

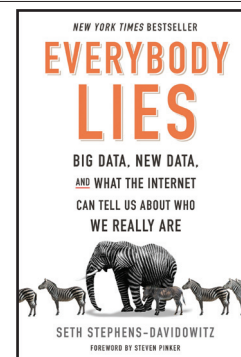
STEPHENS- DAVIDOWITZ, S. *Big Data, New Data, and What the Internet Can Tell Us about Who We Really Are.*

Nueva York: Dey Street Books, HarperCollins Publishers, 2017, 352 páginas.
ISBN 978-0-06-239085-1.

POR MANUEL MARTÍNEZ NICOLÁS

Profesor de la Universidad Rey Juan Carlos

manuel.martinez.nicolas@urjc.es



Pantagruel en el panóptico

Fue el inefable doctor House, aquel médico-detective protagonista de la serie homónima, quien popularizó la demoledora sentencia de que “todo el mundo miente”. Experto en diagnosticar enfermedades raras a partir de síntomas extraños o no usuales, la experiencia le había convencido de que lo último en que podía confiar era en lo que le contaban los propios pacientes, más preocupados –incluso en esas circunstancias– de proteger su reputación que de contribuir a esclarecer el origen de sus males. En cualquier caso, que todo el mundo mienta es algo ya hace tiempo asumido por las ciencias sociales, cuya fuente primordial de datos, las personas, exhiben una tendencia irrefrenable a contar lo que son, piensan, sienten y hacen pasándolo por el tamiz favorecedor de lo que les gustaría ser, pensar, sentir y hacer. Mienten, por tanto, y solo con muchas dificultades pueden los científicos sociales controlar estos fenómenos bien conocidos (desiderabilidad social, espiral del silencio, etc.) a la hora de determinar las actitudes, estados de opinión o comportamientos prevalentes en las sociedades humanas... Y en estas estábamos cuando llegó el Big Data.

Stephens-Davidowitz es doctor en Economía por Harvard, antiguo analista de datos para Google y actualmente articulista especializado en *The New York Times*. Dispone, por tanto, de una sólida formación académica, una amplia experiencia en el trato con los datos digitales y una indudable vocación divulgadora, cualidades que vierte en un trabajo en el que quiere mostrar el potencial del Big Data para la investigación social, celebrando el advenimiento de lo que no tiene reparo en anunciar como una “revolución”. El resultado es un libro vibrante y sugerente, casi siempre riguroso, y a veces excesivo.

Argumenta Stephens-Davidowitz en el arranque de esta celebración que buena parte de lo que sabemos sobre la gente es equivocado, porque lo que creemos saber de ella no tiene más respaldo empírico que aquello que nos cuenta de sí misma, y es bien conocido que todo el mundo miente. Pero en lo

sucesivo ya no necesitamos confiar en lo que la gente nos diga, pues el rastro que deja en buscadores de internet y en redes sociales opera como una suerte de “suero de la verdad digital” que nos revela, por fin, lo que la gente piensa, desea y hace efectivamente. Ese es el potencial de los datos digitales masivos del Big Data: indicadores veraces de lo que a la gente le pasa por la cabeza (casi un “cerebroscopio”, dice Steven Pinker en el prefacio). Datos que son especialmente útiles para sondear en aquellos comportamientos más susceptibles de verse sometidos a las nefastas intermediaciones psico-sociológicas (reputación, desiderabilidad, protección de la propia imagen, etc.) que llevan a la gente a mentir: actitudes discriminatorias (racismo, homofobia, machismo), conductas sexuales, problemas de salud, condiciones estigmatizadas (la enfermedad mental, por ejemplo) o prácticas socialmente reprobadas, ilegalizadas o directamente delictivas (dependencias, aborto, abuso infantil, acoso, etc.). Una fiesta pantagruélica de datos, en fin, que convertiría a estos Big Data en un auténtico panóptico al que poco o nada de lo que sucede en la mente humana podría sustraerse.

Pero ¿en qué consiste exactamente el potencial del Big Data para abrir ese “nuevo camino a la ciencia social en el siglo XXI” que vislumbra Pinker en el prefacio? Al final del capítulo 2 (páginas 53-54), Stephens-Davidowitz da una respuesta sintética a esta cuestión trazando el plan que desarrollará en la parte central del libro (“The powers of Big Data”), donde, *onus probandi*, aporta los resultados de las decenas de estudios que ha realizado con datos digitales, aunque con una fuerte querencia por lo que puede hacerse a partir de los algoritmos de Google.

De entrada, estos datos digitales son “datos honestos”: “En la era predigital, la gente ocultaba sus pensamientos embarazosos a los demás. En la era digital, todavía los oculta a otra gente, pero no a internet en sitios como Google o Pornhub, que protegen el anonimato. Estos sitios funcionan como una especie de suero de la verdad digital” (página 54). Y son “honestos”, sostiene Stephens-Davidowitz, no solo porque el

anonimato esté protegido, sino porque el usuario cuenta con un “incentivo” para no falsear el rastro digital que deja a su paso (búsquedas, visitas a sitios web, etc.): obtener satisfacción a sus demandas. Si estoy deprimido, es posible que no lo confiese en una encuesta, pero no tiene mucho sentido que busque en internet información al respecto si verdaderamente no lo estoy o no quiero ayudar a alguien que lo está. No parece tener vuelta de hoja.

Más sutil e iluminador es el argumento de que el potencial de estos nuevos datos digitales masivos no esté, contra toda expectativa, en que sean “masivos” (*big*), sino en que son “nuevos”. Son “un nuevo tipo de datos”, datos “no convencionales”; anecdóticos, incluso irrisorios, a primera vista, pero disponibles sin más restricción que la imaginación del analista para convertirlos en *indicadores sociales* relevantes. “La revolución del Big Data no va de recoger más y más datos. Va de recoger los datos correctos” (página 62). Los datos correctos (*the right data*), es decir, los datos significativos de entre la inextricable maraña digital disponible. El potencial de esta revolución que profetiza Stephens-Davidowitz (y ejemplifica con sus trabajos) depende de que se active en el científico social esa chispa imaginativa que examina las búsquedas de chistes sobre negros para trazar el mapa del prejuicio racial en Estados Unidos y ponerlo en relación con la distribución territorial del apoyo electoral a Donald Trump. O que indaga en las webs pornográficas para estimar el elevado porcentaje de homosexuales estadounidenses que todavía no se han atrevido a “salir del armario” en los estados más intolerantes. O que rastrea el interés sobre maneras de provocarse un aborto (píldoras, hierbas, sobredosis de vitamina C, etc.) para retar la credibilidad de las cifras oficiales al respecto. O que repara en la frecuencia y distribución de búsquedas del tipo “mi mamá/mi papá me pega” para reconstruir la situación del abuso infantil vinculándola a la de los índices de desempleo en determinados lugares y momentos.

Este es el sentido de la revolución que anuncia aquí Stephens-Davidowitz: elevar estas “evidencias” a la categoría de “datos” significativos para la investigación social. Una revolución cuyo advenimiento va a exigir, probablemente más que nunca, de ese intangible que no deja de rondar en la propuesta de Stephens-Davidowitz: la “imaginación científica”, que en este campo solo puede nutrirse de una rigurosa formación epistemológica que ayude a plantear preguntas pertinentes y maneras audaces de darles respuesta. No sorprende, pues, que sean estos *apóstoles* del algoritmo (desmesurados unos; comedidos otros) los que anden reclamando de nuevo (véase, también, Finn 2018) lo que ya pregonaba hace medio siglo largo Wright Mills (1961): hacer de la “imaginación sociológica” la única actitud válida para afrontar adecuadamente la complejidad del mundo, antes la del analógico, ahora la del digital.

Referencias

FINN, E. *La búsqueda del algoritmo. Imaginación en la era de la informática*. Barcelona: Alpha Decay, 2018 (publicación original: 2017).

WRIGHT MILLS, CH. *La imaginación sociológica*. México: Fondo de Cultura Económica, 1961 (publicación original: 1959).