

El cómputo afectivo como experiencia artística

El cómputo afectivo desarrolla sistemas que interpretan y emulan emociones humanas para crear interfaces más empáticas. Mi obra artística explora estas capacidades mientras cuestiona las implicaciones éticas de sistemas que toman decisiones según señales emocionales como los gestos faciales o sonoros. En este artículo, presento piezas y técnicas que abordan el concepto desde la inteligencia artificial.

Introducción

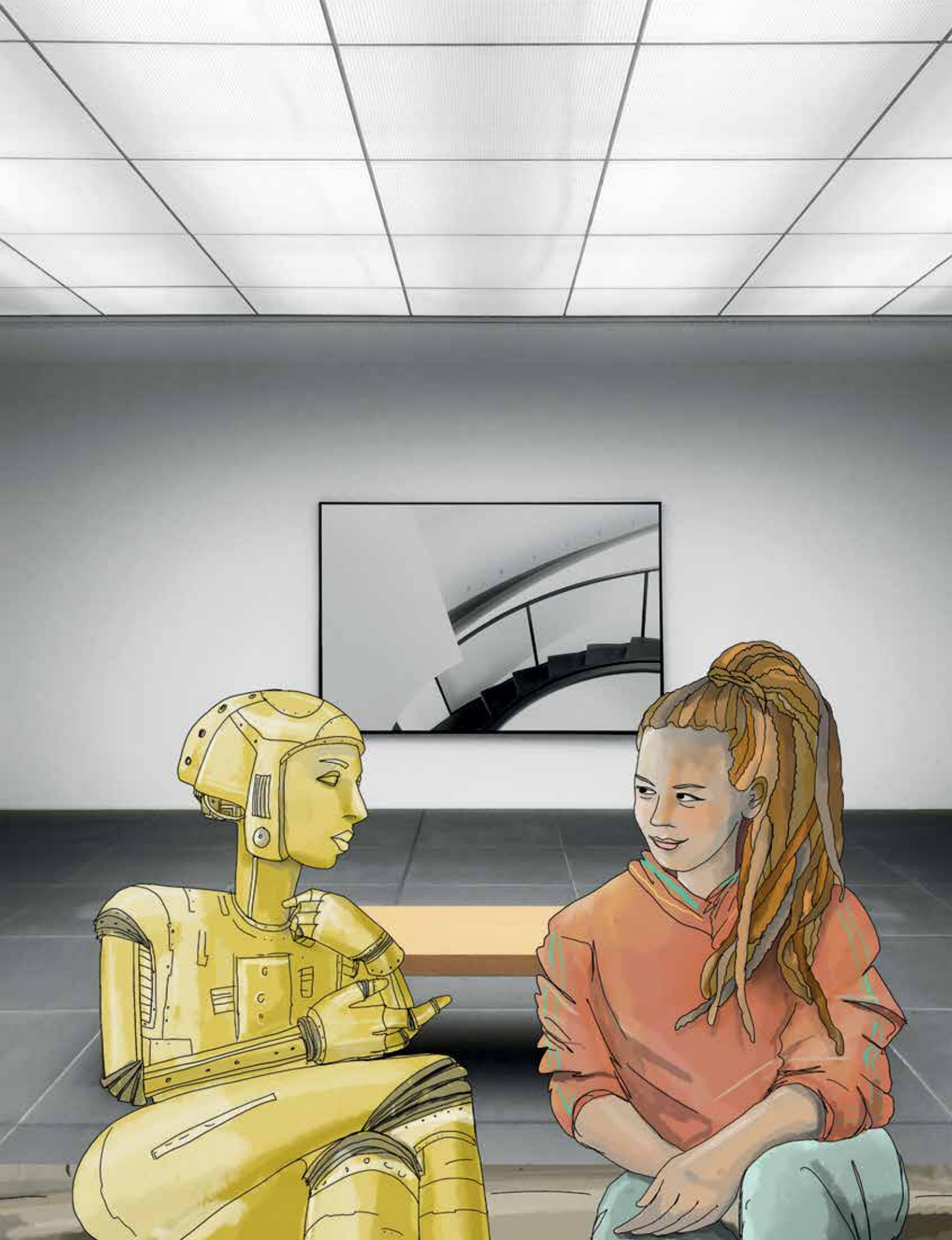
Al ser parte de una generación que vivió la transición de los medios analógicos a los digitales, a lo largo de mi carrera como artista he desarrollado un profundo interés en crear obras que exploren los vínculos entre los humanos y las máquinas. No fue sino hasta hace pocos años que estas prácticas creativas, cuyas piezas aprovechan herramientas tecnológicas, tuvieron reconocimiento de las instituciones de arte contemporáneo en México. Y esto se debe, desde mi perspectiva, principalmente al auge de la inteligencia artificial aplicada a la creación de imágenes y a la popularización de los **NFT**, que transformaron la comercialización del arte digital.

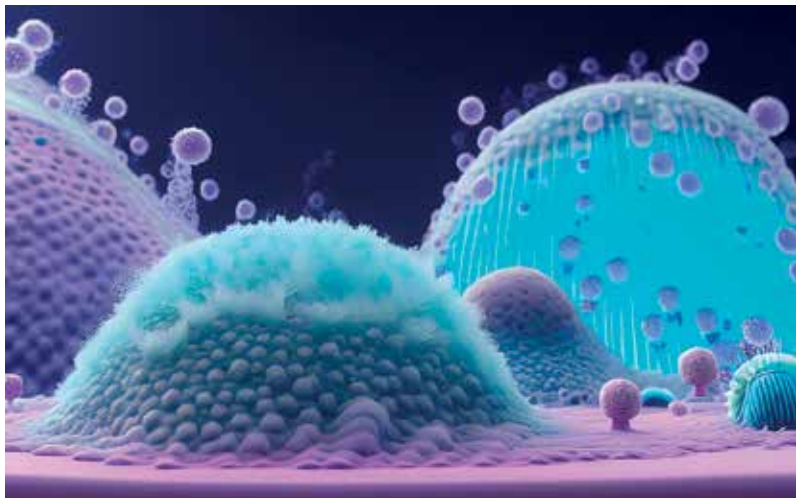
Me considero parte de una segunda generación de artistas electrónicos que exploran las posibilidades de desarrollar sus propios sistemas, tanto de *hardware* como de *software*, como parte integral de sus obras. A partir de la conexión entre arte, ciencia y tecnología busqué, con mis primeras piezas, entablar relaciones entre la experiencia humana, el discurso artístico y el diseño de narraciones interactivas no lineales, mediadas por sistemas digitales autónomos e interfaces electrónicas.

Con la aparición del primer *smartphone* en 2006 y la rápida evolución de los componentes electrónicos, comenzaron a surgir especulaciones sobre el (muy cercano) futuro impacto de las tecnologías digitales en la vida humana. Mi exploración artística se enfocó en cómo evolucionaría la comunicación interpersonal y con entes no humanos, y la manera en que esto impactaría emocionalmente en sus usuarios y, a la larga, en toda la sociedad. Las piezas que describiré en este texto tienen el objetivo común de haber sido configuradas como interfaces

NFT

Del inglés, *Non Fungible Token*, son valores simbólicos en medios digitales que se utilizan para representar la propiedad de objetos únicos.





Anni Garza Lau, *Databiota 1*.

lúdicas en las que el público experimentaba, a través de distintos estímulos, la interacción con sistemas autónomos que apelaban a sus respuestas emocionales. En varias de mis piezas estos sistemas están basados en modelos de inteligencia artificial, y en otras utilizo estas técnicas y tecnologías directamente. Las piezas que describo tienen en común la mediación de una máquina o el diálogo con la misma y las emociones del público. Dichas piezas reflejan la evolución de mi práctica artística del cómputo afectivo a la inteligencia artificial emocional.

■ **Vínculos afectivos: conexiones entre arte y cómputo**

Aunque posiblemente la forma en que apreciamos el arte no se ha modificado demasiado en el último siglo, la sobreexposición a dispositivos digitales con interfaces diseñadas para ser intuitivas y lúdicas, así como el gran volumen de información al que tenemos acceso a través de ellos, ha normalizado y exacerbado la producción de respuestas emocionales al interactuar con objetos discursivos.¹

En mi práctica aprovecho herramientas fabricadas esencialmente para el consumo comercial y me las apropio, desconfigurando y reconfigurando su

¹ Desde imágenes, memes, videos cortos y noticias, hasta productos audiovisuales creados enteramente con inteligencia artificial.

sentido, para construir un discurso propio. Al mezclar los posibles usos y conexiones de dichas herramientas con mis propias ideas y narrativas, surgen experiencias y objetos fuera de lo cotidiano. Como parte de mi investigación artística, uno de los paradigmas que continuamente he explorado con esta metodología es el cómputo afectivo (Picard, 1997).

El cómputo afectivo, o más recientemente llamado inteligencia emocional artificial,² es la rama de las ciencias computacionales que se ocupa del estudio y desarrollo de sistemas que puedan interpretar y emular afectos humanos, mayormente con el propósito de generar empatía en sus usuarios y crear interfaces más naturales para la comunicación entre humanos y máquinas. El término se originó en un artículo de 1995 de Rosalind Picard como parte de una investigación interdisciplinaria entre computación, psicología y ciencias cognitivas en el Massachusetts Institute of Technology. Propone que las máquinas deben ser capaces de reconocer el estado emocional de sus usuarios y modificar sus respuestas para simular empatía. A continuación describo tres piezas en las que utilicé técnicas de inteligencia artificial y que tenían este mismo objetivo, pero como parte de un discurso y experiencia artística e interactiva.

■ ***The Drama Manager*: narración interactiva y agentes**

■ En su libro *On Interactive Storytelling* (2004), Chris Crawford, afamado diseñador de videojuegos, postula la premisa de que la calidad de la interacción en una narración lúdica depende de la cantidad de decisiones en el juego que se les permite tomar a los jugadores y las consecuencias que sus elecciones tienen sobre la forma en que la historia se despliega. Propone que un *gestor de narrativas* o *drama manager* sería capaz de unir eventos de la trama de manera flexible utilizando un narrador digital.

The Drama Manager (2009) fue una obra digital e interactiva que surgió no sólo de mi idea de construir

² Esto debido a que ahora se utilizan, por ejemplo, redes neuronales para procesar patrones en gestos que pueden interpretarse como emociones.

una narración no lineal, sino de encontrar formas en las que los usuarios se relacionaran afectivamente con un personaje (una mujer animada a partir de un video mío en el jardín de mi abuela), teniendo que explorar, a partir de la propia interfaz, sus distintas formas del habla y gestos sonoros, para provocar respuestas emocionales en el personaje, y a la vez reflexionar sobre cómo se sentían al agredir, adular o seducir a un ente virtual (véase la **Figura 1**). La pieza se presentó en Cuadro, salón de artistas emergentes en el Centro Cultural Estación Indianilla, como una proyección en la que transcurría una animación de distintos paisajes. La protagonista, animada con roscopía, paseaba o esperaba pacientemente a que alguien llegara a interactuar con ella, mientras en la esquina izquierda de la pantalla se mostraba, a manera de lo que ahora parecería un *emoji* o *avatar* personalizado, su estado emocional actual. El público se paraba frente a un micrófono y podía producir cualquier sonido (hablar, cantar, gritar, aplaudir, besar, etc.) para comenzar a interactuar.

Para reconocer los patrones en la voz o los sonidos que pueden ser emitidos por el cuerpo, realicé una serie de grabaciones con distintas voces y organicé las pausas, volumen y frecuencias que se repetían al generar cierto tipo de conversación, emitir sonidos guturales o hacer ruidos con la boca y las manos. De esta manera, el programa detectaba cuándo iniciaba una frase y analizaba la duración de cada sonido antes de una pausa. Al producirse un silencio más largo, interpretaba que la intervención había finalizado y calculaba qué tipo de gesto sonoro se había producido.

En colaboración con José Luis García Nava, diseñamos un mapa de emociones para el personaje (véase la **Figura 2**), que contenía círculos distribuidos en cierto espacio dividido por tipos de emociones básicas y complejas, los cuales crecían dependiendo del tipo de estímulo sonoro. Cada círculo tenía un tamaño máximo diferente, para establecer inclinaciones en el carácter del personaje, y aumentaba o disminuía su tamaño dependiendo del tipo de estímulo sonoro detectado en cada momento. Por ejemplo, el círculo de la ira crecía cuando el público le gritaba al personaje. Los círculos, dentro o cerca



Figura 1. Interfaz de *The Drama Manager*. Imagen facilitada por la artista.

de una emoción, actuaban como atractores para un vehículo que representaba la emoción actual. La fuerza de atracción era directamente determinada por el tamaño del círculo. Dependiendo de en qué

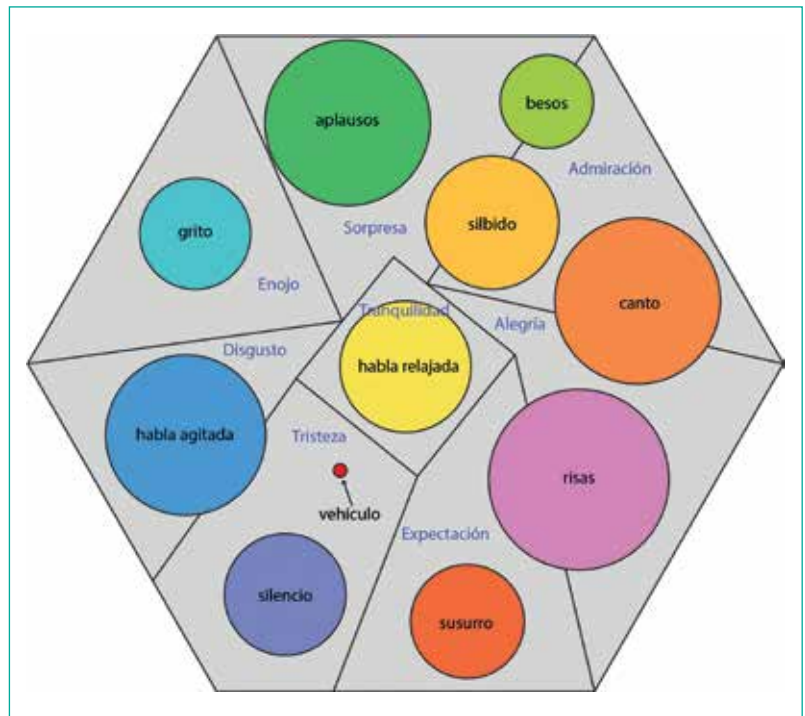


Figura 2. Mapa afectivo en la pieza, basado en la rueda de emociones de Plutchik. Imagen facilitada por la artista.

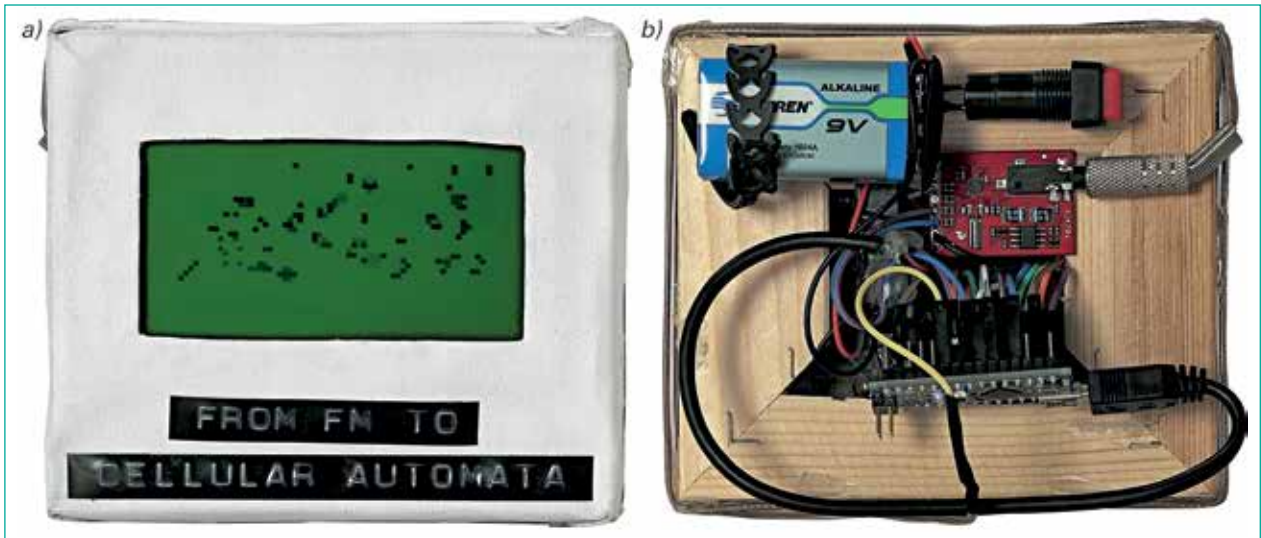


Figura 3. Registro de *From FM to Cellular Automata*. Imágenes facilitadas por la artista.

parte del mapa emocional se encontraba el vehículo, se representaba el estado del personaje, y éste a su vez tomaba decisiones en la narración dependiendo de cómo se “sentía”.

En inteligencia artificial, un agente es una entidad capaz de percibir su entorno, generalmente a través de sensores (en este caso un micrófono), y generar respuestas o tomar decisiones en tiempo real. La respuesta o salida del sistema de la pieza consistía en una segunda capa de decisiones: dependiendo del punto en el que se encontrara en la historia, un cambio emocional podía dirigir la narrativa hacia una nueva línea, aunque pasaba en ocasiones que los estímulos interpretados no eran suficientes para avanzar en la historia y se producían bucles repetitivos de acciones o incluso la falta de acción.

Pantallas LCD (Liquid Crystal Display)

Usan cristales líquidos y una luz de fondo para mostrar imágenes. Son eficientes en energía y comunes en televisores, monitores y dispositivos móviles, pero tienen menor contraste que las OLED.

Los autómatas celulares y el romance

En 2014 fui invitada a realizar una obra en formato pequeño para la Colección Imago Mundi, organizada mundialmente por Luciano Benetton. La colección y el catálogo se llamaron “Mexico: the future is unwritten”.³ Las piezas creadas en lienzos idénticos de 10 × 12 cm se presentaron en Venecia ese mismo año. Al ser un formato tan pequeño, uno de los ma-

yores retos fue pensar en una pieza electrónica que fuera visual y funcionara en relación con el tema del futuro. En gran medida mis propias experiencias en Venecia determinaron que la pieza se conectara con esa ciudad en particular.

From FM to Cellular Automata (2014) fue la pieza resultante de este ejercicio (Figura 3): el lienzo contenía una minipantalla LCD (pantalla de cristal líquido) monocromática en la que los píxeles eran visiblemente grandes, aparentando una cuadrícula en la que imaginé alguna forma de vida sintética. Como parte de mi interés en producir arte generativo, decidí utilizar autómatas celulares para que habitaran de manera activa la pieza. Me basé en el *Juego de la vida* de Conway,⁴ sabiendo que distribuir de manera aleatoria a mis “células vivas” finalmente terminaría por dejar la pantalla vacía, incluí en el circuito (que trabajé en colaboración con el ingeniero electrónico Roberto Garza) una antena de radio que interpretaba las frecuencias de la música de una estación de radio particular de Venecia, que sólo emitía canciones románticas. Dependiendo de las frecuencias de cada canción, nuevas células vivas se agregaban aleatoriamente a la cuadrícula del

³ “México: el futuro no está escrito”.

⁴ El *Juego de la vida* alude a un autómata celular diseñado por el matemático John Horton Conway en 1970. El juego no tiene jugadores y su evolución se determina por un estado inicial sin requerir ninguna intervención adicional.

autómata, asegurándose de que la población nunca se extinguiera.

Los autómatas celulares son modelos matemáticos que representan sistemas dinámicos que evolucionan en pasos. Sus reglas determinan los estados de vivo o muerto de unidades o células en una cuadrícula, y pueden ser bastante sencillas. El *Juego de la vida*, por ejemplo, sólo tiene tres reglas de proximidad que se aplican simultáneamente en cada paso y a cada célula. A pesar de la aparente simplicidad de este sistema, la actualización simultánea (computación en paralelo) y su interacción local (sólo interactúan con las ocho celdas que están a su alrededor), producen interrelaciones que se complejizan rápidamente y a menudo ocurren “comportamientos emergentes”. Aunque el *Juego de la vida* por sí mismo no se considera parte del *reconocimiento de patrones* (una rama de la IA), las simulaciones hechas a partir de autómatas celulares se utilizan a menudo para simular y encontrar patrones en sistemas naturales. Los comportamientos, como la autoorganización, jerarquización o cooperación, que surgen entre los elementos dentro de estos sistemas, pueden verse a su vez como una forma de inteligencia colectiva emergente.

■ Interfaces, identidad y redes neuronales

■ Entre 2019 y 2022 la doctora Lilianha Domínguez y yo exploramos las posibilidades de transmitir, guardar y modificar información identitaria en ADN. El proyecto, nombrado *Selección antinatural*, exploró la idea de **cómputo biológico**, como una forma de mantener información potencialmente sensible en medios que no fueran digitales y al mismo tiempo cuestionar qué es la identidad humana en el contexto de una era digital en la que los sesgos producidos por el reconocimiento facial y otras tecnologías vinculadas a la IA determinan la gobernanza en las sociedades del mundo actual. Como parte del proyecto, extrajimos secuencias de nuestro ADN y las insertamos en plásmidos dentro de bacterias. Posteriormente produjo, con varias técnicas de inteligencia artificial, una pieza digital que continuaba estos cuestionamientos. *Interfaces de la identidad* (2021) se



Anni Garza Lau, *Databiota 2*.

presentó como parte de la exposición conmemorativa del décimo aniversario del Centro de Cultura Digital en la Ciudad de México.

Aunque se trata de una instalación producida por computadora, la pieza utiliza información proveniente de nuestras observaciones al replicar secuencias de ADN humano en bacterias, así como de las mutaciones que sufrieron. En *Interfaces de la identidad*, el espectador se acercaba a un libro, *Un mundo feliz*, de Aldous Huxley, iluminado por una lámpara de lectura (véase la **Figura 4**). Una cámara web capturaba el rostro del visitante y tomaba una foto para ser analizada por un programa escrito en **Python**, que concluía género, etnia, edad y estado emocional de los usuarios. Posteriormente, la pieza mostraba en una proyección la imagen capturada, en la cual los píxeles perdían su orden poco a poco, junto con los resultados del reconocimiento facial y una porción de texto generado por una red neuronal alimentada con los resultados del análisis. En este texto se reconstruía cada vez un fragmento del libro de Huxley y era simultáneamente leído por una voz artificial.

Las diversas capas que se interrelacionan en el sistema de la pieza contraponen la visión actual, forense y política del ADN como conductor de información biológica, con la visión distópica de Huxley, siendo el usuario el objeto de escrutinio para cuestionar los sesgos y cuantificación de la identidad en la era digital. Si bien el reconocimiento de rostro tuvo

Python

Python es un lenguaje de programación de alto nivel, diseñado para ser legible y cercano al lenguaje natural; su versatilidad lo ha convertido en un estándar en web, automatización, IA y ciencia de datos.

Cómputo biológico

Alude al uso de biotecnologías para almacenar información y realizar procesos computacionales a través de estructuras orgánicas como el ADN, a partir de ciertas condiciones (o entornos).



Figura 4. Montaje de *Interfaces de la identidad* en el Centro de Cultura Digital en CDMX. Izq. dibujos de iris y huella digital de la artista (datos biométricos) hechos con bacterias que replican su ADN. Imagen facilitada por la artista.

sus inicios en las décadas de 1960 y 1970, esta área de investigación se consolidó en 1988, con base en la propuesta de Sirovich y Kirby sobre el método de análisis de componentes principales (PCA).⁵ Sobre decir que aunque muchas de las implementaciones para el reconocimiento de rostro sirven para generar avances en educación y salud, en otros casos implican una vigilancia ubicua, usos comerciales como publicidad e incluso el robo de identidad a través de

Deep fakes ▶ *deep fakes.*

Son videos o clips de audio generados por inteligencia artificial que hacen parecer que alguien está diciendo o haciendo algo que nunca hizo; pueden utilizarse para difamar a personas y cometer fraudes.

Además del reconocimiento de emociones y rasgos personales utilizados para generar una corta narración, la pieza emula el crecimiento bacteriano portador de ADN, utilizando los píxeles de la imagen capturada para crear una simulación en la que los colores funcionan como colonias o agentes que hacen una conversión de aminoácidos (A, C, G, T)⁶ a colo-

res (R, G, B, A)⁷. Para la reescritura de una porción del libro *Un mundo feliz*, se utilizó una red neuronal en lenguaje Python y la biblioteca Keras, que fue entrenada durante aproximadamente tres horas con el texto completo del libro. Las redes neuronales recurrentes (RNN) utilizan datos secuenciales (en este caso enunciados), y se distinguen por su “memoria”, ya que obtienen información de entradas anteriores para influir en la entrada y salida actuales. En la pieza existía cierta probabilidad de que una palabra estuviera seguida por otras en particular de acuerdo al entrenamiento. Si bien los enunciados producidos no eran idénticos a los del libro —quizás ni siquiera enteramente legibles en la gramática correcta del español—, la lectura del público los dotaba a menudo de una interpretación que los visitantes tomaban como personal.

Para la lectura realizada por una voz artificial, se entrenó una red neuronal en Python, usando la biblioteca Tacotron2. Grabé y transcribí aproximadamente dos horas y media de audio con mi propia voz, leyendo frases recomendadas para entrenamiento


⁵ Es un tipo de algoritmo para biometría. La técnica estadística utiliza una transformación ortogonal para convertir un conjunto de observaciones de variables posiblemente correlacionadas en un conjunto de valores de variables linealmente no correlacionadas.

⁶ Adenina, citocina, guanina y timina, las cuatro bases nitrogenadas que conforman el ADN.

⁷ Rojo, verde, azul y alfa son los canales de color y transparencia que determinan la apariencia de cada pixel en las pantallas.

de redes neuronales en español, que abarcaran la mayor cantidad de sonidos posibles. La red se entrenó durante 15 horas y sonoramente el resultado fue bastante similar a mis lecturas originales. El *texto a voz* (text-to-speech: TTS, según sus siglas en inglés) es una técnica que consiste en generar un habla que suene natural a partir de una entrada de texto. El motor TTS convierte el texto escrito en una representación fonética y luego convierte ésta en ondas que pueden emitirse como sonido.

Nota final

 Si bien las redes neuronales se han popularizado y su uso extendido en los últimos años gracias a herramientas como Google Colab o la rápida evolución de los procesadores gráficos (GPU, por sus siglas en inglés), existen otras técnicas y métodos en el ámbito de la inteligencia artificial que vale la pena explorar, en particular desde el punto de vista del arte, como herramientas en la creación de piezas que cuestionan el papel de la tecnología en la realidad actual. Debido a sus características generativas, la mezcla de disciplinas, técnicas y narrativas, estos

cruces han promovido, como es mi caso, el desarrollo de piezas e instalaciones que funcionan como sistemas complejos y multidimensionales tanto técnica como conceptualmente.

Anni Garza Lau

Artista visual, Centro Nacional de las Artes.
contacto@annigarzalau.com

Lecturas recomendadas

Canazas, A. P., J. J. Ramos Blaz, P. D. Torres Martínez y X. Jaquehua Mamani (2022), “Sistema de identificación de emociones a través de reconocimiento facial utilizando inteligencia artificial”, *Revista Innovación y Software*, 3(2):140-150.

Jasso, K. y D. Garza Usabiaga (eds.) (2020), *(Ready) media: Hacia una arqueología de los medios y la invención en México*, México, Laboratorio Arte Alameda/Instituto Nacional de Bellas Artes.

Referencias específicas

Crawford, C. (2004), *On Interactive Storytelling*, EUA, New Riders.

Picard, R. (1997), *Affective Computing*, EUA, MIT Press.