

PABLO RODRÍGUEZ

PRÓLOGO DE SANDY PENTLAND, PROFESOR DEL MIT MEDIA LAB



INTELIGENCIA ARTIFICIAL

CÓMO CAMBIARÁ EL MUNDO
(Y TU VIDA)

DEUSTO

Inteligencia artificial

Cómo cambiará el mundo
(y tu vida)

PABLO RODRÍGUEZ

Traducido por Jorge Rizzo



EDICIONES DEUSTO

© 2018 Pablo Rodríguez
© de la traducción Jorge Rizzo, 2018

© Centro Libros PAFP, S.L.U., 2018
Alienta es un sello editorial de Centro Libros PAFP, S. L. U.
Grupo Planeta
Av. Diagonal, 662-664
08034 Barcelona

www.planetadelibros.com

ISBN: 978-84-234-2944-8
Depósito legal: B. 11.473-2018
Primera edición: junio de 2018
Preimpresión: pleka scp
Impreso por Artes Gráficas Huertas, S.A.

Impreso en España - *Printed in Spain*

No se permite la reproducción total o parcial de este libro, ni su incorporación a un sistema informático, ni su transmisión en cualquier forma o por cualquier medio, sea éste electrónico, mecánico, por fotocopia, por grabación u otros métodos, sin el permiso previo y por escrito del editor. La infracción de los derechos mencionados puede ser constitutiva de delito contra la propiedad intelectual (Art. 270 y siguientes del Código Penal).
Diríjase a CEDRO (Centro Español de Derechos Reprográficos) si necesita fotocopiar o escanear algún fragmento de esta obra. Puede contactar con CEDRO a través de la web www.conlicencia.com o por teléfono en el 91.702.19.70 / 93.272.04.47.

Sumario

Prólogo	13
Introducción	19
1. El regreso al futuro de los macrodatos	25
¿Por qué deberías preocuparte?	25
El futuro empieza ahora	32
Un viaje al interior de los datos	37
Historia antigua de los datos	38
La emergencia de la estadística	39
Los primeros días del almacenaje moderno de datos	40
Los inicios de la inteligencia empresarial	41
La aparición de los grandes centros de datos	41
La aparición de internet	42
Las primeras ideas sobre macrodatos	42
La web 2.0 aumenta los volúmenes de datos	43
Se adopta el uso actual del término « <i>big data</i> » (macrodatos) ..	44
El poder de las redes de datos	47
Interconectando Datos: propiedades mágicas	51
Científicos de datos: los gurús de los macrodatos	54

2. El poder de las cosas y sus datos	61
La Internet de las Cosas	61
Cómo acabar con internet	66
Una distribución de datos más rápida	70
La IdC a la velocidad de la luz	78
La Internet de las Cosas y los datos	84
Seguridad para todos	92
3. Inteligencia artificial: humanos contra máquinas ..	99
IA	99
Algoritmos	108
Aprendizaje automático	111
Aprendizaje cognitivo	114
La IA te cambia la vida	120
¿Te da miedo la IA?	126
La IA y los macrodatos	131
4. La IA para el liderazgo y la creatividad	139
Liderazgo y fútbol: el colectivismo y el deporte	141
El fútbol como algoritmo de datos	150
Un laboratorio de datos en el campo	156
El ordenador en la cocina	168
El lenguaje de los sabores	177
La IA y la gastronomía	182
5. Apostarlo todo a la ciencia de datos y IA	193
Un paseo por La Boqueria	193
La IA y el turismo	197
La IA y la seguridad	204
La IA y las finanzas	208
La IA, el empleo y la movilidad	215
6. La IA en busca del bien global	223
Un día en Harvard	223

La IA y la salud mental	226
La salud física y la IA	234
La IA para el bien global	240
7. IA ética, privacidad y el alma de los datos	249
La utopía del ciberespacio	249
La privacidad de los datos	251
Un delicado equilibrio	262
Bancos de datos personales	274
Conclusión	285
Bibliografía	291
Agradecimientos	299

El regreso al futuro de los macrodatos

Cualquier loco inteligente puede hacer las cosas más grandes, más complejas y más violentas. Pero hace falta un toque de genialidad —y mucho valor— para moverse en la dirección opuesta.

E. F. SCHUMACHER

¿Por qué deberías preocuparte?

Durante la parte más productiva de mi trayectoria como científico, pasé unos años fantásticos en Cambridge, la famosa ciudad universitaria junto al río Cam de la que han salido mentes brillantes como Stephen Hawking o Isaac Newton. Cambridge desprende un aire de placidez y la sensación de que hay espacio para que las cosas maduren en su momento, algo que no he encontrado en ningún otro lugar. Es un lugar que ofrece a la mente el espacio necesario para el pensamiento profundo necesario para poder idear las grandes revoluciones tecnológicas. Y fue en Cambridge donde creció y estudió Alan Turing, el científico informático más decisivo de todos los tiempos. También es el lugar donde vive mi amiga la doctora Anastasia Christofilopoulou. Anastasia lleva cuatro años preparando una exposición sobre criptografía y el Proyecto Enigma, que ayudó a poner fin a la segunda guerra mundial y marcó el inicio de una nueva era de datos, algoritmos e inteligencia computacional nunca vista antes.

Vive con su marido Christos Gkantsidis en Ely, donde me invitan a algo tan inglés como el té en una tarde de lluvia, en

pleno otoño, para preparar nuestra reunión familiar anual en algún rincón soleado del Mediterráneo. Ely es un bonito pueblo al norte del condado de Cambridgeshire, con un fantástico paseo junto al río y un animado centro donde se celebra cada año el Festival de la Manzana para la promoción de este fantástico producto nacional. La exposición en la que está trabajando Anastasia es una muestra pionera e interdisciplinaria. *Codebreakers and Groundbreakers* unirá, por primera vez, los notables logros intelectuales y las narrativas paralelas de dos grupos de criptógrafos (*codebreakers*) que trabajaban a la vez, pero de forma independiente: los que intentaban descifrar los códigos de la segunda guerra mundial y los que descifraron el Lineal B, el primer sistema de escritura comprensible de Europa. En una exposición que incluye una máquina Enigma prestada, excepcionalmente, por el GCHQ (el Cuartel General de Comunicaciones y la Agencia de Inteligencia de El Reino Unido), y documentos únicos archivados en la Universidad de Cambridge. La exposición explica la vida y conmemora los logros de criptógrafos de la segunda guerra mundial como Alan Turing o Bill Tutte, así como de Michael Ventris y John Chadwick (también criptógrafo de la base de Bletchley), que descifraron el Lineal B. Es una muestra fascinante, pero yo lo que busco es la historia que desencadenaron los datos, los algoritmos y los descubrimientos en computación que dieron paso a la inteligencia de las máquinas.

Todo empezó en 1951, en Manchester, cuando el M16, servicio de inteligencia británico, interceptó un mensaje que le habían robado a Alan Turing. Alan, conocido profesor de Cambridge, recibió la visita de la policía, que le preguntó por el robo. Le encontraron en casa, pero él mostró una actitud desdeñosa. Los agentes declararon que era una persona insoportable y sospecharon que ocultaba algo.

Años atrás, en 1939, al declararse la guerra, 800.000 personas habían sido evacuadas de Londres. La segunda guerra mundial se había alargado seis años. Las fuerzas aliadas eran limitadas e iban

perdiendo efectivos. Estados Unidos y Canadá intentaban colaborar enviando convoyes de material estratégico.

Alan Turing, que entonces tenía veintisiete años, había llegado a Bletchley Park en el tren, escoltado por agentes de la Marina. Esperaba en el despacho del comandante Denniston. Cuando llegó el comandante, Alan tenía frío y no parecía estar de buen humor. El comandante le preguntó por qué quería trabajar para el gobierno; él respondió que no quería. Mencionó que no le interesaba mucho la política, y el comandante dijo que posiblemente aquélla sería la entrevista de trabajo más corta de la historia. Alan mencionó que no hablaba alemán, pero le dijo al comandante que era uno de los mejores matemáticos del mundo. Se planteaba los códigos alemanes como rompecabezas, que le gustaba solucionar. El comandante le pidió a su secretario que acompañara a Alan a la salida, a lo que Alan respondió con una palabra: «Enigma», dejando claro que estaba al corriente del programa de alto secreto para el que había sido escogido como candidato. Alan explicó que Enigma era el mecanismo de encriptación más grande de la historia, y que si los Aliados conseguían descifrar el código, aquello pondría fin a la guerra. El comandante le dijo que todo el mundo estaba convencido de que Enigma era indescifrable. Alan le pidió que le dejara intentarlo, para que pudieran estar seguros.

Así pues, le permitieron participar en Enigma junto a otros, como Peter Hilton, John Cairncross o Hugh Alexander. Habían conseguido hacerse con una máquina Enigma auténtica que lograron sacar de Berlín, pero no conocían la configuración necesaria para que descodificara los mensajes. Cada noche, a las doce, los alemanes cambiaban la configuración; como interceptaban el primer mensaje cada mañana a las 6.00, los criptógrafos sólo tenían dieciocho horas al día para intentar descifrar el código antes de que volviera a cambiar, cuando tendrían que volver a empezar de cero. Hugh (Matthew Goode), campeón de ajedrez, calculó que eso significaba que había 159 trillones de posibilidades al día. Alan no tenía muy claro que debieran trabajar en equipo; Stewart Menzies, el jefe del MI6, les dijo que en los últimos minutos habían muerto cuatro hombres por no

haber podido descifrar el código, y les ordenó que se pusieran manos a la obra.

Alan le dijo al equipo que todos los mensajes estaban flotando por el aire, al alcance de cualquiera que quisiera hacerse con ellos. El problema era que estaban encriptados y que había 159 trillones de posibilidades. Llevaría 20 millones de años probarlo todo. Una misión imposible. Pero Alan Turing tenía una idea brillante en mente. Estaba convencido de que reuniendo suficientes referencias de mensajes anteriores, descifrando unas cuantas claves sobre cómo se habían encriptado aquellos mensajes y usando sofisticados algoritmos de predicción irían mucho más rápido y podrían descodificar los mensajes cifrados. Si capturaban una gran cantidad de datos y llegaban a comprender cómo se encriptaban los mensajes, podrían intentar predecir lo que escribían los alemanes. ¡Sí, era el inicio de la era de los macrodatos y la inteligencia artificial!

La información ha pasado de ser escasa a superabundante. Eso genera unas enormes ventajas que influyen en nuestro día a día. Los macrodatos y la IA nos ayudan a ahorrar dinero en lo que comemos con programas de fidelidad, páginas web de reintegros y cupones de descuento, todo ello diseñado para reducir la factura semanal de la compra. Lo mismo se puede aplicar al transporte, al recreo, al ocio y a las vacaciones.

Nuestra forma de comprar *online* ha cambiado para mejor gracias a la emergencia de nuevos algoritmos de proceso de datos. La mayoría de cosas que buscamos en internet o en televisión vienen acompañadas de opiniones de cliente y de análisis más detallados de empresas de valoración de productos. Éstos nos dan una imagen mucho más clara de los productos que compramos y de los que han comprado otras personas como nosotros, y sirven para que nos puedan recomendar nuevos productos.

No son sólo las compras en internet las que se han visto modificadas con los macrodatos y la IA; es toda la red. Descubrir nueva información a través de internet nunca ha sido tan fácil. La población global de internet creció un 14,3 por ciento entre

2011 y 2013, y ahora son 3.000 millones de personas los que tienen acceso a la red.

Las redes sociales son un elemento de los macrodatos que ha observado un rápido crecimiento en los últimos años. La constante emergencia de información actualizada a través de Twitter e Instagram da una accesibilidad increíble a noticias y registros de todo el mundo. Cada minuto se envían 204 millones de emails en todo el mundo, y se cuelgan 277.000 tweets y 216.000 posts de Instagram.

Ninguno de nosotros disfruta recibiendo facturas de suministros, pero los macrodatos pueden ayudar a reducir esos costes. Es mucho más fácil rastrear y monitorizar el uso que se hace de la energía para poder configurar temporizadores de calefacción o una presión del agua que nos permita ahorrar en la factura. Los macrodatos y la IA también han simplificado la monitorización del estado del coche en cuanto a kilometraje y consumo de gasolina. Muchos coches cuentan ya con tecnología integrada, como sistemas de navegación por satélite o sensores para el aparcado. Si usas el transporte público para ir a trabajar, dispones de informaciones en tiempo real que te permiten planificar tus viajes por si hay retrasos en el servicio o carreteras o líneas de ferrocarril cortadas por obras.

Las autoridades y los servicios de emergencia usan los macrodatos para poder analizar mejor la delincuencia y los accidentes en todo momento, y la recopilación de datos es cada vez más efectiva. Sin duda, es algo que va a imponerse.

Todo lo que hacemos implica un uso de macrodatos e IA. Desde el momento en que enciendes la tele por la mañana hasta que te acuestas por la noche. Y es una tendencia que va acelerándose cada vez más. En los últimos veinte años, las tecnologías inalámbricas y la conexión a internet se han convertido en algo generalizado, asequible y disponible prácticamente para cualquiera. El uso de teléfonos móviles va aumentando exponencialmente, del 2 por ciento en el año 2000 al 28 por ciento en 2009 y el 70 por ciento en 2017. Hoy en día, personas sin educación formal y con poco que comer se conectan por telefonía móvil de un modo impensable hace sólo treinta años.

Años atrás, cuando la revolución de los *gadgets* tecnológicos empezó a arraigar, la gente solía comprarse dispositivos que mejoraran su calidad de vida —como equipos de música, cámaras fotográficas, sistemas de ocio, enciclopedias, etc.—. Comprar todas esas cosas por separado podía costar miles de dólares. Hoy en día, todo eso viene de serie en un *smartphone* que sale por una fracción de ese precio, y todo lo que no esté incluido en el precio se puede comprar en forma de *app* por menos de lo que cuesta un café.

De este modo, la tecnología se introduce rápidamente en nuestras vidas, cambiándonos para siempre en ciclos cada vez más cortos. Hoy en día tenemos que observar lo que puede suceder en los próximos quince años para comprender lo que ocurrirá dentro de cinco. Ahora mismo, una persona en Papúa Nueva Guinea tiene más posibilidades de comunicación telefónica de las que tenía el presidente de Estados Unidos hace veinticinco años. Y si tiene un *smartphone* con acceso a internet, goza de un mejor acceso a la información del que tenía el presidente hace sólo quince años. En otras palabras, vivimos en un mundo de rápida aceleración de la tecnología y los macrodatos.

En los próximos años presenciaremos el desarrollo de nuevas tecnologías transformadoras —sistemas de computación, computación cuántica, realidad virtual, redes de siguiente generación, inteligencia artificial, redes neuronales, robótica, interfaces de usuario, comprensión de la mente, salud predictiva y ordenadores que igualan la capacidad bruta de procesamiento del cerebro—, lo cual muy pronto hará posible que la gran mayoría de la humanidad experimente muchas cosas que hoy aún están limitadas al reducido grupo de los que tienen una capacidad inmensa de procesamiento de datos.

Lo que significa todo esto es que, por impresionante que resulte la velocidad del cambio tecnológico mundial registrado hasta el momento, aún no hemos visto nada. Cuando todas estas tecnologías (que siguen avanzando a un ritmo exponencial) se combinen, se harán progresos mayores que nunca. El índice actual del progreso tecnológico es más que suficiente para resolver numerosos retos sociales para todos. Esto, combinado con los

macrodatos y la IA, significa que muchas preguntas que hasta ahora permanecían sin responder por fin tendrán respuesta.

Los datos y la IA existen prácticamente en todo lo que usamos y hacemos. Están en teléfonos, coches, carreteras, líneas eléctricas, cursos de agua, contenedores de comida y una cantidad innumerable de otros elementos que nunca habríamos identificado con los ordenadores.

Éstos dicen mucho sobre nuestra conducta colectiva y nuestra sociedad, lo que nos permite hacer cosas increíbles. Por ejemplo, podremos usar los macrodatos y la IA para predecir con suficiente antelación la escasez de alimentos. Variables como los precios de mercado, las sequías, las migraciones, la producción regional anterior o las variaciones estacionales influyen en esta clasificación y en el modelo de aprendizaje de la estructura causal para predecir la probabilidad de que un habitante del entorno rural tenga dificultades para obtener alimentos. Podremos usar los macrodatos y la IA para comprender las trampas de pobreza y cómo salir de ellas, o cuantificar una ola de delincuencia antes de que se extienda en determinando el lugar, el momento y la naturaleza del delito, o decidir qué intervención o serie de intervenciones es más conveniente para mejorar el aprendizaje en las escuelas de países en desarrollo.

Ahora que tenemos todo esto al alcance de la mano, hay una sencilla comprobación que suelo usar: ¿Estás trabajando en algo que pueda mejorar el mundo? ¿Sí o no? El 99,99999 por ciento de la gente responde «no». Yo creo que tenemos que empezar a pensar más en cómo cambiar el mundo y en cómo nos puede ayudar a hacerlo el «alma de nuestros datos y la IA». Se nos presentan numerosas oportunidades para conseguir un gran impacto social, para transformar sociedades, para resolver problemas arraigados como la pobreza, el acceso a la educación o la escasez de comida, eliminando barreras y convicciones que ya no son ciertas (o que van desapareciendo). Así se allanará el camino a las nuevas oportunidades que pueden cambiar el mundo.

El futuro empieza ahora

Por si no lo sabías, si vuelas con frecuencia como yo, sin duda has sido alcanzado por algún rayo mientras volabas. La energía del rayo no atraviesa la cabina, electrocutando a los pasajeros; nueve de cada diez veces se descarga al exterior, sin dejar apenas rastro de la descarga. La estructura del fuselaje y el potente aislamiento actúan como un pararrayos superconductor, canalizando el rayo por el exterior, lejos de los clientes y la tripulación y descargándolo en el suelo a través del tren de aterrizaje.

No obstante, el rayo deja en los aviones pequeñísimos agujeros que cuestan mucho dinero reparar, ya que requieren un meticuloso examen del avión, lo que retrasa la puesta en servicio del mismo. Los rayos que impactan en aviones suelen costar millones a las compañías aéreas, en detección y en reparación. Lo que suelen hacer es recurrir a técnicos expertos, entrenados para encontrar agujeros de 1 centímetro provocados por los rayos en determinados puntos del avión. No obstante, para ello hay que formar primero a estos expertos, ya que podrían pasar cosas por alto, o pueden cometer errores de identificación a causa de la fatiga, en particular cuando un avión llega a última hora de la noche y hay que ponerlo otra vez en servicio lo antes posible.

Actualmente, las aerolíneas han empezado a experimentar con drones que sobrevuelan el avión una vez ha aterrizado. Estos drones usan sofisticados mecanismos de captura de datos de imagen que se transmiten por internet a servidores en la nube para identificar los problemas usando un complejo software de inteligencia artificial. Los drones se comunican a través de la red y envían registros de imagen muy precisos con los que se pueden encontrar los rastros del impacto de los rayos rápidamente, para poder repararlos enseguida, y cometen muchos menos errores que los humanos.

Esto no es más que una mínima muestra de lo que nos espera en el futuro, un futuro que se acelera exponencialmente gracias a los grandes avances tecnológicos que van registrándose, y que cambiarán el mundo gracias a los datos y a la IA.

Los macrodatos y la IA se usan para analizar conjuntos de referencias tan grandes o complejos que no permiten el uso de

las técnicas tradicionales de proceso de datos. Entre los desafíos que ello plantea está el análisis, la captura, el tratamiento de los anteriores, la investigación, la puesta en común, el almacenamiento, la transferencia, la visualización, la consulta, la actualización y la protección de la privacidad de la información.

Con el análisis de conjuntos de macrodatos se pueden encontrar nuevas correlaciones para detectar tendencias de negocio, para prevenir enfermedades, para combatir la delincuencia y mucho más. Ejecutivos, personal médico, publicistas y gobernantes se encuentran periódicamente con dificultades al tratar grandes conjuntos de registros en campos como la búsqueda en internet, las finanzas, la informática urbana o la informática de los negocios. Los conjuntos de datos crecen rápidamente en parte porque cada vez son más baratos y numerosos los detectores móviles, ficheros de registro, cámaras, lectores de identificación por radiofrecuencia (RFID) y redes de sensores *wireless*.

Hace diez años los científicos que analizaban los macrodatos eran una especie rara; hoy en día los datos y la IA son un activo de la mayoría de nuestras empresas y de la sociedad, lo que convierte a los analistas de *big data* y expertos en IA en los profesionales más solicitados en el mundo laboral. Y con el rápido avance de la inteligencia artificial es posible que pronto incluso muchos de ellos se vean reemplazados por máquinas que extraerán conclusiones automáticamente a partir de los datos y la IA. La velocidad a la que están cambiando las cosas es simplemente alucinante.

La capacidad de almacenaje de información mundial per cápita prácticamente se ha ido duplicando cada 40 meses desde los años 1980: en 2012, cada día se generaban 2,5 exabytes ($2,5 \times 10^{18}$) de datos. Sólo en el espacio de referencias de salud esperamos un aumento global del 48 por ciento anual. De hecho, en 2013 el volumen de datos sanitarios era de 153 exabytes. Con la tendencia de crecimiento prevista, para 2020 esa cifra aumentará hasta los 2.314 exabytes.

Para hacernos una idea, si tuviéramos que guardar todos éstos en un montón de tabletas, con las cifras de 2013 ese montón tendría casi 8.800 kilómetros de altura. Siete años más tarde, esa torre crecería hasta superar los 131.000 kilómetros de altura, llevándonos a más de un tercio de la distancia a la Luna.

Y la consecuencia de esto es que el exceso de datos, de macrodatos no procesados, puede tener el mismo impacto que un exceso de tráfico en una ciudad. Con el aumento del número de vehículos en las calles, las grandes áreas metropolitanas se van congestionando. Los atascos se convierten en un problema frecuente, y muchos trabajadores se enfrentan a la posibilidad de no poder llegar a tiempo a sus puestos de trabajo. Del mismo modo, un exceso de registros no procesados puede provocar una pérdida de oportunidades y puntos ciegos en nuestras empresas, o en nuestra vida diaria, que no nos permiten llegar donde queremos llegar.

Básicamente los macrodatos son datos que exceden la capacidad de procesamiento de los sistemas de datos convencionales. Éstos son demasiado grandes, se mueven demasiado rápido o superan las limitaciones de los ordenadores tradicionales. Para sacar un beneficio a estas referencias hay que escoger un modo alternativo de procesarlos.

El uso de *big data*, término de moda en 2012, se ha hecho viable gracias a la aparición de nuevos mecanismos capaces de «domar» el volumen, la velocidad y la variabilidad de estos datos enormes a un precio asequible. En el interior de éstos encontramos patrones e información de gran valor que antes quedaban fuera de nuestro alcance por la cantidad de trabajo necesaria para extraerlos. Las grandes corporaciones como Walmart o Google disponen de este poder desde hace tiempo, pero a un precio enorme. Hoy en día el hardware genérico, el acceso a la nube y el software de código abierto ponen el procesamiento de los macrodatos al alcance de los que menos recursos tienen. Ahora mismo pueden utilizarlos hasta una pequeña empresa emergente creada en cualquier garaje, que por poco dinero puede comprar tiempo de proceso en la nube.

Llamamos macrodatos, o *big data*, a la siempre creciente capacidad que tenemos de recopilar referencias de multitud de fuentes y analizarlos para sacar conclusiones usando algoritmos informáticos avanzados. Hay patrones que los humanos no podemos ver, y que nos permiten comprender mejor las situaciones y las soluciones a los problemas. Los desastres naturales son situaciones tremendas, enrevesadas y complicadas, justo las condicio-

nes perfectas en las que los macrodatos nos pueden ayudar a interpretar el caos. Las enormes cantidades de datos que generamos con los teléfonos móviles, los satélites y las redes sociales pueden ayudarnos a encontrar pistas sobre cómo responder a una situación de la mejor manera posible.

Gran parte del trabajo de desarrollo de los sistemas de *big data* para contribuir a aliviar los efectos de los desastres empezó tras el terremoto de Haití de 2010 y el terremoto de Tohoku (Japón) de 2011 y el posterior tsunami. Japón y Estados Unidos lanzaron un programa de investigación conjunto para reducir los efectos de los desastres naturales que matan a miles de personas cada año y que cuestan miles de millones a la economía global. El año pasado, la National Science Foundation de Estados Unidos y la Agencia de Ciencia y Tecnología japonesa ofrecieron 2 millones de dólares para financiar grupos que trabajaran con soluciones basadas en el uso de datos para afrontar la gestión de desastres.

En el otro extremo de la escala están las iniciativas de *crowdsourcing* creadas desde cero, en que los miembros de la comunidad se unen para ayudar a otros a través de la recopilación y el cotejo de registros. Esto es lo que ocurrió tras el paso del huracán Sandy por Estados Unidos, cuando la colaboración entre estudiantes de instituto permitió crear un mapa online de la zona de Nueva York y Nueva Jersey que mostrase dónde quedaba gasolina.

Tras el paso del tifón Haiyan por las Filipinas, la Cruz Roja internacional y UNICEF colaboraron con voluntarios de todo el mundo para trazar un mapa de los efectos sobre la región y su población. Los cuatro elementos clave de la gestión de desastres son la prevención, la preparación, la respuesta y la recuperación. Los macrodatos ofrecen la posibilidad de trabajar en los cuatro.

Existen muchos ejemplos más de situaciones y modos en que se pueden usar los macrodatos:

- Tu teléfono tiene acceso a un montón de resultados de macrodatos a través de *apps* que permiten orientarte con

mapas, recibir recomendaciones musicales personalizadas de Spotify, comprobar el tiempo que hará mañana, hablar con Siri y otras muchas actividades diarias.

- El filtro de spam de tu gestor de correo es un ejemplo clásico de análisis de macrodatos para determinar qué es spam y qué no lo es, y luego aprende y adapta su inteligencia a tus necesidades al ir adquiriendo experiencia sobre tus preferencias personales. El termostato Nest hace lo mismo, ya que va adaptándose a tu rutina habitual y actúa en consecuencia.
- Cuando buscas un producto en Amazon y luego lo ves como producto «recomendado para ti» en otro sitio web, o cuando Facebook sabe de pronto a quién puede interesarte añadir a tu lista de amigos, están en acción los macrodatos y la inteligencia artificial, trabajando de forma combinada para que las webs intercambien información de lo que te gusta y luego usar esa información para decidir qué mostrarte. Lo mismo ocurre con la sugerencia de Amazon de «Los que han comprado esto también han comprado...».
- Cuando tu avión aterriza, no es un ser humano el que decide a qué puerta debe ir. Ni es una persona la que ha decidido el precio de tu billete; han sido una serie de cálculos con macrodatos.
- Los macrodatos y las búsquedas de internet pueden servir para predecir pandemias de gripe o prever la actividad de la gripe en todo el mundo. El objetivo es intentar hacer predicciones precisas sobre la actividad de la gripe, que pueden usarse posteriormente para crear modelos que prevean la extensión espacial y temporal de una enfermedad.
- El buscador de Google es un cerebro gigante con unos métodos increíblemente sofisticados para establecer una clasificación de las páginas y decidir qué enseñarte en particular. Lo mismo ocurre con el *feed* de Noticias de Facebook.

Y eso no son más que ejemplos del mundo del consumidor. Hoy en día se usan sofisticados sistemas de *big data* en numerosos sectores e industrias como las telecomunicaciones, la banca, los bufetes de abogados, el ejército, las fábricas y las finanzas

(los operadores financieros algorítmicos de alta frecuencia mueven más de la mitad del capital social de los mercados de Estados Unidos), y en los sistemas de expertos como los que ayudan a los médicos a dar diagnósticos, o el famoso Watson de IBM, que contenía suficientes datos como para vencer a los grandes campeones del concurso de televisión *Jeopardy* sin problemas.

Pero los macrodatos también tienen limitaciones. «Te dicen lo que ocurrió en el pasado (o hace un momento), pero no siempre saben lo que ocurrirá en el futuro» (Mayer-Schönberger 2013). Así, se trate de la previsión meteorológica, de terremotos, del calentamiento global, de los resultados de fútbol, de los créditos *subprime* o de la crisis económica mundial, en muchos casos nos cuesta convertir los datos de ayer en predicciones por las que apostar. No obstante, sí podemos convertir el alma de éstos en nueva información con la que aprender y usar mejor nuestra intuición no para acertar siempre, pero sí para equivocarnos menos.

Un viaje al interior de los datos

La era de los macrodatos es una época en la que somos capaces de analizar las informaciones que proporcionan todo tipo de dispositivos conectados a internet (teléfonos, ordenadores, máquinas, coches, etc.) y convertirlos en predicciones y conocimiento procesable, lo que nos convierte en una «inteligencia aumentada». Este nuevo paradigma del conocimiento y el automatismo resultante se encuentran en el sustrato de la transformación económica que se avecina.

Todos estos cambios tecnológicos alterarán la existencia humana de un modo muy particular. Viviremos de un modo mucho más eficiente, y utilizaremos mejor los recursos gracias a los datos, pero sin dejar que ellos nos dominen a nosotros.

La historia de los macrodatos no es larga, pero gran parte de los cimientos sobre los que se apoya se pusieron hace mucho tiempo. Mucho antes de que los ordenadores (tal como los conocemos ahora) estuvieran tan extendidos, entre los estudiosos ya se fraguaba la idea de que estábamos creando un cuerpo de conocimientos cada vez mayor susceptible de ser analizado.

Quizá no siempre lo recordemos, pero nuestra creciente capacidad de almacenar y analizar información es producto de una evolución gradual —aunque desde luego las cosas se aceleraron a finales del siglo pasado, con la invención del almacenamiento digital e internet.

Ahora que sabemos que los macrodatos están destinados a ser objeto de uso común —tal como dijo Bernard Marr en el Foro Económico Mundial (*Breve historia del Big Data*, WeForum 2015)—, echemos un vistazo a la larga historia de pensamiento e innovación que nos ha llevado hasta el nacimiento de la era de los datos, según describe Bernard cronológicamente en dicho foro.

Historia antigua de los datos

- c. 18000 a.C.** Los primeros ejemplos que tenemos de almacenaje y análisis de datos por parte de seres humanos están en los palos de cómputo. El hueso Ishango fue descubierto en 1960 en la actual Uganda, y está considerado una de las primeras pruebas de almacenaje de éstos en la prehistoria. Las tribus paleolíticas hacían muescas en palos o huesos para registrar la actividad comercial o las provisiones. Luego comparaban palos y muescas para realizar cálculos rudimentarios, lo que les permitía hacer predicciones como cuánto les durarían las provisiones de alimentos.
- c. 2400 a.C.** Empieza a usarse en Babilonia el ábaco, primer dispositivo construido específicamente para realizar cálculos. Fue también en esta época cuando empezaron a aparecer las primeras bibliotecas, lo que representa nuestros primeros intentos de almacenaje de registros masivos.
- 300 a.C.-48 d.C.** La Biblioteca de Alejandría es quizá la mayor colección de datos del mundo antiguo: contendría quizá medio millón de pergaminos y todo el conocimiento que habríamos adquirido hasta el momento, prácticamente sobre cualquier tema. Desgraciadamente, se cree que los romanos la destruyeron en el año 48 d.C., durante su invasión, quizá de

forma accidental. En contra de lo que se suele pensar, no todo se perdió: una parte significativa de las colecciones de la biblioteca se trasladó a otros edificios de la ciudad, fue robada o se dispersó por el mundo antiguo.

- c. 100-200 d.C.** Creación del Mecanismo de Anticitera, el computador analógico más antiguo descubierto nunca, probablemente por científicos griegos. Su «CPU» está formada por 30 ruedas dentadas de bronce interconectadas y se cree que fue diseñado para realizar mediciones astrológicas y para predecir el ciclo de los Juegos Olímpicos. Su diseño sugiere que probablemente fuera una versión evolucionada de un mecanismo más antiguo, pero si es así, hasta ahora no se ha descubierto su precursor.

La emergencia de la estadística

- 1663** En Londres, John Graunt realiza el primer experimento de análisis estadístico de datos del que tenemos constancia. Registrando datos sobre mortalidad, teorizó que podía diseñar un sistema de advertencia precoz de la peste bubónica que asolaba Europa.
- 1865** Richard Millar Devens usa el término «inteligencia empresarial» en su *Enciclopedia de Anécdotas Comerciales y Empresariales*, al describir cómo adquirió ventaja sobre sus competidores, el banquero Henry Furnese, que recopiló y analizó de un modo estructurado la información relacionada con sus actividades empresariales, en lo que se cree es el primer caso de una empresa que aplica el análisis de informaciones con fines comerciales.
- 1880** La Oficina del Censo de Estados Unidos tiene un problema. Calcula que tardará ocho años en organizar todos los datos recogidos en el censo de 1880, y prevé que organizar los generados por el censo de 1890 llevará más de diez años. Eso significa que no estarán disponibles hasta después de la elaboración del censo de 1900, que lo habrá dejado obsoleto. En 1881, un joven ingeniero empleado en la Oficina, Herman

Hollerith, crea lo que se dará en llamar el Tabulador Hollerith. Usando tarjetas perforadas, reduce el trabajo de diez años a tres meses y se hace un lugar en la historia como padre de la computación automatizada moderna. La empresa que funda más tarde se llamará IBM.

Los primeros días del almacenaje moderno de datos

1926 En una entrevista publicada en la revista *Colliers*, Nikola Tesla afirma que, cuando la tecnología sin hilos «se aplique perfectamente, toda la Tierra se convertirá en un enorme cerebro, que de hecho es, ya que todas las cosas son partículas de un todo real y rítmico [...] y los instrumentos con los que podremos hacer esto serán sorprendentemente simples en comparación con nuestro teléfono actual. Podrán llevarse en el bolsillo del chaleco».

1928 Fritz Pfleumer, ingeniero germanoaustríaco, inventa un método para almacenar información magnéticamente en una cinta. Los principios que desarrolla aún se usan hoy en día, ya que la gran mayoría de datos digitales se almacenan magnéticamente en discos duros de ordenador.

1944 Fremont Rider, bibliotecario en la Wesleyan University de Connecticut (Estados Unidos), publica un artículo titulado «The Scholar and the Future of the Research Library» («El académico y el futuro de la biblioteca de investigación»).

En uno de los primeros intentos por cuantificar la cantidad de información que se produce, observa que para guardar todas las obras populares de valor que se crean, las bibliotecas estadounidenses tendrían que duplicar su capacidad cada dieciseis años. Eso le llevó a especular que la biblioteca de la Universidad de Yale, en 2040, contendría 200 millones de libros, distribuidos por 9.600 kilómetros de estanterías.