
Nadie ha descifrado el manuscrito Voynich

Autor:

Data de publicació: 17-05-2019

El Manuscrit de Voynich decodificat: el llibre més misteriós del món desxifrat al cap de 600 anys ha afirmat un investigador britànic, El codi del manuscrit de Voynich, el misteriós volum de dibuixos fantàstics i missatges secrets, finalment s'ha pogut decodificar, uns 600 anys després de que fós escrit.

<https://www.youtube.com/watch?v=19vTeK1dYNI>

Ha sido una de las noticias curiosas de la semana: una inteligencia artificial comienza a descifrar uno de los libros más misteriosos de la historia. Pero no. En realidad, si nos fijamos en los detalles de la historia, lo único que ha demostrado esta inteligencia artificial es que frente al manuscrito Voynich está tan perdida como nosotros.

Recapitulemos: entre 1404 y 1438, alguien en algún lugar empezó a escribir un libro en un idioma desconocido y con un alfabeto que nadie había visto nunca. Lleno de ilustraciones de plantas desconocidas, símbolos astrológicos y mujeres desnudas, descifrar “el libro que nadie ha podido leer” hubiera sido un paso gigantesco en la intersección entre criptografía y la inteligencia artificial. Una pena que no haya sido así.

Un mundo lleno de aficionados

Desde que se redescubrió el manuscrito en 1912, los intentos de descifrarlo han sido continuos. Su popularidad ha hecho que muchos profesionales y muchos aficionados sin conocimientos especializados se lancen a la caza de una solución. Kondrak y Hauer, los protagonistas de esta semana, parecen solo un caso más de ellos. He de reconocer que yo había descartado directamente el tema, pero tras leer algunos comentarios, me he decidido volver sobre el artículo original.

Su tesis, que tampoco es original, es que se puede reconstruir el manuscrito para, por pura fuerza bruta, encontrar conexiones con otros textos. Cosas como la frecuencia o la combinación de letras podrían darnos pistas sobre el idioma en el que está escrito.

Y es que, desde los años 60, sabemos que se trata de una lengua natural (o de un código relacionado con una lengua natural) porque cumple la Ley de Zipf, una regularidad empírica que solo se da en las lenguas naturales y que describe la frecuencia aparición de las palabras. Los lenguajes inventados (sobre todo, los lenguajes inventados en el siglo XV)

no la cumplen.

Se han propuesto muchísimos idiomas: árabe, vasco, romaní o, incluso, azteca. Por eso, Kondrak y Hauer decidieron comparar el texto del manuscrito con la Declaración de Derechos Humanos y sus 380 traducciones. Ese fue su primer error. No porque no funcionara (que lo hizo); sino porque el documento se escribió hace más de 500 años. Compararlo con lenguas actuales es, cuanto menos, arriesgado.

Hacerse trampas al solitario

Pero, como digo, su metodología llegó a la conclusión de que el manuscrito está escrito en hebreo. ¿Cómo es posible? Asumiendo ciertas ideas que, desde luego, no están justificadas. Por ejemplo, Kondrak y Hauer recogieron una idea que llevaba años dando vueltas entre los voynichólogos: las palabras son anagramas; es decir, palabras que resultan de la transposición de las letras de las palabras originales. Eso les daba una enorme capacidad de combinación en idiomas como el hebreo antiguo que no usaba vocales.

Esto es interesante, porque lo que dicen Kondrak y Hauer es que el hebreo es la lengua que más se parece (hasta un 80% de las palabras del Voynich puede reordenarse para ser palabras hebreas); pero el siguiente idioma era el malayo. Hebreo y malayo son dos idiomas muy diferentes cuyo único punto de conexión es que el malayo se puede escribir en una versión del alifato árabe (una que tampoco tiene vocales).

Pero han descifrado una frase, ¿no es cierto?

Sí. De hecho, eso es, sin lugar a dudas, lo que ha dado más fuerza a la idea es que han sido capaces de traducir la primera línea del libro. Sin embargo, también hay trampa: en el mismo trabajo explican como tuvieron que cambiar la traducción para que tuviera sentido dado que los primeros resultados no fueron “muy coherentes”.

Shlomo Argamon, un lingüista computacional del Instituto de Tecnología de Illinois, explicaba en Verge que se trata de un método poco riguroso que les da “una gran libertad para hacer este tipo de interpretación impresionista”. Demasiada, de hecho. El trabajo de Kondrak y Hauer es interesante por alguno de los métodos estadísticos que usan, pero su desconocimiento de cosas básicas sobre paleolingüística los han llevado a ver regularidades donde solo parece que hay malas decisiones de investigación.