en la conservación, es decir, en la reutilización de imágenes, textos y sonido. *Estilizamos y conservamos como un medio para crear*.

La música rap, por ejemplo, tal como surgió a principios de la década de 1970 en las comunidades afroamericanas del Bronx, en Nueva York, se considera a menudo la encarnación de la cultura del remix. El género se nutre de la reinterpretación y recontextualización de sonidos, ritmos y letras existentes, haciendo del acto de remezclar no solo un método de creación, sino una declaración de importancia cultural y social. La habilidad del rapero como conservador (curador) de sonido se vuelve primordial, seleccionando y combinando muestras de manera que rindan homenaje a sus orígenes y, al mismo tiempo, traspasen los límites de la expresión musical. Esta práctica refleja una apreciación posmoderna por el bricolaje y la intertextualidad, en el que el significado y el valor de la obra se derivan de su relación con otros textos y sonidos. También vale la pena señalar que, como apunta el rapero y profesor Lupe Fiasco, el rap está estrechamente ligado al avance tecnológico. No es ninguna sorpresa que esa relación siga prosperando gracias a los recientes avances en la inteligencia artificial (IA). Como parte de su colaboración con Google para desarrollar la herramienta TextFx, Fiasco se describe a sí mismo como una «máquina de recopilación de datos».

El estilo y su repetición

Normalmente, las lecturas académicas de la música rap se centran menos en el entretenimiento, y más en su capacidad como forma rica y compleja de expresión cultural y política. Como parte de una cultura de remezcla más amplia, el rap sirve como medio de innovación y crítica cultural, reflejando temas más amplios de resiliencia, identidad y resistencia. Tricia Rose, una investigadora pionera de la cultura hip-hop, sostiene en su libro *Black Noise* (1994) que la práctica del sampleo no es solo una técnica artística, sino un método de preservación y comentario cultural. De manera similar, en *Yo' Mama's Disfunktional!* (1998), el historiador Robin D. G. Kelley posiciona al rap como una forma de contranarrativa, que se consigue mediante el reensamblaje de fragmentos musicales existentes para forjar nuevas expresiones de identidad y resistencia. En *Prophets of the Hood: Politics and Poetics in Hip Hop* (2004), Imani Perry traza la complejidad del hip-hop como movimiento cultural, considerando al rap como algo emblemático en la práctica fundacional de la «citacionalidad», como un «estilo» mediante el cual las referencias a elementos culturales existentes crean significados y opiniones estratificados sobre la sociedad, la política y la cultura. Es importante destacar que, como sostiene Mark Anthony Neal en *What the Music Said* (1999), los estilos de música popular negra desempeñan un papel clave en la configuración de su cultura pública.

En su libro fundamental, *Subcultura. El significado del estilo*, publicado en 1979, Dick Hebdige argumentó con acierto que la «resistencia simbólica» de la moda y la música nunca es solo una cuestión de vestimenta o preferencias musicales, sino que son *formas significativas de comunicación* capaces de reflejar y desafiar tensiones y conflictos sociales más amplios. Sin embargo, como sostiene Hebdige (2004), existe un problema fundamental. Si bien los estilos subculturales son inherentemente políticos —*inicialmente* se proponen subvertir las normas y valores dominantes—, las formas culturales dominantes eventualmente cooptan estos estilos, diluyendo su poder subversivo e incorporándolos a la cultura comercial dominante. Por ejemplo, la representación del punk como una parte *preciada* de la cultura nacional en la ceremonia de apertura de los Juegos Olímpicos de Londres 2012 es un ejemplo claro de tal incorporación (durante su apogeo, el punk era visto como altamente subversivo y claramente antisistema).

En el período de posguerra, el filósofo existencialista Jean-Paul Sartre instó a la práctica del escritor «comprometido», argumentando que los autores (y los creadores en general) pueden tomar *decisiones* dentro del uso de las formas preexistentes, las convenciones establecidas, los géneros y códigos. La producción artística surge de una lucha con lo que ha sucedido antes, lo que a su vez proporciona el «lugar» desde el cual tomar decisiones, como un medio para buscar el cambio y la libertad. En este sentido, la afirmación del «estilo» proporciona un medio para subvertir o burlar las limitaciones de una lengua dominante. Sartre, por ejemplo, posicionó al venerado escritor burgués del siglo XIX, Gustave Flaubert, como un escritor en lucha con su propia identidad de clase (privilegiada), que utilizaba el *estilo* literario como un medio de escape o crítica existencial. La descripción crítica que hace Flaubert de la sociedad burguesa en su novela *Madame Bovary* es un ejemplo de ello.

Roland Barthes, pionero en explorar los ámbitos de la cultura, el texto y el significado más allá de la intención del autor, simpatizaba con la idea de Sartre, pero ofrecía una visión más pesimista y que influyó en escritores como Dick Hebdige. Es importante destacar (aunque difícil de entender) que Barthes intentó distinguir entre *lenguaje* y *estilo* con un tercer término: *escritura*. Su punto era que, tanto el lenguaje como el estilo, son formas que *heredamos*, es decir, no podemos elegirlos. Nacemos en un lenguaje que define una posición subjetiva, y aunque normalmente se considera el estilo una forma de expresión individual, para Barthes se trataba de un medio para definir el estatus del sujeto. El estilo se presenta de muchas maneras y no posee una «gramática» propia, sino que se trata de un conjunto involuntario y encarnado de acciones y afectaciones —algunas de las cuales son, de hecho, hábitos culturalmente compartidos— que aparecen de manera crucial *antes* de la escritura (Barthes, 2011).

Según Barthes, el estilo es «calculable» (en tanto codificado), lo que se opone a la escritura como un medio potencial de lo incalculable (aquello que aún no ha sido definido o codificado). En un marco más amplio, el teórico camerunés Achille Mbembe (2022) aborda conceptos que contrastan lo calculable y lo incalculable, o más específicamente las políticas de visibilidad e invisibilidad. Su trabajo suscita una reflexión crítica sobre lo que se puede cuantificar y lo que escapa a la cuantificación en la vida social y la experiencia humana. En el contexto de los estilos y expresiones culturales, se podría extrapolar la perspectiva de Mbembe para argumentar que, si bien ciertos aspectos de la cultura pueden medirse o analizarse (lo calculable), hay dimensiones de la creatividad, el afecto y el significado humanos que resisten la cuantificación o el análisis computacional (lo incalculable).

En el contexto de la IA y, de hecho, de la cultura digital en general, lo calculable adquiere cada vez más seriedad. El empaquetamiento de estilos y géneros dentro de filtros, aplicaciones y software fáciles de usar parece promover la creatividad y proporcionar medios muy directos e intuitivos de producción cultural, pero, aun así, ofrece una repetición general de estilos, lo que conduce a una reducción de los actos de elección o libertad (en el sentido sartreano). Por ejemplo, la facilidad para seleccionar un sonido de batería o efectos de postproducción en Pro Tools (el estándar de la industria musical para los estudios de grabación) es liberadora en las circunstancias inmediatas de crear una pista, pero traiciona la visión de Barthes sobre el estilo como algo involuntario (como algo *ya* codificado).

El grano de la voz

En su ensayo «El grano de la voz» (2013), Barthes ofrece una forma de entender lo incalculable. Al escribir sobre la recepción de la música (específicamente sobre el vocalista), distingue entre el «feno-canto», que refiere a aspectos de la interpretación culturalmente codificados y comunicables, y el «geno-canto», que representa la presencia corporal y física de la voz misma del cantante, o lo que él llama el «grano». Eso es lo que Barthes

identifica como la cualidad única e irreductible de la voz, ligada a la materialidad del cuerpo que la produce. El *grano de la voz* es el fenómeno, por ejemplo, que nos permite diferenciar entre la versión grabada de una canción y la versión en vivo. Así es como tendemos a asociarnos con la «firma» de un artista específico, su carácter único que, al mismo tiempo, proporciona matices incuantificables a la forma de cantar una canción. Podría decirse que existe una contradicción con la explicación anterior de Barthes del estilo como codificado y el grano de la voz como una forma corporal idiosincrásica y no codificada. Sin embargo, a los efectos que nos ocupan, la cuestión es que sigue existiendo una «presencia» que no puede cuantificarse ni copiarse. Si bien esto podría sonar a visión romántica, el punto de vista de Barthes sigue siendo estructuralista. El «geno-canto» es parte de una estructura general que trabaja en conjunto con el «feno-canto». Fundamentalmente, es un componente que no forma parte de la comunicación de significado (no se basa en la estructura del lenguaje, por ejemplo; no es el medio codificado para representar sentimientos o expresiones). De hecho, el «geno-canto» o el «grano de la voz» es simplemente sonido puro. Además del significado de una canción, lo que escuchamos es el timbre del vocalista individual —una forma de fricción o «ruido»—. Es una cualidad que normalmente designaríamos como única y «analógica», es decir, no reducible a aquello que puede calcularse, repetirse o digitalizarse (codificado como 1 y 0). Es posible hacer una muestra del grano de la voz, pero básicamente se trata de realizar una grabación específica. Por el contrario, la letra y la melodía de una canción se pueden transmitir, transponer y «citar» fácilmente. Haciendo una analogía con la escritura a mano, que tendemos a pensar como irreductible y difícil de compartir, el grano de la voz es propiedad exclusiva del individuo.

Sin embargo, en el contexto de la IA, el reconocimiento de patrones y la informática de alto rendimiento, nuestra comprensión general de lo incalculable se ve muy cuestionada. Las nuevas tecnologías de audio están avanzando en formas que conciben el grano de la voz menos como un relato romántico de presencia única y más como una «unidad» real de significado y medida.

El software Overdub de Descript, por ejemplo, permite a los usuarios clonar su voz usando IA y escribir texto que se reproducirá luego como audio. Es una tecnología utilizada para corregir podcasts o narraciones de videos sin necesidad de volver a grabarlos. En el ala opuesta, el Proyecto Euphonia, de Google, tiene como objetivo ayudar a las personas con discapacidades del habla para comprender y sintetizar patrones de habla deficientes mediante inteligencia artificial (aunque sin modelar directamente la voz de la persona). Un *software* más destacado es Vocaloid, capaz de «imitar» la voz de un cantante para luego extraer el canto introduciendo melodía y letra. Utiliza bancos de voz grabados por actores de doblaje o cantantes y puede crear cantos con varias voces, incluidas las de personajes ficticios y cantantes conocidos. En principio, significa que se puede utilizar *cualquier voz para cantar cualquier canción*. La tecnología se relaciona con implementaciones más amplias de *deepfake*, mediante las cuales se pueden clonar voces reales (generalmente de celebridades) a través de entrenamiento de la IA con horas de audio real.

El uso de la IA para modelar la voz de un cantante abre posibilidades innovadoras para la producción musical, la accesibilidad y el entretenimiento, pero también plantea cuestiones importantes sobre el consentimiento, los derechos de autor y el uso ético (fig. 1). Podemos plantear aquí curiosas posibilidades. Copiar a un cantante como Michael Jackson, por ejemplo, y grabar canciones *jamás escritas* con el estilo único de Jackson (es decir, utilizando el grano de su voz) se vuelve técnicamente factible. Esto plantea numerosas preguntas filosóficas sobre la autenticidad, la creatividad y la esencia de la expresión humana en el arte, pero, no obstante, emerge como una posibilidad real en la que la «presencia» del artista ya no es crucial. Más allá de las complejas discusiones sobre la naturaleza de la producción cultural, surgen preocupaciones sobre la *saturación* en la era digital. Las nuevas capacidades tecnológicas plantean la cuestión de si puede haber «demasiada cultura», es decir, una sobresaturación de producción cultural que los humanos nunca podríamos consumir completamente.

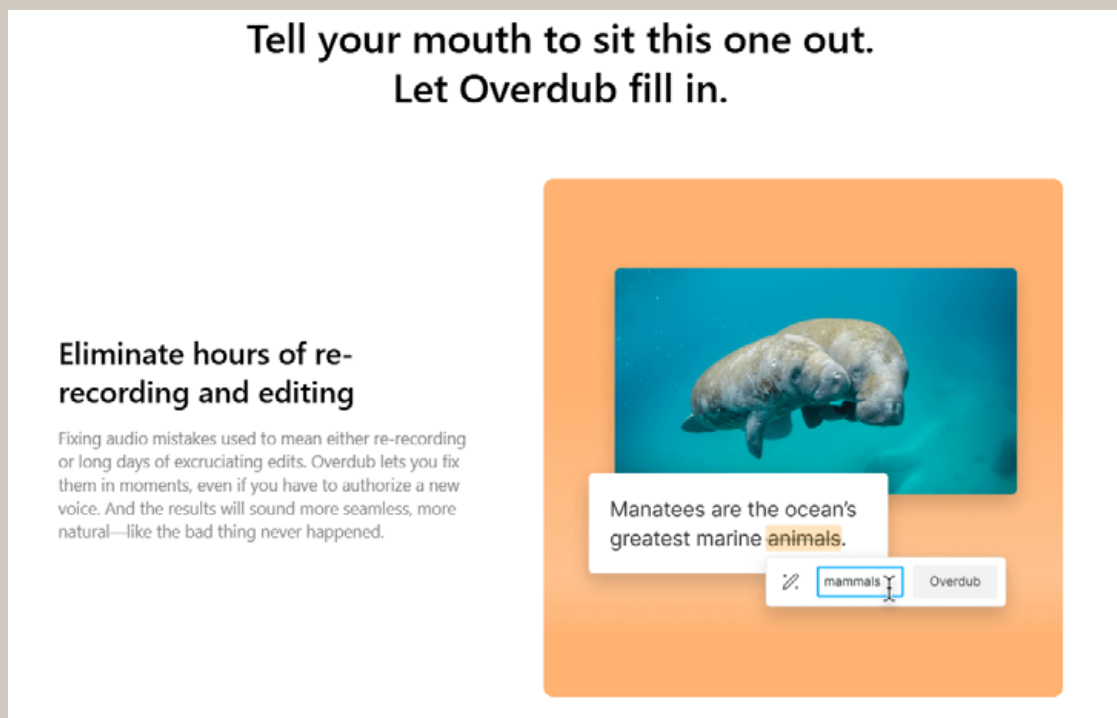


Fig. 1. Overdub, plataforma para clonar voz y sonido

¿Hay demasiada cultura?

En su análisis estructural de la narrativa, Roland Barthes destaca el fenómeno de las historias como «material humano, una clase de cosas que los humanos producen» (Fabbri, 2022, p. 163). Al principio, parece que no podemos imponer ningún tipo de orden a las narrativas: «Hay millones y millones de narrativas, desarrolladas durante un período de tiempo indefinido, cuyos orígenes se desconocen. (...) La narrativa está en todas partes» (Fabbri, 2022, p. 163). Por lo tanto, para la facultad humana, las narrativas y las melodías bien podrían ser infinitas: las matemáticas son demasiado grandes para sujetarlas. Sin embargo, no es lo mismo decir que las narrativas del mundo son infinitas. En *Words and Rules* (1999), Steven Pinker nos proporciona algunos números. Está interesado en la naturaleza combinatoria y recursiva del lenguaje, y tiene en cuenta sus reglas combinatorias.

Digamos que el inglés cotidiano tiene cuatro artículos determinantes (*a*, *any*, *one* y *the*) y diez mil sustantivos. Entonces, la regla para un sintagma nominal permite cuatro opciones para el determinante, seguidas de diez mil opciones para el sustantivo principal, lo que da $4 \times 10\,000 = 40\,000$ formas de pronunciar un sintagma nominal. La regla para una oración permite que estos cuarenta mil sujetos sean seguidos por cualquiera de los cuatro mil verbos, proporcionando $40\,000 \times 4\,000 = 160\,000\,000$ formas de pronunciar las tres primeras palabras de una oración. Luego hay cuatro opciones para el determinante del objeto (640 millones de comienzos de cuatro palabras) seguidas de diez mil opciones para el sustantivo principal del objeto, o $640\,000\,000 \times 10\,000 = 6\,400\,000\,000\,000$ (6,4 billones) de oraciones de cinco palabras. Supongamos que se necesitan cinco segundos para producir una de estas oraciones. Para poner algunos ejemplos, desde *El abandono degradó la abadía* y *El abandono degradó al abad*, pasando por *El abandono degradó al zoólogo*, hasta llegar a *El zoólogo zonificó el zoológico*, esto tomaría un millón de años (Pinker, 1999, p. 16).

Pinker evoca el cuento de Jorge Luis Borges «La biblioteca de Babel» —entendido como «quizás la descripción más vívida del asombroso poder de un sistema combinatorio» (1999, p. 16)—. Según cuenta la historia, en algún lugar de la biblioteca hay un libro que contiene la verdadera historia del futuro (incluida la historia de tu muerte); un libro de profecía que reivindica los actos de cada hombre en el universo; y un libro que contiene el esclarecimiento de los misterios de la humanidad» (1999, p. 16). Por supuesto, incluso después de extinguida la especie humana, la biblioteca (y sus posibilidades combinatorias) permanecerá intacta. Sin embargo, *técnicamente*, explica Pinker:

Borges no necesitaba haber descrito la biblioteca como «infinita». Con ochenta caracteres por línea, cuarenta líneas por página y 410 páginas por libro, el número de libros es alrededor de 10 [elevado a] 1 800 000, o 1 seguido de 1,8 millones de ceros. Se trata, sin duda, de un número muy grande: solo hay 10 [elevado a] 70 partículas en el universo visible, pero es un número finito (1999, p. 16).

Si bien parece infinito, es la naturaleza (aunque masivamente) finita del lenguaje lo que nos lleva a comprender el giro estadístico en el desarrollo de la IA. Si bien a nivel humano no podemos tener una idea de la magnitud de las palabras, oraciones y narrativas, para la informática de alto rendimiento (y potencialmente la computación cuántica aún por llegar) las sumas están dentro de su alcance. Y no solo palabras, sino el cómputo de la enorme variedad de estilos en palabras, sonidos y gestos. Podemos empezar a imaginar una «máquina cultural» capaz de *hacer, consumir y rehacer* todas las posibilidades combinatorias *finitas* de todas las canciones, novelas y películas que jamás se hayan realizado.

Podríamos sentirnos tentados a revisar el teorema del mono infinito, según el cual un mono que escribe aleatoriamente en un teclado durante una cantidad de tiempo infinita producirá, en algún momento, las obras completas de Shakespeare. Una supercomputadora que escriba probabilísticamente durante un período de tiempo *finito* podrá producir las obras completas de Shakespeare. Sin embargo, es importante destacar que la experiencia del tiempo en este caso es muy diferente a la experiencia humana. Si una persona pudiera leer 60 libros en un año (y suponiendo que comenzara a leer a los 6 años y viviera hasta los 81), podríamos estimar que leería alrededor de 4 500 libros. La lectura humana, por supuesto, solo se acumula con el tiempo. Sin embargo, consideremos que, después de un período relativamente corto, es técnicamente posible que una computadora ya haya podido leer todos los libros disponibles para el hombre jamás publicados. Es como si, con una combinación segura, toda la cultura se desbloqueara con un clic y, con base en ello, fuera posible volver a representarla nuevamente, *toda a la vez*.

La última canción de The Beatles

Teniendo en cuenta esta revisión del infinito como finito, volvamos al modelado del «grano de la voz» y a una aplicación específica en el mundo real con el lanzamiento de la canción de The Beatles, «Now And Then», en octubre de 2023, ampliamente auspiciada como el «último» sencillo de la banda.² Originalmente escrita y cantada por John Lennon a finales de la década de 1970, sobrevivió solo como un «boceto» registrado en una grabadora portátil mientras Lennon estaba sentado al piano (con varios sonidos ambientales en la habitación, incluido un televisor). Mucho después de la muerte de Lennon, a mediados de los años noventa, los Beatles restantes, Paul McCartney, George Harrison y Ringo Starr, trabajaron en la canción, pero fue abandonada porque no pudieron aislar satisfactoriamente la voz y el piano de Lennon. Cuatro décadas después de su composición inicial, McCartney y Starr finalmente terminaron la canción. Tras su lanzamiento, la tan promocionada «última» canción de los Beatles recibió una recepción amplia y positiva, descrita, por ejemplo, como «un tributo conmovedor al vínculo de la banda» y como algo «inquietante» (Petridis, 2023).

Entonces, ¿qué fue lo que permitió completar la canción? En resumen, fueron avances en inteligencia artificial. Específicamente, fue trabajo del director Peter Jackson, cuyo equipo empleó un enfoque revolucionario para procesar grabaciones de audio de las sesiones originales de *Let It Be*, de 1969. La técnica, conocida como aprendizaje asistido por máquina (*Machine-Assisted Learning*), en homenaje a Mal Evans (un querido chófer y asistente de los Beatles), abordó el desafío de aislar la voz del ruido de fondo, incluida la música, la charla y otros sonidos presentes durante las sesiones de grabación. El resultado fue la aclamada serie documental *The Beatles: Get Back* (Peter Jackson, 2021) que, entre otras cosas, ayudó a sacar a la luz sorprendentes imágenes de la composición de canciones en tiempo real (el espectador, por ejemplo, puede presenciar a Paul McCartney escribiendo «Get Back» literalmente desde cero).³

Si bien las especificaciones técnicas del aprendizaje asistido por máquina son exclusivas y no se divulgan en su totalidad, existen numerosas metodologías similares implementadas en el arte y la restauración del sonido. El núcleo de la tecnología implica modelos de aprendizaje automático entrenados con grandes can-

² La canción fue estrenada mundialmente por McCartney en su concierto en Montevideo, el 1.º de octubre de 2024.

³ Véase Polonis (2021). «Why All Creatives Must Watch The Beatles Get Back Documentary». Medium. <https://medium.com/the-millennial-republican/why-all-creatives-must-watch-the-beatles-get-back-documentary-f718b1eab92d>

tidades de datos para reconocer y diferenciar varios tipos de señales de audio, ayudando así a identificar las características de la voz humana y distinguiéndola de otros sonidos. En *Get Back* se aplicó esta tecnología a las grabaciones de las sesiones inéditas, conocidas por su mala calidad de audio y sonidos superpuestos. El equipo de Jackson pudo mejorar significativamente la claridad del diálogo, volviendo audibles las interacciones de los miembros de la banda con un detalle sin precedentes. Después de la separación inicial, las pistas de audio se limpian, equilibran y mejoran aún más para revisar la claridad y la calidad. Esto puede implicar reducción de ruido, ecualización y otras técnicas de masterización de audio. El proceso general plantea preguntas complejas sobre el estado del sonido original y su producción final. ¿Lo que recibimos es la *restauración* del audio original o el proceso implica la producción sintética de audio basada en aprendizaje probabilístico? La distinción afecta a debates más amplios en ingeniería de audio. Por ejemplo, en el momento en que un cantante canta frente a un micrófono, ¿estamos capturando solo una aproximación de su voz? En esencia, la respuesta es «sí», pero con el tiempo nos acostumbramos al sonido de la voz grabada hasta el punto de entenderla como un sonido original.

En el caso de *Get Back*, la tecnología buscaba principalmente separar los sonidos. Sin embargo, se puede decir que el proceso involucra aspectos de producción sintética si se requiere que la inteligencia artificial «prediga» o «rellene» información de audio imposible de extraer del ruido de fondo. No se ha detallado completamente hasta qué punto ocurrió esto con los Beatles, lo que genera ambigüedad sobre en qué medida el audio final es una restauración directa o una reconstrucción sintética.

Expresiones del estilo artístico

Un caso similar es la serie documental de Netflix *Los diarios de Andy Warhol*. La serie, de seis capítulos, utilizó tecnología de inteligencia artificial para recrear la voz de Andy Warhol. Como explica Flora Roumpani:

La voz del pionero del arte pop estadounidense se creó utilizando un algoritmo de conversión de texto a voz que incorporó su acento nativo de Pittsburgh. Luego, el actor Bill Irwin grabó las frases, combinando su actuación y la voz digital de IA para crear la voz que se escucha en la serie documental. (...) El sistema de inteligencia artificial utilizado en la serie documental se basa en una arquitectura de red neuronal llamada WaveNet, que fue desarrollada por investigadores de DeepMind de Google. WaveNet es un modelo de aprendizaje profundo que puede generar formas de onda de audio de alta calidad modelando directamente la señal de audio sin procesar. En el caso del proyecto Warhol, el sistema supuestamente fue entrenado con más de 1.000 horas de datos de audio para crear una voz realista y auténtica para el artista (2023, párrs. 2 y 4).

Como se señaló anteriormente, la técnica implicó una superposición de audio generado por IA con actuaciones en vivo, lo que representa una forma novedosa de producción, un ejemplo de cómo la IA está comenzando a usarse en procesos de producción de medios más amplios y establecidos. El director, Andrew Rossi, estaba interesado en adoptar este complejo proceso para «apreciar plenamente la extrema sensibilidad que Andy revela en sus diarios», argumentando que «necesitábamos escuchar las palabras en la propia voz de Andy» (Roumpani, 2023, párr. 3).

Estos ejemplos de restauración de IA en *Get Back*, el sencillo «Now and Then» y *Los diarios de Andy Warhol*, representan usos considerados y elaborados de técnicas de IA. Pero, no obstante, son parte de un dominio más amplio, una nueva perspectiva en la que podemos modelar y replicar el estilo. Hasta ahora, lo significativo es que hemos tendido a no considerar el estilo como algo que puede replicarse o que puede considerarse propiedad exclusiva del *copyright*. La ley de derechos de autor protege principalmente las obras originales de autoría fijadas en un medio tangible de expresión, incluidas obras literarias, musicales, dramáticas y artísticas, entre otras. El concepto de «estilo» es más abstracto y generalmente se refiere a la forma o técnica distinta que utiliza un artista o creador en su trabajo. Dado que el estilo en sí es un conjunto amplio y a menudo intangible de características, no es directamente protegible bajo la ley de derechos de autor como tal.

Las expresiones específicas del estilo de un artista en formas tangibles, como una pintura, una obra escrita o una pieza musical, se pueden proteger. Esto significa que, si bien no se pueden salvaguardar los derechos de autor de un estilo (por ejemplo, el impresionismo, el cubismo, un género musical específico o una forma de escribir), las obras individuales producidas con ese estilo son protegibles. De manera similar, los elementos distintivos que son originales y están fijados de forma tangible dentro de una obra (como la apariencia única de un personaje, un arreglo musical específico o un diseño gráfico particular) pueden preservarse por sus

derechos de autor. Entonces, lo más importante es que, si bien la ley de *copyright* protege la expresión de ideas y activos tangibles, eso no impide que otros creen obras *con el mismo estilo*, siempre que no copien expresiones protegidas de ese estilo.

Aquí es importante la noción de la IA generativa. Los nuevos modelos de IA generativa para texto, imágenes, música y video, si bien se basan en grandes cantidades de obras tangibles anteriores (incluidos materiales protegidos por derechos de autor), no *reproducen* obras existentes. En cambio, *producen* nuevas obras basadas en el «conocimiento» de formas anteriores. La distinción clave es que la IA generativa funciona a través de la «forma» (y el estilo), más que de un contenido específico. Entonces, así como en un tributo se permite copiar el *estilo* de una banda o artista (siempre que se obtenga permiso para tocar las canciones originales), de igual modo la IA generativa es técnicamente capaz de reproducir estilos *ad infinitum*. El tributo, por ejemplo, no pretende ser la banda *original*. Siguiendo la misma lógica, esto permite que otros artistas graben sus propias versiones de las canciones. Se les permite adoptar su propio estilo, si tienen permiso para utilizar la «idea» de la canción original. Sin embargo, el problema que surge con la IA generativa (y el cálculo de aprendizaje profundo en general) es una dimensión completamente nueva en la cuestión del estilo.

Coda: nunca fuimos posmodernos

Cuando los estilos se vuelven calculables, ya no tiene sentido considerarlos dominio *potencialmente* exclusivo de un artista o grupo individual, es decir, como algo incapaz de ser cartografiado o territorializado. Consideremos, por ejemplo, si es factible hacer referencia a la «última» canción de los Beatles. Teniendo en cuenta que los modelos actuales son en gran medida defectuosos, pero centrados en que estamos vislumbrando un futuro completamente nuevo en el que es realmente posible imitar el grano de la voz y, de hecho, todo el sonido y el estilo compositivo de una banda o artista, ¿es válido entonces asegurar que no habrá más canciones de los Beatles? Nunca digas nunca...

Investigadores de la empresa Sony han utilizado el *software* Flow Machines para analizar una base de datos de 13 000 partituras musicales de diferentes géneros de todo el mundo. A partir de esa capacitación, el software escribe sus propias melodías, desarrolladas luego por un compositor humano para convertir el material en una pista musical. A través de este proceso produjeron «Daddy's Car», una canción pop *al estilo de* The Beatles. Como escribe James Vincent:

La pregunta que queda... ¿de qué diablos se trata realmente «Daddy's Car»? Aunque la letra fue escrita por [el compositor humano] Carré en lugar del software, es imposible no detectar las intimidaciones amenazantes de la inteligencia artificial en la composición, accionando presumiblemente a través de Carré como un siniestro titiritero (2016, párr. 3).

La respuesta de Vincent es algo dramática, pero de todos modos refleja una ansiedad y una sospecha generalizadas sobre los nuevos desarrollos de la IA, entre ellos, la perspectiva de que la IA generativa elimine el papel de los artistas y diseñadores. Por supuesto, podríamos argumentar que se expresaron inquietudes similares con los primeros desarrollos de la fotografía, pero finalmente este medio se convirtió en un pilar de la práctica artística. La plataforma de video generativo de OpenAI, Sora, aún no está disponible públicamente, pero, al momento de escribir este artículo, la empresa ha dado acceso a «artistas visuales, diseñadores, directores creativos y cineastas» para evaluar las primeras sensaciones de los usuarios (OpenAI, 2024).

Cabe destacar *Air Head*, un cortometraje sobre un hombre cuya cabeza es un globo amarillo lleno de aire caliente, creado con Sora por la compañía canadiense shy kids. Técnicamente, es impresionante, dado que representa de manera convincente a ese protagonista humano con cabeza de globo (fig. 2); como explica la voz narrativa: «Estoy literalmente hecho de aire caliente». Pero lo que más impresiona es la poesía de la película. Es una pieza ingeniosa y provocativa, un buen ejemplo de lo que pueden hacer los artistas de manera independiente. Como apunta Walter Woodman, director del cortometraje: «Por muy bueno que sea Sora generando cosas que parecen reales, lo que nos emociona es su capacidad para hacer cosas que son totalmente surrealistas» (OpenAI, 2024). En otras palabras, los artistas se centran menos en la tecnología y más en el trabajo que quieren realizar.



Fig. 2. *Air Head*, cortometraje generado con IA por shy kids utilizando Sora, de OpenAI

Quizás estemos al borde de la explosión de una nueva generación de obras de arte y cultura. Y, como ocurre con cualquier «ruptura» epistémica, abundan las grandes preguntas. Si no aprovechamos las nuevas tecnologías de manera creativa, es posible que nos veamos inundados por una masa de «cultura» autogenerativa demasiado grande para que podamos siquiera experimentarla. Sin embargo, este es un escenario improbable, aunque solo sea porque no tiene sentido semejante masa de producción. Es importante reconocer la llegada de una nueva condición cultural, mucho más allá de cualquier cosa que Warhol pudiera haber imaginado.

En su libro *Nunca fuimos modernos*, publicado en 1991, Bruno Latour cuestionaba la idea de separación entre naturaleza y cultura que durante mucho tiempo se pensó como una narrativa central de la modernidad (por ejemplo, el progreso de la ciencia y la tecnología en la superación del mundo natural). Sostenía que esa separación es un mito, ya que la modernidad en realidad representa la purificación simultánea de los objetos (a través de la ciencia y la tecnología) y la proliferación de entidades híbridas (que él llama «cuasi-objetos» o «redes» y, por lo tanto, la ciencia nunca está exenta de lo social y lo cultural). En lugar de considerarnos modernos haciendo una separación artificial, deberíamos reconocer la interconexión de los humanos, las entidades no humanas y los artefactos tecnológicos en lo que Latour denominó el «Parlamento de las Cosas». Es importante destacar que su relato fue una crítica al discurso posmoderno, acusando a los filósofos posmodernos (como Lyotard y Baudrillard, pero también a pensadores como Barthes, Lacan y Derrida) de que su pensamiento simplemente giraba en torno a «mundos de signos» artificiales (Latour, 2022). Por esa razón, los desafiaba con la provocativa afirmación de que «nunca fuimos modernos» que da título a su libro.

Sin embargo, ¿qué pasa si nunca hemos sido posmodernos? En otras palabras, ¿qué sucedería si la capacidad de captar el mundo de los signos a finales del siglo XX (en el momento en que los filósofos posmodernos produjeron sus icónicas obras) fuera enormemente insuficiente para comprender y responder tanto a la idea de las entidades híbridas de Latour como a las visiones posmodernistas sobre el relativismo, la fragmentación, la simulación y la «muerte del autor»? ¿Y si recién ahora empezamos a comprender el calculable «mundo» de cosas, signos y *estilos* que siempre han *persistido*, aunque solo sea fuera de nuestro alcance? Podríamos considerar que las «técnicas» emergentes de la IA generativa, junto con internet, los macrodatos y otras tecnologías digitales, marcan el inicio de una nueva era que trasciende (y hace maleable) el escepticismo y el relativismo del posmodernismo. Esta era podría caracterizarse por una nueva forma de realismo u otra postura filosófica que reconozca el profundo impacto de la tecnología en nuestra comprensión de la verdad, la realidad y la identidad humana. Las distinciones que el posmodernismo estableció entre alta y baja cultura, o entre autor y lector, podrían entenderse como consideraciones menores en un mundo cada vez más irreverente en el que los algoritmos producen arte, literatura y música, lo que a su vez desafía no solo nuestras nociones de creatividad y originalidad, sino también el espacio y el tiempo mismos de hacer arte.

Un pensamiento desconcertante. Por ahora, sin embargo, podemos aferrarnos al estilo en todas sus idiosincrasias. Podemos consolarnos, tal vez, con algunas líneas de «Space Oddity», de David Bowie (1969):

And I'm floating in a most peculiar way
And the stars look very different today

No solo es una canción icónica, sino una nunca replicable mediante mera calculabilidad, ya que cuando Bowie sincopa la palabra *peculiar*, representa en nuestro nombre una naturaleza humana totalmente peculiar, un recordatorio del estilo más infatigable.

Referencias bibliográficas

- Barthes, R. (1985). *The Responsibility of Forms: Critical Essays on Music, Art, and Representation* (Richard Howard, Trad.). New York: Hill and Wang.
- Barthes, R. (2011). *El grado cero de la escritura*. Buenos Aires: Siglo XXI.
- Barthes, R. (2013). *El grano de la voz. Entrevistas 1962-1980*. Buenos Aires: Siglo XXI.
- Cikaj, K. (2023). Why Did Bruno Latour Claim that «We Have Never Been Modern»? *The Collector*. <https://www.thecollector.com/bruno-latour-we-have-never-been-modern/>
- Fabbri, P. (2022) On Narrative: An Interview with Roland Barthes. *Theory, Culture & Society*, 39(7-8), 159-174. <https://doi.org/10.1177/02632764221141819>
- Flow Machine. (2016). AI makes pop music. *Flow Machines*. <https://www.flow-machines.com/history/events/ai-makes-pop-music/>
- Hebdige, D. (2004). *Subcultura. El significado del estilo*. Barcelona: Paidós.
- Jackson, P. (Director). (2021). *The Beatles: Get Back*. Disney+.
- Kelley, R. D. G. (1998). *Yo' Mama's Disfunktional! Fighting the Culture Wars in Urban American*. Boston: Beacon Press.
- Latour, B. (2022). *Nunca fuimos modernos. Ensayos de antropología simétrica*. Madrid: Clave Intelectual.
- Mbembe, A. (2022). *Brutalismo*. Barcelona: Paidós.
- Neal, M. A. (1999). *What the Music Said*. New York: Routledge.
- OpenAI. (2024). Sora: first impressions. *OpenAI*. <https://openai.com/index/sora-first-impressions/>
- Perry, I. (2004). *Prophets of the Hood: Politics and Poetics in Hip Hop*. Durham: Duke University Press.
- Petridis, A. (2 de noviembre de 2023). The Beatles: Now and Then review – ‘final’ song is a poignant act of closure. *The Guardian*. <https://www.theguardian.com/music/2023/nov/02/the-beatles-now-and-then-review>
- Pinker, S. (1999). *Words and Rules: The Ingredients Of Language*. New York: Basic Books.
- Polonis, J. (2021). «Why All Creatives Must Watch The Beatles Get Back Documentary». *Medium*. <https://medium.com/the-millennial-republican/why-all-creatives-must-watch-the-beatles-get-back-documentary-f718b1eab92d>
- Rose, T. (1994). *Black Noise: Rap Music and Black Culture in Contemporary America*. Middletown: Wesleyan University Press.
- Rossi, A. (Director). (2022). *Los diarios de Andy Warhol*. Netflix.
- Roumpani, F. (23 de abril de 2023). Lights, Camera, AI. From Andy Warhol reading his own diaries, to post-production wonders of the Marvel universe. *Medium*. <https://medium.com/the-ai-and-arts-forum/lights-camera-ai-98e6bc4a2978>
- Sartre, J. P. (2004). *¿Qué es la literatura?* Buenos Aires: Losada.
- Vincent, J. (26 de setiembre de 2016). This AI-written pop song is almost certainly a dire warning for humanity. *The Verge*. <https://www.theverge.com/2016/9/26/13055938/ai-pop-song-daddys-car-sony>