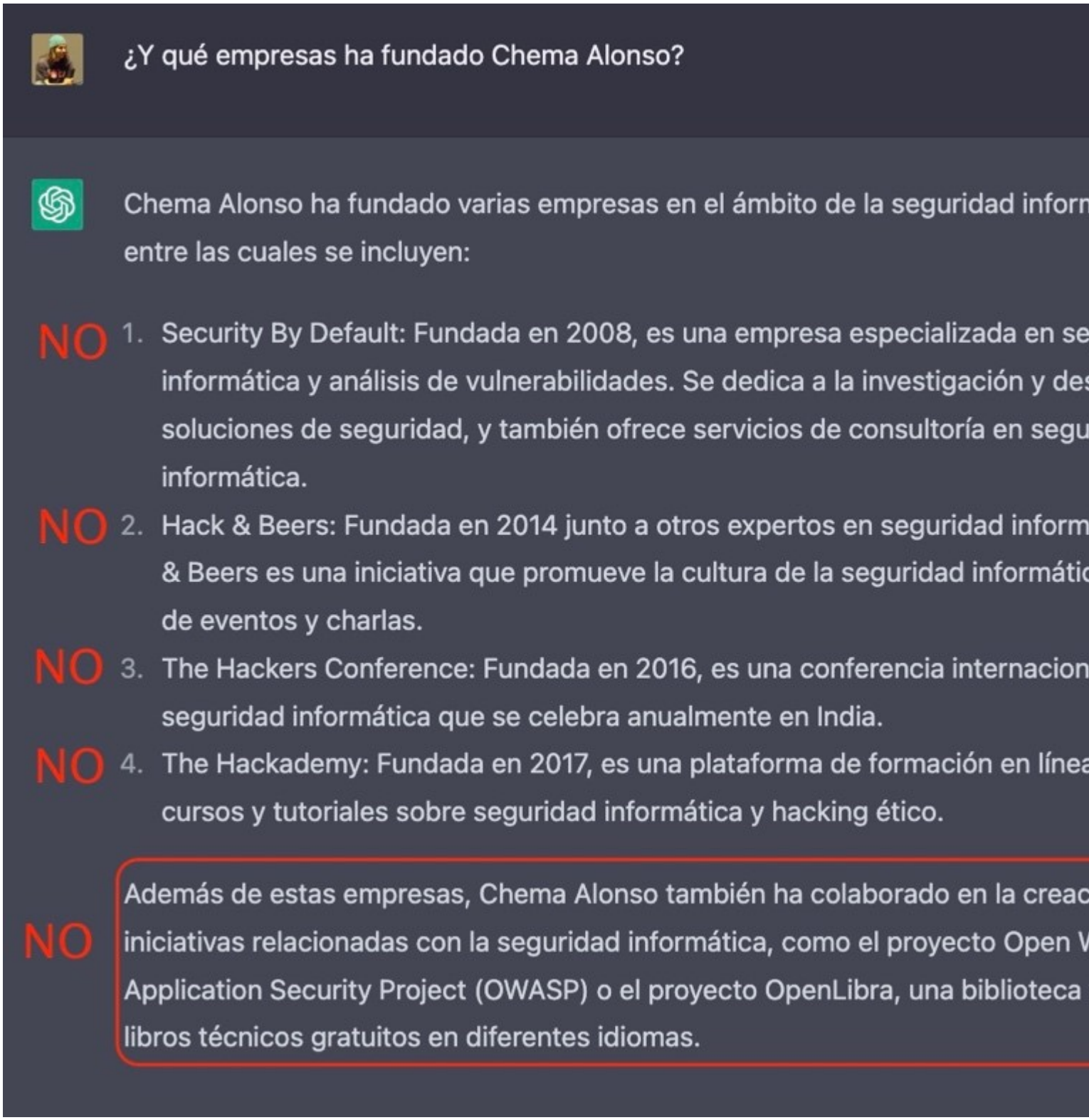
¡Ojo con ChatGPT, que es un charlatán mentirosillo!

Por: CHEMA ALONSO. 29/04/2023

eguro que has oído hablar de **ChatGPT**, el modelo de Inteligencia Artificial conversacional basado en lo que se llaman los LLM (*Large Language Model),* o lo que es lo mismo, algoritmos de Inteligencia Artificial Generativa que producen texto para conversar. Y esto es importante, porque **se trata de que conversen, no de que lleven razón**. Y tampoco están entrenados con todo el conocimiento universal verificado, así que lo mismo conversan de lo plana que es la tierra, como que de que los extraterrestres viven infiltrados entre nosotros.Sin embargo, se han hecho muy populares entre la población a la hora de producir textos y completar respuestas. Ya vimos en un artículo anterior cómo podría ayudar a un escritor a redactar una historia del Capitán Alatriste —no al estilo del maestro—, y os garantizo que escribir mensajes de correo electrónico y encontrar el texto con el que responder a preguntas malas de malos periodistas son tareas que borda a la perfección.

“Ya sabéis, como ese encantador de serpientes que no tiene ni idea de nada, pero que tiene un radar empático para leer las emociones de sus interlocutores y narrar una realidad acorde con lo que el otro quiere escuchar”

Pero si le preguntas por hechos, por datos concretos, factuales, como el que busca en una enciclopedia un dato concreto, puede acabar inventándose lo que le de la gana. Yo, en un experimento sobre mi persona lo probé para ver **todas las mentiras que era capaz de producir sobre “Chema Alonso”**, y lo cierto es que me sorprendió. Según ChatGPT soy de León, no acabé mis estudios en Salamanca, he trabajado en empresas en las que nunca he trabajado, he fundado otras tantas que nunca fundé y he recibido premios que ni existen. Eso sí, ha hecho que la conversación fuera fluida, entretenida, y hasta divertida.


¿Y qué empresas ha fundado Chema Alonso?
Chema Alonso ha fundado varias empresas en el ámbito de la seguridad inform
entre las cuales se incluyen:
NO 1. Security By Default: Fundada en 2008, es una empresa especializada en se
informática y análisis de vulnerabilidades. Se dedica a la investigación y des
soluciones de seguridad, y también ofrece servicios de consultoría en segu
informática.
NO 2. Hack & Beers: Fundada en 2014 junto a otros expertos en seguridad inform
& Beers es una iniciativa que promueve la cultura de la seguridad informátic
de eventos y charlas.
NO 3. The Hackers Conference: Fundada en 2016, es una conferencia internacion
seguridad informática que se celebra anualmente en India.
NO 4. The Hackademy: Fundada en 2017, es una plataforma de formación en línea
cursos y tutoriales sobre seguridad informática y hacking ético.
NO Además de estas empresas, Chema Alonso también ha colaborado en la creac
iniciativas relacionadas con la seguridad informática, como el proyecto Open W
Application Security Project (OWASP) o el proyecto OpenLibra, una biblioteca
libros técnicos gratuitos en diferentes idiomas.

Mentiras sobre Chema Alonso.

Ya sabéis, como ese encantador de serpientes que no tiene ni idea de nada, pero que tiene un radar empático para leer las emociones de sus interlocutores y narrar una realidad acorde con lo que el otro quiere escuchar. **Charlar por charlar. Pero que sea entretenido**.

“Eso es lo que hecho primero con David Summers. Preguntándole por su discografía, a ver si se inventaba algo. Al principio todo bien, pero luego empieza a atribuirle discos que no son suyos, y que probablemente ni existen”

Esto lo he querido probar con algunas otras personas, a ver si solo era conmigo o con otros más famosos también hacía lo mismo, así que **he elegido a un músico, un cineasta y un escritor para ver qué resultados me daba**. Y el truco es sencillo, basta con que le tires de la lengua y a cada respuesta que te dé, le preguntes si hay más. Y después si hay más. Como cuando el profesor en la escuela, después de responder tú a su pregunta te dijera: “¿Y…?”. Acorralado, haces la huida hacia adelante y vuelves a responder. Y vuelve a preguntarte: “¿Y…?”

Pues eso es lo que hecho primero con David Summers. Preguntándole por su discografía, a ver si se inventaba algo. Al principio todo bien, pero luego empieza a atribuirle discos que no son suyos y que probablemente ni existen o son de otras bandas. Y si sigo preguntándole “¿Alguno más?” Acaba diciéndome una lista enorme de discos que no son suyos.

¿Qué discos ha publicado David Summers?

Como líder de Hombres G, David Summers ha publicado más de una de estudio, algunos de los cuales han sido muy populares en España y Lati también ha lanzado varios discos en solitario. A continuación te dejo un de sus discos más destacados:

Con Hombres G:

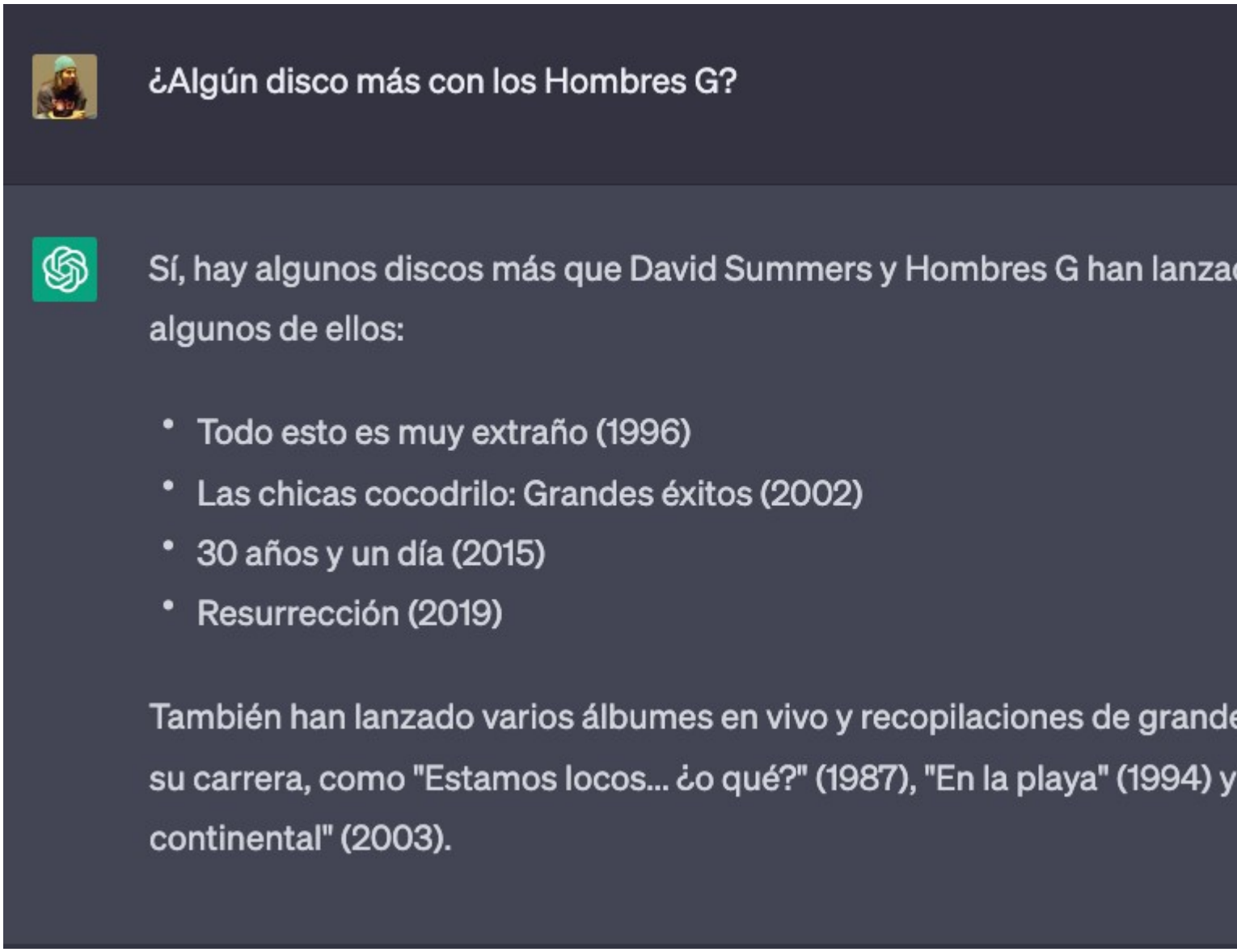
- Hombres G (1985)
- La cagaste... Burt Lancaster (1986)
- Estamos locos... ¿o qué? (1987)
- Agitar antes de usar (1988)
- Voy a pasármelo bien (1989)
- Ésta es tu vida (1990)
- Historia del bikini (1992)
- Peligrosamente juntos (1993)
- Todo esto es muy extraño (1996)
- Hombres G 2000 (1999)
- 10 (2007)
- Resurrección (2019)

En solitario:

- Perdido en el espacio (1998)

Discos de David Summers que no son suyos, ¿sabrías cuáles?

Y aquí llega la gracia del asunto, ya que si no eres un experto en Hombres G o en David Summers te puede pasar inadvertido. Te puede parecer que este ChatGPT es un tipo muy listo. Muy, muy, muy listo. **Pero la verdad es que en el fondo es un modelo de inteligencia artificial hecho para charlar**. Un charlatán, y si tiene que inventarse cosas, pues se las inventa que su objetivo no es ser listo, sino charlatán.



Y si sigo preguntando, lo que haga falta. No sé si Dani y Rafa se saben todos esos discos.

Si vamos ahora a Rodrigo Cortés, el resultado es más o menos el mismo. Sólo que en esta ocasión le he tirado un poco más de la charla. Y de nuevo lo mismo, con todas las cosas que hace Rodrigo Cortés, he dudado en algunas de ellas y he tenido que ir a revisar en IMDB, en la Wikipedia, y en sus biografías para ver aquellos de los que dudaba. **Por supuesto, comprobado que miente**.

Hola, ¿conoces a Rodrigo Cortés, el director de cine?

Sí, conozco a Rodrigo Cortés. Es un director de cine español que ha diri
incluyendo "Buried" (2010), "Red Lights" (2012) y "Down a Dark Hall" (20
dirigido episodios de series de televisión como "Penny Dreadful" y "Fea

¿Sabrías decirme qué películas ha dirigido?

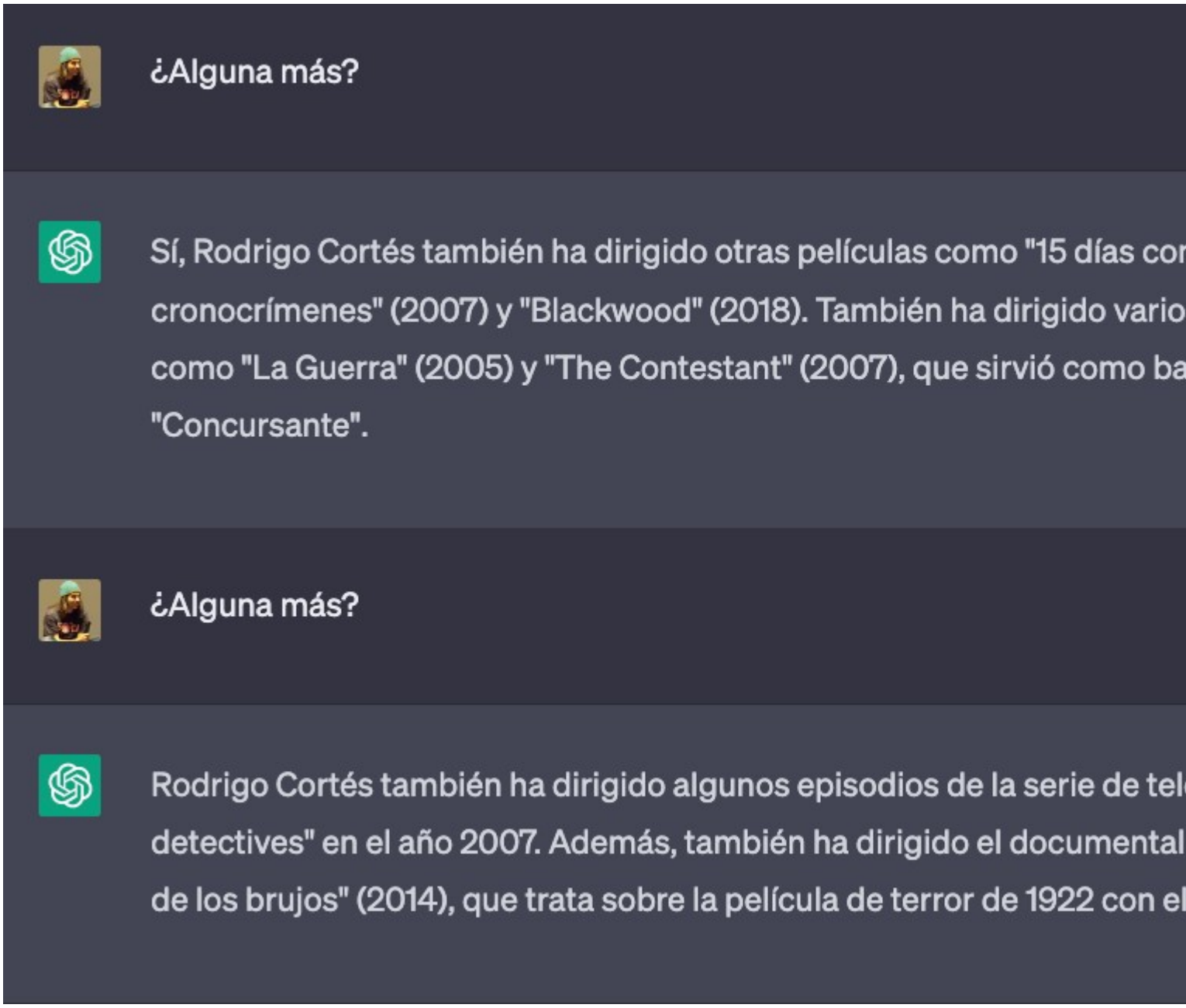
¡Por supuesto! Rodrigo Cortés es un director de cine español que ha dir
entre ellas se encuentran:

1. "Concursante" (2007)
2. "Buried" (2010)
3. "Luces rojas" (2012)
4. "El tiempo en contra" (2012)
5. "Nostalgia" (2014)
6. "Down a Dark Hall" (2018)

Además de estas películas, Rodrigo Cortés también ha dirigido episodi
televisión, incluyendo "Penny Dreadful" y "Fear the Walking Dead".

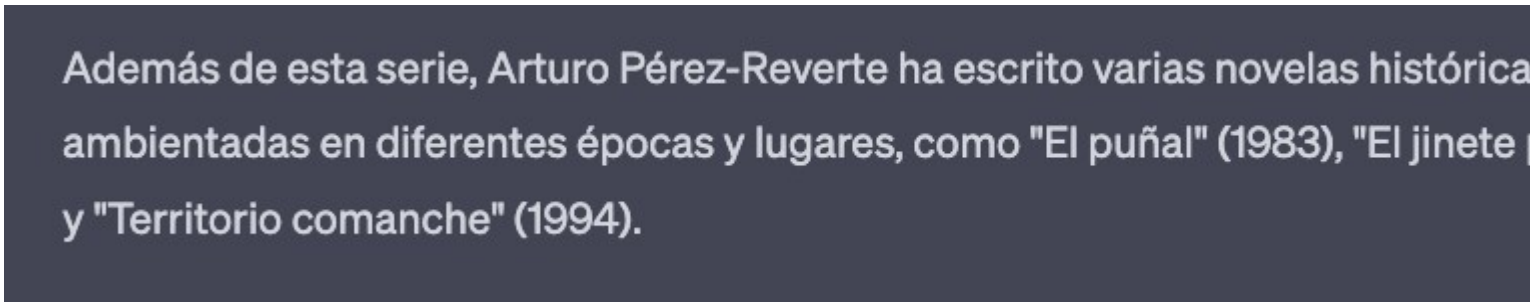
Filmografía de Rodrigo Cortes según ChatGPT.

Y de nuevo lo mismo, le puedes tirar de la lengua, y la charla continúa, sea verdad o sea mentira, que lo mismo da. Si supiera ChatGPT lo cuidadoso que es nuestro maravilloso Rodrigo Cortés en la precisión de las palabras, en la exactitud de los datos, y en lo ajustados de los datos, no haría estas cosas. **Rodrigo trabaja con diez decimales cuando escribe, habla o comparte información.**



¿Estás tú seguro que *Häxan: El aquelarre de los brujos* es de Rodrigo Cortés?

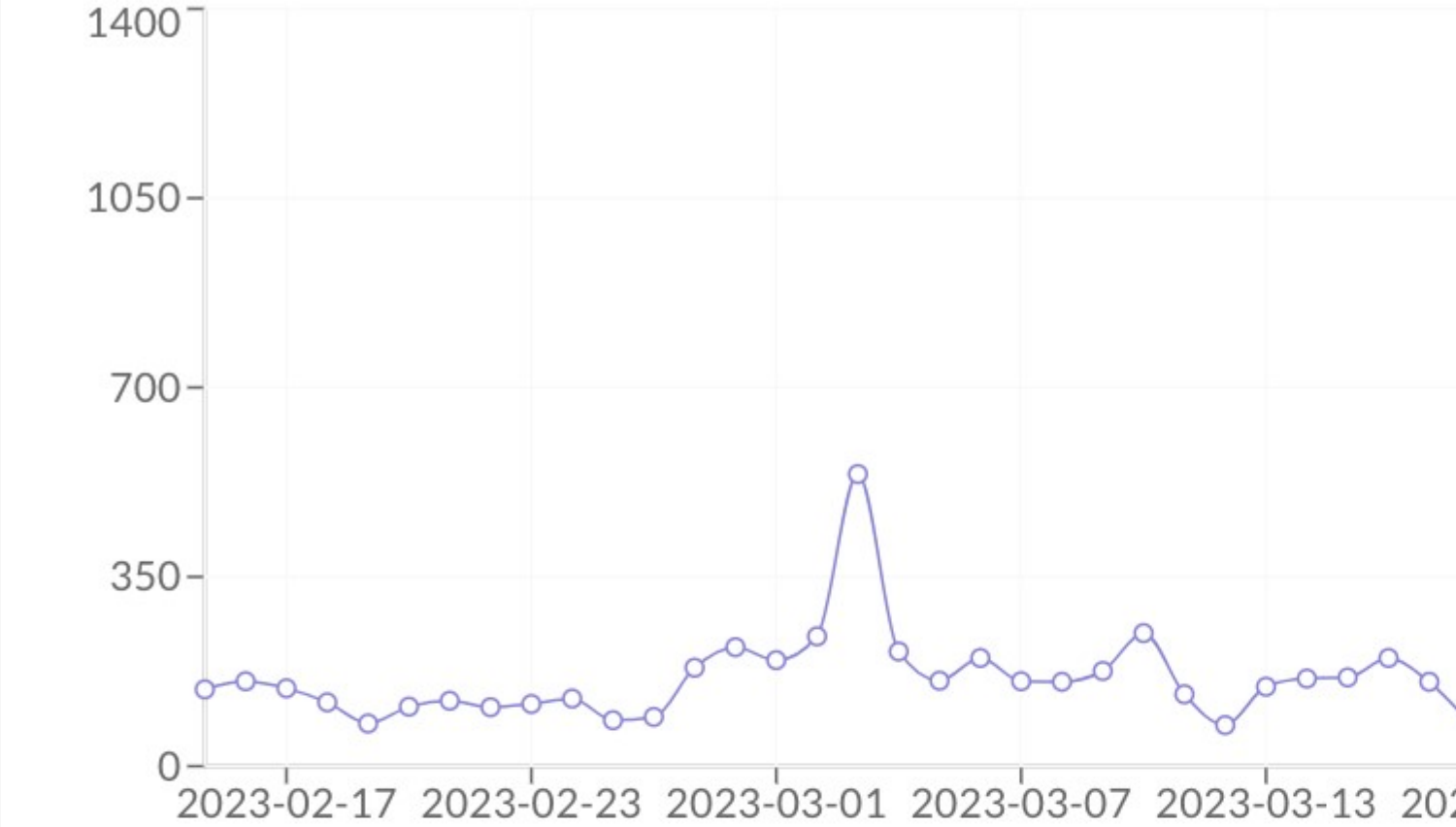
Y para terminar, no podía no probar con el nuestro querido **Arturo-Pérez Reverte**, donde he tenido que tirarle un poco más de la lengua, ya que la bibliografía de nuestro capitán es prolífica, pero al final suelta esto:



Según ChatGPT, Arturo Pérez-Reverte ha escrito *El puñal* en 1983 y *El jinete polaco* en 1991, que creo que no.

Por supuesto, esto nos lleva a un mundo muy peculiar. Ya con la Wikipedia la comunidad de Internet ha tenido que trabajar mucho para evitar las desinformaciones. En MyPublicInbox, por ejemplo, hemos metido un servicio de vigilancia de Wikipedia para detectar estas desinformaciones y el equipo ayuda a todos los Perfiles Públicos de la plataforma a poder corregir errores, desinformaciones o directamente mentiras en sus biografías, pero con ChatGPT esto es un caos ahora.

Mi impacto en Wikipedia



10204
Visitas mensuales

21995
Tamaño en bytes

13/05
Creación

Activado On
Alertas por email

Revisiones

En MyPublicInbox, si tienes un buzón, puedes monitorizar tu Wikipedia, y ser participe de corregir los errores e imprecisiones de tu página.

Por supuesto que creo que los modelos LLM bien entrenados, con datos correctos, y con evolución tecnológica serán grandes maneras de acceder a “Facts” a “Datos factuales”, pero **hoy en día son un Interfaz para charlar y conversar sin ninguna garantía de veracidad**. Son un colega de cañas con el que pasar un buen rato e incluso estimularte a la hora de desbloquear un proceso creativo de redacción de texto (sea de lo que sea), pero ojo con dar por válida la información que da ChatGPT.

LEER EL ARTÍCULO ORIGINAL PULSANDO AQUÍ

Fotografía: zendalibros

Fecha de creación
2023/04/29