



## Nate Silver

# La señal y el ruido

Cómo navegar por la maraña de datos que nos inunda, localizar los que son relevantes y utilizarlos para elaborar predicciones infalibles

## Índice

Portada	
Dedicatoria	
Introducción	
1. Un error de predicción catastrófico	
2. ¿Es usted más listo que los expertos de la tele?	
3. A mí lo único que me importa son las victorias y las derrotas	
4. Durante años nos han dicho que la lluvia era verde	
5. Buscando la señal desesperadamente	
6. Cómo ahogarse en un metro de agua	
7. Modelos de conducta	
8. Cada vez menos, menos y menos erróneo	
9. Rabia contra las máquinas	
10. La burbuja del póquer	
11. Si no puedes con ellos...	
12. Un clima de sano escepticismo	
13. Ojos que no ven, tortazo que te pegas	
Conclusión	
Agradecimientos	
Notas	
Créditos	

Te damos las gracias por adquirir este EBOOK

Visita [Planetadelibros.com](http://Planetadelibros.com) y descubre una nueva forma de disfrutar de la lectura

---

**¡Regístrate y accede a contenidos exclusivos!**

Próximos lanzamientos  
Clubs de lectura con autores  
Concursos y promociones  
Áreas temáticas  
Presentaciones de libros  
Noticias destacadas

---

Comparte tu opinión en la ficha del libro  
y en nuestras redes sociales:



Explora Descubre Comparte

*Para mamá y papá*

## INTRODUCCIÓN

Éste es un libro sobre la información, la tecnología y el progreso científico. También es un libro sobre la competencia, el libre mercado y la evolución de las ideas. Es un libro sobre las cosas que nos hacen más inteligentes que los ordenadores y también un libro sobre el error humano. Éste es un libro sobre cómo aprendemos, paso a paso, a conocer el mundo objetivo y por qué a veces damos pasos en falso.

Éste es un libro sobre predicciones y sobre cómo éstas se encuentran en la encrucijada de todas esas cosas. Es un estudio sobre por qué algunas predicciones se cumplen y otras no. Mi esperanza es que aprendamos a planificar un poco mejor nuestro futuro y a reducir las probabilidades de repetir nuestros errores.

### A MÁS INFORMACIÓN, MÁS PROBLEMAS

La revolución original en la tecnología de la información no llegó con el microchip, sino con la imprenta. El invento de Johannes Gutenberg en 1440 puso la información a disposición de las masas, y eso desencadenó una explosión de ideas de consecuencias involuntarias y efectos impredecibles. Aquella chispa se propagó y desembocó en la Revolución Industrial de 1775,<sup>1</sup> un punto de inflexión en el que nuestra civilización pasó de no haber logrado prácticamente ningún avance científico ni económico durante la mayor parte de su existencia a los niveles de crecimiento y cambio exponenciales a los que hoy estamos acostumbrados.

Aquel evento desencadenó los acontecimientos que propiciaron la Ilustración en Europa y la fundación de la República Americana.

Pero en un primer instante la imprenta tuvo otra consecuencia: cien años de guerra santa. El ser humano se creyó capaz de predecir su suerte, de elegir su destino, a lo que siguió la época más sangrienta de la historia de la humanidad.<sup>2</sup>

Los libros habían existido antes de Gutenberg, pero no se escribían ni se leían de forma generalizada. Al contrario, los libros eran objetos de lujo para el uso exclusivo de la nobleza, producidos de manera individual por los escribas.<sup>3</sup> La tarifa por reproducir un único manuscrito era de un florín (una moneda de oro equivalente a unos doscientos dólares actuales) cada cinco páginas,<sup>4</sup> de modo que un libro como el que el lector tiene entre las manos costaba unos veinte mil dólares. Además, muy probablemente estaría sembrado de errores de transcripción, pues sería una copia de una copia de una copia, y los errores habrían ido mutando y multiplicándose con cada nueva generación.

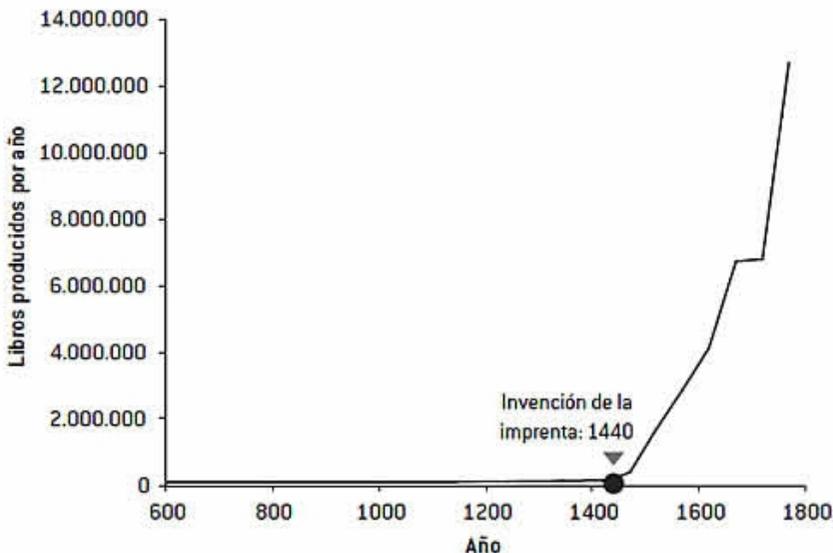
Eso dificultaba enormemente la acumulación de conocimiento. De hecho, evitar que el conocimiento almacenado se redujera requería un esfuerzo heroico, pues con frecuencia los libros se deterioraban a una velocidad mayor a la del escriba que debía reproducirlos. Sobrevivieron varias ediciones de la Biblia, junto con un reducido número de textos canónicos, como los de Platón y Aristóteles, pero a lo largo de los siglos se perdió una cantidad ingente de conocimiento,<sup>5</sup> al tiempo que los limitados incentivos para ponerlos por escrito se mantenían constantes.

La búsqueda del conocimiento se consideraba una tarea inherentemente fútil, cuando no directamente inútil. Si hoy en día experimentamos cierta sensación de transitoriedad debido a la rapidez con la que se suceden los cambios, esa sensación era un motivo de preocupación mucho más

literal para las generaciones que nos precedieron. No había «nada nuevo bajo el sol», por decirlo con un hermoso verso del Eclesiastés, no tanto porque ya se hubiera descubierto todo, sino porque todo acabaría olvidándose.<sup>6</sup>

La imprenta cambió todo eso de forma tan radical como definitiva. Casi de la noche a la mañana, el coste de producir un libro se redujo unas trescientas veces,<sup>7</sup> de modo que un libro que antes costaba veinte mil dólares actuales pasó a costar sólo setenta. La imprenta se propagó con gran rapidez por toda Europa: de la Alemania de Gutenberg llegó hasta Roma, Sevilla, París y Basilea en 1470, y a prácticamente todas las capitales europeas a lo largo de los siguientes diez años.<sup>8</sup> La producción de libros creció de forma exponencial y aumentó unas treinta veces durante el siglo posterior a la invención de la imprenta.<sup>9</sup> El conocimiento humano había empezado ya a acumularse a gran velocidad.

Figura I.1. PRODUCCIÓN DE LIBROS EN EUROPA



Sin embargo, y como también sucedió durante los primeros tiempos de internet, la calidad de la información no era ni mucho menos uniforme. Aunque la imprenta produjo resultados casi inmediatos en lo tocante a la producción de mapas de mayor calidad,<sup>10</sup> la lista de libros más vendidos pronto quedó copada por textos religiosos heréticos y pseudocientíficos.<sup>11</sup> De pronto los errores se podían producir en masa, como sucedió con la conocida como la «Biblia de los pecadores», que contenía la errata más desafortunada de la historia: «Cometerás adulterio».<sup>12</sup> Al mismo tiempo, el contacto con tantas nuevas ideas generó una confusión masiva. La cantidad de información crecía mucho más deprisa que la capacidad humana de decidir qué hacer con ella, o de discernir la información útil de las medias verdades.<sup>13</sup> Paradójicamente, el aumento del conocimiento compartido se tradujo en un aislamiento creciente entre las diversas realidades nacionales y religiosas. Nuestra reacción instintiva ante un «exceso de información» consiste en abordarla de forma selectiva. Así, elegimos las partes que nos gustan e ignoramos el resto, y convertimos en aliados a quienes han hecho las mismas elecciones que nosotros y en enemigos a los demás.

Los clientes más entusiastas de la primera época de la imprenta fueron quienes vieron en el invento un instrumento útil para perseguir sus objetivos evangelizadores. En realidad, *Las noventa y cinco tesis* de Martín Lutero no eran tan radicales; tesis similares se habían debatido un millar de veces antes. Lo verdaderamente revolucionario, tal como escribe Elizabeth Eisenstein, fue que las tesis de Lutero «no quedaron clavadas a la puerta de la iglesia».<sup>14</sup> La imprenta de Gutenberg produjo trescientas mil copias del documento,<sup>15</sup> todo un éxito incluso según los estándares modernos.

El cisma provocado por la Reforma Protestante de Martín Lutero pronto sumió a toda Europa en la guerra. De 1524 a 1648 tuvieron lugar la guerra de los campesinos ale-

manes, la guerra de Esmalcalda, la guerra de los Ochenta Años, la guerra de los Treinta Años, las guerras de Religión de Francia, las guerras confederadas de Irlanda, la Guerra Civil Escocesa y la Guerra Civil Inglesa, muchas de ellas simultáneamente. Tampoco debemos olvidar la Inquisición española, que empezó a actuar en 1480, ni la guerra de la Liga de Cambrai entre 1508 y 1516, aunque estos dos factores tuvieron menos que ver con la expansión del protestantismo. Sólo la guerra de los Treinta Años provocó la muerte de un tercio de la población de Alemania,<sup>16</sup> y el siglo xvii fue probablemente el más sangriento de la historia, un título que sólo puede disputarle la primera mitad del siglo xx.<sup>17</sup>

Y, sin embargo, la imprenta pronto impulsó también el progreso científico y literario. Galileo empezó a compartir sus ideas (censuradas) y Shakespeare a producir sus obras.

Muchas de las obras de Shakespeare giran en torno a la idea del destino, como suele suceder con tantas obras dramáticas. A menudo, la tragedia hay que buscarla en la distancia que separa lo que los personajes querrían lograr y lo que les depara el destino. Según parece, la idea de controlar el propio destino pasó a formar parte de la conciencia humana en la época de Shakespeare, pero no se puede decir lo mismo de las competencias necesarias para lograr dicho objetivo. Por ese motivo, quienes osaban desafiar al destino solían pagarlo con la vida.<sup>18</sup>

*La tragedia de Julio César* explora dichos temas en profundidad. A lo largo de la primera mitad de la obra, César recibe todo tipo de aparentes advertencias, que él llama profecías<sup>19</sup> («guárdate de los idus de marzo»). Dichas predicciones apuntan a que su coronación se convertirá en un baño de sangre. César, naturalmente, ignora esas señales, asegura orgulloso que se refieren a la muerte de otro e interpreta las pruebas de forma selectiva. Al final, César muere asesinado.

«Pero los hombres pueden interpretar las cosas a su manera, contrariamente al fin de las cosas», nos advierte Shakespeare por boca de Cicerón: un buen consejo para quien desee abrirse paso por entre la nueva sobreabundancia de información. Era difícil distinguir la señal del ruido. La historia que nos transmite la información suele coincidir con la que queremos oír y, por lo general, nos aseguramos de que tenga un final feliz.

Sin embargo, si bien *La tragedia de Julio César* abordaba la antigua idea de las profecías (y la asociaba al fatalismo, la adivinación y la superstición), también aludía a una idea más moderna y completamente radical: que podemos interpretar las señales para sacar provecho de ellas. «Los hombres son a veces dueños de su destino», dice Casio con la esperanza de convencer a Bruto para que participe en la conspiración contra César.

La idea según la cual el hombre era dueño de su destino estaba muy en boga. Actualmente usamos las palabras «predicción» y «previsión» casi de forma indistinta, pero en la época de Shakespeare tenían diferentes significados. Una predicción eran las palabras de un adivino, mientras que una previsión se acercaba más a la idea de Casio.

En inglés, el término *forecast* (más próximo al de «previsión») es una palabra de raíces germánicas,<sup>20</sup> mientras que la palabra *predict* deriva del latín.<sup>21</sup> Así, la «previsión» respondía al punto de vista terrenal del nuevo protestantismo, por oposición al punto de vista esotérico del Sacro Imperio Romano. Por lo general, prever algo implica planificar algo en condiciones de incertidumbre y apela a la prudencia, la sabiduría y la diligencia.<sup>22</sup>

Esta idea tiene una serie de complejas implicaciones teológicas,<sup>23</sup> aunque tal vez lo fueran menos para quienes pretendían labrarse una existencia más holgada en el mundo terrenal. Estas cualidades suelen asociarse con la ética de trabajo protestante, que al parecer de Max Weber facili-

tó el advenimiento del capitalismo y la Revolución Industrial.<sup>24</sup> La idea de «prever» las cosas iba íntimamente vinculada a la idea de progreso. Toda la información contenida en aquellos libros debía ayudarnos a planificar mejor nuestras vidas y predecir el curso del mundo en nuestro propio beneficio.

Los protestantes, que habían desatado varios siglos de guerra santa, estaban aprendiendo a utilizar el conocimiento acumulado para cambiar la sociedad. La Revolución Industrial se inició fundamentalmente en los países protestantes y sobre todo en los que gozaban de libertad de prensa, allí donde las ideas religiosas y científicas podían circular sin miedo a la censura.<sup>25</sup>

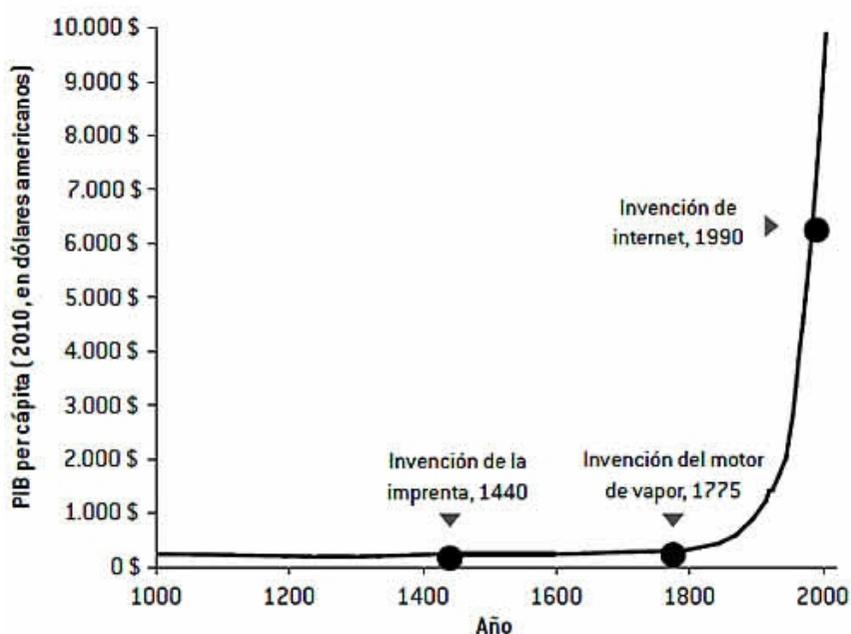
Difícilmente se podrá exagerar la importancia histórica de la Revolución Industrial. A lo largo de la historia de la humanidad, el crecimiento económico había sido de aproximadamente el 0,1 por ciento anual, lo que había permitido un aumento muy gradual de la población, pero no así de los estándares de vida per cápita.<sup>26</sup> Y entonces, de repente, empezó a haber progreso donde antes no lo había habido. El crecimiento económico se disparó mucho más que la tasa de crecimiento de la población, tal como ha sucedido hasta la fecha y a pesar de la crisis económica global.<sup>27</sup>

Así, pues, resulta evidente que la explosión de información provocada por la imprenta nos ha hecho mucho bien. Bastaron 330 años (y millones de muertos en los campos de batalla de toda Europa) para que dichas ventajas cristalizaran.

## LA PARADOJA DE LA PRODUCTIVIDAD

Cualquier incremento de la cantidad de información por encima de nuestra capacidad de procesarla supone un peligro. Los últimos cuarenta años de la historia de la humanidad parecen sugerir que aún tardaremos mucho tiempo en hallar la forma de transformar la información en conocimientos útiles, y que, si no tenemos cuidado, es posible incluso que demos un paso atrás en el intento.

Figura I.2. PIB GLOBAL PER CÁPITA, 1000-2010



El término «era de la información» no es particularmente nuevo. Su uso empezó a extenderse a finales de la década de 1970. Otro término emparentado con éste, «era de la informática», comenzó a utilizarse todavía antes, en 1970 aproximadamente.<sup>28</sup> En torno a esa época, los ordenadores se empezaron a usar de forma más generalizada en laboratorios y entornos académicos, si bien aún no se habían convertido en los aparatos domésticos que son actualmen-

te. En esta ocasión no fueron necesarios trescientos años para que la tecnología de la información empezara a producir beneficios tangibles para la sociedad humana, pero sí quince o veinte.

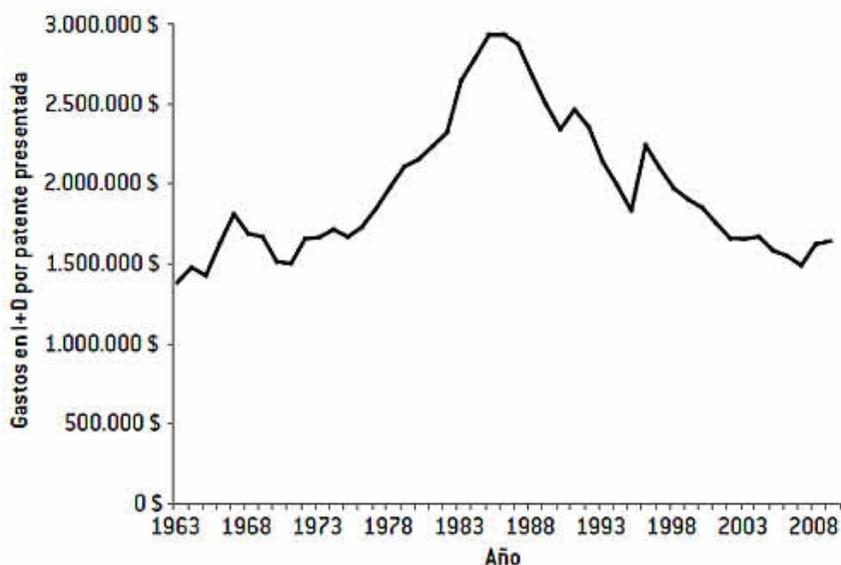
La década de 1970 marcó el punto álgido de la tendencia a «aplicar una cantidad ingente de teorías a una cantidad extremadamente pequeña de datos», tal como en una ocasión me dijo Paul Krugman. Habíamos empezado a utilizar los ordenadores para producir modelos del mundo, pero tardamos un tiempo en comprender que se trataba de máquinas toscas y anquilosadas en los supuestos, y que en realidad la precisión de los ordenadores no era un sustituto de las previsiones fiables. Los años setenta fueron una época en la que, en ámbitos tan dispares como la economía o la epidemiología, se realizaron todo tipo de predicciones, que a menudo no se cumplieron. En 1971, por ejemplo, alguien afirmó que en una década íbamos a ser capaces de predecir terremotos,<sup>29</sup> algo que cuarenta años más tarde aún seguimos igual de lejos de lograr.

De hecho, el *boom* de la informática de las décadas de 1970 y 1980 provocó un descenso temporal de la productividad científica y económica, un efecto que los economistas bautizaron como la «paradoja de la productividad». «La era de la informática se aprecia en todas partes, menos en las estadísticas sobre producción», escribió el economista Robert Solow en 1987.<sup>30</sup> Entre 1969 y 1982, Estados Unidos sufrió hasta cuatro recesiones.<sup>31</sup> Los últimos años de la década de 1980 aportaron una mayor solidez a la economía de Estados Unidos, algo que no se puede afirmar de todos los países del mundo.

El progreso científico resulta más difícil de medir que el progreso económico.<sup>32</sup> Sin embargo, un indicador en este sentido es el número de patentes registradas, especialmente si se compara con la inversión en investigación y desarrollo (I+D). En principio, si producir un nuevo invento re-

sulta más barato, eso significa que estamos utilizando la información de forma inteligente y transformándola en conocimiento. Si, por el contrario, producir un nuevo invento resulta más caro, deberemos asumir que estamos confundiendo el ruido con la señal e invirtiendo nuestro tiempo en seguir pistas falsas.

Figura I.3. GASTOS EN I+D POR PATENTE PRESENTADA



Durante la década de 1960, Estados Unidos gastó aproximadamente un millón y medio de dólares (dato ajustado a la inflación)<sup>33</sup> por cada patente presentada<sup>34</sup> por un inventor estadounidense. Sin embargo, lejos de descender, esa cifra no hizo más que crecer durante los albores de la era de la información, y en 1986 alcanzó un máximo de aproximadamente tres millones de dólares, el doble que veinticinco años antes.<sup>35</sup> Finalmente adoptamos una visión más realista de lo que la nueva tecnología podía hacer por nosotros y, en la década de 1990, la productividad de las investigaciones empezó a crecer de nuevo. Nos metimos

en varios callejones sin salida, pero poco a poco los ordenadores empezaron a facilitarnos la vida cotidiana y favorecer la economía. Las historias sobre predicciones suelen ser historias de progresos a largo plazo, a los que se llega tras retrocesos a corto plazo. Muchas empresas que parecen predecibles a largo plazo frustran nuestros planes inmediatos.

## PROMESAS Y ESCOLLOS DEL *BIG DATA*

El término de moda en la actualidad es *Big Data* («macrodatos»). IBM calcula que en 2012 generamos 2,5 trillones de *bytes* de datos diarios, y que más del 90 por ciento de la información disponible se ha creado durante los últimos dos años.<sup>36</sup>

Este crecimiento exponencial de la información se ve a menudo como una especie de panacea, algo parecido a lo que sucedió con los ordenadores durante la década de 1970. En 2008, Chris Anderson, editor de la revista *Wired*, escribió que la ingente cantidad de datos terminaría por convertir las teorías e incluso el método científico en algo prescindible.<sup>37</sup>

Éste es un libro con un enfoque radicalmente favorable a la ciencia y la tecnología, y en mi opinión es un libro muy optimista. Sin embargo, la tesis que defiende es que las premisas de este tipo son extremadamente erróneas: los números no tienen forma de hablar por sí mismos. Somos nosotros quienes hablamos por ellos, quienes los imbuimos de significado. Como César, tenemos la capacidad de interpretarlos de maneras que nos resultarán favorables, pero que se separan de la realidad objetiva.

Las predicciones basadas en datos pueden acertar, pero también pueden fallar. Y cuando negamos nuestro papel en el proceso de predicción, las probabilidades de error