

THE CONVERSATION

Academic rigour, journalistic flair



Shutterstock / Martin of Sweden

ChatGPT escribe un poema al estilo de Lorca: el resultado muestra sus debilidades

Published: February 6, 2023 6.39pm GMT

Jorge Franganillo

Profesor de la Facultad de Información y Medios Audiovisuales, Universitat de Barcelona

Javier Guallar

Profesor de la Facultad de Información y Medios Audiovisuales, Universitat de Barcelona

El siguiente poema es obra de *ChatGPT*. Le pedimos que lo escribiera con el estilo de Federico García Lorca, en lengua española y con rima consonante. Y *ChatGPT* no se negó. La inteligencia artificial acepta cualquier desafío. El resultado, sin embargo, muestra las muchas debilidades que aún tienen las máquinas cuando se ponen a crear:

*La ansiedad se adueña de mi ser
al ver que el autobús se aleja sin mí.
Un destino que no puedo alcanzar,
una oportunidad que se desvanece.*

*El tiempo se escurre entre mis dedos
como el agua de un río impetuoso
y me siento atrapado en un laberinto
donde el futuro se desdibuja.*

*La tristeza me invade en la soledad
que me ha dejado ese vehículo.
La esperanza se desvanece, y me sumerjo
en el abismo de la desesperación.*

*La lucha es inútil, el autobús se ha ido
y me quedo atrás, con un destino incierto,
una oportunidad perdida,
una tragedia que me atormenta.*

Aunque no es más que una burda aproximación, sin trazas de Lorca, el resultado no deja de ser sorprendente. Asusta lo que la herramienta puede llegar a hacer.

Parece magia que un sistema de inteligencia artificial sea capaz de aprender y luego de aplicar lo aprendido en aquello que se le pida. En efecto, *ChatGPT* es un sistema de chat basado en el modelo lingüístico GPT-3, que se caracteriza por su capacidad para generar texto coherente y natural. Escribe poesía, cartas, resúmenes, diálogos, código fuente... El poema de Lorca exigía más de lo que *ChatGPT* es capaz de hacer por ahora, pero otros encargos los realiza con razonable solvencia.

***ChatGPT* y la autoría no humana**

Desde su lanzamiento, a finales del pasado noviembre, *ChatGPT* ha conectado con el gran público porque es eficiente y fácil de usar. Basta con hacerle preguntas o darle indicaciones como se haría en una conversación normal. El *chatbot* ofrece una respuesta razonable, casi sobre cualquier disciplina, a partir del conocimiento depositado en Internet. Resuelve casi cualquier problema y redacta la solución como si fuera una persona más o menos experta.

 Respuesta de ChatGPT a una consulta formulada en lenguaje natural.

Respuesta de *ChatGPT* a una consulta formulada en lenguaje natural. Author provided

A primera vista resulta difícil distinguir un texto artificial de uno humano. Esto abre la puerta a nuevos desafíos éticos: la desinformación, la proliferación de contenido de baja calidad y la pérdida de confianza en las comunicaciones escritas. Y desata cuestiones, que van más allá de lo filosófico, sobre la autoría no humana y los derechos de autor. ¿A quién hay que atribuir el mérito de un artículo, o de un poema, impulsado por una inteligencia artificial?

En lo educativo, la irrupción de esta herramienta ha hecho saltar las alarmas de escuelas y universidades, a las que ha cogido desprevenidas. Las posibilidades que abre entrañan enormes riesgos. El principal temor es el plagio. Será difícil saber si un escrito es de creación propia o si, por el contrario, es el producto de una inteligencia artificial. Los trabajos académicos quedan así en entredicho, y ciertas voces vaticinan cambios en la manera de evaluar al estudiante.

Asombra, pero no hace magia

ChatGPT resulta versátil y produce textos de aspecto natural, con una sorprendente fluidez narrativa y de conversación, pero presenta limitaciones. Una de ellas, que el propio *chatbot* admite, es que sus datos acaban en 2021 y, como no tiene acceso a Internet, carece de información actualizada. Es una barrera hoy insalvable en una rápida comparación con cualquier motor de búsqueda.

Por otra parte, no basta con que el sistema sea competente, sino que además debe actuar con ética. Un modelo generativo puede reproducir estereotipos culturales, étnicos o de género que están en los datos con los que se ha entrenado. Pero el sistema, si se calibra y se reentrena, evita perpetuar sesgos perjudiciales y mejora así la calidad de los resultados.

Otra limitación de *ChatGPT* es que no siempre se apoya en fuentes fiables o evidencias robustas. Es esto, y no la aparente coherencia del texto, lo que determina la veracidad o la solidez de un argumento. El rigor no está entonces garantizado, y la información generada se debe contrastar con fuentes fiables y personas expertas en la materia.

Démosle cierto margen de error

La inteligencia artificial es válida entonces para situaciones que admiten cierto margen de error, incluso algún disparate. Pero no lo es para cuestiones críticas, tales como un trabajo científico, una asesoría legal o financiera, o una consulta médica. Produce una engañosa ilusión de pensamiento racional, pero no razona ni dispone de conocimiento fiable sobre el mundo. No entiende, en un sentido humano, nada de lo que escribe.

Como *ChatGPT* se ha programado para ser, ante todo, conversacional, no es necesariamente veraz. Las respuestas, aunque elocuentes y hasta convincentes, son a veces incorrectas o absurdas, porque se inventa hechos, personas, datos o fuentes. Nos puede enredar con el aplomo de un mentiroso compulsivo. Y ni siquiera se puede afirmar que mienta, ya que no cuenta con un modelo de verdad.

 Los mejores recetarios en español para cocinar con insectos, según *ChatGPT*. Nadie se moleste en buscarlos; estos libros no existen.

Los mejores recetarios en español para cocinar con insectos, según *ChatGPT*. Nadie se moleste en buscarlos; estos libros no existen. Author provided

En el ámbito literario, el modelo GPT-3 es capaz de simular el estilo de un autor. Capta la esencia del estilo, el léxico y hasta la “atmósfera” de los textos. Pero debe nutrirse de un *corpus* lo bastante amplio y entrenarse en tareas concretas. Para el poema que abre este artículo, le pedimos a *ChatGPT* que imitase el estilo emotivo y lírico de Federico García Lorca, característico por su lenguaje sensorial y poético. Le exigimos, además, rimar los versos en consonante. Pero también aquí erró el tiro por completo.

Tiene explicación: rimar versos es un proceso complejo que requiere comprender a fondo las reglas fonéticas y gramaticales del idioma y la relación entre palabras. Y aunque *ChatGPT* está programado para manejar varios idiomas, no se entrenó para generar rimas en español.

Rumbo al futuro

La tecnología de modelos lingüísticos avanza inexorable, y todo apunta a que GPT-4, la próxima generación de GPT, mejorará la capacidad para comprender el lenguaje y para generar contenido variado. A pesar de sus flaquezas, a lo largo de los próximos meses observaremos una adopción generalizada de estos sistemas, que serán un negocio muy lucrativo. El contenido automático será cada vez más habitual.

De hecho, se están haciendo esfuerzos para favorecer su integración en distintos entornos. Es un ejemplo la interfaz de búsqueda *Perplexity AI*. Aún en fase embrionaria, combina GPT-3 con *Bing* para resolver consultas cuyas respuestas se vinculan con las respectivas fuentes de información. *ChatGPT*, por su parte, llegará pronto al servicio *Azure* de Microsoft. Y hay indicios de que Google, el gigante de las búsquedas, también moverá pieza.

La naturaleza sintética de los textos artificiales, según se ha visto, es difícil de detectar. Y lo será cada vez más, por lo que cobrarán importancia las herramientas para identificarlos. Existen desarrollos incipientes para distinguir los textos humanos de aquellos escritos por modelos generativos. Pero su eficacia es relativa y deben combinarse con el sentido común y con cierta pedagogía para abordar esos escritos con pensamiento crítico.