

En busca del libro que queremos leer

José Gordon

¿Cómo saber si el libro que está en las manos de un buen editor interesará profundamente a los lectores? En el texto “Constelaciones de libros”, Gabriel Zaid subraya la dimensión del reto que enfrenta la oferta editorial:

Hay más libros en los cuales detenerse que estrellas en una noche en alta mar. En esa inmensidad, ¿cómo puede un lector encontrar su constelación personal, esos libros que mueven su vida a conversar con el universo? Y ¿cómo puede un libro, entre millones, encontrar sus lectores?

En la Universidad de Stony Brook en Nueva York, la profesora del Departamento de Computación, Yejin Choi, piensa que las herramientas de la ciencia pueden ayudar a resolver este problema. Dice la investigadora: “Predecir el éxito de los trabajos literarios plantea un dilema masivo tanto para los editores como para los escritores”. En este marco, dice Choi,

examinamos la conexión cuantitativa entre el estilo de escritura y la literatura que tiene éxito. Basados en novelas de diferentes géneros, investigamos el poder predictivo de la estilometría estadística para discernir las obras literarias que han logrado una buena recepción. Hemos identificado los elementos de estilo que son prominentes en esta escritura.

De acuerdo con estos estudios, publicados por la Asociación de Lingüística Computacional, si una novela pasa por el análisis de una máquina con los algoritmos adecuados, se puede predecir en un 84 por ciento de los casos si va a ser popular o no. Matthew Sparkes plantea que la “estilometría estadística” es una técnica que exa-

mina matemáticamente el uso de las palabras y la gramática. Los algoritmos que así se generan toman en cuenta diversas variables tales como: “interesabilidad” (así le llaman: “interestingness”), novedad, estilo de escritura y la forma con que nos atrapa la historia.

El campo de estudio se definió mediante los libros que se descargan en Internet en el llamado Proyecto Gutenberg, que ofrece de manera gratuita más de 42 mil libros. Algunos ejemplos de los más descargados: *Hojas de hierba*, de Walt Whitman; *Los miserables*, de Victor Hugo; *La metamorfosis*, de Kafka; *Ulises*, de James Joyce y *Los hermanos Karamazov*, de Dostoievsky.

Así, se analizaron 800 libros que forman parte del catálogo del Proyecto Gutenberg y también se incluyeron algunos de los libros menos valorados en Amazon como *El símbolo perdido*, de Dan Brown. El algoritmo los filtró correctamente con el porcentaje de bateo ya referido: le atinó 84 por ciento de las veces.

El análisis del léxico, de la sintaxis y de los patrones discursivos que caracterizan los estilos de escritura de las novelas que resultaron ser las más descargadas reveló el uso frecuente de conjunciones tales como “y” y “pero”, así como un gran número de sustantivos y adjetivos. Los verbos suelen describir procesos de pensamiento tales como “reconoció” o “recordó”. En contraposición, las que fueron menos descargadas tienen un porcentaje mayor de adverbios y verbos que describen acciones y emociones tales como “quiso”, “tomó”, “prometió”, “lloró” o “aplaudió”. Utilizan más palabras que son clichés tales como: “amor”, “sin aliento” y “herido”.

Yejin Choi plantea que la literatura altamente calificada tiene una correlación

opuesta con la legibilidad, apela a cierto grado de dificultad en la lectura. La investigadora aclara que estos hallazgos, más que causas, implican correlaciones:

Nosotros conjeturamos que la complejidad de las obras literarias que han tenido mucho éxito (*creo que la palabra reconocimiento sería más adecuada*), requiere una complejidad sintáctica que va en contra de la legibilidad.

El secreto: evitar los lugares comunes. Por eso aunque la novela de Dan Brown aludida tuvo una buena venta comercial, fue mal valorada por los lectores.

¿Se podrán usar los algoritmos de la doctora Choi para evaluar lo que sería bueno publicar? No hay que apresurarnos tanto. Creo que hay factores culturales de narrativa, de prestigio literario y editorial que también juegan en la ecuación y que no fueron tomados en cuenta en estos estudios. Un dato interesante es que los investigadores reconocen que hay una variable oculta que aún no pueden definir. Le llaman suerte. Tal vez tiene que ver con el encuentro azaroso entre el libro y el lector. Eso ya nos lo había dicho Gabriel Zaid sin necesidad de algoritmos:

La adivinación, la suerte, tienen un papel decisivo. Uno se resiste a creer qué improbable es encontrar un libro en una librería o biblioteca: el que busca porque sabe que existe, el que busca sin saber si existe, el que ni siquiera sabe que busca, hasta que lo ve.

Hay lectores que están en la búsqueda milagrosa de su autor y editores que están en la búsqueda milagrosa que permita propiciar ese encuentro. **U**